Performance Management in Local Government: The Application of System Dynamics to Promote Data Use

Carmine Bianchi
Professor of Business and Public Management
Department of European Studies
University of Palermo, Italy
Phone: +39091587400
carmine.bianchi@unipa.it

William C. Rivenbark
Professor and Director of MPA Program
School of Government
University of North Carolina at Chapel Hill
Phone: 919-962-3707
rivenbark@unc.edu

Abstract

One possible avenue to enhance performance management in the public sector is the application of system dynamics, where modeling organizational systems and using simulation techniques are used for understanding the behavior of complex systems. The purpose of this paper is to demonstrate how system dynamics can be used to enrich performance management in local government, focusing specifically on how the development of conceptual and simulation system dynamic models can foster a common shared view of the relevant system among stakeholders to overcome factors that limit data use. Responding to this purpose, we present a normative case study on how key drivers can be used to foster a shared view of the residential refuse collection system for supporting policy and process changes. A major finding from our research, however, is that performance management cannot overlook the broader forces of citizenship outcomes as described by Wichowsky and Moynihan (2008).
Introduction

While some scholars have suggested that the forces of New Public Management are now in decline (Hughes, 2003), a clear success story of this reform is the use of performance measurement in the public sector for tracking the outputs and outcomes of service delivery. In fact, research has demonstrated that well-managed performance measurement systems are critical for accurately and systematically demonstrating operational accountability in governmental organizations (Rivenbark, 2007). However, we must be cautious in how we think about the use of performance measurement systems because of their two distinct parts as described by de Lancer Jules and Holzer (2001). The first part involves adoption, where public officials develop performance measures, track them over time, and report them on a periodic basis. The second part involves implementation, where public officials actually use performance information to make policy and process changes for improving service delivery. The problem is that success is clearly more associated with adoption rather than implementation.

Ammons and Rivenbark (2008) addressed this issue by studying the patterns of implementation from fourteen municipalities associated with the North Carolina Benchmarking Project. While the authors concluded that the record of these municipalities actually using performance data remains modest, certain factors did emerge that promoted the move from adoption to implementation of performance measurement systems. They included the focus on the higher-order measures of efficiency and effectiveness, the willingness to benchmark against other organizations, and the need to imbed performance measures within other management systems. A more recent study also suggests that managerial involvement in the Government Performance and Results Act (GPRA) of 1993 and the Program Assessment Tool (PART) has produced relatively few aspects of performance information use in federal agencies (Moynihan
and Lavertu, 2012). This research, however, identified a number of organizational factors that increase the likelihood of using performance data, including leadership commitment to results (Behn, 1991; Kotter and Heskett, 1992; Moynihan and Ingraham, 2004), learning routines led by supervisors, and the ability to link measures to actions.

One possible avenue to enhance performance management in the public sector— which is the term used for the implementation of performance measurement as described by de Lancer Jules and Holzer (2001)— is the application of system dynamics, where modeling organizational systems and using simulation techniques are used for understanding the behavior of complex systems. This line of inquiry builds on the research of Ghaffarzadegan, Lyneis, and Richardson (2011), Andersen et al, (2009), and Kim at al. (2013) where small system dynamics models were used to enhance public policy and decision-making. The advantage of using this approach is placing performance measures into the broader context of the relevant system (in respect to observed behaviors), responding to the reality that even simple policy and process changes to impact specific outputs and outcomes are not likely to be that “simple” in organizations (Bianchi, Winch, and Tomaselli, 2008). There also is another possible advantage to the approach. Rather than looking for factors that promote data use, system dynamics may give insights to factors that prevent data use.

The purpose of this paper is to demonstrate how system dynamics can be used to enrich performance management in local government, focusing specifically on how the development of conceptual and simulation system dynamics models can foster a common shared view of the relevant system among stakeholders to overcome factors that limit data use. We begin this paper by describing the background of a residential refuse collection program in a North Carolina municipality, including how specific performance measures were used to make changes in
service delivery. After discussing the methodology of using system dynamics to enhance performance management, we present our case on how key drivers can be used to foster a shared view of the residential refuse collection system for supporting policy and process changes. In other words, our goal is to show how a systems approach can help public officials move from performance adoption to performance implementation. However, a major outcome of our research is that the practice of performance management must take into account citizenship outcomes as described by Wichowsky and Moynihan (2008) before certain policy decisions can be made on the grounds of service efficiency and effectiveness. We conclude this paper by identifying several possibilities of how system dynamics can be used to improve the quality of performance management in local government and the aptitude of policy makers to link outputs with outcomes.

**Background**

The city of Winston-Salem is located in the central-part of North Carolina, with a current population of 230,345. The city, which operates under the council-manager form of government, embraced performance measurement in the 1980s as part of an effort to expand operational accountability of service delivery. The budget director of the city, however, wanted to take advantage of benchmarking by comparing the city’s performance and cost data against other municipalities in North Carolina. At the 1994 winter conference of the North Carolina Local Government Budget Association, the budget direct recommended that members of the association form a consortium to develop a process for comparing service performance and cost data on a systematic basis (Jones, 1997). Another meeting was held in 1995—involving representatives from several of the state’s larger municipalities along with staff members of the
School of Government—to create the North Carolina Benchmarking Project (Rivenbark, 2001). This initial group also made the major decision to use a cost accounting model to capture the full-cost of service delivery.

The first performance and cost data report was published by the School of Government in 1997, comparing data from seven of the larger municipalities in the state (Asheville, Cary, Durham, Greensboro, Raleigh, Wilmington, and Winston-Salem) across seven service areas (residential refuse collection, household recycling, yard waste, police patrol, police investigations, emergency communications, and street maintenance). The budget director’s desire to compare performance and cost data from Winston-Salem against other municipalities immediately paid dividends. The first report showed that tons collected per FTE (full-time equivalent) in the service area of residential refuse collection was 307.2, which was much lower than the overall average of 687.6 (Few and Vogt, 1997). One explanation of this outcome is that Winston-Salem was one of three municipalities participating in the benchmarking project that used backdoor garbage pickup rather than curbside collection. However, when Winston-Salem compared itself only to backdoor providers, the city was still behind in tons collected per FTE (Ammons, 2000).

During the FY 1997–1998 annual budget process, staff members from the budget department and the sanitation division met with the city manager to discuss options for closing this performance gap in the service area of residential refuse collection. The decision was made not to renew a private contract for garbage collection in an annexed area of the city and to adjust current city routes to provide coverage within existing resources. After adjusting for the additional tipping fees at the landfill, this policy change increased tons collected per FTE by approximately 30 percent and produced an annual net savings of $395,000 (Rivenbark, Ammons,
and Roenigk, 2005). While this was deemed a major success story of the North Carolina Benchmarking Project, the performance and cost data report for FY 2000–2001 showed that Winston-Salem remained extremely inefficient when compared to the other participating municipalities in the service area of residential refuse collection. One reason for this outcome is that several of the other municipalities had embraced technology to increase tons collected per FTE and to decrease cost per ton collected and cost per collection point (Rivenbark and Pizzarella, 2002).

In April 2004, the city faced a budget deficit of approximately $8,000,000 for FY 2004–2005. Using information from the FY 2002–2003 performance and cost data report, budget staff members presented information to the city council that showed Winston-Salem’s residential refuse collection continued to be the most inefficient operation as compared to the other fourteen municipalities now participating in the benchmarking project. The analysis included that the city could save approximately $450,000 per year by moving from backyard collection to a voluntary curbside collection program and could save approximately $1,800,000 per year by moving to a fully automated curbside collection project. Before this presentation was made, the budget staff members had already presented their analysis to the city’s citizen efficiency review committee and conducted a web-based survey to garner support for curbside collection. The city council, however, decided to approve voluntary curbside collection, with approximately 15,500 households agreeing to participate in the new program (Rivenbark, Ammons, and Roenigk, 2005).

The city’s FY 2010–2011 adopted budget contained nine strategic budget objectives, which were adopted by council several years prior, to guide the preparation and adoption of the annual budget. One of the nine objectives addresses economy operations, which mandates that
city operations should be continuously reviewed to create cost savings for the taxpayers. The problem is that city’s residential refuse collection program remains extremely inefficient—even with the change to the voluntary curbside collection program. The city’s cost per ton of $136 and cost per collection point of $90 remain above the fourteen municipality averages of $98 and $84, respectfully, as shown in the FY 2010–2011 performance and cost data report (Roenigk, 2012). The report also shows tons collected per FTE of 524 for Winston-Salem as compared to the fourteen municipality average of 1,406.

While the city finally made the decision to mandate curbside collection in October 2010 for implementation during the following fiscal year, the city has possessed the trend and benchmark data to clearly justify this transition since 1997. The question is what prevented the city from making this transition in 2004 that would have produced an annual savings of $1,800,000 or 22.5 percent of the budget deficit faced during that time? The performance and cost data had been collected and analyzed from an internal and external perspective, the city’s citizen efficiency review committee had been consulted, and a web-based survey had shown that citizens were open to this policy change. One possibility is that an understanding of the complete system was not shared among all stakeholders involved, which prevented the performance and cost data from being used to drive change.

Methodology

One approach to overcome the myopic view of relying on a handful of performance indicators to facilitate change is system dynamics, which is adopted to map system structure to capture and communicate an understanding of behavior driving processes and the quantification of the relationships to produce a set of equations that form the basis for simulating possible system
behaviors over time. The underlying principle is that if process structure determines system behavior, and system behavior determines organization performance, then the key to developing sustainable strategies to maximize performance is acknowledging the relationship between processes and behaviors and managing the leverage points. It is possible to identify two converging streams of research regarding the application of system dynamics to performance management.

A resource-based view of performance management

A first research stream can be associated to the resource-based view of organizations, where decisions aimed to affect organizational performance focus on strategic resources (Wernerfelt, 1984; Mahoney and Pandian, 1992). According to the dynamic resource-based view (Morecroft, 2007; Warren, 2002), strategic assets are modeled as stocks of available tangible or intangible factors in a given time. Their dynamics depend on the value of corresponding inflows and outflows. Such flows are modeled as “valves” on which decision makers can act through their policies, in order to influence the dynamics of each strategic resource, and therefore—through them—performance indicators (Bianchi, 2010).

Earlier studies adopting a dynamic resource-based view of organizations (Bianchi and Winch, 2005; 2009) have confirmed that the management of strategic resources, and more specifically the maintenance of an appropriate balance between such assets, is the key to sustainable development. The emerging models all center on the building up and decline of key core assets, including workers, equipment, population, workload, perceived service quality, and financial resources. Each of the strategic resources can to some extent be controlled in isolation of the others; however, where there is not balanced growth or coherence in the assets, then
organizations or territorial areas where they operate, will likely be unable to grow to achieve their own potential, or might grow in a non-sustainable way. The two common features in strategic resource management as shown in figure 1 are the requirement for consistency between strategic assets and the need to actively manage each strategic asset to maintain balance.

[Figure 1 about here]

Therefore, each strategic resource should provide the basis to sustain and foster others in the same system. For instance, both workers and equipment provide capacity, which affects perceived service quality. This affects territory attractiveness, which, in turn, influences population dynamics. A change in the population that a municipality must serve will affect workload and perhaps the stock of available financial resources, and eventually capacity and service. The feedback loops underlying the dynamics of the different strategic resources imply that the flows affecting such resources are measured over a time lag. Therefore, understanding how delays influence strategic resources and achieved results becomes a key-issue to manage performance in dynamic complex systems.

Another key-issue suggested by a dynamic resource-based view of performance is the need to adopt a broad enough perspective in order to understand the driving forces affecting achieved results. This implies that the number and range of stakeholders involved in making decisions influencing strategic resource dynamics—and, therefore, the relevant system’s performance—are often located in several organizational units and institutions in a given territorial area. Such implication is particularly relevant when performance management concerns the outcomes generated by public policies and the sustainability of performance output indicators measured in the short-run.
A dynamic view of performance management

A second research stream, focused on dynamic performance management, takes its own premises from the literature that has demonstrated the lack of relevance of conventional financially-focused planning & control systems (Johnson and Kaplan 1987; Kaplan and Norton 1996). Such systems are no longer able to provide information that can support: dynamic complex management, measurement of intangibles, detection of delays, understanding linkages between short- and long-term, and setting proper system boundaries in strategic planning. To cope with such problems, the balanced scorecard (BSC) has been used by many organizations. The two main concepts underlying the BSC framework can be synthesized as follows: Organizational performance cannot be managed by only focusing on end-results and performance cannot be measured only in terms of finance. It also must include the customer, the process, and the learning & growth quadrants.

However, even the BSC presents some conceptual and structural shortcomings (Linard et al., 2002; Sloper et al., 1999). In particular, it does not support an understanding of how end-results can be affected by performance drivers, how performance drivers can, in turn, be affected by the use of policy levers aimed to influence strategic resource accumulation and depletion processes, and how the flows of strategic assets are affected by end-results. In order to provide decision makers with proper lenses to interpret such phenomena, to understand the feedback structure underlying performance, and to identify alternative strategies to adopt to change the structure for performance improvement, system dynamics modeling has been used (Kaplan and Norton 1996; Richmond 2001; Ritchie-Dunham 2001; Bianchi, 2012; Bianchi & Montemaggiore 2008).
Figure 2 illustrates how the end-results provide an endogenous source in an organization to the accumulation and depletion processes affecting strategic resources. In fact, they can be modeled as *in-* or *out-*flows, which change over a given time span the corresponding stocks of strategic resources, as a result of actions implemented by decision makers. For instance, liquidity (strategic resource) may change as an effect of cash flows (end-result); image and credibility of an organization towards citizens (strategic resource) may change as an effect of their satisfaction (end-result). There also are interdependencies between different strategic resources: image may affect the capability of an organization to get funds from different stakeholders. Furthermore, both image and financial resources may affect its capability to recruit skilled human resources and keep them.

Organizational growth can be sustainable if the rate at which end-results change the endowment of corresponding strategic resources is balanced. This implies that management is able to increase the mix of strategic resources and this increase is not obtained by reducing the endowment of the wider strategic resources in the territory. End-results can be measured over a sequential chain and positioned on several layers. “Last layer” end-results are those changing the endowment of strategic resources that cannot be purchased in the market. To affect the results positioned on this “last layer,” further layers must be identified. For example, cash flows can be affected by the current income and net working capital flows. These more detailed financial measures are, in turn, affected by non-monetary, end-results. So, activity volumes affect revenues and the net working capital flow. They also affect purchase volumes, which impact on purchase costs and (through purchases on credit and the change in inventory) on the net working
capital flow. Therefore, activity volumes can be located on a first layer of end-results. Such results can be affected through performance drivers.

*Competitive performance* drivers are associated to critical success factors in the competitive system. They can be measured in relative terms—as a ratio between the organizational performance perceived by citizens and a benchmark—or target. Such a denominator must be gauged in relation to perceived past performance, users’ expectations, or even (if relevant) competitors’ performance. Also *social performance drivers* can be measured in terms of ratios between organizational strategic assets and a target, which can mostly be expressed in terms of either stakeholder’s expectations or perceived past organizational performance. For example, a social performance driver could be referred to the ratio between the actual and planned number of perceived undertaken social initiatives.

*Financial performance drivers* also must be measured in relative terms. For instance, the debts-to-total investments ratio often affects the change in company solvency perceived by funders. Such driver is the ratio between two stocks. Efficiency measures affecting operational costs can be gauged in terms of ratios as well. For example, the employee’s time per unit of workload is an expression of the ratio between two stocks—employees (unit of measure: persons) and workload (unit of measure: widgets per week), multiplied by a constant (working hours per people per week).

**The Winston-Salem Case**

A major difficulty in developing models is the lack of explicit data that can support the identification of possible system structures that might explain recorded performance data. We started by modeling population and collection points of the residential refuse collection system
for the city of Winston-Salem, where patterns over the period 1997–2010 were well defined by available benchmarking records. The decision to start the analysis from such variables also was associated to the fact that an increase in the collection points determines an increase in the workload and—other conditions being equal—an increase in collection costs. However, a decrease in the quality of garbage collection services or an increase in taxation—due to higher garbage collection costs—may dampen the city’s attractiveness, and therefore may contribute to increase (though with a delay) population outflows as shown in figure 3.

[Figure 3 about here]

Both population and collection points were identified as strategic resources in the system because they correspond to factors that primarily affect the performance of the sanitation division and the city as a whole. The dynamics of collection points were modeled through co-flows depending on population accumulation and depletion rates and on an input parameter, which is called “average household size.” This parameter was estimated based on a ratio between population and collection points using historical data. The upper section of figure 3 highlights how an outcome perspective should characterize the sanitation division’s policies. In planning capacity and service provision, setting performance standards and evaluating results, the division should interact with other divisions in the same municipality with a view to understand different factors impacting workload (total garbage to collect) and the driving forces affecting population flows. Relevant to this point, though not explicitly linked to sanitation services, are policies regarding: housing, infrastructure, education, health care, enterprise services, and public safety.

We now must explore the factors impacting the outflow depleting the “total garbage to collect” stock (tons collected per time) to explain how the sanitation division’s performance affects population. Such factors will affect service efficiency (municipal financial needs) and
effectiveness (the perceived quality of provided service). Figure 4 shows how tons collected per week should be tracked as an end-result that is affected by the performance driver “tons collected per working hour” (auxiliary variable) and by the “working hours per week” (strategic resource stock).

[Figure 4 about here]

The auxiliary “tons collected per working hour” gauges workers productivity, which is in turn affected by the level of investments in service automation (strategic resource) and by another performance driver: the backyard collection ratio. This is the percentage of collected garbage from backyards on the total (ratio between two stocks). As the ratio increases, worker productivity decreases because less tons will be collected per working hour and less tons of garbage will be collected each week unless more staff is hired. We also must be careful when making causalities. For instance, the “working time saturation index” (the ratio between the working hours per week stock and a maximum working hours per week input parameter) cannot be considered as a performance driver. Likewise, figure 5 shows that the “tons collected per 1,000 population” is an index that is affected by, but does not affect organizational performance.

[Figure 5 about here]

While performance drivers are ratios between a current state (resource) and a benchmark, which affect performance, usually through a normalized graph function, performance indexes are synthetic measures of the quality or state of the system. They do not affect performance. Implying that an improvement in such indexes generates an improvement in other variables underlies inverting between causes with effects.

However, a high percentage of garbage collected from backyards on total could be identified as a driver of satisfaction. This is also a major driver of inefficiency. Such a
phenomenon could generate delays in garbage collection, which would counterbalance the positive effects on citizen satisfaction that backyard collection would generate. In order to avoid such counterproductive effects, the sanitation division could accept the burden of hiring more staff, but this would increase its operating costs and might generate more taxation, which could decrease citizen satisfaction. Likewise, using more automated machinery might increase the yield per staff unit, but higher investments and depreciation costs might increase the municipality’s financial needs. Though such trade-offs may look quite straightforward on a conceptual viewpoint, they are usually difficult to explore by divisional decision makers if a static and non-systemic view of performance is adopted. Dynamic performance management modeling can foster deeper insights on such trade-offs and may support decision makers in framing the complexity characterizing the systems where they operate.

Another performance driver, which indirectly affects the tons collected per week, can be referred to as the quality and consistency of incentive planning systems. Again, this is modeled as an auxiliary variable affected by investments into an intangible strategic resource stock. The design of better human resource planning and incentive systems will be likely to contribute to increase individual productivity. A higher individual productivity would contribute (given a stock of available workers) to increase the actual total working hours per week, which would increase the tons of garbage collected per week. Therefore, a same result could be achieved through a multitude of combined policies. On this regard, dynamic performance management simulation models can support scenario analysis.

In order to track performance under an effectiveness profile (the perceived quality of provided service), the measurement system should have to focus a set of non-monetary indicators that might contribute to evaluate the extent to which the city’s sanitation services are aligned
with user expectations (competitive performance). On this regard, as previously commented, the perceived time to serve households could be considered as a strategic success factor. The time to serve households can be modeled as an auxiliary variable, and more specifically as a ratio between the “total garbage to collect stock” and its outflow as shown in figure 5.

[Figure 6 about here]

A change in such ratio does not immediately affect the municipality’s reputation towards its service users (perceived time to serve households). In fact, an information delay smoothes changes in the actual time to serve households because it may also take several weeks by citizens to change their mindset (strategic resource) about the provided service. A change in such mindset will contribute to change in the long-run the overall citizen satisfaction rate, which will generate an impact on the population stock. In this model, it is implied that lack of efficiency and effectiveness in residential refuse collection would contribute to determine—in the long-run—a decision of more citizens to move to another municipality.

But while system dynamics has the potential of helping local governments move from performance adoption to performance implementation, which includes tracking citizen satisfaction rates around specific areas of service delivery, local officials cannot forget about the broader citizenship outcomes of political participation, social capital, civic belonging, and self-worth as addressed by Wichowsky and Moynihan (2008) in their effort towards modeling performance. The authors maintained that an exclusive focus on mission-based goals and objectives, including performance measures of output, efficiency, and outcome, fails to incorporate community and democratic values. The case of Winston-Salem clearly demonstrates that attempting to move from backyard to curbside collection based on the intricacies of
performance measurement and benchmarking was not enough to persuade the city council to approve mandatory curb-side collection until October 2010.

Regardless of its inefficiency, the budget director of Winston-Salem described back-door collection as part of the community’s culture and identity (Rowe, 2013). He also described that the percentage of citizens that moved to voluntary curb-side collection after it was approved in 2004 did not understand the historical context of backyard collection for the broader population. Therefore, local officials were confronted with a conflict between span of accountability and span of support as described by Simons (2005). The span of accountability was always high given the use of financial and nonfinancial measures to show a need for change. The span of support, however, was low because of a lack of shared understanding between the short-term and long-term payoffs for making this policy decision, which requires leadership and planning strategies to specifically address citizenship outcomes (Im and Lee, 2012).

Conclusion
This paper illustrates how system dynamics can be used to help local officials move from the adoption of performance measurement to the implication of the management tool, which is commonly referred to as performance management. Given that performance measurement has become a professional norm in local government, our profession has an obligation of actually using performance data for accountability and for service improvement given the amount of resources needed to maintain a well-managed performance measurement system. Therefore, it is imperative that academics and practitioners continue to look for approaches to help performance management become a professional norm in local government.
In applying system dynamics to the service area of residential refuse collection system for the city of Winston-Salem, North Carolina, several benefits of using this methodology were identified that can play an important role in improving the quality of organizational performance management systems and that can help local officials actually use performance data for making critical policy decisions. They include framing trade-offs in time and space associated with alternative scenarios, understanding how the accumulation and depletion of strategic assets are impacted by different policy levers, and determining how performance drivers affect end-results. System dynamics also can help local officials with establishing goals and objectives, linking the political to the managerial level, and focusing attention on selecting relevant targets and on evaluating results. The evaluation process is critical, where system dynamics is used to capture intangibles, system delays, and non-linear relationships and is used to avoid common errors of causalities between performance indexes and performance drivers. The question remains, however, would Winston-Salem have moved from backyard to curbside residential refuse collection in 2004 if the city would have embraced a system’s view rather than focusing on a handful of performance measures? Part of this answer would depend on whether or not citizenship outcomes were accounted for within the broader context of system dynamics modeling.

One final observation is that system dynamics does not represent the final step in moving toward an environment of performance management. It does place an important tool in our toolbox in addressing what factors promote data use in local government and what factors prevent data use. The hope is that other municipalities employ system dynamics modeling to make critical organizational changes by creating a shared view among all stakeholders when performance indicators clearly identify major inefficiencies in local government processes.
References


Figure 1: A Dynamic Resource-Based View of Organizations
Figure 2: A Dynamic Performance Management View
Figure 3. Mapping Population and Collection Points

- Population
  - Accumulation rate
  - Depletion rate

- Collection points
  - Accumulation rate
  - Depletion rate

- Average household size

- New Garbage to collect per collection point per week

- Total Garbage to collect
  - Accumulation rate
  - Tons collected per time

- Service efficiency
- Service effectiveness

Graphs show:
- Population
- Collection points (households)
- Tons collected/year
Figure 4. Modeling Worker Productivity
Figure 5. Performance Indexes
Figure 6. Modeling Service Levels