Does Performance Management Lead to Better Outcomes? Evidence from the U.S. Public Transit Industry

Although performance management processes are widely assumed to be beneficial in improving organizational performance in the public sector, there is insufficient empirical evidence to back this claim. In this article, the authors examine the impact of performance management practices on organizational effectiveness in a particular segment of the public transit industry in the United States. The analysis utilizes original survey data on performance management practices comprising both strategy formulation and performance measurement in 88 small and medium-sized local transit agencies in conjunction with comparative outcome data drawn from the National Transit Database maintained by the Federal Transit Administration. The results provide evidence that more extensive use of performance management practices does in fact contribute to increased effectiveness in this segment of the transit industry.

Over the past 20 years, some of the biggest questions of public management have revolved around performance management, the process of establishing goals for an agency or program, and then using performance information to help manage effectively so as to achieve those goals (Behn 1995; Moynihan and Landuyt 2009; Moynihan and Pandey 2010). In the public sector, performance management usually refers to engaging in strategic planning to establish a direction and major goals, setting more specific objectives and perhaps targets at multiple levels in the organization, and then using performance measurement to help focus on achieving them (Poister 2010). Thus, Moynihan defines performance management as “a system that generates performance information through strategic planning and performance measurement routines and that connects this information to decision venues” (2008, 5). Indeed, these are the defining characteristics of the Government Performance and Results Act of 1993 and its successor, the Government Performance and Results Modernization Act of 2010, in the federal government as well as in the performance regimes in many state and local governments.

Expectations that performance management should contribute to higher levels of organizational performance and, in particular, better outcomes is almost universal among proponents and critics of the performance movement (Ammons 1995; Behn 2003; Epstein 1984; Halachmi and Bouckaert 1996; Hatry 2006; Heinrich and Marschke 2010; Moynihan 2008; Poister 2003; Radin 2006; Wholey and Hatry 1992). Indeed, most authorities on performance management agree with an early report published by the National Academy of Public Administration (1994) contending that performance measures should not be viewed as an end in themselves but rather as vehicles for improving performance. While many authors point to a variety of purposes of performance measurement, improved performance, especially regarding outcomes, is always central to the discussion (Ammons 1995; Epstein et al. 2005; Kopczynski and Lombardo 1999; Poister 2003; Van Dooren, Bouckaert, and Halligan 2010; Wholey and Hatry 1992).

While interest in performance management is largely predicated on the assumption that it will lead to performance improvement, empirical research testing this relationship has been relatively sparse to date. A few “celebrated” cases of effective performance management have been documented, such as CompStat in New York City (Bratton and Malinowski 2008) and CitiStat in Baltimore (Behn 2007), and a growing number of case studies demonstrate great potential for using performance information to manage for results (Ammons and Rivenbark 2008; Barnow and Smith 2004; de Lancer Julnes et al. 2007; Gore 1993; Holzer and Yang 2004; Sanger 2008). Nevertheless, Ammons and Rivenbark contend that “most claims of performance measurement’s value in influencing decisions and improving services tend to be broad and disappointingly vague,” and they point out that “hard evidence documenting performance measurement’s impact on management decisions and service improvements is rare” (2008, 305). Rigorous large-sample evaluations of the impact of performance management impact on management decisions and service improvements are rare, and thus, it is difficult to provide robust evidence for or against the claim that performance management leads to better outcomes.
management approaches on performance are still lacking (Andersen 2008; Boyne 2003).

This article presents the results of research on the effect of performance management practices involving strategic planning and performance measurement on outcomes generated by a fairly homogenous subset of public transit agencies in the United States, essentially small and medium-sized bus systems operating in service areas where they are the sole provider or public transit service. Earlier, we analyzed the effect of strategy formulation on transit system performance and found that the use of formal strategic planning practices has a beneficial effect on outcomes in this public service industry (Poister et al., forthcoming). In this piece, we expand the model to examine whether performance measurement as the monitoring and evaluation component as well as the agenda-setting strategic planning component of performance management both lead to better outcomes. Using measures generated from an original survey of transit managers on strategy formulation and performance measurement, this analysis tests the effect of these two critical elements of performance management on conventional measures of transit system outcomes drawn from the National Transit Database compiled by the Federal Transit Administration. The findings provide evidence that performance management contributes to superior performance in small and medium-sized public transit systems in the United States.

Performance Management and Organizational Performance

The logic underlying the assumption that performance management will lead to stronger performance rests on the importance of both goal clarification and performance monitoring in managing for results. When clear goals are established concerning strategic planning and outcomes through strategic planning, monitoring outcomes on a systematic basis will generate useful information that can be utilized by program managers and others, helping them make better decisions and manage programs more effectively, which, in turn, will lead to improved performance (Behn 2003; Hatry 2006; Marr 2012; Poister 2003; Walker, Damanpour, and Devece 2011; Wholey and Hatry 1992).

Barriers to Performance Improvement

Although this logic seems sound in theory, there are many reasons why these relationships might not hold up in practice. First, some public programs simply do not lend themselves readily to performance measurement (Bouckaert and Balk 1991; Leeuw 2000; Radin 2006) because of complexity and/or goal ambiguity or outputs and outcomes that are difficult to measure and difficult to control (Jennings and Haist 2004; Wilson 1991). Second, performance measures may not generate useful information because they lack validity and reliability, are not timely and actionable, are descriptive rather than prescriptive, and focus on past performance without offering implications regarding strategy and execution (Hatry 2006; Heinrich 2007; Marr 2009; Poister 2003).

Even when programs readily lend themselves to measurement, the lack of utilization of performance information has long been recognized as a critical barrier to the efficacy of performance management regimens (de Lancer Julnes and Holzer 2001). This may be attributable to a lack of administrative flexibility, an organizational culture that is not performance oriented, a lack of leadership commitment to results, or low levels of public service motivation, among other reasons (Moynihan and Levartu 2012; Moynihan and Pandey 2010). Furthermore, Moynihan (2008) points out that in contrast to purposeful use of the data by public managers and decision makers to improve performance, some actors in these performance regimes are likely to make passive use of performance data—minimally complying with procedural requirements without actually working to improve performance—or political use or even perverse use of performance information to support self-serving interests, which might actually work against achieving the kinds of results a program is intended to produce.

Moreover, the challenges of improving the performance of many kinds of public programs can be monumental, regardless of whether good performance data are available or utilized (Funnell and Rodgers 2011; Rossi, Lipsky, and Freeman 2004; Weiss 1972). Many programs operate in difficult environments or deal with intractable problems. Thus, more effective strategies, more surgical interventions, or more innovative service delivery systems for generating the desired outcomes may not be readily apparent, particularly after the “low-hanging fruit”—obvious opportunities for improving performance that can be readily implemented—have been picked. Even when agencies have engaged in comparative performance measurement and benchmarking processes, best practices may not have been identified, validated, or disseminated (Keehley and Abercrombie 2008).

The logic underlying the assumption that performance management will lead to stronger performance rests on the importance of both goal clarification and performance monitoring in managing for results.

In addition, resource constraints often present barriers to performance improvement (Boyne 2003) by limiting training and development activities, discouraging innovation and experimentation with newer programmatic approaches, or preventing implementation of new strategies aimed generating increased outcomes. Moreover, organizational characteristics might impede efforts to improve performance as well. For example, an agency might have extensive formal rules, procedures, and red tape that may stifle experimentation in new approaches to improving performance (Behn 2003; Boyne 2003; Brewer and Walker 2010). Or the organizational culture might not be supportive of performance improvement, which may be overridden by concerns with nonmission priorities such as equity, due process, and representative bureaucracy (Brewer and Selden 2000; Moynihan 2008). Some organizations may not be able to leverage their influence on external stakeholders, other organizations inside or outside government, or networks whose engagement and cooperation would be essential to realize performance improvements (DeGroff et al. 2010).

Finally, an organization may not have the kinds of performance-oriented management systems, such as performance budgeting processes, performance contracting and/or grant management, process improvement, or customer service processes, that may be required to make
meaningful use of the performance data. Other organizations may not have “soft” human resource management practices that emphasize the importance of job satisfaction and morale in place, leaving employees less than fully committed to achieving organizational outcomes (Boyne 2003; Brewer and Selden 2000). Furthermore, some agencies may not have the kind of transformational leadership that can increase mission valence and mobilize employees’ efforts and attention in support of performance improvement (Brewer and Selden 2000; Wright, Moynihan, and Pandey 2012).

Previous Research
Surprisingly, given the centrality of the performance movement in the field of public management in the United States and elsewhere, objective research on the efficacy of performance management has been relatively sparse. With respect to the impact of strategy formulation, the research to date has examined the effects of both logical incrementalism and formal strategic planning on performance. Very briefly, logical incrementalism, as advocated by Quinn (1978), emphasizes ongoing strategy formulation and revision based on experimentation and learning; negotiation with a mix of other critical actors and stakeholders; accommodation of the imperatives of personal, political, and power relationships; and adaptation to changes in the environment as new challenges and opportunities present themselves, sometimes unexpectedly. Formal strategic planning, on the other hand, emphasizes clarification of mission and vision for the future, setting clear goals, developing strategies and strategic initiatives designed to accomplish those goals, and establishing commitments to action plans for implementing those strategic initiatives (Berry and Wechsler 1995; Bryson 2004; Eadie 1983; Niven 2003; Nutt and Backoff 1993).

The results of this research have been mixed. In one of these studies, Andrews et al. (2009) included measures representing logical incrementalism and an absence of strategy, as well as formal strategic planning based on local officials’ responses to survey items concerning strategy development in local government authorities in Wales. Their dependent variable was a standardized performance index based on numerous performance indicators across a range of services that are reported to the National Assembly for Wales. The results of the ordinary least squares models indicated that logical incrementalism and absence of strategy both had negative impacts on performance. However, these authors found no evidence that formal strategic planning had an effect on performance in either direction.

A second study by Walker et al. (2010) is similar to the foregoing research except that the subjects of this study consisted of 101 English local authorities. The dependent variable in this case was an aggregate measure of performance, the Core Service Performance index, which was computed by the British Audit Commission by combining performance scores across seven service areas that were weighted by their relative importance in the budget. In contrast to Andrews et al. (2009), this study found no evidence that logical incrementalism had an impact on performance. However, the results of this research did suggest that formal strategic planning had a positive impact on performance, but only when past performance was not included in the model.

Following this research in Great Britain, our initial analysis of the relationship between strategy formulation and performance in U.S. public transit agencies focused on the effect of a blended approach using both logical incrementalism and formal strategic planning, as well as each of these approaches separately. We found that although logical incrementalism on its own was associated with poorer performance, both formal strategic planning and a blended approach were positively related to measures of transit system outcomes (Poister et al., forthcoming).

Some elements of performance measurement, such as target setting, have also been subjected to study. Research by Boyne and Gould-Williams (2003), for example, found that setting targets actually had a negative impact on the perception of performance improvement in Welsh local authorities. In this study, Boyne and Gould-Williams interviewed more than 1,000 local government officers and elected officials in Wales on their perceptions of the extent to which performance on several dimensions had improved as a result of the Best Value planning regime installed by the central government. The results showed that both setting targets and the number of targets established (up to a maximum of six) were associated with perceptions of lower performance on all seven performance measures tested, and they were statistically significant for three of them. Thus, setting targets and implicitly monitoring performance measures to determine whether they were being attained was negatively associated with perceived performance. Boyne and Gould-Williams speculated that the negative relationship might have been caused by target setting creating ceiling effects or by goal ambiguity generated by setting numerous targets.

A somewhat similar study of 147 local English education authorities by Boyne and Chen (2007) used more objective measures of performance rather than perceived performance and produced very different results. This research found that pupils’ examination scores increased as the number of targets (with a maximum of six) set by the local education authorities increased. According to Boyne and Chen’s analysis, student grades on English, mathematics, and science were increased significantly by setting performance targets, and the more of these targets that were set, the more these scores might increase.

Going beyond target setting, a recent study by Walker, Damanpour, and Devece (2011) across a range of service areas in 17 English local government authorities represents one attempt to link performance management as a holistic process to organizational performance. The elements included in the performance management construct in this study comprised such factors as indicators of ownership and understanding of organizational mission and goals, specification of appropriate performance measures and targets, devolution of control to service managers, and taking corrective action when results deviated from plans, while organizational performance was measured by aggregated core service performance scores constructed by the British Audit Commission. In both a multiple regression model incorporating the performance management variable as well as a measure of management innovation in these local service agencies, and in a structural equation model in which performance management mediated the effect of innovation, the authors found that performance management exerted a positive effect on performance.

Two other studies have been published in the public management literature on the effects of performance management initiatives on academic achievement in public schools. In the first of these,
Andersen (2008) assessed the collective impact of such instruments as written goals, written evaluations or feedback on results, and management by objectives on test scores of several thousand ninth graders in Danish public schools. The results of multilevel regression models for various years show weak positive effects of the performance management index on test scores, which are statistically insignificant, that give way to statistically significant but negative effects on test scores when an inequity index is taken into account.

Accountability is another component of performance management. A study examining the effects of accountability systems tied to standardized testing in 30 U.S. states in conjunction with legislation from the federal No Child Left Behind Act found no evidence that stronger systems led to improvements in students’ educational achievement. In fact, the results suggest that states with stronger accountability systems showed less improvement in student outcomes as measured by change in math and reading scores (Patrick and French 2011).

Thus, the fairly meager empirical research to date on the impact of performance management practices has produced decidedly mixed results, with four of the eight studies employing cross-sectional analysis across comparable public organizations showing positive effects on performance. Therefore, the question of the effectiveness of performance measurement or management practices in actually improving performance is very much still at issue. Given the huge investment of time, effort, and resources in performance management regimes in the field of public management at large, and the long-standing but largely unsubstantiated expectations that they will contribute to tangible improvements in the performance of public agencies and programs, further research examining the effect of a wide array of performance management approaches on performance itself is clearly required.

The Research Framework

Building on our earlier analysis of the effect of strategy formulation on performance (Poister et al., forthcoming), the present study adds performance measurement into the mix, complementing strategic planning as a critical element of a public organization’s overall performance management process. While strategic planning—establishing outcome-oriented goals and strategy initiatives to achieve them—is the agenda-setting component of the process, performance measurement is the monitoring and evaluative component, which entails assessing performance on an ongoing basis, identifying performance problems when they arise, and signaling the need to take corrective action to get performance back on track when problems do arise. Our principal research question is whether performance measurement, along with formal strategic planning, leads to better outcomes.

Clearly, a complex web of causal factors surrounds the issue of implementing and utilizing performance management practices in government, and we have much to learn about their efficacy in practice (Moynihan 2008). Notwithstanding all of the barriers to effective performance management outlined earlier, as well as the “stochastic processes” that influence organizational performance or the lack of same in government (Behn 2003), we believe that, on balance, the use of performance management practices embodying strategic planning and performance measurement should yield some net benefit in terms of performance improvement.

First, strategic planning is concerned with optimizing the “fit” between an organization and the environment in which it operates, and for public agencies, that involves strengthening performance in providing services to the public. The most direct linkage between performance management and performance improvement is provided by Locke’s goal-setting theory, which holds that because behavior reflects conscious goals and intentions, employees’ efforts and thus performance in organizations will be influenced by the goals assigned to them (Fried and Slowick 2004). Furthermore, Latham (2004) explains that goal setting drives performance because it diverts energy and attention away from goal-irrelevant activities toward goal-relevant efforts and energizes people to put forth greater effort to achieve goals. This is especially important in the public sector, where problems of goal ambiguity can impede performance (Chun and Rainey 2005). Thus, setting goals, objectives, or targets regarding organization or program performance will keep the organization focused on priorities, outcomes, and results and thereby lead to improvements in performance (Ammons 2008; Kelly 2003; Poister 2003; Van Dooren, Bouckaert, and Halligan 2010). Strategic planning establishes goals and objectives, many of which are likely to be performance related, and creating and implementing strategies designed to achieve them would be expected to lead to improved performance (Bryson 2004; Koteen 1989; Niven 2003; Poister 2010; Walker, Damanpour, and Devece 2011).

Hypothesis 1: Transit agencies that engage more fully in formal strategic planning processes will be more likely to generate intended outcomes.

While logical incrementalism on its own may not lead to, and may even detract from, performance because it does not necessarily focus on desired outcomes in a disciplined manner (Andrews et al. 2009; Walker et al. 2010), integrating the interactive dimension of logical incrementalism with the more formal rational approach embodied in strategic planning may well maximize organizational effectiveness (Camillus 2006; Papadakis and Barwise 1998; Quinn 1978). Other observers have noted that the value of strategic planning is enhanced when it is conducted within a larger context or with purposeful incrementalism (Barzelay and Campbell 2003; Poister 2010) and that it is important for public agencies to be opportunistic as well as deliberate in formulating strategy (Bryson 2004).

Hypothesis 2: Transit agencies that rely more heavily on logical incrementalism on its own will be less likely to generate intended outcomes.

Hypothesis 3: Transit agencies that blend formal strategic planning with logical incremental decision making will be more likely to generate intended outcomes.
The logic underlying the assumption that performance management will lead to higher levels of performance, though not often examined critically, is quite straightforward. According to Behn (2003), performance measurement systems signal the dimensions of an agency’s performance that are important and deserve or require attention, motivate managers and employees to work harder and effectively to achieve higher performance levels, help managers and others to learn and understand how performance can be improved, and enable managers and higher authorities to control the behavior and actions of employees and others under their purview to ensure that their attention and efforts are focused on improving performance.

The importance of performance management as a motivator is also crucial (Behn 2003). Reflecting the needs of achievement theory (McClelland 1965), many employees are naturally motivated to be successful in their jobs and thus will do their best to “look good” and hit targets on the measures rather than fall short of them, contributing to stronger performance (Poister 2003). Using performance management systems to provide them with feedback on their performance will thus motivate them to try to improve performance (Rainey 2009). Others will be more strongly motivated by expectations of rewards (Vroom 1964), and thus performance management systems that tie rewards to individuals or organizational units based on documented performance provide powerful incentives to improve performance (Halachmi and Bouckaert 1996; Poister 2003; Rodgers and Hunter 1992; Swiss 1990).

In addition, the data generated by monitoring systems and their associated analyses should be expected to provide useful information that will enable managers to make better decisions and manage more effectively (Ammons 2008; Hatry 2006; Wholey and Hatry 1992). Public agencies are goal-seeking systems in which feedback in the form of performance information is used to improve performance (Mowitz 1980). This will play out through their strategic management, resource allocation, program management, service delivery, human resource development, procurement, organizational and individual-level performance management, process improvement, and other processes to improve performance (Poister 2003). When problems arise, the feedback provided by performance data can help top executives and line managers take corrective action on a timely basis, which is essential to getting things back on track and improving performance in the long run (Walker, Damanpour, and Devece 2011).

Higher-level authorities use performance measures to hold agencies, organizational units, and managers and employees accountable for maintaining and improving performance, and this can provide powerful incentives to do just that (Cavalluzzo and Itner 2004; Wang 2002). The provision of timely performance information helps close the information gap between principals, who are often removed from the immediate work context, and their agents, who may have a fair amount of discretion as to how they work and what they try to accomplish (Ostrom 1999; Wang 2002). Through regular feedback from the performance data, higher-level managers can become more familiar with the situation “on the ground” and intervene as necessary to ensure that lower-level managers and employees focus their attention and effort on high-priority work and desired results (Armstrong 2000; Halachmi and Bouckaert 1996; Kelly 2003). Monitoring data on program quality, efficiency, outputs, and outcomes allows these principals to achieve greater accountability, which should also lead to higher levels of performance (Bernstein 2001; Moynihan 2008; Poister 2003).

Thus, in tandem with developing strategic goals and objectives, monitoring performance and utilizing performance information should enable public managers to make better decisions and manage programs and agencies more effectively, as performance routines help keep the focus on performance, motivate managers to achieve goals, provide insights regarding performance problems and potential solutions, and help hold lower-level managers and organizational units accountable for maintaining or improving performance (Behn 2003; Hatry 2002; Poister 2003; Wholey and Hatry 1992).

Hypothesis 4: Transit agencies that engage more fully in performance measurement will be more likely to generate intended outcomes.

We test these hypotheses by examining performance management practices and specific outcome data in a set of urban public transit agencies in the United States. Public transit systems have often been used in research designed to analyze the performance implications of variables that are of interest to public management scholars, such as size (Brown and Thompson 2008; Perry and Babitsky 1986), institutional structure (Leland and Smirnova 2009; Perry and Babitsky 1986), and privatization (Giuliano 1980; Iseki 2008; Leland and Smirnova 2009; McCullough, Taylor, and Wachs 1998; Perry and Babitsky 1986).

As service delivery systems that produce readily observable and, to some extent, controllable outputs and outcomes, public transit agencies would seem to be good candidates for the effective implementation of performance management systems (Jennings and Haist 2004), and because they operate in competitive marketplaces, they may face more pressure to achieve high performance levels than many other public services. Thus, if managing performance through strategic planning and performance measurement does indeed hold potential for improving outcomes, we might well expect that relationship to manifest itself in the public transit industry.

Methodology

This analysis of the effect of performance management on outcomes utilizes indicators of strategy formulation and performance measurement in the transit industry created from responses to an original survey of a sample of transit managers in the United States, along with conventional measures of transit system performance drawn from the National Transit Database (NTD) compiled on an ongoing basis by the Federal Transit Administration.

Sample Selection and Survey Methodology

To ensure comparability among transit agencies, we used the NTD to identify a subset of transit agencies in the United States that were similar in terms of scope of services and operating environment to ensure a fair amount of comparability among the agencies surveyed. First, the smallest agencies—those with fewer than nine vehicles in...
operation—were eliminated from our sampling frame because they are exempted from reporting to the NTD, and thus performance data are not available for them. Second, we focused on bus transit systems and eliminated agencies that operate rail service because these agencies have institutional structures, infrastructure, operating characteristics, and service demands that are very different from agencies that operate only bus systems, in some cases along with demand-responsive services. Third, agencies with overlapping service areas were also eliminated in order to ensure that the performance of any given agency in our sample was not affected by the presence of another transit agency operating in the same service area or market (Kittleson et al. 2003; Ryus 2010).

The last two criteria also served to eliminate from our sample transit agencies in the largest metropolitan areas in the United States, such as those in New York City, Washington, D.C., Chicago, San Francisco, and Atlanta, among others, which operate much larger systems than most transit agencies in the country. After applying these parameters, our population consisted of 236 transit agencies, which can generally be characterized as agencies that operate small and medium-sized bus transit systems in urbanized areas where they are the lone providers of public transit services. This definition of our population of interest is consistent with long-standing practice in research regarding performance in the transit industry (Boschken 1992; Leland and Smirnova 2009; Perry and Babitsky 1986; Zullo 2008).

Thus, the exclusion of the smallest and largest transit systems was not based directly on size but rather resulted from the fact that the smallest systems do not report operating data to the NTD, on the one hand, and from decisions concerning transportation mode and overlapping service areas, on the other. These decisions are in keeping with a recent report on comparing performance in the transit industry, which stated that high-level performance comparisons should focus on agencies of similar size, characteristics, and operating conditions to ensure that they are “similar, but not identical” (Ryus 2010, 17). The report went on to say that transit agencies should not be compared against others operating in qualitatively different kinds of contexts or providing relatively uncommon types of service.

Our definition of the population of interest in this study is also consistent with long-standing practice in research regarding performance in the transit industry. As Boschken acknowledged, “it is conventional in transit to group agencies for analytic purposes by their organization size and production mode (rail vs. bus)” (1992, 277). Several studies have focused solely on transit agencies of particular sizes (Benjamin and Obeng 1990; Leland and Smirnova, 2009; Ripplinger and Brand-Sargent 2010; Williams 1979). In addition, numerous studies have sharpened their focus to include only bus systems in their analyses (Anderson 1983; Cubukcu 2006; De Borger and Kerstens 2000; Fielding, Babitsky, and Brenner 1985; Iseki 2008; Lave and Pozdena 1977; Zullo 2008). On balance, then, although it eliminated some potentially interesting transit agencies, our approach provided a set of public transit systems that was more homogeneous with respect to size, mode, and operating environment without constraining variation in terms of performance management practices or performance itself.

Data on performance management practices were solicited using an online survey of the chief executive officers of these 236 transit agencies that was conducted from October to December 2009. Contact information for the executive directors was gathered from agency Web sites and/or phone calls made to the transit agencies. The survey was initiated with an e-mail containing a link to the online survey site that was sent to the chief executives/officers/managing directors of transit agencies. While the survey clearly was targeted individually to the chief executive officers, with no pretense of anonymity, the introductory materials provided a guarantee of confidentiality of their responses in the form of an assurance that the survey results would be reported only in the aggregate. Almost half of the 236 selected transit agencies responded (103 or 43 percent). Those agencies that did respond provided a good cross-section of transit systems around the United States. Only 11 of the 50 states were not represented, and four of those states had no transit agencies in the population of interest.

In responding to these survey items, transit managers were asked to base their answers on strategy development and performance measurement in their agency during the period from 2004 to 2008. This allowed us to obtain a reading on their performance management practices over a period of four years immediately preceding the point when the current recession began to take hold in the second half of 2008. Assuming that the hypothesized beneficial impact of performance management on performance might not occur immediately, but rather accumulates over time, examining performance management practices from 2004 to 2008 helps establish causality chronologically, as our performance data pertain to 2008. It is also important to note that the 2004–8 time frame avoids a potential problem of real effects of performance management being suppressed by the effects of the higher unemployment rates borne of the recession, which contributed to lower levels of transit ridership nationally in 2009 and beyond than might have been the case under a constant cause system.

The survey responses were then merged with outcome-related dependent variables drawn from the NTD. While we might have solicited the information needed to compute the measures of performance from the survey in addition to asking about management practices, it was preferable to draw the performance data from the NTD, for two reasons. First, although the NTD data are also self-reported by the transit systems, they are expected to be more careful in entering their data according to uniform reporting specifications because they are subjected to quality assurance reviews by the Federal Transit Administration, thus providing a higher level of reliability for comparison purposes. In addition, relying on the NTD for the performance data allowed us to keep the number of survey questions to a reasonable level in order to encourage a higher response rate.

Looking at the performance measures, then, because ridership measures are usually considered to represent important outcomes in the transit industry (Kain and Liu 1999; Karlaftis and McCarthy 1997; Kittleson et al. 2003), we focused first on the number of unlinked passenger trips per capita in each transit system’s service area as an indicator of a community’s propensity to use public transit. In addition, because transit system effectiveness is also measured in terms of system productivity, relating ridership to service outputs (Fielding,
Babitsky, and Brenner 1985; Ryus 2010), we focused on the ratio of unlinked passenger trips to vehicle miles operated as a second dependent variable.

Analysis

We used ordinary least squares multiple regression models to test our hypotheses with respect to each of the two transit outcome measures, passenger trips per capita and passenger trips per vehicle mile operated. Two strategy formulation variables, logical incrementalism and formal strategic planning, were included in these models as influencing the outcomes independently, and the blending of these two approaches was represented by their interaction term. In addition, a measure representing the extent to which these transit agencies use performance measures was included as a primary independent variable in these models. Each of our strategy formulation variables was constructed using responses to four items included in the survey of public transit managers, and the performance measurement variable was based on five items in the survey. We used principal component analysis to estimate factors for each of these performance-management-related variables. The appendix presents the survey items used to construct each of these indices, as well as their Cronbach’s alpha statistics, which are acceptably high.

We utilized raw data from the NTD to create the outcome measures that constitute the ultimate dependent variables in the regression models to be estimated: passenger trips per capita and passenger trips per vehicle mile operated. These data pertain to performance in fiscal year 2008, the year at the end of the period to which the survey questions regarding strategy formulation and performance measurement referred, 2004–8. In addition, 2008 is the most recent year in which transit system performance would have been largely unaffected by the current recession.

We also included a measure of past performance in each of these models to account for other factors that might be expected to influence performance. Most of the prior research testing various aspects of Meier and O’Toole’s management model (e.g., Andrews et al. 2009; Nicholson-Crotty and O’Toole 2004; O’Toole and Meier 1999; Walker et al. 2010) included measures of past performance to reflect the basic autoregressive nature of the performance of most public sector organizations. As a surrogate measure representing (to some extent) the net effect of an unknown number of contextual factors that are likely to influence transit system performance, past performance constitutes an important control variable. Thus, in the present study, each model constructed to test the influence of performance management on a particular measure of transit system performance in fiscal year 2008 included a measure of that same performance indicator in fiscal year 2004 as a control variable. In addition, to take differences in transit system characteristics into account, we also included as control variables measures of population density, automobile ownership, institutional structure, and privatization of service delivery (Giuliano 1980; Iseki 2008; Leland and Smirnova 2009; McCullough, Taylor, and Wachs 1998; Perry and Babitsky 1986).

Results

Table 1 displays descriptive statistics for all of the variables included in the regression models. Two aspects of these data are of particular interest regarding the transit outcome measures. First, there is widespread variation in these measures across the transit systems in our sample. With respect to passenger trips per capita, the standard deviations in both 2004 and 2008 are nearly as large as their respective means. Second, the period in question was generally a good one for the transit industry; passenger trips per capita increased 21 percent on average from 2004 to 2008, while the number of passenger trips per vehicle mile increased almost 10 percent over the same period.

The results of the regression analysis are shown in table 2; both models in the table were estimated with 88 cases, as 15 cases were lost because of missing values in one or more of the required survey items or some element of the NTD data that was needed to compute the outcome measures of interest. Looking at table 2, while the effect of logical incrementalism on passenger trips per capita is negative as hypothesized and statistically significant, the effect of strategic planning, though positive as hypothesized, does not reach statistical significance. However, the combined approach utilizing both formal strategic planning and logical incrementalism, represented as their interaction term, does have a positive and statistically significant effect on this outcome measure. Briefly, the positive effect of strategic planning on passenger trips per capita is substantially stronger in transit agencies that also engage more fully in logical incrementalism, while the negative impact of logical incrementalism on performance diminishes noticeably in agencies that engage more fully in formal strategic planning. In addition, the performance measurement variable has a positive and statistically significant effect on passenger trips per capita as hypothesized.

The results of the analysis focusing on passenger trips per vehicle miles operated, shown in table 2, are similar, although they are

<table>
<thead>
<tr>
<th>Variables</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logical incrementalism</td>
<td>–3.30</td>
<td>2.18</td>
<td>.009</td>
<td>1.00</td>
</tr>
<tr>
<td>Formal strategic planning index</td>
<td>–2.50</td>
<td>1.82</td>
<td>–0.003</td>
<td>1.00</td>
</tr>
<tr>
<td>Formal strategic planning and logical incrementalism interaction</td>
<td>–.920</td>
<td>5.43</td>
<td>.580</td>
<td>1.08</td>
</tr>
<tr>
<td>Performance measurement</td>
<td>–3.40</td>
<td>1.73</td>
<td>.000</td>
<td>1.00</td>
</tr>
<tr>
<td>Passenger trips per vehicle mile (2008)</td>
<td>.130</td>
<td>3.09</td>
<td>1.01</td>
<td>.610</td>
</tr>
<tr>
<td>Passenger trips per vehicle mile (2004)</td>
<td>.120</td>
<td>2.84</td>
<td>.920</td>
<td>.500</td>
</tr>
<tr>
<td>Percentage of households not owning a car 2008</td>
<td>3.30</td>
<td>13.30</td>
<td>7.097</td>
<td>2.08</td>
</tr>
<tr>
<td>Population density (2007)</td>
<td>40.75</td>
<td>6,281.56</td>
<td>1,884.096</td>
<td>1,241,277</td>
</tr>
<tr>
<td>Proportion of service contracted out (2008)</td>
<td>.000</td>
<td>1.00</td>
<td>.279</td>
<td>.356</td>
</tr>
<tr>
<td>Independent agency or government unit (2008)</td>
<td>.000</td>
<td>1.00</td>
<td>.528</td>
<td>.502</td>
</tr>
</tbody>
</table>

Table 1 Descriptive Statistics
different in two important respects. First, the effect of formal strategic planning is still positive, but it reaches statistical significance in this model. Second, the positive effect of performance measurement on passenger trips per vehicle mile operated reaches a higher level of statistical significance than it does in the model focusing on passenger trips per capita. Thus, the positive effects of formal strategic planning and performance measurement are both stronger in terms of statistical significance with respect to system productivity—outcomes in ratio to outputs produced—than basic effectiveness as measured by outcomes in ratio to a contextual variable of primary importance.

To substantiate the reliability of these models, given control variables that could possibly be highly correlated with other control or primary independent variables, we checked for multicollinearity. In the models for both performance measures, however, the lowest tolerance level across the nine independent variables exceeded 0.50, while the highest variance inflation factor values were below 2.0, providing no indication of multicollinearity in the data. In addition, we ran three fixed-effects models to determine whether our models might be confounded with geographic differences around the country, but the results of this sensitivity analysis generally sustained our original results. 2

Turning to the issue of practical significance, the magnitude of the estimated impact generated by increased strategic planning and performance measurement activity on ridership would be quite meaningful to most public transit agencies. For example, as engagement in formal strategic planning increases by one unit on that index—approximately equivalent to the difference between strongly disagreeing and disagreeing or between being neutral and agreeing with the four survey items regarding strategic planning—transit ridership on those 88 systems tends to increase by slightly less than one additional passenger trip per capita per year. Given that the mean average annual passenger trips per capita is 13.14, increasing engagement in strategic planning by one step on that scale could be expected to lead to a ridership increase of approximately 7.4 percent (0.967/13.14). Similarly, the effect of a one-unit increase in the formal strategic planning activity scale would produce a 7.2 percent increase (0.073/1.01) in the number of passenger trips per vehicle mile operated by the “typical” transit agency in our sample.

The effects of engaging in performance measurement are even more pronounced. As engagement in the use of performance measures increases by one unit on that scale, the number of annual passenger trips per capita tends to increase by 1.71. Thus, a one-unit increase in performance measurement activity would be expected to generate a 13 percent increase in ridership on the typical transit system carrying 13.14 annual passenger trips per capita. Furthermore, with the same one-unit increase in performance measurement activity, ridership also tends to increase by an additional 0.103 passenger trips per vehicle mile, which for the average transit system carrying 1.01 passenger trips per vehicle mile operated translates into an increase of slightly over 10 percent. Given their heavy reliance on ridership to generate resources both through fare box revenue and performance-based federal and state grants, it is safe to assume that all or at least most of the transit systems in our sample would consider ridership increases on the order of 7.5 percent, 10 percent, and 13 percent to be substantial increases and worthy targets. And, obviously, if transit agencies increased their engagement in strategic planning and/or performance measurement by more than a single unit on those scales, the models estimated here would suggest even larger ridership increases.

**Conclusions**

These results provide at least limited evidence that both strategic planning and performance measurement, the principal components of performance management in public organizations, do contribute to improved performance in small and medium-sized transit systems in the United States. The overall finding is that engaging more proactively in what have become conventional performance management practices is indeed associated with gains in outcomes among these systems. Within the confines of a simple cross-sectional research design, but one that largely controls for interference from external and contextual factors with the use of past performance and other control variables, we conclude that engaging more fully in these performance management practices leads to benefits in the form of increased effectiveness and system productivity in these public transit systems that are by no means trivial.

Governments at all levels in the United States, as well as those in many other countries, have invested considerable time, energy, and resources on performance management practices in the belief that they will trigger or otherwise contribute to improved performance in producing the desired kinds of results. Yet this assumption that performance management will lead to better outcomes is still largely untested by systematic research, and the few studies that have been published in the public management literature to date have produced mixed results. Our analysis adds to this small but crucial literature stream a finding that engaging more fully in performance management practices is associated with stronger outcomes among small and medium-sized urban public transit systems operating bus service.

The current study is the first to examine the two major components of the classical approach to performance management in the public sector together in terms of their effects on agency performance. This analysis indicates that both have positive effects on transit system ridership. Interestingly, the effects of performance measurement activity outweigh the effects of formal strategic planning on

---

**Table 2 Regression Results**

<table>
<thead>
<tr>
<th></th>
<th>Passenger Trips per Capita</th>
<th>Passenger Trips per Vehicle Mile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logical incrementalism</td>
<td>-2.099***</td>
<td>-0.187***</td>
</tr>
<tr>
<td>Formal strategic planning</td>
<td>0.967</td>
<td>0.073*</td>
</tr>
<tr>
<td>Formal strategic planning * logical incrementalism</td>
<td>1.216**</td>
<td>0.076**</td>
</tr>
<tr>
<td>Performance measurement</td>
<td>1.709*</td>
<td>0.103**</td>
</tr>
<tr>
<td>Percentage of households not owning a car 2008</td>
<td>0.069</td>
<td>0.043**</td>
</tr>
<tr>
<td>Population density</td>
<td>0.00005</td>
<td>0.00002</td>
</tr>
<tr>
<td>Proportion of service contracted out</td>
<td>-2.406*</td>
<td>-0.076</td>
</tr>
<tr>
<td>Independent agency or government unit</td>
<td>-1.23</td>
<td>-0.009</td>
</tr>
<tr>
<td>Past performance</td>
<td>1.083***</td>
<td>0.770***</td>
</tr>
<tr>
<td>R2/Adjusted R2</td>
<td>880/876</td>
<td>880/759</td>
</tr>
</tbody>
</table>

* Significant at the .05 level; ** significant at the .01 level; *** significant at the .001 level (one-tailed test).
ridership to a noticeable degree. Because increasing or at least maintaining ridership is an aspiration shared by virtually all public transit systems in the United States, strategic planning will not usually be required to clarify this primary goal, but it still can be very useful in targeting particular markets and developing strategies for building ridership, while performance measurement may then be even more important on an ongoing basis for keeping the focus on ridership gains and ensuring that they are achieved.

While this research finds that performance measurement and formal strategic management, the two principal components of performance management, can contribute to improved outcomes produced by the kinds of small and medium-sized public transit agencies observed in our study, these results may not necessarily be generalized beyond the scope of our analysis. In particular, with respect to substantially larger transit systems, agencies operating rail service as opposed to or in addition to bus service, and/or systems operating in metropolitan areas in which other transit agencies are providing similar and perhaps competing services (i.e., transit systems operating in much more complex environments), the efficacy of performance management might be substantially reduced. On the other hand, such systems might be characterized by greater need as well as a wider range of options for effective performance management, and the relation between utilization of such practices and actual performance could be even more dramatic than in the transit systems that were the focus of our study. Unraveling what this relationship might look like would require further research focusing directly on these larger and more complex transit systems.

Furthermore, within the larger framework of the full range of public services, public transit systems in the United States are not likely to provide the most strenuous test of the ability of performance management to influence performance. Applying criteria developed by Jennings and Haist (2004), because transit systems generate outputs and outcomes that are readily observable and have a relatively high degree of control over the outputs that they produce, and because the causal relationships that connect transit outputs to outcomes are fairly well known, these agencies may be particularly good candidates for utilizing performance management practices to strengthen performance. The extent to which the results obtained in this research would apply in less tangible program areas, such as human services, is not clear at this point. Thus, more research is needed along these lines to gauge the usefulness of performance management in a wide range of public programs and operating environments in order to determine the conditions under which it is likely to be more effective or less effective, as well as to begin to pinpoint more specific practices that may facilitate or impede the efficacy of performance management in generating desired organizational outcomes.

Notes
1. Unlinked passenger trips is a measure of passenger boardings on transit vehicles. When a passenger begins a trip on one vehicle but then transfers to a second vehicle in order to reach his or her destination, that counts as two unlinked trips. Unlinked passenger trips is the standardized measure of ridership reported by transit agencies to the NTD.
2. In a fixed-effects model controlling for the 36 states represented in our data, the coefficients remained quite similar to those in the original model, although only four of the seven associations that reached statistical significance in the original model did so in the state fixed-effects model because of the attrition of degrees of freedom out of a sample of only 88 cases to begin with. However, in a fixed-effects model controlling for the 10 federal regions, the levels of statistical significance as well as the coefficients were nearly identical to those in the original model, and in a five regions fixed-effects model, all eight of the principal associations of interest were statistically significant, as opposed to seven in the original model, and the error probabilities were somewhat reduced for some of the other significant associations.

References
Washington, DC: International City/County Management Association Press.


---


