

**Do They All Perform Alike? An Examination of Perceived
Performance, Citizen Satisfaction and Trust with U.S. Federal Agencies**

Accepted for publication in *International Review of Administrative Sciences*

Forrest V. Morgeson III*

Research Scientist

American Customer Satisfaction Index

Ann Arbor, MI

Email: fmorgeson@theacsi.org

*Primary contact

Claudia Petrescu

Professor of Political Science

Eastern Michigan University

Ypsilanti, MI

Email: cpetrescu@emich.edu

Abstract:

What drives citizen satisfaction and trust with U.S. federal government agencies? Are these determinants constant across agencies, or do they differ? In this paper, we examine elements of citizen perceived performance as determinants of satisfaction and trust in federal agencies using a multi-year, cross-sectional, multi-agency sample of respondents. Focusing on six high-incidence federal agencies experienced by a large cross-section of American citizens, we observe differences in the determinants of satisfaction and trust, which include perceptions of the quality of the services experienced and the information provided, demographic factors, citizen expectations, and e-government adoption. We discuss these differences, with a particular focus on the implications of these findings to the practice of performance benchmarking. Given that one central objective of a variety of recent federal initiatives has been to better measure bureaucratic performance with the goal of improving the quality of services delivered to citizens through benchmarking, these findings highlight the complexity of government efforts to realize this goal. In conclusion, suggestions for improving cross-agency benchmarking are provided.

Keywords: Citizen satisfaction; Trust; Performance measurement; Performance benchmarking; American Customer Satisfaction Index (ACSI)

INTRODUCTION

What determines citizen satisfaction and trust with U.S. federal government agencies? Are these determinants constant across agencies, or do they differ? What are the implications of the answers to these questions for those attempting to compare agency performance or seek agency process improvements through performance benchmarking? Given the growing importance of performance measurement among U.S. federal agencies, and especially the practice of performance benchmarking of citizen satisfaction data (and related data measuring citizen perceptions of government performance), a more thorough and nuanced understanding of what drives citizen perceptions across diverse government agencies is vital. Seeking this more complete understanding is the central objective of this paper.

In this paper, we examine determinants of citizen satisfaction and trust in U.S. federal government agencies using a multi-year, multi-agency, cross-sectional sample of survey responses from Americans – data from the American Customer Satisfaction Index (ACSI) – and structural equation modeling statistical techniques. Focusing on six high-impact, high-incidence U.S. federal agencies experienced by a large cross-section of American citizens – the Internal Revenue Service (IRS), Medicare, Medicaid, the National Parks Service (NPS), the Social Security Administration (SSA), and the Department of Veterans Affairs (VA) – we observe several cross-agency differences in the determinants of satisfaction and trust, which include government performance as citizen perceptions of the quality of the services experienced and the information provided, demographic factors, citizen expectations, and e-government adoption. We discuss these differences, and conclude with a discussion of the relevance of these

findings to the increasingly important practice of performance benchmarking as a means for both determining relative agency performance and seeking critical process improvements.

Our findings suggest that the use of satisfaction data for the purposes of performance measurement, by elected officials and regulatory agencies (as a means of judging agency performance) or by senior executives and public managers (as a strategy for realizing improvements in agency performance), must be done carefully to avoid drawing incorrect conclusions about how well agencies are performing, or the recommended strategies for improving citizen perceptions of this performance and citizen trust. In short, we argue that differences in the determinants of citizen satisfaction and its outcomes complicate the practice of cross-agency performance benchmarking. In conclusion, suggestions for improving the practice of identifying benchmarking partners and thereby conducting more effective performance benchmarking are provided.

BACKGROUND AND RESEARCH QUESTIONS

Governments at all levels – along with a growing number of public, private and non-profit organizations – are now engaged in performance measurement and management. Performance measurement, which can be defined as the “regular measurement of the results (outcomes) and efficiency of services or programs,” comes in a variety of shapes and sizes (Hatry 2006, 3). While some government organizations continue to focus solely on internal measures of performance (e.g. performance-based budgeting, internal agency “scorecard” measurement of total workload and production, or the achievement of key strategic goals, missions and objectives, etc.) (Gilmour & Lewis

2005; Poister & Streib 1999), increasingly external performance measures are also being employed, typically taking the form of citizen satisfaction surveys or related survey data measuring citizen perceptions of government performance (Kelly & Swindell 2002; Van Ryzin et al. 2004; Wang & Gianakis 1999).

While a significant number of local and state governments in the United States have for some time measured citizen satisfaction with basic service provision, such as policing, waste disposal, and fire protection (Beck et al. 1987; Kelly & Rivenbark 2003; Kelly & Swindell 2002; Miller & Kobayashi 2000; Ostrom 1973; Van Ryzin et al. 2004), satisfaction measurement has also become commonplace among federal agencies in the United States. Driven in part by the New Public Management movement and its strong focus on performance measurement and management (Box 1999; Jones & Thompson 1999), as well as a group of relatively recent directives that mandate this type of measurement within federal agencies (such as the *Government Performance and Results Act* and the “Program Assessment Rating Tool”), it is now typical for federal agencies to analyze and publicize the results of satisfaction surveys (Behn 2003; Dull 2006; GPRA 1993; Meier & O’Toole 2009; Milakovich 2003; Osborne & Gaebler 1992; Stalebrink 2008).

The justifications for and uses of citizen satisfaction data among federal departments and agencies are many, and the basic purposes driving this practice have been understood differently (Behn 2003). Measurement of satisfaction data is valuable, some have argued, because it provides citizens with the information needed to hold government officials responsible for their performance, thereby facilitating greater democratic accountability (Kelly 2005; Kelly & Swindell 2002; Roch & Poister 2006).

Satisfaction measurement also provides a feedback loop between government agencies and citizens similar to that of the free market, a communication network that is otherwise lacking for most government agencies (Fornell et. al. 2005; Wholey & Hatry 1992). Moreover, satisfaction measurement provides public managers with the ability to gauge citizen perceptions of the quality of services (and other goods) delivered by their agency, and to acquire the information needed to reform and retool sub-optimal processes. Finally, measures of satisfaction are sometimes regarded as useful indicators or determinants of general trust in government among the population, as a proxy for public support for the political system as a whole (Bouckaert & Van de Walle 2003; Yang & Holzer 2006).

The diversity of justifications for citizen satisfaction measurement notwithstanding, perhaps the most significant rationale for this activity is its usefulness in *performance benchmarking*, providing the data needed for comparing results and outcomes across units of government or distinct government agencies (Aksim & Christophersen 2007; Ammons 1999; Bell & Morey 1994; Bruder & Gray 1994; Coe 1999; Folz 2004; Yasin 2002). Performance benchmarking, an idea borrowed from the quality management practices of the private sector that aims “to identify competitive targets which render the weak points of the benchmarking organization visible and to establish means of improvement,” is now viewed as an essential tool for government as well (Kouzmin et. al. 1999, 123). Indeed, the Office of Management and Budget’s (OMB) guidance on the “Program Assessment Rating Tool” – an initiative adopted during the administration of George W. Bush aimed at improving government service

quality and efficiency – explicitly encourages “agencies to provide the means by which their programs may be benchmarked internationally or across agencies” (OMB 2008, 74).

At its core, performance benchmarking is deemed an important government activity primarily due to its usefulness in realizing two goods. First, the ability to benchmark citizen satisfaction data across government agencies provides the means for *differentiating levels of performance* (or incremental improvement in performance, etc.) across these agencies, and the ability to render decisions about reform priorities or budget appropriations on this basis (Osborne & Plastrik 2000). That is, the ability to compare agencies in terms of citizen satisfaction performance assists public officials involved in oversight functions in distinguishing the strong-performing from the poor-performing agencies. Furthermore, the ability of public managers to benchmark satisfaction data can facilitate *inter-agency learning and organizational improvement* (Keehley et. al. 1996). From this perspective, the availability of satisfaction data helps agencies identify their most successful counterparts and – through a critical analysis of process differentiation – adopt the best practices responsible for satisfaction success. By comparing themselves with agencies providing superior satisfaction, so it is thought, agencies can improve their own satisfaction.

Yet the perceived benefits attained through performance benchmarking of citizen satisfaction data notwithstanding, there are difficulties with efforts to interpret this data, difficulties that must be seen to confound the benchmarking-related objectives mentioned above. Stating the problem simply, distinct government agencies and programs serve very different segments of the public, perform very different missions, and administer very different programs. However, whether employed by administrators searching for

best practices or by elected representatives attempting to critically compare service performance across government units, a considerable degree of *inter-agency homogeneity* must be assumed. That is, for performance benchmarking to serve its purpose, those doing it must assume that agencies are similar – that the selected “benchmarking partners” are comparable – in how they interface with their citizen-customers and in what drives satisfaction among these disparate groups of citizens.

As they relate to the practice of performance benchmarking, these assumptions and the problems they introduce raise several questions: What determines citizens’ satisfaction and trust across distinct federal government agencies? Are these determinants constant across agencies, or do they differ? Most critically, if these determinants do differ, what are the implications for those attempting to compare cross-agency performance or seek process improvements through benchmarking? Without answers to these basic questions, any conclusions drawn from performance benchmarking could be dismissed as inapplicable, due to inherent and incommensurable differences between agencies that may make poor benchmarking partners. Indeed, the benchmarking of agency performance data – regarding satisfaction or any measure of citizen perceptions of performance, for that matter – will likely face criticisms regarding the comparability of selected benchmarking partners, criticisms that could challenge the core value of the benchmarking process.

Surprisingly, relatively little is known about the determinants of citizen satisfaction and trust across various types of U.S. federal agencies. The extant literature on citizen satisfaction in the U.S has thus far focused largely on local government services (Kelly & Rivenbark 2003; Kelly & Swindell 2002; Miller & Kobayashi 2000;

Van Ryzin et al. 2004), while a handful of studies have examined the relationship between performance, satisfaction and trust (at various levels of government) in other national contexts (Bouckaert & Van de Walle 2003; Bouckaert, Van de Walle & Kampen 2005; Heintzman & Marson 2005; Van de Walle & Bouckaert, 2003; Vigoda-Gadot 2007; Vigoda-Gadot & Yuval 2003). This paper intends to fill this gap in the literature, and represents (to our knowledge) a first attempt at outlining the determinants of citizen satisfaction and trust in U.S. federal government agencies using a multi-year, cross-agency sample of respondents. As such, results from this analysis should provide essential guidance to those engaged in cross-agency satisfaction benchmarking at the federal level.

The remainder of this paper proceeds as follows. In the next section, we outline a conceptual model of government performance grounded in the “performance-satisfaction-trust” linkage now common in both the literature and in measurement practice. This model includes measures of citizen satisfaction, trust, and perceived government performance (service quality, information delivery), among other variables. The following section describes the data – provided by the American Customer Satisfaction Index – we will use to test our models. Thereafter, we describe the statistical methods used to analyze the data, and present the results of our analysis. The methods employed include structural equation modeling of the aggregate dataset, as well as sub-sample modeling of each of the agencies mentioned earlier, facilitating the inter-agency comparisons central to this study. Finally, the concluding section considers the implications of our results, specifically focusing on their relevance for the growing number of federal agencies both collecting citizen satisfaction data (and other external

performance measurement data) and using this data to make vital operational decisions through performance benchmarking.

GOVERNMENT PERFORMANCE, CITIZEN SATISFACTION AND TRUST

Performance measurement through satisfaction measurement is an area of intense interest for many organizations. The private sector took the lead in developing conceptual models of consumer-oriented performance measurement (for some examples, see: Fornell et al. 1996; Oliver 1997; Szymanski & Henard 2001), but these models have increasingly been adapted and applied to citizen experiences with government, and used to investigate the formation of citizen attitudes and citizen satisfaction with this class of services (Bouckaert & Van de Walle 2003; Bouckaert, Van de Walle, and Kampen 2005; Fornell et al. 1996; Heintzman & Marson 2005; James 2009; Morgeson 2010; Morgeson et al. 2010; Van de Walle & Bouckaert, 2003; Van Ryzin et al. 2004; Van Ryzin 2007; Vigoda-Gadot & Yuval 2003). Much like the private sector, the language and proverbs of satisfaction measurement pervade many government organizations today.

While conceptual and empirical models of citizen satisfaction with government services continue to vary (i.e. model specifications defining what influences the formation of satisfaction attitudes, the ideal proxy for or measure of satisfaction, the primary outcomes of citizen satisfaction, and so forth), a majority of academic work on the topic adopts some version of what can be called the “performance-satisfaction-trust” model (Bouckaert & Van de Walle 2003; Bouckaert, Van de Walle & Kampen 2005; Fornell et al. 1996; Heintzman & Marson 2005; James 2009; Morgeson 2010; Morgeson et al. 2010; Van de Walle & Bouckaert, 2003; Van Ryzin et al. 2004; Van Ryzin 2007;

Vigoda-Gadot & Yuval 2003). This linkage, represented in Figure 1 below, adopts a simple yet theoretically concise and well-tested approach to explaining the determinants and outcomes of citizen satisfaction, identifying citizen perceptions of government performance as the primary determinants of citizen satisfaction, and in turn positions satisfaction as the main predictor of outcomes such as citizen trust.

--FIGURE 1 ABOUT HERE--

In what follows, we outline a conceptual model grounded in this performance-satisfaction-trust approach, based on both earlier findings from the literature and the data at our disposal (described in the next section) – a model we test below using structural equation modeling (SEM) techniques. However, given the relatively advanced state of these types of models in the literature, we eschew some theoretical background when defining this model, and proceed to a discussion of the key variables we include in the model for analysis. Nevertheless, the model we intend to test is represented graphically in Figure 2 below, and described in greater detail in what follows.

--FIGURE 2 ABOUT HERE--

Government Performance

Government performance includes many things. Indeed, one perennial difficulty in measuring government performance across agencies, units, or levels of

government is found in defining attributes of performance that apply equally well across them all (Van Ryzin 2007). Nevertheless, a few generic aspects of government performance apply to most citizen experiences with government. For instance, the improved and transformed delivery of *customer service* has been recognized as a focus for improving government performance writ large (and particularly administrative performance) for some time (Osborne & Gaebler 1992; Poister & Henry 1994; Swiss 1992; Wagenheim & Reurink 1991). It is now generally agreed – by both academics and practitioners – that a central feature of “better government” is the provision of better customer service.

In our model, as in much of the earlier literature, we identify customer service as a multi-dimensional construct (Heintzman & Marson 2005; Morgeson 2010; Morgeson et al. 2010; Van Ryzin et al. 2004; Vigoda-Gadot & Yuval 2003). Specifically, we propose two core (but general) elements of government performance as customer service commensurate with both theory and the findings of prior research: the citizen’s perceptions of (1) the quality of the services received, and (2) the information disseminated by government agencies. Operationalized more concretely in our model (see below), we broadly define these attributes of agency performance as the ease of obtaining services and the timeliness and efficiency of the delivery of those services, and the ease of accessing information and the clarity of information accessed from an agency. For both variables we would anticipate positive relationships with satisfaction, with perceptions of stronger service quality leading to stronger citizen satisfaction (Fornell et al. 1996).

In addition, as determinants of satisfaction we include two variables not specifically indicative of government performance, but strongly likely to co-vary with perceptions of government performance and satisfaction. First, we include a measure of prior citizen expectations – that is, what citizens anticipate receiving from an agency prior to their experience – as a predictor of satisfaction in the model. A large body of literature suggests (and has found) a connection between expectations of an experience with government – expectations that tend to “frame” the experience – and a range of resultant citizen perceptions (James 2009; Morgeson & Mithas 2009; Morgeson 2010; Morgeson et al. 2010; Reisig & Chandek 2001; Roch & Poister 2006; Van Ryzin 2004; Van Ryzin 2006; Van Ryzin 2007; Van Ryzin et al. 2004). More specifically, a positive relationship between expectations and satisfaction – as expectations increase, so too does satisfaction – has been found across different types of government services (DeHoog, Lowery & Lyons 1990; Fornell et al. 1996; James 2009; Kelly & Swindell 2002; Roch & Poister 2006; Van Ryzin 2006), and we would likewise anticipate a positive relationship in this context.

Second, we include in the model as a predictor of satisfaction a variable relating to the communications mode through which the citizen interacted with a federal agency. More specifically, this variable tests the relationship between the citizen’s adoption of e-government, as opposed to traditional modes of contact (i.e. telephone, mail, visiting an office, etc.), and the citizen’s resultant satisfaction. E-government, a mode of citizen-to-government contact founded in information and communications technologies, has been seen as a means for transforming

government-citizen interactions dramatically, and thus is likely to impact citizen satisfaction with federal government (Chadwick & May 2003; Morgeson & Mithas 2009; Morgeson 2010; Morgeson et al. 2010; Thomas & Streib 2003; Tolbert & Mossberger 2006). We propose a positive relationship between e-government and citizen satisfaction, with e-government users expressing significantly higher satisfaction than off-line users.

Finally, we include four demographic predictors of satisfaction, both as control variables and as possibly significant predictors of satisfaction in their own right – age, education, income and gender. Because federal agencies often serve very different segments of the population (e.g. some agency customer groups are older, less affluent and less well educated than others, such as those of the Social Security Administration), it is essential to control for demographic factors, particularly when examining an aggregate, cross-agency model (as we do in our statistical analysis below) where demographic differences are likely to be more pronounced. Furthermore, prior research has shown that demographic factors impact both experienced and reported satisfaction across a variety of contexts (Bryant & Cha 1996; DeHoog et. al. 1990; Van Ryzin et al. 2004), and are likely to do so vis-à-vis citizen satisfaction with federal government as well.

Citizen Satisfaction

As described and represented graphically above, within performance-satisfaction-trust models of government performance, satisfaction is typically positioned as the key moderating variable. Inarguably, satisfaction represents a

predominant concept in government performance measurement research, with studies of citizen satisfaction with government services having proliferated in recent years (James 2009; Morgeson & Mithas 2009; Morgeson 2010; Morgeson et al. 2010; Roch & Poister 2006; Van Ryzin et al. 2004a; Welch et al. 2005). While different measures of satisfaction have been developed, if measured as a multi-dimensional, cumulative concept – as is done here – it is typically employed to reflect the sum-total of the consumer’s sense of fulfillment with his or her experience (Fornell et al. 1996). In most models, therefore, satisfaction is introduced as a mediator between core consumer characteristics and their attitudes regarding government performance and their resulting (or future) perceptions and behaviors, and we adopt this convention here too.

Trust and Confidence

Finally, we turn to outcomes of performance and satisfaction. In most of the prior literature on the topic, citizen trust is viewed as the most crucial outcome of citizen satisfaction (Heintzman & Marson 2005; Roch & Poister 2006; Van de Walle & Bouckaert, 2003; Van de Walle, Van Roosbroek & Bouckaert 2008; Van Ryzin 2004; Van Ryzin 2007; Van Ryzin et al. 2004; Vigoda-Gadot & Yuval 2003). From this perspective, and similar to increased revenues, positive word-of-mouth, or customer loyalty for private sector companies, trust is the essential value thought to be produced by increased citizen satisfaction (Van Ryzin 2007).

As such, we include two distinct perspectives on trust in federal government as outcomes in our model, as trust is a concept that is open to a wide variety of

different definitions and interpretations (Thomas 1998). First, we include a measure of trust in the federal government in Washington D.C. overall. This high-level perspective on trust, which focuses on citizen perceptions of all of the institutions of the federal government in Washington, is a common metric within studies investigating trust in government (Morgeson & Mithas 2009; Morgeson 2010; Morgeson et al. 2010; Tolbert & Mossberger 2006; Welch et al. 2005). The second measure of trust included is the respondent's confidence that the specific federal agency experienced will do a good job providing services in the future. That is, rather than focusing on trust in the federal government as a whole, this particularized perspective on confidence focuses on a single federal agency and the citizen's confidence that that agency will perform well delivering services in the future.

In our model, we identify one determinant of the measure of confidence in the agency experienced, citizen satisfaction. Satisfaction is hypothesized to positively influence confidence in an agency, as citizens more satisfied with their experience with an agency should also be more confident in that agency's future performance. Moreover, we include two predictors of the generalized measure of trust in Washington overall. For basically the same reasons, we hypothesize a positive relationship between satisfaction and trust in the federal government as a whole; citizens who are more satisfied with their experiences with a federal agency should have both more confidence in that particular agency and greater trust in the government overall. We also propose a positive relationship between confidence in an agency and trust in Washington. The reason for this relationship is that feelings

about aspects or parts of an entity should influence perceptions of the entire entity (Hetherington 1998). Similarly, we suggest that experiences with particular agencies and confidence in those agencies will positively determine generalized trust in the federal government as a whole.

DATA, STATISTICAL METHODS, AND RESULTS

The central objective of this study is to compare determinants (i.e. citizen perceptions of government performance) of citizen satisfaction and citizen trust across multiple U.S. federal agencies. More specifically, we are interested in testing the conceptual model outlined above both in the aggregate and across the distinct federal agencies included in our sample, and in examining similarities and differences in the results. This endeavor should provide insight into the existence and/or extent of the differences in what drives satisfaction across agencies, and the potential complications these differences (if they do exist) pose. While a variety of statistical methods would suffice in achieving this objective, here we employ structural equation modeling (SEM) techniques. We begin by describing the data we utilize, followed by the results from the various elements of SEM modeling – as well as some of its features and advantages over alternative methods – below.

Data

The data we use to test the model outlined above was collected as part of the American Customer Satisfaction Index (ACSI) study.¹ As one element of measuring national consumer satisfaction in the United States, each year the ACSI interviews a

¹For more on the ACSI project, visit: www.theacsi.org.

sample of approximately 1500 Americans regarding their experiences with federal government services, utilizing a general and standardized survey instrument. To maximize the sample available for analysis, we utilize data from two annual ACSI studies, the first collected in July and August of 2008, and the second in July and August of 2009. Random-digit-dial probability sampling and multiple call-back and refusal conversion techniques were used to identify respondents and collect a representative sample. Computer-assisted telephone interviewing (CATI) (administered by an experienced professional market research firm) was the interviewing method used.

Regarding the survey, all of the substantive items included in the questionnaire (i.e. on information, services, satisfaction, trust, etc.) were generalized for applicability across different types of federal agency experiences. Prior to interviewing, respondents were screened for recent personal experience with a federal agency before being deemed eligible to participate in the study and complete the survey. The respondents in the sample indicated having some direct experience with a federal government department, agency or program over the past year (excluding the U.S. Postal Service), and were asked about that specific experience during interviewing. Those indicating contact with more than one federal agency over the prior 12 months were asked to identify and respond to the survey items in regards to the agency they contacted “most recently.” Within the sample, interviewees identified a total of 57 distinct federal agencies or departments, ranging from the Department of Agriculture to the Department of Veterans Affairs. Table 1 summarizes the questions included in the survey that we use to operationalize the model.

--INSERT TABLE 1 HERE--

The 2008 and 2009 ACSI federal government samples include nearly 3000 cases/citizen interviews in total, although we eliminate roughly one-third of these cases prior to analysis. Because we are interested in examining similarities and differences in citizen experiences across particular federal agencies, and because not all agencies were identified as often by respondents in the sample (i.e. some agencies were “low incidence” and thus had very small samples, while conversely others were overrepresented in the sample), a majority of the agency-level samples were too small to include in our analysis. For purposes of statistical reliability, we only include those federal agencies in the final sample that had at least 100 respondents (in total) across 2008 and 2009. After removing the remaining cases from the dataset, the final sample includes 1884 respondents across six federal agencies. Table 2 provides descriptive statistics for the full sample for each variable included in our analysis.

--INSERT TABLE 2 HERE--

Finally, Table 3 shows the distribution of the final sample by federal agency, while Appendix A includes abbreviated descriptive statistics for each agency sub-sample. Below, we describe our statistical analysis of the data and the results of this analysis.

--INSERT TABLE 3 HERE--

Latent Constructs

Given the model outlined above and the data at our disposal, our structural equation models will include both observed and unobserved (or latent) factors. More specifically, we specify multivariable constructs for three of the variables included in the model in Figure 2 – “Service,” “Information,” and “Satisfaction.” The survey items included in these latent variables are identified in Tables 1 and 2 above (via the standard “ η ” notation). Therefore, as is the convention in SEM, we begin by examining the properties of these three latent constructs, both as they appear in the “overall” model (which analyzes the pooled sample from all six agencies), and for each agency sub-sample analyzed separately.

Table 4 provides statistics for the “Satisfaction,” “Information” and “Service” latent constructs for all seven models (i.e. the aggregate and six sub-sample models), with standardized factor loadings (λ 's), explained variance for each item (R^2), and a measure of construct validity (Cronbach's alpha).

--INSERT TABLE 4 ABOUT HERE--

The standardized factor loadings and the R^2 statistics indicate that across all models and all three latent variables, the manifest variables load strongly and significantly on the constructs, providing evidence of convergent validity. Furthermore, the Cronbach's alpha statistics, ranging across the models and latent variables from $\alpha = 0.796$ to $\alpha = 0.930$, indicates strong construct validity; an alpha value greater than

0.700 is generally considered sufficient (and 0.800 very good) for establishing construct validity in SEM (Hair 1998).

Structural Equation Models

Turning next to the estimation of our full structural models, including both the observed and unobserved latent variables outlined in our conceptual model earlier, we obtained maximum likelihood (ML) estimates of the model coefficients. To do so we utilize the full information maximum likelihood estimation (FIML) method, a technique that estimates multiple ML equations simultaneously and includes information on the conditional covariances among all model variables in the parameter estimates (Arbuckle 2006). The FIML technique provides reliable and robust estimates for both continuous and interval-level variables, while also accommodating models containing both observed and unobserved variables (Arbuckle 2006). Although OLS estimators should prove similar to ML estimators for large, multivariate normal samples (Bollen 1989), two conditions descriptive of our data, the FIML procedure has been shown to produce more efficient estimators for data with missing observations (Enders 2001; Enders & Bandalos 2001). Given that several of our variables have some missing observations (with the most extreme case being the income variable), the FIML approach is preferable to alternative methods. The zero-order correlation matrices for all observed model variables are provided in Appendix B.

Results for the FIML analysis, including goodness-of-fit measures, standardized direct effects, standard errors, significance of the coefficients, and measures of explained variance (squared multiple correlations or R^2), first for the pooled-sample model and then

(in abbreviated form) for the agency sub-sample models, are presented below in Tables 5 and 6.

--INSERT TABLES 5 AND 6 ABOUT HERE--

Results

We begin with a discussion of the results in Table 5. First, the data show a good fit to the model, providing empirical support for the conceptual model we have specified. The root mean square error of approximation statistic indicates a good model fit (RMSEA=0.045); an RMSEA below 0.05 is interpreted to represent a “close” model fit, with a lower score better still (Arbuckle 2006). Further, each of the baseline measures of fit, which compare our specified model to a perfectly-fitting (or saturated) model, register above the 0.90 threshold, reflecting an adequate to very good fit (NFI=0.979; RFI=0.958; IFI=0.984; TLI=0.966; CFI=0.984) (Arbuckle 2006). Finally, the model has a favorable χ^2 to degree of freedom ratio (4.808), with a ratio of 5.00 or less considered satisfactory.

Turning now to analysis of the standardized direct effects provided in Table 5, and beginning with the four demographic factors included in the model, only the education estimate is a significant predictor of satisfaction at the $p < 0.05$ level ($\beta = -0.036$), although its effect is relatively small. That is, while age, income and gender have no significant impact on citizen satisfaction with a federal agency, education has a significant negative effect, with better-educated citizens expressing lower satisfaction with their experience with a federal agency. These results suggest that the composition of

an agency citizen-customer cohort – at least at the aggregate level – has little to do with federal agency satisfaction, outside the small effect of education.

Furthermore, both e-government usage and prior expectations have significant effects, although one of these variables shows a directionally different relationship than anticipated. Expectations are significantly and positively related to satisfaction ($\beta = 0.177$), showing that citizens with higher expectations tend to experience higher satisfaction. Thus much like the findings from studies regarding other levels of government and types of government services, it would appear that expectations positively frame satisfaction judgments with federal government, as we hypothesized (James 2009; Morgeson et al. 2010; Roch & Poister 2006; Van Ryzin 2004; Van Ryzin 2006; Van Ryzin et al. 2004). However, contrary to our expectations e-government usage is found to have a significant but *negative* effect on satisfaction ($\beta = -0.042$). These results suggest that citizens who adopt e-government (rather than a traditional off-line means for interacting with an agency) are not experiencing the stronger satisfaction that most advocates of e-government have predicted. While not an unprecedented finding (Morgeson et al. 2010), one we discuss more fully below when considering the agency sub-sample models, it certainly challenges the mostly-accepted narrative within federal government that the move towards e-government is a means for immediately improving citizen perceptions of government services.

The final two determinants of satisfaction, the latent “Service” and “Information” variables, show the anticipated relationships. That is, both are significant, strong, and positive predictors of satisfaction, with Service ($\beta = 0.581$) showing a comparatively stronger effect than Information ($\beta = 0.224$). These results, which we discuss more fully

in reference to cross-agency benchmarking, suggest that while the provision of easily obtainable and clear information is important, easily obtainable and timely and efficient services are most vital in determining citizen satisfaction for federal agencies, at least according to citizens. Thus in the aggregate and in general, federal agencies should focus on the provision of services (primarily) and information (secondarily) as a means to optimize citizen satisfaction, rather than simply attempting to manage expectations or shift citizens to alternative modes of contact (i.e. e-government).

Focusing now on the two outcomes of satisfaction – confidence in the agency experienced and trust in the government as a whole – the results show that citizen satisfaction is in fact a significant predictor of both confidence in an agency ($\beta = 0.797$) and trust in the government in Washington overall ($\beta = 0.093$). As anticipated, citizens who experience higher satisfaction with a federal agency both have more confidence in that agency's future performance and have greater trust in the federal government overall. Further, confidence in an agency has a strong effect on trust in Washington ($\beta = 0.332$), suggesting that satisfaction can both directly and indirectly (through higher confidence in particular agencies) improve overall trust in the government. If nothing else, these results would seem to support the underlying purpose of performance measurement among federal agencies: measuring and managing the satisfaction of citizens and the determinants of satisfaction provides an effective tool for improving citizen trust in government.

We turn now to the estimates for the agency sub-sample models in Table 6, with a particular interest in comparing these results across agencies. Generally speaking, while some similarities do exist, these results suggest that citizen satisfaction is *determined by*

different factors across different federal agencies, with differences also appearing in the model outcomes.

First, and similar to the aggregate model, the demographic variables have little effect across the sub-sample models, with almost no evidence to indicate that the composition of the customer population dramatically impacts satisfaction for any agency included in the sample. Similarly, e-government adoption is a predominantly insignificant factor across these models, with only a single agency – NPS – showing a significant (but again, unexpectedly negative) relationship between e-government adoption and satisfaction. While these findings continue to undermine the notion that e-government is already drastically improving satisfaction with federal agencies, it does show that the negative relationship between e-government use and citizen satisfaction is limited to only this one agency (at least within our sample) rather than being a systematic phenomenon.

Furthermore, for three agencies – IRS, SSA and Medicare – prior expectations are found to be a positive and significant predictor of satisfaction, while for the other three – NPS, Medicaid, and VA – expectations is an insignificant factor. This finding suggests that while for some agencies the management of citizens' prior expectations is an important tool in optimizing and understanding citizen satisfaction, for others it is less so.

Most importantly, at least from the perspective of those seeking strategies for improving agency satisfaction and an understanding of what drives satisfaction within and across agencies, differences in the strength of the “Service” and “Information” latent variables across the models are noteworthy. For three agencies – Medicare, Medicaid and VA – the “Service” predictor is particularly strong, while the “Information” latent variable is an insignificant determinant of satisfaction. Thus for these agencies, the ease

of obtaining services and the timeliness and efficiency of the delivery of those services is paramount in determining citizen satisfaction. Given the primary citizen-facing mission of these agencies – to deliver free benefits (and more specifically, predominantly health care benefits and related services) to needy or otherwise eligible citizens – the predominance of service delivery in this instance is not surprising.

For two other agencies – SSA and NPS – both “Services” and “Information” are significant determinants of satisfaction, but “Services” dominates in terms of relative importance. For these two agencies, both the delivery of services and information are important to citizen satisfaction, but the ease of obtaining services and the timeliness and efficiency of the delivery of those services is more important. Lastly, for the IRS model “Services” and “Information” are both significant determinants of satisfaction, and the two are more roughly equal in their impact on satisfaction. This finding is not unexpected either; as the IRS provides a service as a regulatory process that demands the efficient delivery of information vital to the citizen experiencing this process (i.e. information about the tax-filing process), this finding makes sense. Taken on the whole, these results suggest that while some agencies can focus almost exclusively on the delivery of timely, efficient and easily accessible services (and little on information provision), others must focus more equally on the two activities to improve citizen satisfaction.

Finally, for all of the sub-sample models satisfaction is a significant – and similarly strong – predictor of confidence in an agency. As expected, agencies that deliver stronger satisfaction also create citizen-customers who are more confident they will receive good service in the future. Moreover, while confidence in an agency is a significant and strong predictor of trust in the government overall for most of the sub-

sample models (with the exception of VA), for only the IRS sub-sample does satisfaction positively and significantly determine trust in the government in Washington. This final result suggests that improvements in citizen satisfaction with their experiences with one prominent and high-profile federal agency – the IRS – will yield the greatest benefits in terms of citizen trust in federal government overall, a result that should not be lost on the federal government as it struggles with low levels of citizen trust.

Taken together, these results show that citizens demand different things from different federal agencies, and that citizen satisfaction is determined in different ways across these agencies. This result is perhaps unsurprising; given that the agencies examined here run the gamut from primarily benefits-delivering agencies (SSA, Medicare, Medicaid, VA) to regulatory enforcement agencies (IRS), and other types of service-deliverers (NPS), it is fair to assume that citizens would require different things from these agencies. Yet these results confirm this intuition and show that different agencies do not “perform alike,” and in some sense cannot and should not perform alike. In turn, this conclusion complicates the practice of performance benchmarking among agencies as a means for both judging individual agency performance and discerning best practices across agencies; after all, comparing “mostly different” entities is a more complex task, especially when cross-agency learning is the goal of the comparison. In effect, these results indicate that benchmarking must always be done within the context of a clear understanding of how citizen perceptions differ across agencies. We will elaborate on this finding and its implications in the concluding section.

STUDY LIMITATIONS AND CONCLUSIONS

The findings in this study provide an overview of the relationship between determinants of satisfaction – or citizen perceptions of government performance – and outcomes of satisfaction including both trust and confidence in federal government, at both an aggregate level and among several important federal agencies. Most significantly, these results provide insight into the differences in what determines satisfaction across federal agencies, indicating that citizens interacting with different agencies are likewise satisfied differently. These findings carry significant implications for federal agencies involved in citizen-centered performance measurement and management, and particularly those involved in performance benchmarking.

But before proceeding to consider the implications of our findings for benchmarking, a few caveats regarding some limitations of this study, and how these limitations might be overcome in future research, are worth mentioning. First, our results are based on two samples of citizens collected in 2008-2009 who experienced a relatively small number of federal government agencies and programs, and we must therefore be careful not to generalize to the entire universe of experiences with the (very diverse) assortment of federal government services based on these results. Additionally, our conceptual and empirical model, while grounded both in theory and in prior research, does not include the full range of possible determinants (or outcomes) of satisfaction, and therefore alternative explanations for our findings cannot be entirely eliminated. Future research looking at a longer time period, a more diverse assortment of federal agencies, and a larger cohort of variables would be useful in overcoming these shortcomings.

These limitations notwithstanding, this study reveals several interesting findings about how citizen perceptions of government performance influence satisfaction across diverse federal agencies. Yet perhaps the most important implications of the findings in this study – at least for the practitioner engaged in the practice – relate to performance benchmarking. To be sure, these findings must be seen to complicate the practice of *identifying benchmarking partners*, both for those doing so internal to agencies and seeking best practices and process improvements, and for those external to agencies attempting to compare relative agency performance. In short, if citizen perceptions of government performance drive satisfaction differentially across federal agencies, as this study suggests, an inappropriate selection of benchmarking partners could lead to both incorrect conclusions about how agencies are performing relative to one another (by external observers), or about strategies for improving processes/services and optimizing citizen satisfaction (for internal actors).

Take, for example, a simple exercise derived from the results presented here. These results suggest that among the agencies included in the sample, in terms of the overall satisfaction variable, the IRS scores lowest on satisfaction while NPS scores highest (see Appendix C). A simplistic interpretation of this result derived from performance benchmarking by elected officials or policymakers interested in knowing “who is doing better” could lead to an incorrect – and potentially damaging – conclusion: that NPS’s superior satisfaction performance is grounds for selecting this agency as a model for seeking improvements in satisfaction with the IRS. Yet a more thorough analysis of the determinants of satisfaction and a list of

recommended actions based on these results, like that found below in Table 6, shows that citizen-customers of these two agencies are satisfied in fundamentally different ways. As a result, using NPS as a guide for improving satisfaction with the IRS could have little (or worse still, a negative) effect.

--INSERT TABLE 6 ABOUT HERE--

However, the recognition of these shortcomings should not, in our estimation, lead to a wholesale rejection of the usefulness or practical applicability of benchmarking across federal agencies. Rather, they provide a cautionary tale about the importance of the careful selection of benchmarking partners. Indeed, difficulties in identifying appropriate organizational benchmarking partners have long been recognized, particularly in the public sector (Ammons 1999; Bell & Morey 1994; Bruder & Gray 1994; Coe 1999; Folz 2004; Yasin 2002). As Folz (2004, 210) succinctly points out, “In practice, however, scholars and practitioners concur that it is difficult to make performance comparisons... and to find benchmarking partners that employ and collect data on performance measures that are truly comparable, reliable, and useful for making decisions about service improvements.” The findings from this paper would seem to reinforce this point vis-à-vis performance benchmarking of citizen satisfaction data across federal agencies.

Yet the kind of analysis undertaken above, we would suggest, not only reveals the difficulties in performance benchmarking across federal agencies, but

also provides one solution for the complex process of selecting benchmarking partners. That is, one possible alternative technique for benchmark partner selection can be found in what might be called “benchmark partner selection by determinants.” On this approach, the selection of a benchmarking partner would be based not only on straightforward considerations – Is the agency a top performer? Is data available for comparative analysis? – but first and foremost on an analysis of similarities in determinants of citizen satisfaction, similar to what we have done here. Through this type of benchmark partner selection, agencies most similar in how they interact with and satisfy citizens could be determined, and thus most appropriate benchmarking partners selected.

In conclusion, in this study we have attempted to illuminate both what determines citizen satisfaction across diverse federal agencies and how differences in these determinants might influence the practice of performance benchmarking. Based on these results, it is recommended that performance benchmarking be preceded by a cross-agency examination of determinants of satisfaction, using something like the model we have proposed in this paper. Benchmarking citizen satisfaction data in isolation, without a thorough understanding of how satisfaction is differently determined, could potentially result in the errant choice of benchmarking partners, and thus in misleading or incorrect conclusions. But with the added insight of a model such as this, those involved in the oversight of federal agencies can best utilize performance benchmarking to judge how (and why) some agencies perform better than others in terms of citizen satisfaction, and which agencies might provide a useful and applicable model for seeking best practices capable of improving agency performance.

REFERENCES

- “Program Assessment Rating Tool Guidance No. 2008-01,” Office of Management and Budget. January 29, 2008. Accessed online at: http://www.whitehouse.gov/omb/assets/performance_pdfs/part_guid_2008.pdf.
- Aksim, J., A. Johnsen, and K.A. Christophersen (2007). “Factors behind Organizational Learning from Benchmarking: Experiences from Norwegian Municipal Benchmarking Networks,” *Journal of Public Administration Research and Theory*, 18(2).
- Ammons, D.N. (1999). “A Proper Mentality for Benchmarking,” *Public Administration Review*, 59(2), 105-109.
- Arbuckle, J. (2006). *AMOS 7.0 User’s Guide*. Spring House, PA: AMOS Development Corporation.
- Beck, P., H. Rainey, K. Nicholls, and C. Traut (1987). “Citizen Views of Taxes and Services: A Tale of Three Cities,” *Social Science Quarterly*, 68(2), 223–43.
- Behn, R. (2003). “Why Measure Performance? Different Purposes Require Different Measures,” *Public Administration Review*, 63(5), 586-606.
- Bell, R.A. and R.C. Morey (1994). “The Search for Appropriate Benchmarking Partners: A Macro Approach and Application to Corporate Travel Management,” *Omega*, 22(5), 477-490.
- Bollen, K.A. (1989). *Structural Equations with Latent Variables*. N.Y., N.Y.: John Wiley & Sons.
- Bouckaert, G. and S. Van de Walle (2003). “Comparing Measures of Citizen Trust and User Satisfaction as Indicators of “Good Governance,” *International Review of Administrative Sciences*, 69(3), 329-343.
- Box, R.C. (1999). “Running Government Like a Business: Implications for Public Administration Theory and Practice,” *American Review of Public Administration*, 29(1), 19-43.
- Bruder, K.A. and E.M. Gray (1993). “Public Sector Benchmarking: A Practical Approach,” *Public Management*, 76(9), S9-S14.
- Bryant, B.E. and J. Cha (1996). “Crossing the Threshold.” *Marketing Research*, 8(4), 20-28.
- Chadwick, A. and C. May (2003). “Interaction between States and Citizens in the Age of the Internet: “e-Government” in the United States, Britain and the European

Union,” *Governance: An International Journal of Policy and Administration*, 16(2), 271-300.

Coe, C. (1999). “Local Government Benchmarking: Lessons from Two Major Multigovernment Efforts,” *Public Administration Review*, 59(2), 110-123.

DeHoog, R., D. Lowery, and W. Lyons (1990). “Citizen Satisfaction with Local Governance: A Test of Individual, Jurisdictional and City-Specific Explanations,” *Journal of Politics*, 52, 807-837.

Dull, M. (2006). “Why PART? The Institutional Politics of Presidential Budget Reform,” *Journal of Public Administration Research and Theory*, 16(2), 187-215.

Enders, C.K. (2001). “The Performance of the Full Information Maximum Likelihood Estimator in Multiple Regression Models.” *Educational and Psychological Measurement*, 61(5), 713-740.

Enders, C.K. and D.L. Bandalos (2001). “The Relative Performance of Full Information Maximum Likelihood Estimation for Missing Data in Structural Equation Models.” *Structural Equation Modeling*, 8(3), 430–457.

Folz, D.H. (2004). “Service Quality and Benchmarking the Performance of Municipal Services,” *Public Administration Review*, 64(2), 209-220.

Fornell, C., D. Van Amburg, F. Morgeson and B. Bryant. *The American Customer Satisfaction Index at 10 Years*. 2005. Ann Arbor, MI: The Stephen M. Ross School of Business.

Fornell, C., M.D. Johnson, E.W. Anderson, J. Cha and B.E. Bryant (1996). “The American Customer Satisfaction Index: Nature, Purpose and Findings.” *Journal of Marketing*, 70(1), 3-14.

Gilmour, J.B. and D.E. Lewis (2005). “Assessing Performance Budgeting at OMB: The Influence of Politics, Performance, and Program Size.” *Journal of Public Administration Research and Theory*, 16(2), 169-186.

Hatry, H.P. (2006). *Performance Measurement: Getting Results*. Washington, D.C.: Urban Institute Press.

Heintzman, R., and B. Marson. 2005. People, Service and Trust: Is There a Public Sector Value Chain?” *International Review of Administrative Sciences*, 71(4), 549-575.

Hetherington, M.J. (1998). “The Political Relevance of Political Trust,” *The American Political Science Review*, 92(4), 791–808.

James, O. (2009). "Evaluating the Expectations Disconfirmation and Expectations Anchoring Approaches to Citizen Satisfaction with Local Public Services," *Journal of Public Administration Research and Theory*, 19(1), 107-123. 107-123

Jones, L.R., and F. Thompson (1999). *Public Management: Institutional Renewal for the Twenty-First Century*. Stamford, CT: JAI Press.

Keehley, P.S., S. Medlin, S. MacBride and L. Longmire (1996). *Benchmarking for Best Practices in the Public Sector*. San Francisco, CA: Jossey-Bass.

Kelly, J. (2005). "The Dilemma of the Unsatisfied Customer in a Market Model of Public Administration," *Public Administration Review*, 65(1), 76-84.

Kelly, J. and D. Swindell (2002). "A Multiple Indicator Approach to Municipal Service Evaluation: Correlating Performance and Citizen Satisfaction across Jurisdictions," *Public Administration Review*, 62(5), 610-621.

Kelly, J.M. and D. Swindell (2002). "Service Quality Variation across Urban Space: First Steps towards a Model of Citizen Satisfaction," *Journal of Urban Affairs*, 24(3), 271-288.

Kelly, J.M., and W.C. Rivenbark (2003). *Performance Budgeting for State and Local Government*. Armonk, N.Y.: M.E. Sharpe.

Kouzmin, A., E. Loffler, H. Klages and N. Kakabadse (1999). "Benchmarking and Performance Measurement in Public Sectors," *International Journal of Public Sector Management*, 12(2), 121-144.

Meier, K.J. and L.J. O'Toole Jr (2009). "The Proverbs of New Public Management: Lessons From an Evidence-Based Research Agenda," *The American Review of Public Administration*, 39(1), 4-22.

Milakovich, M.E. (2003). "Balancing Customer Service, Empowerment and Performance with Citizenship, Responsiveness and Political Accountability," *International Public Management Review*, 4(1), 61-83.

Miller, T.I., and M. Kobayashi (2000). *City Surveys: How to Do them, How to Use Them, and What They Mean*. Washington, D.C.: ICCMA.

Morgeson III, F.V. and S. Mithas (2009). "Does E-Government Measure up to E-Business? Comparing End-User Perceptions of U.S. Federal Government and E-Business Websites," *Public Administration Review*, 69(4), 740-752.

Morgeson III, F.V. (2010). "Comparing Determinants of Website Satisfaction and Loyalty across the e-Government and e-Business Domain." *Electronic Government: An International Journal*, Forthcoming.

Morgeson III, F.V., D. VanAmburg and S. Mithas (2010). "Misplaced Trust? Exploring the Structure of the E-Government-Citizen Trust Relationship." *Journal of Public Administration Research & Theory*, Forthcoming.

Oliver, R.L. (1997). *Satisfaction: A Behavioral Perspective on the Consumer*. N.Y., N.Y.: McGraw-Hill Irwin.

Osborne, D. and T. Gaebler (1992). *Reinventing Government*. Reading, MA: Addison-Wesley Press.

Osborne, D. and P. Plastrik (2000). *The Reinventor's Fieldbook: Tools for Transforming your Government*. San Francisco, CA: Jossey-Bass.

Ostrom, E. (1973). "The Need for Multiple Indicators in Measuring the Output of Public Agencies," *Policy Studies Journal*, 2, 85-91.

Poister, T.H. and G. Streib (1999). "Performance Measurement in Municipal Government: Assessing the State of the Practice," *Public Administration Review*, 59(4), 325-334.

Poister, T.H. and G.T. Henry (1994). "Citizen Ratings of Public and Private Service Quality: A Comparative Perspective," *Public Administration Review*, 54, 155-160.

Roch, C.H. and T.H. Poister (2006). "Citizens, Accountability and Service Satisfaction: The Influence of Expectations," *Urban Affairs Review*, 41(3), 292-308.

Stalebrink, O.J. (2008). "National Performance Mandates and Intergovernmental Collaboration: An Examination of the Program Assessment Rating Tool (PART)," *The American Review of Public Administration*, 39(6), 619-639.

Swiss, J.E. (1992). "Adapting Total Quality Management (TQM) to Government," *Public Administration Review*, 52(4), 356-362.

Szymanski, D.M. and D.H. Henard. (2001). "Customer Satisfaction: A Meta-Analysis of the Empirical Evidence," *Journal of the Academy of Marketing Science*, 29(1), 16-35.

Thomas, C.W. (1998). "Maintaining and Restoring Public Trust in Government Agencies and Their Employees," *Administration & Society*, 30 (2), 166-93.

Thomas, J.C. and G. Streib (2003). "The New Face of Government: Citizen-Initiated Contacts in the Era of E-Government," *Journal of Public Administration Research and Theory*, 13(1), 83-102.

Tolbert, C.J., and K. Mossberger (2006). "The Effects of E-Government on Trust and Confidence in Government," *Public Administration Review*, 66 (3), 354–69.

Van Ryzin, G.G. (2004). "Expectations, Performance, and Citizen Satisfaction with Urban Services," *Journal of Policy Analysis and Management*, (23)3, 433-448.

Van Ryzin, G.G., D. Muzzio, S. Immerwahr, L. Gulick, and E. Martinez. (2004). "Drivers and Consequences of Citizen Satisfaction: An Application of the American Customer Satisfaction Index Model to New York City," *Public Administration Review*, 64(3), 331-341.

Van Ryzin, G.G., D. Muzzio and S. Immerwahr (2004). "Explaining the Race Gap in Satisfaction with Urban Services." *Urban Affairs Review*, 39(5), 613–32.

Van Ryzin, G.G. (2006). "Testing the Expectancy Disconfirmation Model of Citizen Satisfaction with Local Government," *Journal of Public Administration Research and Theory*, (16)4, 599-611.

Van Ryzin, G.G. (2007). "Pieces of a Puzzle: Linking Government Performance, Citizen Satisfaction and Trust." *Public Performance & Management Review*, 30(4), 521–535.

Van de Walle, S. and G. Bouckaert (2003). "Public Service Performance and Trust in Government: The Problem of Causality," *International Journal of Public Administration*, 26(8-9), 891-913.

Van de Walle, S. and G. Bouckaert (2007). "Perceptions of Productivity and Performance in Europe and the United States," *International Journal of Public Administration*, 30(11), 1123-1140.

Van de Walle, S., S. Van Roosbroek, and G. Bouckaert (2008). "Trust in the Public Sector: Is There Any Evidence for a Long-Term Decline?" *International Review of Administrative Sciences*, 74(1), 47-64.

Vigoda-Gadot, E. (2007). "Citizens' Perceptions of Politics and Ethics in Public Administration: A Five-Year National Study of Their Relationship to Satisfaction with Services, Trust in Governance, and Voice Orientations," *Journal of Public Administration Research and Theory*, 17(2), 285-305.

Vigoda-Gadot, E. and F. Yuval (2003). "Managerial Quality, Administrative Performance and Trust in Governance Revisited," *International Journal of Public Sector Management*, 16(7), 502-522.

Wagenheim, G.D. and J.H. Reurink (1991). "Customer Service in Public Administration." *Public Administration Review*, 51(3), 263-270.

Wang, X. and G.A. Gianakis (1999). "Public Officials' Attitudes Toward Subjective Performance Measures," *Public Productivity & Management Review*, 22(4), 537-553.

Welch, E. W., C. C. Hinnant and M. J. Moon (2005). "Linking Citizen Satisfaction with E-Government and Trust in Government," *Journal of Public Administration Research and Theory*, 15(3), 371-391.

Wholey, J.S. and H. Hatry (1992). "The Case for Performance Monitoring," *Public Administration Review*, 52(6), 604-610.

Yang, K. and M. Holzer (2006). "The Performance-Trust Link: Implications for Performance Measurement," *Public Administration Review*, 66(1), 114-126.

Yasin, M.M. (2002). "The Theory and Practice of Benchmarking: Then and Now," *Benchmarking: An International Journal*, 9(3), 217-243.

Figure 1: Performance-Satisfaction-Trust Model



Figure 2: Conceptual Model

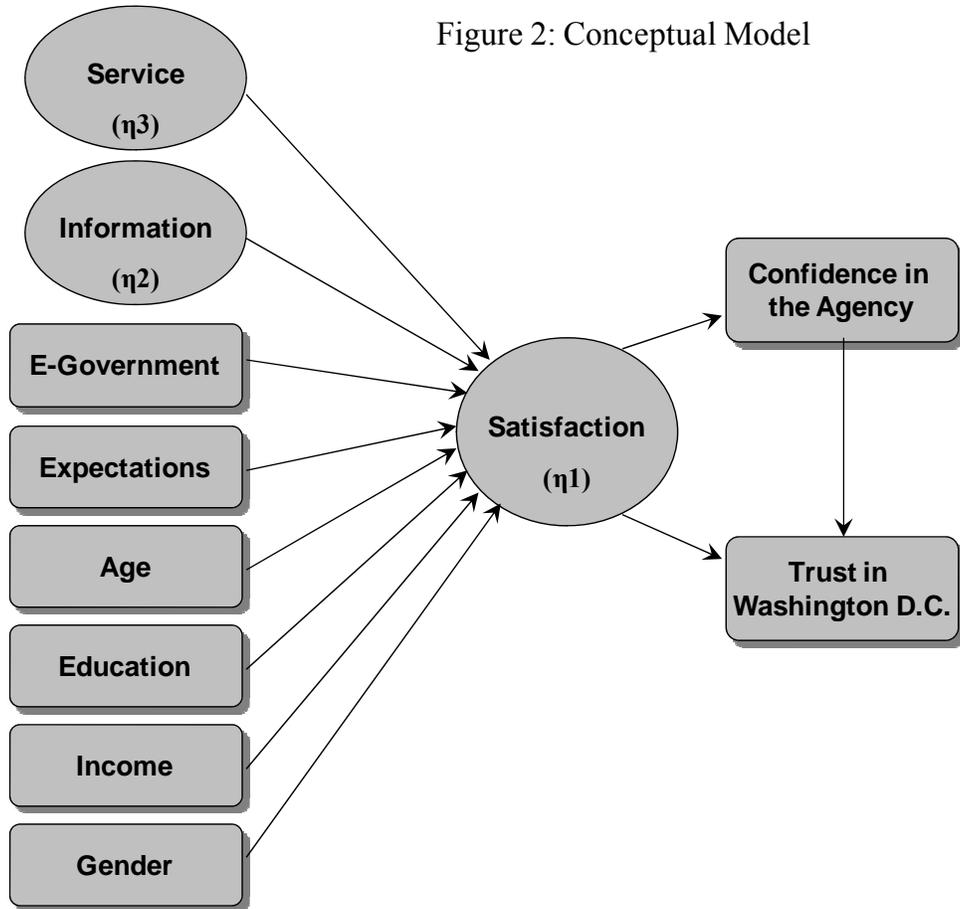


Table 1: Questions and Question Wording (Abbr.) in ACSI Sample

1. Age	What is your age, please?
2. Education	What is the highest level of formal education you completed?
3. Income	What was your total annual family income in [prior year]?
4. Gender	[Gender measured by observation/interviewer recognition]
5. E-Government	Did your experience involve using a website, email, talking by telephone, receiving printed materials by mail, visiting the agency's office, or receiving a check or a benefit?
6. Expectations	How would you rate your expectations of the overall quality of services from the (AGENCY/DEPARTMENT)?
7. Ease obtaining (Information, η_2)	How difficult or easy was it to get information about the (AGENCY/ DEPARTMENT)'s services?
8. Clarity (Information, η_2)	Was the information about (AGENCY/DEPARTMENT)'s services clear and understandable?
9. Timeliness and efficiency (Service, η_3)	How timely and efficient was the (AGENCY/DEPARTMENT) in providing the services you wanted?
10. Ease obtaining (Service, η_3)	How difficult or easy was it to obtain services from the (AGENCY/ DEPARTMENT)?
11. Overall satisfaction (Satisfaction, η_1)	First, please consider all your experiences to date with the (AGENCY/ DEPARTMENT)'s services. How satisfied are you with the (AGENCY/DEPARTMENT)'s services?
12. Confirmation to expectations (Satisfaction, η_1)	Considering all of your expectations, to what extent have the (AGENCY/ DEPARTMENT)'s services fallen short of your expectations or exceeded your expectations?
13. Close to ideal (Satisfaction, η_1)	How well do you think the (AGENCY/DEPARTMENT) compares with that ideal organization?
14. Confidence	How confident are you that the (AGENCY/DEPARTMENT) will do a good job providing the services that you used in the future?
15. Trust	Generally speaking, how much of the time do you think you can trust the government in Washington?

*The “ η ” notations identify variables included in latent variables included in our structural models in the next section. η_1 = Observed variables included in the "Satisfaction" latent variable; η_2 = Observed variables included in the “Information” latent variable; η_3 = Observed variables included in the “Service” latent variable.

Table 2: Descriptive Statistics for all Variables (Full Sample)*

	N	Min.	Max.	Mean	SD
Age (years)	1839	18	99	59.82	15.07
Education (dichotomous; 0=less college degree; 1=college grad or more)	1881	0	1	0.464	0.499
Income (dichotomous; 0=less than \$40K; 1=\$40K or more)	1587	0	1	0.589	0.492
Gender (male=1)	1883	0	1	0.364	0.481
Contact Type (e-government (website or email) =1)	1811	0	1	0.158	0.365
Expectations	1848	1	10	7.168	2.406
Overall Satisfaction (η_1)	1875	1	10	7.801	2.429
Confirmation to Expectations (η_1)	1847	1	10	7.103	2.501
Close to ideal agency (η_1)	1708	1	10	6.718	2.696
Ease of obtaining information (η_2)	1844	1	10	7.861	2.486
Clarity of information (η_2)	1860	1	10	7.610	2.509
Timeliness and efficiency of services (η_3)	1853	1	10	7.745	2.651
Ease of obtaining services (η_3)	1841	1	10	7.707	2.693
Confidence in Agency	1849	1	10	7.056	2.674
Trust in Washington	1834	1	10	4.316	2.490

Table 3: Sample Size by Agency*

	Sample
Internal Revenue Service (IRS)	510
Medicaid	118
Medicare	422
National Parks Service (NPS)	193
Social Security Administration (SSA)	505
Veterans Affairs (VA)	136
<i>Total</i>	<i>1884</i>

*Including both 2008 and 2009 data

Table 4: Latent Variable Estimates (Overall and by Agency)*

	Overall	IRS	Medicare	NPS	SSA	VA	Medicaid
Satisfaction (η_1):							
Overall Satisfaction (λ)	0.920	0.929	0.896	0.879	0.903	0.923	0.934
Confirmation to Expectations (λ)	0.875	0.853	0.867	0.824	0.877	0.898	0.907
Close to ideal agency (λ)	0.794	0.792	0.731	0.758	0.776	0.906	0.733
Overall Satisfaction (R^2)	0.846	0.863	0.804	0.773	0.816	0.852	0.872
Confirmation to Expectations (R^2)	0.766	0.728	0.752	0.679	0.769	0.806	0.823
Close to ideal agency (R^2)	0.630	0.628	0.535	0.574	0.602	0.821	0.538
Cronbach's α	0.894	0.884	0.868	0.853	0.890	0.930	0.886
Information (η_2):							
Ease of obtaining information (λ)	0.846	0.845	0.854	0.871	0.805	0.847	0.884
Clarity of information (λ)	0.843	0.821	0.771	0.801	0.874	0.934	0.816
Ease of obtaining information (R^2)	0.715	0.713	0.729	0.758	0.648	0.717	0.781
Clarity of information (R^2)	0.711	0.675	0.595	0.642	0.764	0.873	0.666
Cronbach's α	0.833	0.818	0.796	0.816	0.827	0.882	0.837
Service (η_3):							
Timeliness and efficiency of services (λ)	0.891	0.867	0.893	0.861	0.918	0.886	0.837
Ease of obtaining services (λ)	0.907	0.913	0.896	0.901	0.927	0.906	0.811
Timeliness and efficiency of services (R^2)	0.794	0.752	0.797	0.742	0.843	0.785	0.701
Ease of obtaining services (R^2)	0.823	0.834	0.802	0.812	0.859	0.821	0.658
Cronbach's α	0.893	0.881	0.883	0.874	0.919	0.894	0.807

*Estimates for each latent variable reflect Standardized Factor Loadings (λ), Squared Multiple Correlations (R^2) and Cronbach's Alpha (α)

Table 5: Structural Equation Model Estimation Results (Overall Model)

	Satisfaction (η_1)	Confidence	Trust
Age	-0.001		
	(0.002)		
Education	-0.036*		
	(0.066)		
Income	-0.018		
	(0.074)		
Gender	0.001		
	(0.063)		
E-Government	-0.042**		
	(0.090)		
Expectations	0.177***		
	(0.016)		
Information (η_2)	0.224***		
	(0.045)		
Service (η_3)	0.581***		
	(0.039)		
Satisfaction (η_1)		0.797***	0.093*
		(0.020)	(0.043)
Confidence			0.332***
			(0.035)
Squared Multiple			
Correlation	0.798	0.634	0.168

***Significant at $p < .001$; **Significant at $p < .01$; *Significant at $p < .05$

Estimates for each independent variable (going down the far left column) reflect Standardized Direct Effects and Standard Errors (in parantheses)

Model Fit Statistics: $X^2(59) = 283.70$, $P = 0.00$; $X^2/df = 4.808$; RMSEA = 0.045

Baseline Model Comparisons: NFI = 0.979; RFI = 0.958; IFI = 0.984; TLI = 0.966; CFI = 0.984.

Table 6: Structural Equation Model Estimation Results (Overall and by Agency)

	Overall	IRS	SSA	NPS	Medicare	Medicaid	VA
Satisfaction (η1):							
Age	-0.001	-0.013	0.006	0.011	-0.009	-0.117	-0.008
Education	-0.036*	-0.012	-0.060*	0.040	-0.025	-0.163	-0.022
Income	-0.018	0.004	-0.003	0.025	-0.059	0.055	0.025
Gender	0.001	-0.046	-0.002	-0.010	0.040	-0.078	0.023
E-Govt	-0.042**	-0.034	0.037	-0.097*	0.041	0.018	-0.053
Expectations	0.177***	0.215***	0.147***	0.089	0.247***	0.128	0.051
Service (η3)	0.581***	0.441***	0.589***	0.585***	0.680***	1.119**	0.827***
Information (η2)	0.224***	0.328**	0.229**	0.281*	0.063	-0.325	0.125
Confidence:							
Satisfaction (η1)	0.797***	0.865***	0.730***	0.742***	0.779***	0.776***	0.850***
Trust:							
Satisfaction (η1)	0.093*	0.246**	0.100	0.017	0.137	-0.105	0.199
Confidence	0.332***	0.220**	0.393***	0.249*	0.251**	0.513***	0.137

***Significant at $p < .001$; **Significant at $p < .01$; *Significant at $p < .05$

Estimates for each independent variable (in bold down the far left column) for each model reflect Standardized Direct Effects

Table 7: Benchmarking Matrix: Recommended Focal Points to Improve Federal Agency Satisfaction and Trust

	Customer Type							
	(i.e. demographic					Satisfaction	Satisfaction	Confidence
	differences)	E-Government	Expectations	Service	Information	to Confidence	to Trust	to Trust
Overall	No	Yes, improve	Yes	Yes	Yes, but 2nd to Service	Yes	Yes	Yes
IRS	No	No	Yes	Yes	Yes, equal to Service	Yes	Yes	Yes
SSA	No	No	Yes	Yes	Yes, but 2nd to Service	Yes	No	Yes
NPS	No	Yes, improve	No	Yes	Yes, but 2nd to Service	Yes	No	Yes
Medicare	No	No	Yes	Yes	No	Yes	No	Yes
Medicaid	No	No	No	Yes	No	Yes	No	Yes
VA	No	No	No	Yes	No	Yes	No	No

APPENDIX A

Mean variable scores for agency sub-samples

	IRS	Medicare	NPS	SSA	VA	Medicaid
Age (years)	52.212	70.308	50.423	64.558	60.939	49.664
Education (dichotomous; 0=less college degree; 1=college grad or more)	0.575	0.458	0.668	0.329	0.463	0.239
Income (0=less than \$40K; 1=\$40K or more)	0.737	0.541	0.870	0.431	0.603	0.311
Gender (male=1)	0.425	0.268	0.332	0.319	0.728	0.263
Contact Type (e-government (website or email)=1)	0.318	0.049	0.383	0.053	0.085	0.052
Expectations	6.432	7.380	8.218	7.239	7.459	7.198
Overall Satisfaction (η 1)	6.760	8.226	8.870	8.089	7.859	7.695
Confirmation to Expectations (η 1)	6.196	7.460	7.948	7.381	7.142	7.129
Close to ideal agency (η 1)	5.647	7.098	7.967	7.011	7.065	6.404
Ease of obtaining information (η 2)	7.122	8.144	8.832	8.106	7.674	7.559
Clarity of information (η 2)	6.671	7.676	8.979	7.940	7.799	7.466
Timeliness and efficiency of services (η 3)	7.251	8.029	8.910	7.872	7.104	7.171
Ease of obtaining services (η 3)	7.065	8.354	8.804	7.674	7.164	7.129
Confidence in Agency	6.461	7.039	8.281	7.210	7.159	6.896
Trust in Washington	4.279	4.257	4.700	4.335	3.962	4.371

APPENDIX B

Zero-order Correlations for all Observed Model Variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Age														
2. Education	-0.039													
3. Income	-0.177**	0.397**												
4. Gender	-0.004	0.092**	0.132**											
5. E-Government	-0.296**	0.166**	0.180**	0.039										
6. Expectations	0.150**	-0.067**	-0.082**	-0.048*	-0.032									
7. Ease obtaining information (η_2)	0.120**	-0.033	-0.016	-0.075**	-0.017	0.468**								
8. Clarity of information (η_2)	0.127**	-0.021	0.008	-0.065**	-0.043	0.476**	0.713**							
9. Timeliness and efficiency (η_3)	0.138**	0.033	0.028	-0.080**	0.029	0.462**	0.654**	0.636**						
10. Ease of obtaining services (η_3)	0.173**	0.020	0.046	-0.077**	-0.006	0.465**	0.676**	0.646**	0.807**					
11. Overall satisfaction (η_1)	0.166**	-0.063**	-0.025	-0.082**	-0.055*	0.564**	0.640**	0.652**	0.709**	0.732**				
12. Confirmation to expectations (η_1)	0.175**	-0.068**	-0.019	-0.076**	-0.076**	0.492**	0.594**	0.611**	0.682**	0.683**	0.816**			
13. Close to ideal (η_1)	0.165**	-0.068**	-0.023	-0.068**	-0.039	0.504**	0.556**	0.584**	0.603**	0.607**	0.704**	0.701**		
14. Confidence	0.065**	-0.061**	-0.023	-0.035	-0.008	0.504**	0.532**	0.582**	0.600**	0.593**	0.727**	0.685**	0.703**	
15. Trust	-0.013	0.018	-0.021	-0.035	-0.003	0.267**	0.235**	0.273**	0.293**	0.274**	0.316**	0.302**	0.326**	0.409**