Implementation of sustainable public procurement in local governments: a measurement approach

Lily Hsueh, Stuart Bretschneider and Justin M. Stritch
School of Public Affairs, Arizona State University, Phoenix, Arizona, USA, and
Nicole Darnall
School of Sustainability, Arizona State University, Tempe, Arizona, USA

Received 18 September 2019
Revised 14 February 2020
15 May 2020
Accepted 22 July 2020

Abstract

Purpose – Assessing and measuring the extent of organization-level policy implementation has received little scholarly attention, especially in the areas of local governments' procurement and environmental protection. To rectify the paucity of research in this area, this paper adopts Leonard-Barton's (1988) conceptualization of the misalignment between the (policy) innovation and the organization and draws on an original survey of local government finance, environment and public works departments in a representative sample of US cities with at least 25,000 residents to develop a strategy for measuring the extent of the implementation of a sustainable procurement policy (SPP) in local governments.

Design/methodology/approach – The authors demonstrate through the construction of a composite index that in order for a SPP to be fully implemented, standards and rules, routines and decision-making criteria need to exist to reconcile any technical, infrastructural and decision-making misalignments between the new policy and an organization's pre-existing norms and routines.

Findings – The authors empirically assess and demonstrate that the paper's proposed composite measure of policy implementation is robust to multiple specifications and measurement reliability and construct validity tests.

Originality/value – Whereas the existing literature from political science and policy science has tended to focus on higher levels of implementation in government through a complex hierarchical system, this paper underscores the importance of the policy implementation at the organizational level. Moreover, the authors contribute methodologically by our development of a strategy to measure the extent of the implementation of a SPP by local governments.

Keywords Policy implementation, Organizations, Local governments, Cities, Sustainable procurement policy, Sustainable public procurement, Sustainable public purchasing

Paper type Research paper

Introduction

Local governments are often required and responsible for implementing federal and state policies (Carpenter and Krause, 2015; Krause, 2011b; Mullin and Daley, 2010; Peterson *et al.*, 2010; Rabe, 2006, 2007; Vig and Kraft, 2016). Local governments also initiate and implement their own policies in the absence of federal and state requirements (e.g. Gray *et al.*, 2012; Ji and Darnall, 2018; Krause, 2011b; Krebs and Pelissero, 2010). Both types of policy adoption and implementation pathways have received much attention from prior policy science and public administration scholars (e.g. Chatwin *et al.*, 2019; Carpenter and Krause, 2015; Krause, 2011a). By contrast, assessing and measuring the extent of organization-level policy implementation has received less scholarly attention, especially in the areas of local government's procurement and environmental protection (Hsueh and Darnall, 2017; Krause, 2011b). This

Funding: This research was funded by a grant from the V.K. Rasmussen Foundation.

International Journal of Public Sector Management
Vol. 33 No. 6/7, 2020
pp. 697-712

0951-3558
DOI 10.1108/IJPSM-09-2019-0233

gap is likely because of the absence of cross-organizational data that are both specific and broad enough for the task, in addition to the lack of conceptual and measurement strategies for operationalizing an organization's degree of policy implementation. This paper attempts to rectify both challenges in the context of the implementation of a sustainability procurement policy in local governments.

A sustainable procurement policy (SPP) attempts to modify existing procurement processes, rules and structures to improve sustainability and reduce adverse environmental and social impacts of local government. About a quarter of US local governments have adopted SPPs, yet of these, about two out of five local governments report that they struggle to implement SPP fully (Darnall *et al.*, 2017). This reported variation illustrates how enhancing the sustainability of public procurement is likely to be implemented differently across local governments. We make use of this type of variation to measure the degree of implementation of SPPs across local governments.

Toward this end, we adopt Leonard-Barton's (1988) characterization of misalignments or mismatches between a new technology—policy or otherwise—and the organization at the time of initial trial or use of the innovation [1]. This mismatch between a new policy and the organization must be reconciled if the implementation is to succeed [2]. Leonard-Barton categorizes implementation misalignments into three types: technical, infrastructural and organizational performance criteria incongruence. First, technical misalignment refers to fact that existing technical specifications of the organization's production processes need to be altered to fit the new technology. Second, an organization's infrastructure for product or service delivery is manifested in the organization's routines; these need to be altered to fit the new policy. Finally, in order for the new policy to become an organization's regular function and practice there needs to be new criteria for decision-making. This paper adopts Leonard-Barton's conceptualization of misalignments as a strategy for operationalizing implementation success and demonstrates the efficacy of this approach by using it to better understand local government implementation of SPPs.

This paper draws on an original representative survey of local government finance, environment and public works departments in US cities with at least 25,000 residents. Through the construction of a composite index, we demonstrate that in order for a SPP to be fully implemented, there needs to be the existence of standards and rules, routines and decision-making criteria that reconcile any technical, infrastructural and decision-making misalignments between the new policy and an organization's pre-existing norms and routines. We make a novel contribution to the implementation literature by empirically establishing how the misalignments between a new policy and the pre-existing organization are reflected in the implementation of the new policy. In doing so, we provide a nuanced way of thinking about the measurement of policy implementation that might prove useful in areas of policy research beyond SPP.

The rest of the paper is structured as follows. The next section describes public procurement in local governments and introduces SPP as a new policy that many local governments have initiated and began to implement. The following section characterizes the components of policy implementation—namely, standards and rules, routines and decision-making criteria—that must exist within an organizational setting for an SPP to be fully implemented in local governments. We next operationalize SPP implementation with our development of a composite index that measures the extent to which local governments enact environmental sustainability standards and rules, routines and decision-making criteria into their procurement practices. In this section, we also present an empirical assessment of measurement reliability and construct validity. We conclude with a summary of the strengths and weaknesses of our measurement strategy and a discussion of this paper's implications for conceptualizing and measuring policy implementation more generally before suggesting directions for future research.

Local governments' public procurement and sustainable procurement

Public procurement is the process by which public organizations acquire goods, services and supplies from outside sources (Snider and Rendon, 2012; Thai, 2001). Procurement in local governments is an ideal case for studying organization level policy implementation for several reasons. First, it is an important and ongoing activity of government that is subject to a high degree of formalization. It involves significant expensing of tax payer dollars, and, thus, must adhere to substantial political and bureaucratic oversight (Pandey *et al.*, 2011; Thai, 2001). Oversight helps ensure transparency and accountability in the purchasing process (Telgen *et al.*, 2007). Any policy changes associated with procurement, therefore, must be compatible with pre-existing rules and routines.

Second, public procurement is a decision-based process. Each purchase reflects a purchasing professional's choice that is bound by formal or informal criteria. Purchasing choices must meet the needs and demands of the public organization while attending to limited operating budgets. In order for purchasing professionals to meet their service priorities, they must rely on existing decision-making processes and structures to accomplish their work (Johnson *et al.*, 2003; McCue and Gianakis, 2001).

Finally, while a major focus of public procurement is the efficient operation of the government, it is also an area that has historically been used to support a wide array of social objectives (Ingraham *et al.*, 2003). These broader objectives are typically designed to improve social equity, enhance economic development and address other policy objectives (McCrudden, 2004; Qiao *et al.*, 2009). For example, local governments often enact policies to favor veteran-, women- and minority-owned businesses (Fernandez *et al.*, 2013; Marvel and Resh, 2015; McGrann, 2014; Smith and Fernandez, 2010). Public procurement is also used to encourage local economic development (Brammer and Walker, 2011) via purchasing from locally-owned businesses (Nijaki and Worrel, 2012).

Local governments' sustainable procurement policy

SPP attempts to modify existing procurement processes, rules and structures to improve sustainability and reduce adverse environmental impacts of local government. It introduces environmental criteria into public procurement processes (Burchard-Dziubinska and Jakubiec, 2012; Darnall *et al.*, 2018) that may include reducing energy and water consumption, greenhouse gas emissions, solid waste and other factors. A city's SPP may include formal policy approaches such as enacting ordinances, executive orders, resolutions and administrative directives. It may also include less formal policy approaches such as adding sustainable purchasing requirements to existing sustainability plans or energy conservation policies.

While many US local governments have adopted SPPs, a significant portion reports that they have failed to implement SPP fully (Darnall *et al.*, 2017). This reported variation in local governments' SPP implementation illustrates how the "greening" of public procurement varies across organizations. We suggest that these implementation variations create a unique setting to measure the degree of policy implementation and more specifically, the degree of local governments' SPP implementation.

Public procurement rules and standards, routines and decision-making criteria

We draw on Leonard-Barton's (1988) conceptualization of the misalignments between the (policy) innovation and the organization to argue and show empirically—in our development of a measurement strategy for the degree of implementation of an SPP in local governments—that policy implementation is an expression of the way in which organizations enact rules and standards, establish routines and generate decision-making criteria. When rules and standards, routines and decision-making criteria are established for the new policy, any technical, infrastructural and decision-making misalignments between the new policy and an organization's pre-existing norms and routines are reconciled.

Rules and standards are a set of documented rights, duties and procedures that reach a wider set of organizational actors in a consistent manner (Pugh *et al.*, 1963). In the procurement context, these are specific product or service requirements that must be satisfied prior to purchase. They tend to determine the acceptance and/or evaluation of competitive bids and help ensure that the final purchase meets a specific need that otherwise might not be met if the purchase price was the sole decision criterion. Examples include technical specifications which articulate precise requirements related to a product's manufacturing content and composition, performance, use, safety/operation and other factors. Technical specifications encompass descriptive requirements but also quantitative indicators, such as minimum performance levels.

Routines are the patterns by which organizational actors undertake certain activities (Pentland and Rueter, 1994). Local governments typically develop purchasing routines for numerous product categories, including chemical products, road construction/maintenance services, general construction services, electrical products, information technology hardware and services, professional services, transportation/fuels and wood and paper products. For instance, chemical cleaning products tend to cause harmful health impacts to employees who use them. As such, local governments may be more likely to develop purchasing routines that seek to reduce those impacts. Similarly, electrical products can involve significant energy costs during their use which may motivate local governments to develop purchasing routines that seek to minimize those costs.

Purchasing routines in local government are either routine or non-routine. Routine purchases typically occur frequently and thus have more formalized patterns for procurement that require little oversight. Examples include office supplies or general electronic purchases. By contrast, non-routine purchases lack formalized purchasing patterns because they occur infrequently. For this reason, they also typically involve more oversight and involve higher levels in the approval process. Examples include purchases of aviation monitoring systems and heavy construction equipment.

Decision-making criteria are the factors considered in the purchasing decision. Balancing multiple criteria increases tension within the purchasing process because optimization across all criteria is generally not possible (Leonard-Barton, 1988). Purchasing officers, therefore, must consider the trade-offs of meeting some purchasing criteria and achieving second-best (or less) for others. Procurement cost is a common decision-making criterion. It refers to the dollar amount associated with a purchase. As local governments face consistently constrained budgets, purchasing cost tends to have significant importance in purchasing decisions. Other decision-making criteria include environmental criteria, such as reducing disposal costs, minimizing packaging waste, increasing recyclability or reuse and reducing greenhouse gas impacts.

In our operationalization, these three characteristics – rules and standards, routines and decision-making criteria – are not independent of each other and typically interact in a variety of ways. For example, a decision regarding the construction of a new building may involve substantial technical specifications (e.g. standards) which could also include environmental specifications. By contrast, decisions to purchase office supplies are likely to involve few (if any) technical specifications and limited decision-making criteria.

Importantly, when rules and standards, routines and decision-making criteria are established to match the new policy with the existing organization, the implementation of the new policy would be fully realized. Toward this regard, the focus of the rest of the paper is to develop an empirical approach to operationalize the extent of local governments' implementation of an SPP with respect to the existence of rules and standards, routines and decision-making criteria. This assumes a number of conditions. First, the focus is at the organizational level. Second, the presumption is that over time, an organization's implementation of a new policy like SPP occurs at different rates but is reflected in the current rules, routines and decision processes at any given point in time. The extent to which

policy implementation aligns with that which is implied by a new policy captures the extent to which the organization has "implemented" the new policy regime.

Measuring sustainable procurement policy implementation

Data source and description

We operationalize SPP implementation as the extent to which local governments enact and make use of standards and rules, establish routines and generate decision-making criteria to embed environmental sustainability into their procurement practices. We draw on data from an original survey ("Sustainable Procurement in US Cities") of 1,825 finance, environment and public works departments. The survey was sent to a sample of department directors in 791 US cities with 25,000 residents or more. In comparing these 791 cities to data obtained from the US Census Bureau for all local governments with 25,000 residents or more, there were no statistical differences between the sample and the broader population when considering number of residents, median family income and geographic distribution by state.

The survey was co-sponsored by International City/County Management Association (ICMA). In targeting directors of finance, environment and public works departments, the survey sought perspectives from individuals who were either knowledgeable about the purchasing process, were affected significantly by purchasing, or had detailed information about the city's environmental management processes.

The survey consisted of 37 questions that addressed city-level purchasing activities, city-level environmental sustainability policies/practices, department-level policies/practices, department structure and culture and professional/personal information. Within these broader areas, questions covered topics such as the structure of purchasing decisions, access to information, vendor roles and the influence of stakeholders. Appendix contains the survey instrument.

The survey was distributed online over a period of eight weeks in the spring of 2016. City directors received an initial letter informing them of the survey. Several days later, directors received an email containing a link to a Qualtrics-based survey. Non-respondents received up to four email reminders, two postcard reminders and two phone call reminders.

The final sample consisted of 608 city directors who completed the survey for a response rate of 33.3%. We received responses from at least one director in 61.1% (483) of the 791 cities that received the survey. The final sample consisted of 47.0% finance directors, 15.0% environmental managers and 36.5% public works directors.

Measures

Our measure of the extent of SPP implementation was a linear composite index using 16 survey items. Each survey item assessed some aspect of rules, procedures and decision criteria associated with sustainable procurement and used a five-point Likert scale that ranged from 1 for "not important" to 5 for "very important." Survey items were dispersed throughout the survey instrument but are organized here into three broad categories: standards and rules, routines and decision-making criteria. While we recognize that each of these grouping is useful in theory, the strong likelihood that items correlate across groups supports treating our composite index as a unidimensional construct. Operationally, when a response is relatively larger than another response it suggests that the rule, procedure or decision criteria being considered is perceived as being more important, and thus, signifies that current practice is more in line with the new policy.

SPP rules and standards (1 item). To measure SPP standards and rules, we used one item from a series of items associated with the following question; "How important are technical specifications to your department when managing the following aspects of purchases?" The specific item we used refers to the importance of using technical specifications for "managing environmental sustainability concerns."

SPP routines (8 items). Items associate with SPP routines came from one question that asked, "Within your department, how important are environmental sustainability concerns

to the purchase of the following types of products and services?" The question presented eight items, each one representing a specific type of products or services that are routinely procured by cities: chemical products, road construction/maintenance services, general construction services, electrical products, information technology hardware and services, professional services, transportation/fuels and wood and paper products.

SPP decision-making criteria (7 items). To measure decision-making criteria, we used data items drawn from two survey questions. The first question asked, "There are many different criteria that may be considered in the purchase of a product or service. In thinking about your department's purchasing criteria, how important is each of the following characteristics of a product or service?" Two items from this question used in the measure were associated with the importance of greenhouse gas (GHG) emissions reductions and environmental impacts.

The second question asked, "How important or unimportant are the following considerations in your department's purchasing decisions?" The measure included five items characterizing a purchase: environmental sustainability of products/services offered, disposal costs, packaging waste reduction, recyclability or reuse and GHG emissions reduction.

There is empirical evidence that there are three distinct mechanisms involved in measuring the extent of SPP implementation, namely, the existence of rules and standard, routines and decision-making criteria. Table 1 reports the results from using a standard iterated principle factor analysis based on three factors and a promax oblique rotation to facilitate interpretation (Thurstone, 1947). Oblique rotation relaxes the requirement of standard orthogonal rotations that factors be completely uncorrelated. Consequently, the oblique rotation process tries to balance interpretation while allowing some correlation between factors. The correlations between the three factors provided in Table 1 range from 0.049 to 0.104 which are low. These empirical results support the idea that there are multiple processes that affect implementation success. However, our objective is to generate a single measure that enables enhanced study of implementation success. As such, in the next section, we develop a single composite index and focus on evaluating its reliability and validity.

SPP implementation index

We constructed the SPP Implementation Index by first summing each item score over all 16 items. Since each item score ranged between 1 and 5, the sum of the composite index ranged between 16 and 80. The average score was then calculated by dividing by the total number of survey items. Respondents were also given the option of answering "Don't Know."

Table 1.
Promax oblique rotation of three factor solution for iterated principle factor analysis

Variable	Factor 1	Factor 2	Factor 3	Uniqueness
Chemical products	0.6799	0.3742	0.2105	0.3534
Road construction	0.7810	0.3252	0.2153	0.2379
General construction services	0.8210	0.3028	0.2436	0.1750
Electrical products	0.7652	0.2990	0.2745	0.2498
IT hardware and services	0.7273	0.2733	0.2933	0.3103
Professional services	0.6647	0.2555	0.2836	0.4124
Transportation/fuel	0.7183	0.3626	0.3228	0.2484
Woods and paper products	0.6443	0.3927	0.3523	0.3065
Environmental sustainability of the products	0.4229	0.6528	0.4519	0.1908
Disposal costs in purchasing decisions	0.3389	0.7124	0.1437	0.3570
Reducing packaging waste in purchasing decisions	0.3750	0.7296	0.3341	0.2154
Recyclability or reuse in purchasing decisions	0.3655	0.7020	0.3165	0.2735
Reducing GHG impacts in purchasing decisions	0.3924	0.5934	0.4702	0.2729
Technical specifications in environmental sustainability'	0.3221	0.3756	0.5986	0.3969
GHG reductions in purchasing criteria	0.2811	0.2437	0.8175	0.1933
Environmental impact in purchasing criteria	0.2560	0.2648	0.7795	0.2567

Responses were excluded if there was at least one missing value or a "Don't Know" response for an item. The use of a "Don't Know" category typically generates more missing responses but it increases the quality of response and reduces refusals (Shoemaker *et al.*, 2002).

Table 2 reports the summary statistics of the survey items and the result of our composite index. A low score indicates that a local government unit has low SPP implementation, whereas a high score indicates higher SPP implementation. In the parlance of Leonard-Barton's implementation misalignments, a lower index score indicates more misalignment, whereas as a higher index score indicates less misalignment between the new policy and the organization. In the context of SPP more misalignment indicates continued reliance on older rules, routines and decision-making criteria. Accordingly, the higher score suggests that an organization is closer to fully implementing a new policy compared an organization with a lower score.

Non-response bias, measurement reliability and construct validity

Non-response bias and sample size

Non-response bias typically arises within survey data when some items are missing or are coded as "Don't Know." Table 3 shows that for the 597 survey responses in our sample, some level of non-response exists with only 348 cases containing responses to all 16 items. Table 3 describes each case in terms of the number of items where respondents indicate "Don't Know." While some responses to items indicated "Don't Know" or were missing, slightly less than half of the sample exhibited one or more "Don't Know" and about 85% of the sample had 4 or less items with a "Don't Know" response. Eleven cases were dropped due to one or more items having missing values.

Variable	*N*	Mean	Std. dev.	Min	Max
Standards and rules					
Importance of...					
Technical specifications in managing environmental sustainability concerns	512	2.72	1.27	1.00	5.00
Routines					
Importance of environmental concerns in purchases for...					
Chemical products	525	3.42	1.30	1.00	5.00
Road construction	503	3.13	1.31	1.00	5.00
General construction services	513	3.03	1.28	1.00	5.00
Electrical products	499	3.05	1.27	1.00	5.00
IT hardware and services	497	2.91	1.25	1.00	5.00
Professional services	523	2.70	1.31	1.00	5.00
Transportation/fuel	520	3.11	1.31	1.00	5.00
Woods and paper products	524	3.00	1.25	1.00	5.00
Decision-making criteria					
Importance of...					
GHG reductions in purchasing criteria	559	2.73	1.22	1.00	5.00
Environmental impact in purchasing criteria	560	3.07	1.22	1.00	5.00
Environmental sustainability of the products offered in purchasing decisions	560	2.47	1.22	1.00	5.00
Disposal costs in purchasing decisions	567	2.96	1.21	1.00	5.00
Reducing packaging waste in purchasing decisions	543	2.55	1.25	1.00	5.00
Recyclability or reuse in purchasing decisions	486	2.00	1.13	1.00	5.00
Reducing GHG impacts in purchasing decisions	520	2.51	1.28	1.00	5.00
SPP implementation index	348	2.83	0.99	1.00	4.94

Table 2. Summary statistics of survey items and the SPP implementation index

Number of "Don't know" items by observation	Freq.	Percent	Cum.
0	348	58.29	58.29
1	67	11.22	69.51
2	36	6.03	75.54
3	40	6.70	82.24
4	19	3.18	85.43
5	10	1.68	87.10
6	9	1.51	88.61
7	12	2.01	90.62
8	18	3.02	93.63
9	2	0.34	93.97
10	3	0.50	94.47
11	9	1.51	95.98
12	5	0.84	96.82
13	7	1.17	97.99
14	4	0.67	98.66
15	8	1.34	100.00
Total	597	100.00	

Table 3.
Number of "Don't know" items by observation

In order to understand the potential implications of item non-response we conduct two standard analyses. The first analysis regresses a set of variables likely to identify selection differences on whether an observation contained one or more "Don't Know" responses. The main variables we consider are the functional role of the respondent (i.e. finance, environment and public works), city size, city wealth and demographic diversity. The second analysis uses the same independent variables but examines the number of "Don't Know" responses. This second analysis employs the negative binomial model after confirming problems of over dispersion in the data [3].

The results reported in Table 4 identify only one systematic source of nonresponse bias. Both models find that none of the city-level factors create selection bias, although public works directors generally had fewer "Don't Know" responses than directors in the finance or environmental units. Thus, there is the potential for measurement bias associated with being from a finance or environmental department. Our analysis of reliability and validity will therefore consider the

	Model 1[a]	se[b]	Model 2[a,c]	se[b]
Department				
Finance	−0.19	0.25	−0.04	0.19
Public works	−0.52**	0.26	−0.79***	0.22
Total population	0.00	0.00	0.00	0.00
Median family income	0.00	0.00	0.00	0.00
Unemployment	−0.00	0.04	0.04	0.03
Families in poverty	0.01	0.03	0.02	0.02
Number of Black residents	0.00	0.00	0.00	0.00
Constant	−0.52	0.67	−0.19	0.62
lnalpha constant			1.29***	0.09
N	588		588	

Note(s): *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

[a]Model 1 is an OLS regression on the incidences of non-response. Model 2 is a negative binominal model on the number of "Don't Know" responses

[b]Robust standard errors

[c]The likelihood-ratio test of $\lambda = 0$ (i.e. no over dispersion) can be rejected with a χ^2 value of 1393.33 with probability $\geq \chi^2$ is virtually 0

Table 4.
Regression models for potential non-response selection and "Don't know" items

overall SPP implementation index for all types of respondents and also two sub-groups, public works managers versus the combination of finance and environmental unit managers.

A related issue that derives from the item non-responses is sample size. Excluding all cases associated with one or more item non-response reduces the sample significantly. A standard approach to missing values is mean imputation. Mean imputation can distort the distribution of a variable, leading to an underestimation of the standard deviation since it assumes that the mean does not change with the inclusion of "Don't Know" responses (Little and Rubin, 2014). Moreover, mean imputation can distort relationships between variables by "pulling" estimates of the correlation toward zero (OECD, 2008). Despite these concerns there are reasons to make use of mean imputation particularly when considering correlational assessments of construct validity. Since mean imputation biases correlations toward zero, estimates will be conservative and if strong, that much more compelling. Thus, as we proceed to empirically assess the reliability and construct validity of our implementation index, we consider the actual data, breakouts of the data by the two subgroups described above and the impacts of a larger sample using mean imputation, also broken out for the two subgroups.

Measurement reliability and construct validity

To assess measurement reliability and construct validity we consider a number of perspectives. First, to account for possible selection effects associated with public works managers, we consider three distinct versions of the SPP implementation index constructed around respondents: all respondents, only public works respondents and both environmental and finance manager respondents. Second, to consider the effect of sample size on construct validity associated with the empirical sample, we use an augmented alternative specification for each of the three indexes that accounts for the mean imputation of missing items from the "Don't Know" responses. These mean imputations are *case specific* in that the final score is based on the mean score generated from the reduced number of items, thus accounting for some of the variance associated with individual characteristics, such as respondent type.

Table 5 reports a summary of the Cronbach's alpha analysis, including the inter-item correlations for the three versions of the SPP Implementation Index. The Cronbach's alpha estimates are similar across the three indexes as are most of the item-test, item-rest and alpha components. There are, however, some differences associated with the index built solely on public works directors' responses. This is not surprising given our previous analyses, but the differences suggest that, overall, public works managers are less likely to generate relatively high scores based on this approach to measuring implementation of an SPP. Specifically, their responses on technical specifications and decision making criteria (i.e. environmental impact, disposal cost in purchasing decisions and reducing packaging waste in purchasing decisions) are weaker relative to the indexes associated with the full dataset and the environmental and finance department split sample. The general implication is that the overall index and the sub-index for environmental and finance managers is likely biased upward, though for these data the magnitude does not seem to be large. Despite this, all three measures show strong internal consistency and overall reliability.

Construct validity

To determine how well the SPP implementation indexes measure the degree of local governments' SPP implementation we need to assess their construct validity. Construct validity analysis assesses the extent to which a proposed measure behaves with respect to other valid construct measures. Though construct validity is often subdivided into different forms (e.g. predictive, concurrent and discriminant) assessing it is typically implemented through some form of correlational analysis (Trochim, 2006). We consider two types of outcomes that relate with higher- vs. lower-levels of SPP implementation. The first is an objective measure that considers the extent to which managers make use of environmental information as inputs to the

Table 5.
Cronbach's alpha analysis

Survey items	SPP implementation index N	Item-test Correlation	Item-rest Correlation	Alpha	SPP implementation index - Public works N	Item-test Correlation	Item-rest Correlation	Alpha	SPP implementation index - Env/Finance N	Item-test Correlation	Item-rest Correlation	Alpha
Standards and rules (Importance of. . .)												
Technical specifications in managing environmental sustainability concerns	512	0.713	0.661	0.957	202	0.655	0.599	0.951	305	0.742	0.691	0.961
Routines (Importance of environmental concerns in purchases for. . .)												
Chemical products	525	0.789	0.752	0.956	211	0.766	0.727	0.949	306	0.805	0.770	0.959
Road construction	503	0.808	0.773	0.955	210	0.812	0.778	0.948	285	0.809	0.776	0.959
Construction services	513	0.826	0.795	0.955	210	0.826	0.793	0.947	295	0.827	0.797	0.959
Electrical products	499	0.818	0.786	0.955	200	0.765	0.723	0.949	291	0.847	0.820	0.959
IT hardware and services	497	0.789	0.754	0.955	182	0.771	0.733	0.948	307	0.799	0.766	0.959
Professional services	523	0.767	0.726	0.956	207	0.782	0.740	0.949	308	0.758	0.718	0.960
Transportation/fuel	520	0.832	0.801	0.955	204	0.811	0.777	0.948	308	0.840	0.810	0.959
Woods and paper products	524	0.827	0.797	0.955	203	0.813	0.781	0.948	313	0.842	0.814	0.959
Decision-making criteria (Importance of. . .)												
GHG reductions in purchasing criteria	559	0.854	0.827	0.954	211	0.823	0.792	0.948	339	0.871	0.846	0.958
Environmental impact in purchasing criteria	560	0.708	0.661	0.958	212	0.625	0.571	0.952	339	0.748	0.705	0.961
Environmental sustainability of the products offered in purchasing decisions	560	0.817	0.783	0.955	211	0.816	0.779	0.948	340	0.820	0.789	0.959
Disposal costs in purchasing decisions	567	0.785	0.748	0.956	214	0.741	0.701	0.949	344	0.809	0.774	0.959
Reducing packaging waste in purchasing decisions	543	0.824	0.793	0.955	208	0.780	0.744	0.948	327	0.844	0.816	0.959
Recyclability or reuse in purchasing decisions	486	0.703	0.660	0.957	194	0.701	0.656	0.950	288	0.702	0.660	0.961
Reducing GHG impacts in purchasing decisions	520	0.722	0.669	0.958	204	0.713	0.664	0.950	311	0.723	0.666	0.961
Test scale				0.958				0.952				0.962

purchasing process. The second measure accounts for the differences we would expect to see between those cities with a formal SPP and those without one.

Each department director was provided with a list of ten inputs typically used in purchasing decisions. These inputs included use of vendor lists, use of prior contracts, access to a bid process, online information search system, environmental impact information, life-cycle cost information, prior approval requirements from unit level management, prior approval from higher level management, ecolabel information and preferred green product lists. Three of the ten items were related to sustainability, namely, environmental impact information, ecolabel information and preferred green product lists. This list of inputs was provided for three different types of purchasing decisions related to (1) routine, low-cost purchases (RLC) such a paper and office supplies, (2) routine, high-cost purchases such as computer equipment or vehicles (RHC) and (3) non-routine, high-cost purchases (NRHC). Routine low- and high-cost purchases are common and have well-established patterns. By contrast, non-routine, high-cost purchases are less common, have less established purchasing patterns and involve more oversight. Examples within this category are very diverse and can include major retrofitting of technology (e.g. updating street lighting), building new structures and one-time purchases of enterprise level software. High-cost purchases (of all sorts) typically involve technical specifications. Respondents were asked to check all inputs that they currently employ in their purchasing decisions. We then calculated the proportion of "green" inputs to all inputs considered by that respondent.

Table 6 provides summary statistics for our six alternative SPP implementation indexes, three of which are based on using raw data only and three of which are based on using raw and imputed values. Table 6 also reports summary statistics for the three outcome measures for the proportion of "green" inputs. Note that the typical average share of "green" inputs used in purchasing was between 9% and 14% with the lowest share associated with routine, low-cost purchases. This result offers some face validity such that increases in costs and complexity are related to increases in using more inputs (i.e. a more complex decision process), including sustainability criteria.

Table 7 describes the correlations between the SPP implementation indexes and the percent shares of green inputs to the purchasing process. All are positive and statistically significant even when sample sizes are small. This is consistent with our theoretical expectation. As anticipated, managers in cities that are further along in their implementation of an SPP have higher SPP index values and make use of "green" inputs within their routine low, routine high and non-routine high cost purchasing decisions. Not surprisingly, correlations using the sample data are larger than those associated using the imputed data; these findings identify a known problem associated with mean imputation. That said, there is

SPP implementation indexes	*N*	Mean	Std. dev.	Min.	Max.
Data					
Overall	348	2.83	0.99	1.00	4.94
Env/finance	203	2.83	1.01	1.00	4.81
Public works	142	2.81	0.96	1.00	4.94
Data/imputed					
Overall (Imputed)	597	2.81	1.01	1.00	5.00
Env/finance (Imputed)	367	2.79	1.04	1.00	5.00
Public works (Imputed)	221	2.85	0.95	1.00	5.00
Percent share of green criteria in purchasing					
Routine, low-cost purchases (RLC)	519	0.09	0.18	0.00	1.00
Routine, high-cost purchases (RHC)	577	0.11	0.15	0.00	0.67
Non-routine, high-cost purchases (NRHC)	576	0.14	0.18	0.00	1.00

Table 6. Summary statistics for SPP implementation indexes and percent share of green criteria in purchasing

still a significant and moderately-sized positive correlation. While this type of analysis does not identify bias, it is an indication of our measure's predictive and criterion validity. One interesting finding here is that for non-routine purchases environmental/finance directors had a higher correlation than did public works directors. This may be driven by finance directors responsibility for oversight in the administration and purchase of most non-routine, high-cost products and services.

Our second validity analysis focuses on how well the SPP Implementation Index differentiates between respondents in cities that have a formal SPP from those lacking a formal SPP policy. We employ a difference in means approach to examine this issue (see Table 8). Cities with a formal policy had a significantly higher SPP Implementation Index score than those without a formal policy. Interestingly though, cities that lacked a formal SPP policy exhibit some level of SPP implementation prior to the enactment of a formal policy. This finding is consistent with informal implementation processes advanced by street-level bureaucrats (Moulton and Sandfort, 2017; May and Winter, 2009; Hill, 2003). There are many potential explanations for this finding including open system influences from external interest groups. As such, our findings suggest that formal policy adoption may not be a prerequisite for informal implementation. Thus, our framework for measuring policy implementation may have applicability to settings that lack a formal policy.

Conclusion

Assessing and measuring the extent of organization-level policy implementation has received little scholarly attention, especially in the areas of local government's procurement and environmental protection (Hsueh and Darnall, 2017; Krause, 2011b). To address this concern, this paper's contribution is to operationalize Leonard-Barton's (1988) conceptualization of the implementation misalignments between a (policy) innovation and the organization. Our research draws on an original representative survey of local governments in the US to develop a strategy that measures local governments' degree of SPP implementation. We use a

Table 7.
Correlation of implementation indexes with percent share of green criteria in purchasing

SPP implementation index	% Share of Green Criteria in Routine, low-cost purchases	Routine, high-cost purchases	Non-routine, high-cost purchases
Overall	0.352***	0.349***	0.306***
Env/finance	0.320***	0.317***	0.342***
Public works	0.380***	0.383***	0.256***
Overall (Imputed)	0.291***	0.295***	0.227***
Env/finance (Imputed)	0.300***	0.280***	0.270***
Public works (Imputed)	0.254***	0.300***	0.157***

Note(s): *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Table 8.
SPP implementation index mean differences: Formal SPP vs. No Formal SPP

SPP implementation index	Formal SPP *N*	Mean	Std. dev.	No formal SPP *N*	Mean	Std. dev.	Difference in means *t*-stat
Overall	103	3.25	0.09	222	2.61	0.06	−5.59***
Env/finance	66	3.18	0.12	127	2.65	0.09	−3.48***
Public works	35	3.38	0.14	94	2.55	0.10	−4.64***
Overall (Imputed)	167	3.26	0.07	357	2.58	0.05	−7.51***
Env/finance (Imputed)	109	3.23	0.09	217	2.59	0.07	−5.52***
Public works (Imputed)	55	3.35	0.11	135	2.58	0.08	−5.34***

Note(s): *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

composite index to demonstrate that implementation success is a function of the existence of standards and rules, routines and decision-making criteria that align the organization's activities with the objectives of the new policy. This measure is robust to multiple specifications and reliability and validity tests.

Our operationalization of the implementation of an SPP in local governments is sufficiently broad in that it is responsive to both the implementation of a *formal policy* as well as an *informal policy*. While the former is expected, the latter is an important contribution to the literature in that even in the absence of a formal SPP policy, organizations could still implement informal policy elements to "green" the procurement process.

A potential limitation of the paper is that constructing a measurement strategy requires contextual understanding, which makes the framework highly applicable to a specific policy setting. Applying this approach to other policy settings, such as emergency management or forest stewardship, may therefore require knowledge of each setting's pre-existing relevant rules, routines and decision-making systems, as well as the implied changes to them that may be required by a new policy. That being said, a key takeaway of this paper for policy and administration scholars when evaluating policy implementation in other policy contexts is to measure the extent of an organization's rules, routines and decision criteria in the generation of the changes that are required to align a new policy with the organization.

Moreover, we recognize that Leonard-Barton's misalignment framework was developed with the implementation of technology innovation as opposed to the implementation of policy. However, the implementation of many soft technologies is consistent with what we might think of as policy—or the result of policy themselves. For example, management by objective (MBO) or program budgeting (PPBS) are soft technologies but are also policies.

This paper's framework may not apply to more complex implementation settings involving multiple organizations and levels, such as the implementation of a federal policy across multiple agencies, states and local governments. This is particularly relevant in a federal system like the US where state governments may fundamentally resist federal efforts. We also acknowledge all the limitations associated with the empirical use of cross-sectional survey dataset, such as the issues of response and non-response bias noted above. While there is some potential for common source bias, the construct validity tests have applied objective outcome measures, which limit the potential of such problems.

Our approach captures a snapshot in history of the multiple units at one point in time. This static frame for measuring policy implementation is intentional, although a potential shortcoming. The rationale for our "static" approach relates the need to establish the relevant elements and concepts of what constitute policy implementation within organizations before adding complexity, particularly complexity related to institutional and organizational change. By focusing our approach and assessing data from a national representative survey, this research inches closer to addressing Goggin's (1986) criticisms that scholars frequently draw on a limited number of observations of implementation processes and identify too many variables to explain variations in implementation outcomes. However, prospective research would benefit from repeatedly measuring the degree to which rules, routines and decision-making criteria move away from their initial conditions (i.e. prior to the adoption of the new policy) toward those implied by the new policy. Although it is difficult to collect these data, such an analysis would offer important about policy implementation processes over time. Our hope is that this research provides a foundation for such an investigation.

Future research should also consider how the implementation of a new policy—SPP or otherwise—occurs over time and how it reconciles pre-existing rules, procedures and decision-making criteria with new requirements. Toward this end, our paper offers a conceptual and empirical roadmap for implementation scholars. Abstracting from the rich policy context of this paper, our objective is to establish a reliable and valid outcome measure for the extent of policy implementation. Such a measure can facilitate our understanding of

the implementation process, particularly at the organization level. Toward that end, the initial finding presented here that non-adopters have some level of implementation, suggest a number of new theoretical frameworks. For example, the adoption decision itself is likely endogenous to the institutional forces which generate informal, pre-policy adoption levels of implementation. These findings may also allow us to link theories of organizational change to implementation. Our measure of policy implementation, in particular, may then be relevant for prospective research, serving as a dependent or independent variable for future studies that assess policy implementation in organizations, building on existing research on the success or failure of policy implementation across scales of government and governance.

Notes

1. Hereon forward, technology and innovation—policy or otherwise—will be used interchangeably.
2. Similarly, in the public policy literature Moulton and Sandfort (2017, p. 148) conceptualize the reconciliation between (policy) innovation and organization to be the "degree of intervention alignment with other program processes and technologies." In the implementation science literature, for May *et al.* (2009) and May (2013), this is part of the "normalization process" of a new policy or practice into everyday activities of an organization.
3. The likelihood-ratio test of $\lambda = 0$ (i.e. no over dispersion) can be rejected with a χ^2 value of 1393.33 with probability $\geq \chi^2$ is virtually 0.

References

Brammer, S. and Walker, H. (2011), "Sustainable procurement in the public sector: an international comparative study", *International Journal of Operations and Production Management*, Vol. 31 No. 4, pp. 452-476.

Burchard-Dziubinska, M. and Jakubiec, T. (2012), "Green public procurements (GPP) as an instrument of implementation of sustainable development. Analysis of the experience of the Łódź region local government", *Comparative Economic Research*, Vol. 15 No. 3, doi: 10.2478/v10103-012-0015-9.

Carpenter, D. and Krause, G.A. (2015), "Transactional authority and bureaucratic politics", *Journal of Public Administration Research and Theory*, Vol. 25 No. 1, pp. 5-25, doi: 10.1093/jopart/muu012.

Chatwin, M., Arku, G. and Cleave, E. (2019), "Defining subnational open government: does local context influence policy and practice?", *Policy Sciences*, Vol. 52, pp. 451-479, doi: 10.1007/s11077-018-09347-7.

Darnall, N., Stritch, J.M., Bretschneider, S., Hsueh, L., Duscha, M., Iles, J., No, W., Suarez, J. and Burwell, C. (2017), *Advancing Green Purchasing in Local Government*, Arizona State University, Center for Organization Research and Design, Sustainable Purchasing Research Initiative, Phoenix.

Darnall, N., Stritch, J.M., Hsueh, L. and Bretschneider, S. (2018), "A framework for understanding sustainable public purchasing", *Academy of Management Proceedings*, No. 1, 15677.

Fernandez, S., Malatesta, D. and Smith, C.R. (2013), "Race, gender, and government contracting: different explanations or new prospects for theory?", *Public Administration Review*, Vol. 73 No. 1, pp. 109-120, doi: 10.1111/j.1540-6210.2012.02684.x.

Goggin, M. (1986), "The 'Too Few Cases/Too Many Variables' problem in implementation research", *Western Political Quarterly*, Vol. 39 No. 2, pp. 328-347, doi: 10.1177/106591298603900210.

Gray, V., Hanson, R.L. and Kousser, T. (2012), *Politics in the American States: A Comparative Analysis*, Sage Publication, Washington, DC.

Hill, H.C. (2003), "Understanding implementation: street-level bureaucrats' resources for reform", *Journal of Public Administration Research and Theory*, Vol. 13 No. 3, pp. 265-282, doi: 10.1093/jpart/mug024.

Hsueh, L. and Darnall, N. (2017), "Introduction: alternative and nonregulatory approaches to environmental governance", Virtual Issue on Alternative and Nonregulatory Approaches to Environmental Governance, *Journal of Public Administration Research and Theory*, available at: https://academic.oup.com/jpart/pages/environmental_governance_vi.

Ingraham, P.W., Joyce, P.G. and Donahue, A.K. (2003), *Government Performance: Why Management Matters*, Johns Hopkins University Press, Baltimore, Maryland, available at: https://jhupbooks.press.jhu.edu/content/government-performance.

Ji, H. and Darnall, N. (2018), "All are not created equal: assessing local governments' strategic approaches towards sustainability", *Public Management Review*, Vol. 20 No. 1, pp. 154-175, doi: 10.1080/14719037.2017.1293147.

Johnson, P.F., Leenders, M.R. and McCue, C. (2003), "A comparison of purchasing's organizational roles and responsibilities in the public and private sector", *Journal of Public Procurement*, Vol. 3 No. 1, pp. 57-74, doi: 10.1108/JOPP-03-01-2003-B003.

Krause, R.M. (2011a), "Symbolic or substantive policy? Measuring the extent of local commitment to climate protection", *Environment and Planning C: Government and Policy*, Vol. 29, pp. 46-62.

Krause, R.M. (2011b), "Policy innovation, intergovernmental relations, and the adoption of climate protection initiatives by US Cities", *Journal of Urban Affairs*, Vol. 33 No. 1, pp. 45-60, doi: 10.1111/j.1467-9906.2010.00510.x.

Krebs, T.B. and Pelissero, J.P. (2010), "Urban managers and public policy: do institutional arrangements influence decisions to initiate policy?", *Urban Affairs Review*, Vol. 45 No. 3, pp. 391-411, doi: 10.1177/1078087409342693.

Leonard-Barton, D. (1988), "Implementation as mutual adaptation of technology and organization", *Research Policy*, Vol. 17 No. 5, pp. 251-267, available at: http://www.sciencedirect.com/science/article/pii/0048733388900066.

Little, R.J.A. and Rubin, D.B. (2014), *Statistical Analysis with Missing Data*, John Wiley & Sons.

Marvel, J.D. and Resh, W.G. (2015), "Bureaucratic discretion, client demographics, and representative bureaucracy", *The American Review of Public Administration*, Vol. 45 No. 3, pp. 281-310, doi: 10.1177/0275074013492008.

May, C. (2013), "Towards a general theory of implementation", *Implementation Science,*, Vol. 8, pp. 18, doi: 10.1186/1748-5908-8-18.

May, P.J. and Winter, S.C. (2009), "Politicians, managers, and street-level bureaucrats: influences on policy implementation", *Journal of Public Administration Research and Theory*, Vol. 19 No. 3, pp. 453-476, doi: 10.1093/jopart/mum030.

May, C.R., Mair, F., Finch, T., MacFarlane, A., Dowrick, C., Treweek, S., Rapley, T., Ballini, L., Ong, B.N., Rogers, A., Murray, E., Elwyn, G., L_egar_e, F., Gunn, J. and Montori, V.M. (2009), "Development of a theory of implementation and integration: normalization process theory", *Implementation Science*, Vol. 4, pp. 29, doi: 10.1186/1748-5908-4-29.

McCrudden, C. (2004), "Using public procurement to achieve social outcomes", *Natural Resources Forum*, Vol. 28 No. 4, pp. 257-267, doi: 10.1111/j.1477-8947.2004.00099.x.

McCue, C.P. and Gianakis, G.A. (2001), "Public purchasing: who's minding the store?", *Journal of Public Procurement; Boca Raton*, Vol. 1 No. 1, pp. 71-95.

McGrann, K. (2014), "Benign neglect: veteran-owned small business in federal procurement today", *Veterans Law Review*, Vol. 6, p. 187.

Moulton, S. and Sandfort, J.R. (2017), "The strategic action field framework for policy implementation research", *Policy Studies Journal*, Vol. 45 No. 1, pp. 144-169, doi: 10.1111/psj.12147.

Mullin, M. and Daley, D.M. (2010), "Working with the state: exploring interagency collaboration within a federalist system", *Journal of Public Administration Research and Theory*, Vol. 20 No. 4, pp. 757-778, doi: 10.1093/jopart/mup029.

Nijaki, L.K. and Worrel, G. (2012), "Procurement for sustainable local economic development", *International Journal of Public Sector Management*, Vol. 25 No. 2, pp. 133-153, doi: 10.1108/09513551211223785.

OECD, E.C (2008), *Handbook on Constructing Composite Indicators: Methodology and User Guide*, OECD Publishing, Paris.

Pandey, D., Agrawal, M. and Pandey, J.S. (2011), "Carbon footprint: current methods of estimation", *Environmental Monitoring and Assessment*, Vol. 178 Nos 1-4, pp. 135-160, doi: 10.1007/s10661-010-1678-y.

Pentland, B.T. and Rueter, H.H. (1994), "Organizational routines as grammars of action", *Administrative Science Quarterly*, Vol. 39 No. 3, pp. 484-510, doi: 10.2307/2393300.

Peterson, P.E., Rabe, B.G. and Wong, K.K. (2010), *When Federalism Works*, Brookings Institution Press, Washington, DC.

Pugh, D.S., Hickson, D.J., Hinings, C.R., Macdonald, K.M., Turner, C. and Lupton, T. (1963), "A conceptual scheme for organizational analysis", *Administrative Science Quarterly*, Vol. 8 No. 3, pp. 289-315, doi: 10.2307/2390971.

Qiao, Y., Thai, K.V. and Cummings, G. (2009), "State and local procurement preferences: a survey", *Journal of Public Procurement*, Vol. 9 Nos 3/4, pp. 371-410.

Rabe, B. (2006), "Race to the top: the expanding role of US. State renewable portfolio standards", *Sustainable Development Law and Policy*, Vol. 7, p. 10.

Rabe, B.G. (2007), "Beyond Kyoto: climate change policy in multilevel governance systems", *Governance*, Vol. 20 No. 3, pp. 423-444, doi: 10.1111/j.1468-0491.2007.00365.x.

Shoemaker, P.J., Eichholz, M. and Skewes, E.A. (2002), "Item nonresponse: distinguishing between don't Know and refuse", *International Journal of Public Opinion Research*, Vol. 14 No. 2, pp. 193-201, doi: 10.1093/ijpor/14.2.193.

Smith, C.R. and Fernandez, S. (2010), "Equity in federal contracting: examining the link between minority representation and federal procurement decisions", *Public Administration Review*, Vol. 70 No. 1, pp. 87-96, doi: 10.1111/j.1540-6210.2009.02113.x.

Snider, K.F. and Rendon, R.G. (2012), "Public procurement: public administration and public service perspectives", *Journal of Public Affairs Education*, Vol. 18 No. 2, pp. 327-348.

Telgen, J., Harland, C. and Knight, L. (2007), "Public procurement in perspective", in Knight, L., Harland, C., Telgen, J., Thai, K.V., Callender, G. and McKen, K. (Eds), *Public Procurement: International Cases and Commentary*, Routledge, Abingdon, Oxfordshire, pp. 16-24, doi: 10.4324/9780203815250-9.

Thai, V. (2001), "Public procurement re-examined", *Journal of Public Procurement*, Vol. 1 No. 1, pp. 9-50.

Thurstone, L.L. (1947), *Multiple-factor Analysis; a Development and Expansion of the Vectors of Mind*, University of Chicago Press, Chicago, Illinois.

Trochim, W.M. (2006), *The Research Methods Knowledge Base*, 2nd ed., Conjoint.ly, Sydney, available at: http://www.socialresearchmethods.net/kb/.

Vig, N.J. and Kraft, M.E. (Eds) (2016), *Environmental Policy: New Directions for the Twenty-First Century*, 9th ed., CQ Press; SAGE Publications, Washington, DC.

Appendix

Appendix is available online for this article.

Corresponding author

Lily Hsueh can be contacted at: lily.hsueh@asu.edu