Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Energy Efficiency in Small and Medium-Sized Manufacturing Firms: Order Effects and the Adoption of Process Improvement Recommendations

Suresh Muthulingam, Charles J. Corbett, Shlomo Benartzi, Bohdan Oppenheim

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 15, No. 4, Fall 2013, pp. 596–615
ISSN 1523-4614 (print) | ISSN 1526-5498 (online)
http://dx.doi.org/10.1287/msom.2013.0439

Energy Efficiency in Small and Medium-Sized Manufacturing Firms: Order Effects and the Adoption of Process Improvement Recommendations

Suresh Muthulingam
Samuel Curtis Johnson Graduate School of Management, Cornell University, Ithaca, New York 14853, muthulingam@cornell.edu

Charles J. Corbett, Shlomo Benartzi
Anderson School of Management, University of California, Los Angeles, Los Angeles, California 90095
{charles.corbett@anderson.ucla.edu, benartzi@ucla.edu}

Bohdan Oppenheim
Frank R. Seaver College of Science and Engineering, Loyola Marymount University, Los Angeles, California 90045, boppenheim@lmu.edu

In many manufacturing operations, profitable energy efficiency opportunities remain unexploited. Although previous studies have tried to explain the underinvestment, we focus on how the way in which a portfolio of opportunities is presented in a list affects adoption decisions. We use information on over 100,000 energy-saving recommendations made to more than 13,000 small and medium-sized manufacturing firms under the Industrial Assessment Centers program of the U.S. Department of Energy. We find that adoption rates are higher for initiatives appearing early in a list of recommendations. This sequence effect is consistent and large: simply moving a recommendation one position lower has the same effect on average as increasing up-front implementation cost by at least 17% from the average value. Given this impact of sequence on adoption of individual recommendations, we utilize variations within our data to examine how various sequencing approaches affect adoption at the portfolio level. Sequences in which recommendations are listed from best to worst payback achieve higher potential energy savings given the investments in energy efficiency made by the firms. We also observe a choice overload effect at the portfolio level, but the magnitude of this effect is small.

Key words: process improvement; energy efficiency; behavioral operations; order effects; econometric analysis; empirical research; energy-related operations; environmental operations
History: Received: April 1, 2012; accepted: March 4, 2013. Published online in *Articles in Advance* June 14, 2013.

1. Introduction

Energy efficiency has been recognized since the early 1970s as an often profitable endeavor to reduce energy consumption in many manufacturing operations. However, many profitable energy efficiency opportunities in industrial plants still remain unexploited (Bernstein et al. 2007). The substantial literature examining this apparent paradox (Jaffe and Stavins 1994a, DeCanio 1998, Charles 2009) tends to examine individual opportunities in isolation, putting forward various reasons why one particular seemingly profitable opportunity fails to be pursued. In practice energy efficiency opportunities are usually presented as sets, typically in the form of a list. Instead of asking why managers invest the total amount that they do, we examine how the structure of the list affects firms' adoption decisions.

Energy efficiency is closely aligned with operations management (OM) for two broad reasons. First, energy constitutes a significant share of manufacturing input. In the United States, industry accounted for a third of the 23.4 quadrillion British thermal units (BTUs) of energy consumed in 2010 (U.S. Energy Information Administration 2012). Furthermore, onsite losses were estimated at 32% of the energy input to plants (U.S. Department of Energy Industrial Technologies Program 2004), which is one of the drivers for the U.S. Department of Energy (DOE) to undertake several initiatives to promote energy efficiency. Second, improving energy efficiency in manufacturing typically involves process improvements, such as modification or replacement of equipment, improved management of existing systems, and minimization of waste or resource usage. However, adoption of energy efficiency has not received much attention in the OM literature to date; e.g., the review by Kleindorfer et al. (2005) mentions energy a few times, but does not cite any work that focuses on energy efficiency. Aflaki et al. (2012) propose a conceptual framework linking energy efficiency to sustainable operations, and briefly discuss how myopia, excessively high discount rates, complexity, and ambiguity contribute to lower adoption of energy efficiency projects than expected. Here we take the level of (under)investment as given and focus on how firms choose between projects conditional on their overall investment, as a function of the structure of the list of recommendations. In doing so, we contribute to the OM literature by our focus on energy efficiency and by introducing a significant but previously overlooked behavioral factor that affects OM decision making in practice.

Our data come from the Industrial Assessment Centers (IAC) program of the DOE, which provides free energy audits to small and medium-sized manufacturing firms in the United States. Our work has been informed by close interaction with the IAC program. One of the authors of this paper was the director of an IAC in California and led 125 assessments from 2001 to 2007. We also interviewed or visited five firms assessed under the IAC program. At the end of our work, we interviewed eight other IACs. Moreover, two authors of this study collaborated with the IAC at San Diego State University (SDSU) as part of a different project, which provided further insight into the entire assessment process and subsequent adoption decisions by participating firms.

The IAC program has been in existence since 1976 and is estimated to have provided cumulative energy savings of 1,714 trillion BTUs by 2007 (U.S. Department of Energy Industrial Technologies Program 2009). The energy efficiency assessments are done by faculty and students from accredited engineering schools (Muller et al. 2004). The recommendations usually have attractive rates of return, and their average payback period is just over a year. A former IAC director (one of the authors) illustrates how easily substantial savings can be achieved: "A quarter-inch diameter hole in a compressed air system implies $5,000 per year in wasted energy costs." Despite this, many energy efficiency recommendations are not implemented. From 1981 to 2006, less than half of the identified energy savings have been implemented.

Many studies indicate that a significant proportion of energy efficiency opportunities still remain unexploited (Expert Group on Energy Efficiency 2007). Several reasons have been proposed in the literature including market-failure and non-market-failure explanations (Jaffe and Stavins 1994b), organizational and institutional factors (DeCanio 1998), technology adoption and learning by using (Mulder et al. 2003), a real options framework (Dierderen et al. 2003), and complexity of regulation (Mueller 2006). More recently, researchers have suggested that behavioral factors may play a greater role in explaining the low adoption of energy efficiency opportunities (Charles 2009, Allcott and Mullainathan 2010). We add to this behavioral perspective by facilitating a better understanding of why managers make the specific choices that they do when faced with a list of opportunities, conditional on their overall investment level. Even though we do not address underinvestment directly, understanding managers' choice behavior is essential to ultimately mitigating their underinvestment in energy efficiency.

In this paper, we investigate (non)adoption of over 100,000 energy-saving recommendations made to more than 13,000 manufacturing firms. It is known that economic attributes (payback, cost, savings) influence the adoption rates of individual recommendations, as documented by Anderson and Newell (2004) using the same database. We focus specifically on two noneconomic factors: the sequence in which the recommendations are presented and the total number of recommendations. This enables us to understand how the overall portfolio of recommendations, and specifically their list structure, influences managerial behavior. The manager's decision problem can be thought of as closely related to the classic knapsack problem, where typically the way items in the knapsack are labeled is assumed to be irrelevant. We find that the sequence in which recommendations are presented has a large and consistent effect on managers' choices.

Although it may not appear unexpected that recommendations listed earlier are more likely to be implemented, it is still an open question in the psychology literature whether and when such a primacy effect exists (Carney and Banaji 2012, Mantonakis et al. 2009). Moreover, such an effect has not been documented before in an operations management context, where one would expect such biases to be less strong than in individual or consumer judgment contexts. The effect is quite large and consistent: simply moving a recommendation one position lower in an average assessment has the same effect as increasing upfront implementation cost by at least 17% from the average value. Although the magnitude and consistency of this effect are interesting findings in themselves, they do not immediately prescribe how to order a list of recommendations to maximize overall adoption. Should attractive recommendations be listed first to maximize their adoption, or should less attractive recommendations be listed first, in the hope that they will be adopted while the more attractive recommendations will be adopted anyway despite appearing toward the end of the list? We find that the former ordering is more effective: firms adopt 3.38% more of potential energy savings given their chosen investment level when recommendations with fastest payback are listed first. We also find that overall adoption of energy savings falls by 1.11% with each additional recommendation included, which suggests the impact of the total number of recommendations is low.

This study makes several contributions to the OM literature and the choice literature. First, it is one of the first papers to explicitly examine order effects in OM. Second, we demonstrate that this effect is consistent and large in our context. Third, we highlight how several guidelines for sequencing energy efficiency recommendations in a list can increase or decrease overall adoption. Fourth, our study of these behavioral issues uses actual field data, which connects our work particularly closely to practice.

The rest of this paper is organized as follows. In §2, we present the hypotheses. In §3, we describe the data and the measures used in our list-level analysis. In §4, we present our methodology and results for the list-level analyses. In §5, we discuss the implications of our results at a portfolio level. In §6, we conclude with discussions and limitations.

2. Hypotheses

Our hypotheses draw on the growing body of behavioral operations literature. Gino and Pisano (2008, p. 681) point out that "many questions regarding how common biases studied in behavioral decision research affect operating systems and processes remain unanswered." Our study responds to their call and examines behavioral considerations in decision-making processes within OM contexts. We first develop two hypotheses that focus on the effect of the sequence and length of the list on adoption of individual recommendations. Later, in §5, we examine implications for the portfolio of recommendations.

The first factor we explore is the sequence in which recommendations are listed. This is because firms in the IAC program are provided a list of recommendations from which they can choose any combination of recommendations to implement (Muller et al. 2004), and several studies in the literature have highlighted that choice or judgment is influenced by the order of presentation. Two effects have been proposed that link the order of presentation to choices or judgments. The first effect is the primacy effect, which indicates that information presented early in a sequence has a higher effect on judgment or choices. The second effect is the recency effect, which suggests that information presented later in a sequence has a higher impact on judgment or choices (Anderson 1971). The impact of primacy and recency effects on choices made has been examined in several settings, such as decision makers selecting from a collection of paintings (Li and Epley 2009) or from a collection of wines (Mantonakis et al. 2009), choosing a salesperson or a team to join, choosing a bubblegum, and paroling a convicted criminal (Carney and Banaji 2012). However, the literature finds varying support for primacy and recency effects.

The experiments done by Mantonakis et al. (2009) and Carney and Banaji (2012) are probably most relevant for our study because decision makers are presented options sequentially and asked to choose the best option after observing all options. Both studies predict and find primacy effects and put forth several reasons to support their prediction. For instance, Mantonakis et al. (2009) point out that decision makers need to form summary impressions of each option to facilitate evaluation. They claim that this process could lead to primacy effects because of first-impression effects, biased processing of later information (driven by a confirmatory mindset) or simple attention decrement. Similarly, Carney and Banaji (2012) claim that primacy effects will be observed in their context because things experienced first are better remembered, create stronger associations, influence impressions more decisively, and persuade more effectively. Whereas the evidence regarding primacy effects is consistent in these studies, the findings related to recency effects are mixed. Mantonakis et al. (2009) find that when expert decision makers evaluate longer sequences, they persist in looking for the best option and effectively compare the memory of past options with the current option in longer sequences, which leads to recency effects. However, Carney and Banaji (2012) do not find recency effects even in experiments optimized to yield recency effects.

Our decision context differs from those of Mantonakis et al. (2009) and Carney and Banaji (2012) in many ways. They focus on choices by consumers and individuals rather than firms; their decision makers can only choose a single option rather than any number of options; their choice has to be made relatively quickly and with little or no conscious analysis, unlike our situation, where firms can analyze each option in detail. Finally, we use large-scale field data rather than experiments. For each of these reasons, it is not evident whether their results would carry over to our context. However, their theoretical arguments related to primacy effects are more likely to apply to our decision context. Moreover, given the mixed evidence for recency effects in the literature and the fact that firms assessed in the IAC program do not have energy efficiency experts (Muller et al. 2004), it is less likely that recency effects will be observed in our context. Therefore, we hypothesize the following:

Hypothesis 1. *Recommendations that occur earlier in an assessment will have higher adoption rates than recommendations that occur later in an assessment.*

The second factor we explore is the number of recommendations in an assessment. The OM literature has found that managerial decisions deviate from profit maximization when the situations are inherently complex (Deshpande et al. 2003, Keizers et al. 2003). In our context, we expect to see this reflected in lower adoption rates for recommendations that appear in longer lists. This effect is referred to as "choice overload" and has been examined mainly in the context of consumers or individual decision makers. Iyengar and Lepper (2000) find that consumers provided a wide array of choices (24 flavors of jam) are less likely to make a purchase than those given limited choices (6 flavors of jam). Benartzi and Thaler (2007) find similar phenomena in the context of planning retirement savings. More recently, scholars have investigated what moderates this effect. Chernev (2003a, b) demonstrates that in situations with some novelty and complexity, decision makers with clearly articulated attribute preferences are more likely to choose from sets with more choices than those without clear preferences. He claims that decision makers without clear preferences face the more challenging task of evaluating alternatives while simultaneously developing the criteria for evaluation. Gourville and Soman (2005) find that "overchoice" is also moderated by the type of alternatives provided. Decision makers prefer larger choice sets when alternatives vary along a single compensatory attribute (engine size of a car). However, when alternatives vary along multiple noncompensatory attributes (sunroof vis-à-vis leather interiors), decision makers exhibit "choice overload" due to the increased cognitive effort required to process all relevant information.

The context of the IAC program involves manufacturing firms, which differs from the above studies that focus mainly on individual consumers. Furthermore, in the IAC program firms can adopt more than one recommendation from the set provided. Despite these differences, we expect the findings from the "choice overload" literature to carry over. This is because, in the IAC program, firms are typically provided a set of recommendations that differ along several attributes, and though the firm may have a clear view of the financial criteria, the other attributes are less directly comparable, (e.g., whether it pertains to a manufacturing process, supplier practices, direct labor, etc.), which increases the cognitive efforts involved in evaluating the recommendations. Therefore, cognitive effort increases with the number of recommendations provided. Consequently, we predict that the adoption rate of any specific recommendation will decrease, ceteris paribus, as the list it appears in becomes longer.

Hypothesis 2. *Adoption rates of individual recommendations will fall as more recommendations are made in the same assessment.*

We test our hypotheses on the adoption of individual recommendations in §4. Subsequently, we undertake several additional tests to assess the robustness of our results. In §5, we explore the implications of our findings for adoption at the portfolio level. In §3, we discuss the data used in our analyses and define the measures used to examine adoption of individual recommendations.

3. Data and Measures

3.1. Data and Context

The DOE's IAC program funds a network of universities to conduct free energy assessments for small and medium-sized manufacturing firms. Over 50 universities have participated in the program at various times since it started in 1976. In fiscal year 2010, the budget for the IAC program was $3.87 million, and 386 assessments were performed (U.S. Department of Energy 2011).

Firms are chosen based on multiple criteria. Plants whose products fall within Standard Industrial Classification (SIC) codes 20–39, that are within 150 miles of the host IAC, that have annual sales below $100 million, fewer than 500 employees, annual energy bills between $100,000 and $2 million, no technical staff whose primary duty is energy analysis are eligible for assessment (Muller et al. 2004). A small number of larger firms were also assessed, based on special request from the DOE.

Firms may either contact the IAC requesting an assessment, or the IAC may directly contact potential firms. Once a firm agrees, the IAC team collects information on the current energy usage. Next, an IAC team visits the firm's manufacturing facility; the visit entails interviews with plant management, plant tours, and collection of operational data. One of the authors, a former IAC director, indicated that sometimes it was surprisingly easy to identify opportunities: "In some plants we hear a constant hiss which indicates a compressed air leak." Other recommendations are identified by analyzing operational data. As the former IAC director said, "In one plant we saw excess flash (extra material) on parts made using injection molding. Using specific heat values for the molding material we found that they were using around 40 times the energy required." After the visit, the team provides a written report with specific recommendations involving energy, waste streams, and productivity. The DOE classifies recommendations by Assessment Recommendation Code (ARC) into 25 major categories and over 600 subcategories. After six to nine months, the IAC asks the firm which recommendations have been or will be implemented in the next year. The IAC tracks which recommendations have been adopted over a period of two years. A database with information on all recommendations and assessments done since 1981 is maintained on a public website hosted by the Center for Advanced Energy Systems at Rutgers University. The assessment-level information includes plant demographics such as annual sales, employees, plant area, production hours, energy consumed, manufacturing sector, date of assessment, etc. The recommendation-level information includes expected savings, implementation costs, payback in years (calculated as implementation costs divided by annual savings), energy conserved, implementation status, the ARC number, and the order in which it appeared in the assessment. Details on the IAC database and the assessment process are available in *The DOE Industrial Assessment Database Manual* (Muller et al. 2004).

3.1.1. Synopsis of Observations from Our Interviews and Interactions with the IAC Program. One of the authors of this paper interviewed managers of five firms assessed by an IAC in California and visited one of them. The firms were in the dyeing, electronic products, aluminum casting, metal plating, and smelting industries and had 20–300 employees. The most common recommendations were to replace an existing HVAC unit with a high-efficiency model and to use a more efficient light source.

We observed that adoption of recommendations was driven not only by implementation costs, savings, and technology type, but also by, among others, whether the recommendations were easy or hard to implement and whether manpower was available for implementation. Furthermore, two of the authors interacted with the IAC at SDSU on a different project, which included looking at reasons for adoption or nonadoption of recommendations. In addition, at the end of our study, we also interviewed eight other IACs to understand how they sequence their recommendations. We found that there are no common guidelines that govern how the IACs present their recommendations. Three of the IACs did not follow any explicit sequencing rules, three sought to present recommendations with higher savings earlier in the assessment, one listed recommendations with shorter payback earlier, and one grouped recommendations by technology type. Even for the IACs that claimed they followed certain guidelines in sequencing the recommendations, when we looked at their assessments, we found that they did not always follow the stated guidelines. None of the IACs expected order effects to influence implementation decisions. One IAC director stated, “I find it hard to believe that they (clients) are going to ignore a recommendation based on the summaries if it is not listed first. Firms implement based on economic factors.”

3.1.2. Data Used for Analyses. We use data from 1981 to 2006. We exclude more recent data because in some instances data collection may be incomplete. We adjust all monetary figures for inflation, scaling to year 2006 U.S. dollars using the Producer Price Index WPUSOP3000 series for finished goods from the Bureau of Labor Statistics (2008). We exclude 2,824 recommendations that are not from 1981–2006, 4,723 that do not have information on implementation status, 778 with payback longer than nine years, 44 that show costs but no savings, and 8 that show negative implementation costs. We also exclude 434 firms with sales over $100 million or with SIC codes outside 20–39. In all cases, our conclusions do not change if we include these observations. Our main analysis is based on 89,299 recommendations. Because the identity of the firms in the IAC database is confidential, we are unable to obtain firm-level data on profitability, budgets, etc.

Tables 1 and 2 provide descriptive statistics and correlations for our data. The average estimated implementation cost and annual savings across all recommendations are $19,118 and $17,791. The average payback period is just over a year. Firms adopted 50.16% of all recommendations.

One may ask whether the IAC teams’ estimates of costs and savings are accurate. The DOE has assessed the IAC program at various times using third parties. Martin et al. (1999) evaluated audits done in 1997 and found that the direct savings realized are in line with the projected savings. The five firms we interviewed also vouched for the accuracy of the recommendations.

3.2. Key Variables Used in List-Level Analysis

The dependent variables for our list-level analyses are indicators that represent whether the recommendation is adopted or not.

3.2.1. Independent Variables. The key independent variables used in our list-level analyses are as follows.

Serial Position. We use the actual serial position of the recommendation as it appears in the assessment to test Hypothesis 1 on order effects.

Number of Recommendations. This is the total number of recommendations made in an assessment and is used to test Hypothesis 2 on choice overload.

Managerial Attention Required. This 0–1 variable indicates whether recommendations require high or low managerial attention. The classification was done independently by two of the authors, one a former director of an IAC and the other a former operations consultant who has worked for over a decade on projects similar in nature to the IAC assessments. Of the 684 distinct recommendation types defined

Table 1 Descriptive Statistics

Variable	Mean	SD	Minimum	Maximum	Number
Adopted[a]	0.5016	0.50	0	1	89,299
Payback (years)	1.06	1.29	0	9	89,299
Implementation Cost (US$)	19, 117.74	237, 804.30	0	34,643,628	89,299
Annual Savings (US$)	17, 790.80	113, 238.70	1.12	8,519,905	89,299
Number of Recommendations	8.37	3.03	1	29	89,299
Serial Position	4.69	2.97	1	29	89,299
Annual Sales (US$)	30,961,110	26,361,626	0[b]	155,426,368	89,299
Employees	164.27	139.59	0[b]	3,200	89,299
Annual Energy Cost (US$)	628,994	1,054,627	0[b]	33,914,308	89,299
Knapsack Gap	0.11	0.18	0	0.90	10,232[c]

Notes. Statistics are based on data for the 89,299 recommendations, representing 12,269 assessments. Monetary figures are in 2006 U.S. dollars.

[a] *Adopted* equals 1 if the recommendation is implemented and 0 otherwise.

[b] Data are missing and coded as 0 for (1) *Annual Sales* (745 records), (2) *Employees* (101 records), and (3) *Annual Energy Costs* (37 records). All of the analyses have also been done by removing the missing data, and the results of the study are still valid.

[c] Data for 2,017 assessments are not used for these analyses because no recommendations were implemented. All the analyses have been done by including these data, and the results of the study are still valid.

by the IAC program, 155 were identified as requiring low managerial attention and 79 as requiring high attention. The remaining 450 types could not be unambiguously classified. The kappa statistic measure of interrater agreement is 0.85, which is quite high. Landis and Koch (1977) suggest that a kappa statistic of above 0.81 represents almost perfect agreement. Our procedure is conservative because we use only recommendations that can be clearly classified as requiring low or high attention. A sample list of recommendations that are classified as requiring low and high managerial attention is provided in Table 3.

3.2.2. Control Variables. *Economic Characteristics of a Recommendation.* We follow Anderson and Newell (2004) and use six variables to control for the economic characteristics of a recommendation: $\ln(Payback)$, $[\ln(Payback)]^2$, $\ln(Cost)$, $[\ln(Cost)]^2$, $\ln(Savings)$, and $[\ln(Savings)]^2$. Cost is one-time implementation costs, for equipment, installation, training, etc. Savings represent the expected annual savings. Payback is defined as costs divided by savings, so lower values indicate shorter time to recoup up-front investments. Following Anderson and Newell (2004), we normalize payback, cost, and savings so that their mean equals 1 to ease interpretation of the coefficients. In line with Anderson and Newell (2004), we use the logarithmic form because it improves model fit, but the linear form provides similar results.

Recommendation Type. We include indicator variables for each of the 25 different mutually exclusive major categories of recommendation type based on the first two digits of the ARC code, to control for the underlying heterogeneity among the recommendations.

Variance of Payback of a Recommendation Type. Managers may believe a recommendation's returns are uncertain, based on their past experience with similar recommendations, discussions with managers in other firms, or even looking at historical data in the IAC database. To control for this perceived uncertainty, we use the variance of payback of recommendation type i across all firms that got recommendation type i, computed as $\sum_{j\in J(i)}[(Payback)_{ij} - (Average\ Payback)_i]^2$, where $J(i)$

Table 2 Correlations

	Correlations (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) *Adopted*[a]	1.00									
(2) *Payback* (years)	−0.13	1.00								
(3) *Implementation Cost* (US$)	−0.04	0.12	1.00							
(4) *Annual Savings* (US$)	−0.04	0.00	0.54	1.00						
(5) *Number of Recommendations*	0.00	−0.05	0.00	0.01	1.00					
(6) *Serial Position*	−0.04	−0.01	−0.01	0.00	0.51	1.00				
(7) *Annual Sales* (US$)	−0.01	−0.01	0.04	0.07	0.08	0.04	1.00			
(8) *Employees*	−0.01	−0.01	0.01	0.04	0.08	0.04	0.50	1.00		
(9) *Annual Energy Cost* (US$)	−0.01	−0.01	0.11	0.16	0.10	0.05	0.33	0.25	1.00	
(10) *Knapsack Gap*	−0.01	−0.03	−0.02	−0.01	0.13	0.06	0.02	0.02	−0.01	1.00

[a] *Adopted* equals 1 if the recommendation is implemented and 0 otherwise.

Table 3 Select List of Recommendations That Need Low or High Managerial Attention

ARC code	Description of recommendations that need low managerial attention
2.7142	Utilize higher efficiency lamps and/or ballasts
2.4236	Eliminate leaks in inert gas and compressed air lines/valves
2.4221	Install compressor air intakes in coolest locations
2.4111	Utilize energy-efficient belts and other improved mechanisms
2.2511	Insulate bare equipment
2.4231	Reduce the pressure of compressed air to the minimum required
2.7143	Use more efficient light source
2.7135	Install occupancy sensors
2.1233	Analyze flue gas for proper air/fuel ratio
2.7261	Install timers and/or thermostats
ARC code	**Description of recommendations that need high managerial attention**
2.1311	Replace electrically operated equipment with fossil fuel equipment
2.4141	Use multiple speed motors or adjustable frequency drive for variable pump, blower and compressor loads
2.2434	Recover heat from air compressor
2.1123	Install automatic stack damper
2.2411	Use waste heat from hot flue gases to preheat combustion air
2.2531	Resize charging openings or add movable cover or door
2.1222	Install turbulators
2.4131	Replace oversize motors and pumps with optimum size
2.3415	Use a fossil fuel engine to cogenerate electricity or motive power; and utilize heat
2.5194	Redesign process

represents all firms that were given recommendation i. This variable is not a perfect measure of perceived uncertainty because it also captures the underlying heterogeneity of the firms, but as long as there is some recommendation-specific component to this overall variance, this measure will be correlated with the uncertainty associated with a recommendation type.

Assessment Year. We use indicator variables to identify the year of the assessment.

Assessment Quarter. Stern and Aronson (1984) indicate that expenses that fit into the present budget cycle require fewer approvals. The 1998 Survey of Small Business Finances by the Federal Reserve Board finds that for nearly 85% of small firms in the United States, the fiscal year coincides with the calendar year. Consequently, to capture the impact of budgetary cycles, we use indicator variables to identify the specific calendar quarter in which the assessment was done.

IAC Control. We use indicator variables to identify which IAC undertook the assessment, which also serve as surrogate control for the state in which the assessed firm is located.

SIC Control. We use indicator variables for each firm's two-digit SIC code.

Other Firm-Level Control. We use sales, number of employees, and the plant area (in square feet) as additional controls for firm-level effects.

4. Methodology and Results for the List-Level Analyses

We test Hypotheses 1 and 2 using models that build on Anderson and Newell (2004), who estimate a conditional logit model and find, as one would expect, that the initial costs, savings, and the payback of a recommendation have a significant effect on the adoption of the recommendation. We enhance their model by including two additional key variables, the serial position of each recommendation and the total number of recommendations in an assessment, as well as several additional controls. However, one cannot simply add these variables to the Anderson–Newell models, for two reasons. First, the serial position of the recommendations in an assessment may be endogenous, so we use probit instrumental variable (IV) models. Second, we can no longer use firm-level fixed effects with these new variables, so we need additional firm-level controls such as sales, number of employees, and two-digit SIC codes. All analyses were done using STATA version 10.1.

4.1. Probit and Instrumental Variables Probit Models

We use an indicator variable Y_{ij} that equals 1 if recommendation i in assessment j is adopted and 0 otherwise. The resultant choice problem is defined by the latent variable model:

$$Y_{ij}^* = \alpha + \mathbf{M}_{ij}\boldsymbol{\beta} + V_{ij}\gamma + \mathbf{T}_{ij}\boldsymbol{\rho} + S_{ij}\zeta + N_j\eta + \mathbf{C}_j\boldsymbol{\chi} + \varepsilon_{ij}, \tag{1}$$

$$\varepsilon_{ij} = \delta_i + \mu_j + \acute{\varepsilon}_{ij}, \tag{2}$$

where Y_{ij}^* is the net benefit of adopting recommendation i in assessment j; $\mathbf{M}_{ij}$ is the vector of economic characteristics; V_{ij} is the variance of payback; $\mathbf{T}_{ij}$ is the vector of recommendation type dummies; S_{ij} represents the serial position of recommendation i in assessment j; N_j is the number of recommendations in assessment j; and the matrix $\mathbf{C}_j$ includes controls for the specific IAC, two-digit SIC codes, sales, number of employees, the year of assessment and the calendar quarter in which the assessment was done. The error terms ε_{ij} are decomposed into three parts. The first part is δ_i, which represents recommendation type-related unobserved characteristics and is partially controlled for by including indicators for recommendation type and the variance in payback. The second part is μ_j, which represents assessment-related unobserved characteristics and is partially controlled for by including indicators for specific IAC, two-digit SIC codes, firm size, and firm-level variables such as sales and number of employees. The third part, $\acute{\varepsilon}_{ij}$, is related to the recommendation and firm-specific unobserved characteristics.

Decision makers will adopt a recommendation only if the benefits from adopting are positive, and thus the probability that a recommendation is adopted is

$$\begin{aligned}\text{Prob}[Y_{ij}=1] &= \text{Prob}[\alpha+\mathbf{M}_{ij}\boldsymbol{\beta}+V_{ij}\gamma+\mathbf{T}_{ij}\boldsymbol{\rho}\\ &\qquad +S_{ij}\zeta+N_j\eta+\mathbf{C}_j\boldsymbol{\chi}+\varepsilon_{ij}>0]\\ &= F(\alpha+\mathbf{M}_{ij}\boldsymbol{\beta}+V_{ij}\gamma+\mathbf{T}_{ij}\boldsymbol{\rho}\\ &\qquad +S_{ij}\zeta+N_j\eta+\mathbf{C}_j\boldsymbol{\chi}),\end{aligned} \tag{3}$$

where F is the cumulative distribution function for ε_{ij}. If ε_{ij} follows a standard normal distribution, we have the probit model (Maddala 2003).

Model (1) treats the serial position of the recommendation as exogenous. However, our interviews suggest that some IACs do (or claim to) follow various guidelines in deciding on how to list the recommendations. Some IACs may place recommendations they consider attractive at the top of the list to increase their chance of implementation, or they may place the attractive recommendations later so that other recommendations come earlier in the assessment and get a higher probability of implementation. Because the IACs are evaluated partly on the number of recommendations implemented, they have an incentive to present the recommendations in a manner that will increase adoption. If the IAC's assessment of the attractiveness of each recommendation were captured by the observable variables, then we could use model (1) to obtain consistent results. However, if the IAC's assessment of attractiveness is not observable, the effect of attractiveness will be captured in the error terms, in which case the serial position is correlated with the error term and is therefore endogenous in the model. Although none of the 10 IACs that we interacted with explicitly said they list more "attractive" recommendations earlier, some did say that they list them at least partly based on savings or payback, so it is possible that some of the IACs do to some extent order recommendations based on attractiveness. A Wald test for exogeneity does confirm that the serial position is endogenous. We address this problem by using two instruments that are related to the serial position of the recommendation, but are otherwise unrelated to the error terms (Wooldridge 2002).

We follow the approach suggested by Wooldridge (2002), as also used in Olivares and Cachon (2009), who analyze the impact of competition on inventory levels. They observe that competition may be correlated with unobserved consumer characteristics, which could lead to biased estimates. They use measures of market population as instruments. They indicate that population will be correlated with competition, but not with the unobserved consumer characteristics, and hence measures of population are valid instruments. Analogously, our first instrument is based on the order in which the recommendations appear in the ARC manual. The ARC manual groups recommendations based on the engineering categories of recommendations such as combustion systems, thermal systems, electrical power, and so forth. We use the ARC code to sequence the recommendations made to a firm so that the recommendation with the lowest ARC code is given the first rank, and so forth. The assessors use the ARC codes to report their recommendations to the IAC database, so their listing of recommendations may partly follow the sequence in the ARC manual. In this case, our instrument based on the order of the recommendation in the ARC manual will be correlated with the serial position of a recommendation in an assessment, but not with the unobserved attractiveness of a recommendation. Consequently the ranking based on the order in the ARC manual can serve as a valid instrument for the serial position, just as measures of population serve as valid instruments for competition in Olivares and Cachon (2009).

The second instrument is related to the propensity with which each IAC makes a recommendation. We follow Cachon and Olivares (2010), who analyze the impact of production flexibility on finished goods inventory levels. They claim that there may be a mechanical relationship between production flexibility and the dependent variable. To address this endogeneity, they use production flexibility of other models produced in the same plant as instruments. Analogously, we compile the frequency with which each IAC makes a particular recommendation across all assessments. We use this to rank the recommendations made to a specific firm so that the recommendation with the highest frequency is ranked first, etc. The resulting ranking is a reflection of the IAC's familiarity with specific recommendations, which may be related to the way they present the recommendations. This ranking is based on the IAC's interaction with all firms it has assessed, and as such it is not related to the preferences of a specific firm and hence should not be correlated with the error term. One possible concern with this instrument may be that IACs are likely to recommend initiatives that have higher probability of adoption among all firms. To address this, we estimate the variance of δ_i relative to $\acute{\varepsilon}_{ij}$, using a linear mixed model with random effects incorporated at the recommendation and assessment levels to evaluate the variance components. We find that the estimated variance for δ_i, the recommendation type-related unobserved characteristics, is 0.01, much smaller than that for $\acute{\varepsilon}_{ij}$, the recommendation- and firm-specific unobserved characteristics, which is 0.12.

We use the following instrumental variable probit model to address the endogeneity:

$$Y_{ij}^* = \alpha + \mathbf{M}_{ij}\boldsymbol{\beta} + V_{ij}\gamma + \mathbf{T}_{ij}\boldsymbol{\rho} + S_{ij}\zeta + N_j\eta + \mathbf{C}_j\boldsymbol{\chi} + \varepsilon_{ij}, \tag{4a}$$

$$S_{ij} = \mathbf{M}_{ij}\Pi_\beta + V_{ij}\Pi_\gamma + \mathbf{T}_{ij}\Pi_\rho + N_j\Pi_\eta + \mathbf{C}_j\Pi_\chi + \mathbf{I}_j\Pi_\omega + \nu_{ij}. \tag{4b}$$

In this model, the variable serial position S_{ij} is endogenous, as opposed to model (1) where S_{ij} is exogenous. The linear projection in Equation (4b) represents the reduced form equation for the endogenous explanatory variable S_{ij}, where $\mathbf{I}_j$ represents the vector of instruments. We assume the error terms in (4a) and (4b) are normally distributed and are orthogonal to all regressors. We follow Anderson and Newell (2004) and estimate a "payback" model where we use the variables $\ln(Payback)_{ij}$ and $[\ln(Payback)_{ij}]^2$ for the vector $\mathbf{M}_{ij}$, and similarly a "cost–benefit" model with $\ln(Cost)_{ij}$, $[\ln(Cost)_{ij}]^2$, $\ln(Savings)_{ij}$, and $[\ln(Savings)_{ij}]^2$.

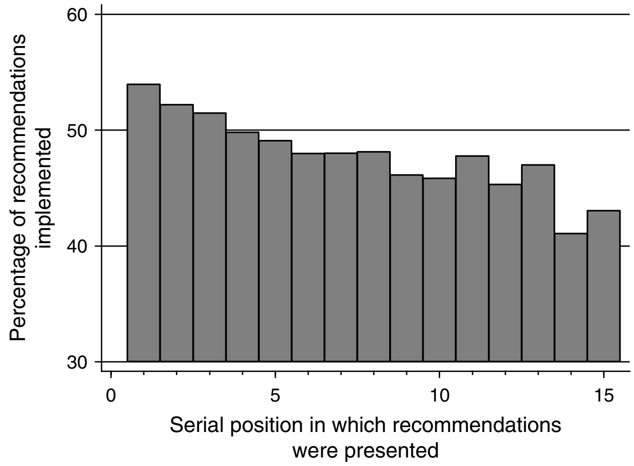
To validate the instruments, we ran an ordinary least squares (OLS) regression of the variables for serial position of the recommendation on the instruments related to the ARC code and the IAC propensity to make a type of recommendation. The R^2 value is 0.23, which is comparable to similar values reported in the literature. Evans and Schwab (1995), in a paper that uses a similar instrumental variable probit methodology as we do, report an R^2 of 0.16 when they regress their endogenous variable, *Catholic school*, on their instrument *Catholic religion*. We also ran an ordered probit model, and the z-statistics for the instrument related to ARC code and the IAC propensity to make a type of recommendation are 91.90 and 70.64, respectively, and both are significant at $p < 0.001$. Therefore, the chosen instruments are valid determinants of the serial position of the recommendations for the model (4a).

Although we find statistical evidence of endogeneity in our data and robustly validate our instruments, we also heed to Murray (2006, p. 130), who cautions, "[One] can never entirely dispel the clouds of uncertain validity that hang over instrumental variable analyses...." Consequently, wherever possible, we evaluate our models with both exogenous and endogenous S_{ij}. Table 4 presents the results for the four models, with exogenous and endogenous S_{ij} for the payback and the cost–benefit models.

4.2. Results

With respect to Hypothesis 1 on order effects, the average adoption rate falls for recommendations that occur later in the assessment (as shown in Figure 1), from over 50% for the earliest recommendations to around 40% for the last ones. Furthermore, the coefficient of the serial position of the recommendation is negative and significant at $p < 0.001$ across all models in Table 4. This supports Hypothesis 1 that the probability of adoption falls as the recommendation occurs later in the assessment. In Table 4 for the "Probit" and "IV probit" cost–benefit models, if we consider an average assessment and move a recommendation one position lower in the assessment, then on average its probability of adoption will fall by 0.0102 and 0.0549, respectively. Even in the more conservative probit model, this means that moving a recommendation from 1st to 10th position would reduce adoption likelihood by 10 percentage points. Given that this is not a controlled experiment, we cannot provide an exact estimate of the magnitude of the serial position effect. However, given that statistically and anecdotally there is evidence that serial position is endogenous, we believe the actual magnitude of the serial position effect lies somewhere between that suggested by the probit and the IV probit models. To have the same effect on adoption likelihood, implementation costs would have to increase by \$3,337 and \$30,699 from the average cost of \$19,118 for the probit and IV probit cost–benefit models, respectively (keeping savings constant). The calculations were done by computing the average drop in probability of adoption when a recommendation is moved one position lower in an average assessment and computing the average increase in cost that results in the same drop in probability of adoption when the recommendation is retained at the existing position. Marginal effects were computed using the approach provided in Greene (2008, p. 780). These results suggest that the serial position effect is substantial. Moreover, the coefficient of the serial position is over five times larger

Figure 1 **Adoption Rate vs. Serial Position of Recommendation in the Assessment**



Note. A drop in adoption rates of over 13% is observed between recommendations that occur in the 1st versus 15th position in an assessment.

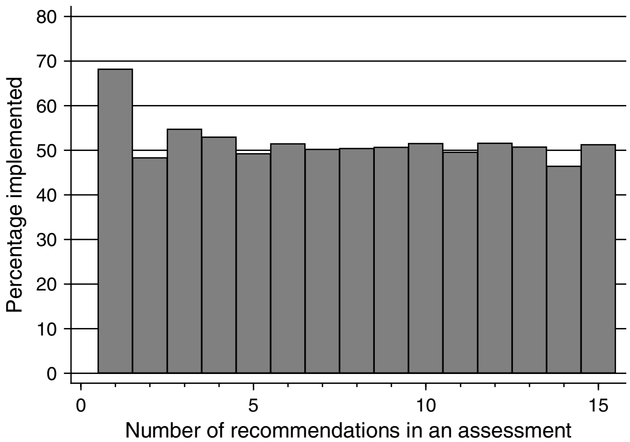
Table 4 Instrumental Variables Probit Estimates of Adoption of Recommendations

Dependent variable: *Adopted* (equals 1 if recommendation is implemented, 0 otherwise)				
	Payback models		Cost–benefit models	
	Probit	IV probit	Probit	IV probit
ln(*Payback*)	−0.1483***	−0.1043***		
	(0.006)	(0.015)		
ln(*Payback*)^2	−0.0167***	−0.0147***		
	(0.002)	(0.002)		
ln(*Cost*)			−0.1643***	−0.1445***
			(0.007)	(0.009)
ln(*Cost*)^2			−0.0097***	−0.0082***
			(0.001)	(0.001)
ln(*Savings*)			0.0796***	0.0156
			(0.008)	(0.014)
ln(*Savings*)^2			−0.0009	0.0023
			(0.002)	(0.002)
Serial Position	−0.0207***	−0.1720***	−0.0268***	−0.1488***
	(0.002)	(0.036)	(0.002)	(0.022)
Number of Recommendations	0.0478+	0.6588***	0.0685**	0.5576***
	(0.025)	(0.146)	(0.025)	(0.091)
Variance of Payback	−0.0791***	−0.1209***	−0.0657***	−0.0755***
	(0.008)	(0.012)	(0.007)	(0.007)
Sales	−0.0023	−0.0001	0.0066	0.0236+
	(0.014)	(0.013)	(0.014)	(0.014)
Energy Costs	−0.0077	−0.0025	0.0024	0.0155**
	(0.005)	(0.005)	(0.005)	(0.006)
Employees	0.0000	0.0000	0.0000	0.0001
	(0.000)	(0.000)	(0.000)	(0.000)
Assessment in 1st Quarter	0.0386+	0.0415*	0.0385+	0.0405*
	(0.021)	(0.020)	(0.021)	(0.021)
Assessment in 2nd Quarter	0.0213	0.0236	0.0217	0.0244
	(0.020)	(0.019)	(0.020)	(0.020)
Assessment in 3rd Quarter	0.0153	0.0222	0.0166	0.0232
	(0.022)	(0.020)	(0.022)	(0.021)
Constant	0.2521	0.5816+	−0.8829+	0.4745
	(0.311)	0.334	(0.458)	(0.336)
Other controls				
Recommendation type (no. significant at $p < 0.05$ out of 25 recommendation types)	5	2	22	9
IAC centers (no. significant at $p < 0.05$ out of 45 IAC centers)	38	35	38	34
Years (no. significant at $p < 0.05$ out of 26 years)	0	0	0	0
SIC code (no. significant at $p < 0.05$ of 20 groupings of two-digit SIC codes)	0	0	0	0
Observations	76,070	76,070	76,070	76,070
Firms (assessments)	12,236	12,236	12,236	12,236
Log-pseudolikelihood	−49,737***	−224,636***	−49,658.7***	−221,033***
Exogeneity Wald statistic	—	14.48		26.78

Notes. Data pertain to recommendations made by IAC centers from 1981 to 2006. The estimation method is maximum likelihood. Standard errors are clustered at the assessment level and reported using robust clustered variance covariance matrix. One IAC center and its 12 related recommendations were dropped from the full sample because all the recommendations were not adopted. A total of 13,187 recommendations were dropped because they had payback equal to 0 and the logarithmic form for payback is not defined. Including these recommendations in a model without logarithmic transformation does not change the inferences we derive from this model. The IV probit models use instrumental variables to instrument the serial position of a recommendation (using sequence generated based on the ARC manual and the propensity with which each IAC makes recommendations). Standard errors are in parentheses.

$^+p < 0.1$; $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Figure 2 **Adoption Rate vs. Number of Recommendations in an Assessment**



Note. On average, approximately half of the recommendations are implemented, irrespective of the number of recommendations made to a firm (except for assessments with a single recommendation).

for the IV probit models than for the probit models. This indicates that the impact of the serial position in the probit models could be understated due to the endogeneity.

An additional interesting observation is the difference in magnitude between the coefficients for costs and those for savings. Anderson and Newell (2004) find that the effect of $1 in up-front costs is 40% greater than that of $1 in annual savings, in contrast to earlier work (not using the IAC data) that found that costs have three (Jaffe and Stavins 1995) or eight (Hassett and Metcalf 1995) times the effect of savings. In our models, correcting for the serial position, we find that costs have between two (probit models) and eight (IV probit models) times the effect of savings in Table 4, a much larger effect than reported by Anderson and Newell (2004) and closer to the range reported in earlier studies.

For Hypothesis 2, we see in Figure 2 (which does not control for payback, etc.) that nearly 50% of the recommendations are implemented irrespective of the number of recommendations made to a firm. Furthermore, we see that the coefficient of the number of recommendations made is positive in all models, and significant at $p < 0.01$ in three of four models in Table 4. This does not support Hypothesis 2, that adoption rates fall as more recommendations are made in an assessment.

4.3. Robustness Checks

We did several additional tests to assess the robustness of our results. Because the recommendation order variable is also related to the total number of recommendations, a possible concern may be that the order effect is partly due to the total number of recommendations. We performed two additional tests to address this concern. First, we redid the analysis related to Table 4 with a normalized measure of serial position instead of the absolute serial position. (In other words, we normalize the serial position of the recommendation within the assessment so that the mean value is 1.) Second, we formed groups of all assessments with the same total number of recommendations and redid the probit instrumental variables analysis within each group. The results of both analyses (available in Tables I and II of the online supplement, available at http://dx.doi.org/10.1287/msom.2013.0439) support the inference that the sequence of recommendations is significant.

Another potential concern is that there could be some unobservable recommendation-specific characteristics that drive adoption. To address this, we selected those recommendations that had another recommendation with the same three- or four-digit ARC code in the same assessment and estimated the probit models within each subsample. The results of these analyses are provided in models (1)–(4) in Table 5, where we see that the coefficient of serial position is negative and significant at $p < 0.05$ in all models. As a further refinement, from the sets of recommendations matched by four-digit ARC code within an assessment, we select sets where the later recommendation has lower costs and higher savings than the earlier matched recommendation. These are sets where the later recommendations financially dominate the earlier recommendations. Within this subsample, we estimate our probit models of adoption using *Relative Serial Position*, defined as the relative rank of recommendations matched by four-digit ARC code within an assessment, instead of *Serial Position*. The results of these tests are shown in models (5) and (6) in Table 5, where the coefficient of *Relative Serial Position* is negative and significant at $p < 0.05$. Even though the later recommendation financially dominates the earlier one and is technically nearly identical (same four-digit ARC code), the earlier recommendation is significantly more likely to be adopted. We believe these results provide compelling evidence for the impact of order effects on adoption.

A different potential concern may be that adoption decisions are driven mainly by the managerial attention required for implementing a recommendation. To address this, we include an indicator variable, H_{ij}, for whether recommendations require high or low managerial attention:

$$Y_{ij}^* = \alpha + \mathbf{M}_{ij}\boldsymbol{\beta} + V_{ij}\gamma + \mathbf{T}_{ij}\boldsymbol{\rho} + S_{ij}\boldsymbol{\zeta} + N_j\eta + H_{ij}\omega + \mathbf{C}_j\boldsymbol{\chi} + \varepsilon_{ij}, \tag{5a}$$

$$S_{ij} = \mathbf{M}_{ij}\Pi_\beta + V_{ij}\Pi_\gamma + \mathbf{T}_{ij}\Pi_\rho + N_j\Pi_\eta + H_{ij}\Pi_\omega + \mathbf{C}_j\Pi_\chi + \mathbf{I}_j\Pi_\omega + \nu_{ij}. \tag{5b}$$

Table 5 Probit Estimates of Adoption of Recommendations with Same Three- or Four-Digit ARC Code

Dependent variable: *Adopted* (equals 1 if implemented, 0 otherwise)

	Three-digit ARC match		Four-digit ARC match		Dominant recommendations later for four-digit ARC match	
	Payback (1)	Cost–benefit (2)	Payback (3)	Cost–benefit (4)	Payback (5)	Cost–benefit (6)
ln(*Payback*)	−0.1549*** (0.013)		−0.1496*** (0.018)		−0.1327 (0.085)	
ln(*Payback*)^2	−0.0236*** (0.004)		−0.0161** (0.005)		−0.0271 (0.025)	
ln(*Cost*)		−0.2121*** (0.017)		−0.2162*** (0.023)		−0.3737** (0.128)
ln(*Cost*)^2		−0.0196*** (0.003)		−0.0186*** (0.004)		−0.0541** (0.019)
ln(*Savings*)		0.0820*** (0.018)		0.1110*** (0.024)		−0.0347 (0.091)
ln(*Savings*)^2		0.0027 (0.003)		0.0048 (0.005)		−0.0437* (0.020)
Serial Position	−0.0199*** (0.004)	−0.0277*** (0.005)	−0.0146* (0.006)	−0.0219** (0.007)		
Relative Serial Position (*for matched recommendations*)					−0.251* (0.099)	−0.2715** (0.101)
Number of Recommendations	0.0065 (0.044)	0.0320 (0.044)	−0.0236 (0.062)	0.0004 (0.063)	−0.5251* (0.262)	−0.5521* (0.264)
Variance of Payback	−0.1616*** (0.021)	−0.1430*** (0.021)	−0.0982** (0.035)	−0.0798* (0.035)	−0.1048 (0.162)	−0.048 (0.164)
Sales	0.0257 (0.024)	0.0365 (0.024)	0.0405 (0.034)	0.0501 (0.034)	0.2911 (0.191)	0.2733 (0.195)
Energy Costs	−0.0008 (0.008)	0.0062 (0.009)	−0.0055 (0.011)	0.0005 (0.011)	−0.0174 (0.055)	0.0248 (0.057)
Employees	−0.0001 (0.000)	−0.0001 (0.000)	0.0000 (0.000)	0.0000 (0.000)	−0.0002 (0.001)	−0.0001 (0.001)
Constant	0.4826 (1.077)	0.4716 (1.097)	1.1208 (0.926)	−1.7255*** (0.492)	0.375 (1.496)	1.778 (1.569)
Other controls						
Recommendation, IAC, year, SIC, quarter	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,664	16,664	7,987	7,987	482	482
Log-pseudolikelihood	−10,826***	−10,792***	−5,167.14***	−5,147.14***	−259.16	−250.62

Notes. Data pertain to recommendations made by IAC centers from 1981 to 2006. The estimation method is maximum likelihood. Standard errors are clustered at the assessment level and reported using robust clustered variance covariance matrix. Three-digit (four-digit) ARC match represents analyses for recommendations that have the same three-digit (four-digit) digit ARC within an assessment. *Relative Serial Position* (*for matched recommendations*) indicates recommendations matched on four-digit ARC such that the matched recommendation appearing later in the assessment has lower implementation costs and higher savings compared to the matched recommendation appearing earlier in the assessment. Standard errors are in parentheses.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

We use the same approach as for models (4a) and (4b) to evaluate models (5a) and (5b), but restrict our analyses to recommendations that were clearly identified as requiring either high or low managerial attention. The results in Table 6 show that the coefficient of serial position is negative and significant at $p < 0.001$ across all models, providing further support for Hypothesis 1.

Our results so far are consistent with an earlier-is-better bias. This may be in part because decision makers find it challenging to evaluate multiple recommendations on various attributes and instead adopt a simplifying heuristic of looking at the sequence. Tversky et al. (1988) point out that decision makers may find it difficult to trade off one attribute against another when they evaluate alternatives on multiple attributes, and instead may resolve conflict by selecting alternatives superior on the most important attribute. In the context of the IAC program, the economic characteristics of a recommendation can be considered as the most important attribute considered by the decision makers (Anderson and Newell 2004). Often firms and decision makers have well-defined criteria on economic characteristics, such as implement projects with payback periods less than three years. Such well-defined preferences are similar

Table 6 Instrumental Variables Probit Estimates of Adoption of Recommendations (Only for Recommendations Classified as Requiring Low or High Managerial Attention for Adoption)

Dependent variable: *Adopted* (equals 1 if recommendation is implemented, 0 otherwise)

	Payback models		Cost–benefit models	
	Probit	IV probit	Probit	IV probit
ln(*Payback*)	-0.1261^{***} (0.007)	-0.0653^{***} (0.016)		
ln(*Payback*)^2	-0.0127^{***} (0.002)	-0.0110^{***} (0.002)		
ln(*Cost*)			-0.1586^{***} (0.010)	-0.1282^{***} (0.014)
ln(*Cost*)^2			-0.0113^{***} (0.002)	-0.0081^{***} (0.002)
ln(*Savings*)			0.1021^{***} (0.011)	0.0098 (0.027)
ln(*Savings*)^2			0.0000 (0.002)	0.0004 (0.002)
Serial Position	-0.0268^{***} (0.002)	-0.2141^{***} (0.035)	-0.0266^{***} (0.003)	-0.1608^{***} (0.034)
Number of Recommendations	0.0588^{*} (0.028)	0.7760^{***} (0.136)	0.0581^{*} (0.029)	0.5680^{***} (0.129)
Variance of Payback	0.0091 (0.009)	-0.0658^{***} (0.017)	0.0141 (0.009)	-0.0145^{***} (0.012)
Sales	0.0113 (0.016)	0.0170 (0.014)	0.0111 (0.016)	0.0341^{*} (0.016)
Energy Costs	−0.0052 (0.006)	0.0006 (0.005)	−0.0032 (0.006)	0.0130^{+} (0.008)
Employees	0.0000 (0.000)	0.0000 (0.000)	0.0000 (0.000)	0.0001 (0.000)
Assessment in 1st Quarter	0.0190 (0.025)	0.0213 (0.022)	0.0194 (0.025)	0.0211 (0.024)
Assessment in 2nd Quarter	0.0104 (0.024)	0.0167 (0.022)	0.0111 (0.024)	0.0176 (0.023)
Assessment in 3rd Quarter	−0.0036 (0.025)	0.0015 (0.023)	−0.0027 (0.025)	0.0022 (0.024)
High Managerial Attention	-0.5644^{***} (0.027)	-0.4744^{***} (0.041)	-0.5446^{***} (0.028)	-0.4551^{***} (0.041)
Constant	−0.3807 (0.432)	1.3103 (0.951)	−0.6335 (0.382)	−0.0605 (0.445)
Other controls				
Recommendation type (no. significant at $p < 0.05$ out of 25 recommendation types)	14	0	15	6
IAC centers (no. significant at $p < 0.05$ out of 45 IAC centers)	34	34	34	34
Years (no. significant at $p < 0.05$ out of 26 years)	3	6	4	4
SIC code (no. significant at $p < 0.05$ of 20 groupings of two-digit SIC codes)	0	0	0	0
Observations	50,033	50,033	50,033	50,033
Firms (assessments)	12,055	12,055	12,055	12,055
Log-pseudolikelihood	−32,388	−145,948	−32,365.9	−143,182
Exogeneity Wald statistic	—	20.37		13.79

Notes. Data pertain to recommendations made by IAC centers from 1981 to 2006. The estimation method is maximum likelihood. Standard errors are clustered at the assessment level and reported using a robust clustered variance–covariance matrix. The IV probit models use instrumental variables to instrument the serial position of a recommendation (using the sequence generated based on the ARC manual and the propensity with which each IAC makes recommendations). Standard errors are in parentheses.

$^{+}p < 0.1$; $^{*}p < 0.05$; $^{***}p < 0.001$.

to the ideal point as defined by Chernev (2003a, b). If the variation in the economic characteristics of recommendations in an assessment is low, then the decision maker still faces the cognitive challenge of choosing between similar options. By contrast, if variation in economic characteristics of recommendations in an assessment is high, then the decision maker faces the cognitively simpler task of choosing among dissimilar options. The reduced cognitive requirements in such situations will lower the impact of behavioral factors (such as looking at the sequence) on adoption. We use the coefficient of variation of payback within an assessment as an indicator of how distinct the recommendations are, and predict that assessments with a higher coefficient of variation will display weaker serial position effects. We examine this in two ways. First we split the sample into four quartiles by coefficient of variation of payback. For each group, we estimate the "cost–benefit" models (4). The results in Table 7 show that the magnitude and significance of serial position decreases as the payback of the recommendations is more widely dispersed. For example, in model (1) of Table 7, the coefficient of serial position is -0.1820 and significant at $p < 0.001$, whereas in model (4) the coefficient of serial position has a lower magnitude of -0.0536 and is not significant. Second, we include an interaction term between the serial position and the coefficient of variation of payback in the cost–benefit models (4), resulting in models (5) and (6) of Table 7. In the interaction model (6), the coefficient of the interaction term is positive and significant at $p < 0.001$, which indicates that the impact of serial position is mitigated for assessments that have higher coefficients of variation of payback.

Mantonakis et al. (2009) predict that longer sequences will display a recency effect when the decision maker can select one option, because an item seen earlier will have to beat more alternatives to be the overall winner than an item seen later. Given that our decision makers can choose multiple options and can spend more time going back and forth through the list, we did not expect to find a recency effect in our context. We did check by including indicator variables for the last few (up to three) recommendations in an assessment, but found no evidence to support recency effects. (These results are available in Table III of the online supplement.)

A potential concern related to Hypothesis 2 is that the effect of the number of recommendations is comingled with the serial position variable. To address this, we formed groups of all recommendations with the same serial position and estimated probit models of adoption rates within each group separately. The results in Table 8 show that the coefficient of the number of recommendations is not significant in all models, providing further evidence against Hypothesis 2. Another mechanism that could have the same effect as choice overload is the presence of budget constraints. We do not observe the budgets for the firms assessed in the IAC program because their identities are kept confidential. Consequently, we use each firm's four-digit SIC code to identify the average industry profitability and operating cash availability in the year in which the assessment was done, and use these as controls in our analysis. (These results are available in Tables IV and V of the online supplement.) Our results remain essentially the same, which suggests that the impact of budget constraints may be distinct from that due to choice overload, though additional firm-level data would be needed to truly assess the effect of budget constraints.

5. Implications of Results at a Portfolio Level

Though we find compelling evidence that sequence affects adoption of individual recommendations, a priori it is not obvious how one should sequence a list of recommendations to maximize adoption at the portfolio level. On the one hand, the IAC could sequence recommendations with desirable economic characteristics earlier to increase their prospect of implementation, or, on the other hand, the IAC could sequence recommendations with less attractive economic characteristic earlier to enhance their prospect of adoption assuming that recommendations with more attractive economic characteristics will be adopted anyway even if they appear at the end of the list. Our context involves two counteracting forces, and their overall impact may depend on the specific assessment. To examine how the sequence of an entire list affects adoption at the portfolio level, we consider five different sequencing rules and use the natural variation in our data to identify assessments that obey one of them and investigate their overall adoption rates. We are not aware of any theory that predicts how the sequence of an entire list affects adoption at a portfolio level. The portfolio management and capital budgeting literatures point out that, though there are many sophisticated techniques to select portfolios, in practice simpler rules based on the net present value, profitability index, or payback are used (e.g., Bierman and Smidt 2007, De Reyck et al. 2005). De Reyck et al. (2005) find that over 90% of firms used payback to assess projects in their portfolio management decisions, so we use increasing order of payback as one of our sequencing rules. Two other factors highlighted in those literatures are cash flow considerations and the complexity of projects. Related to cash flow, we use increasing order of costs and decreasing order of savings as two of our sequencing rules. Related to complexity, we look at two sequencing rules with all recommendations needing high or

Table 7 Instrumental Variables Probit Estimates of Adoption of Recommendations for Grouping by *Coefficient of Variation of Payback* and for Interaction of *Serial Position* with *Coefficient of Variation of Payback*

Dependent variable: *Adopted* (equals 1 if recommendation is implemented, 0 otherwise)

	Cost–benefit models					
	Groups by coefficient of variation of payback					
	Smallest (1)	(2)	(3)	Highest (4)	Main effect model (5)	Interaction model (6)
ln(*Cost*)	−0.1408***	−0.1299***	−0.1308***	−0.1171***	−0.1456***	−0.1181***
	(0.0149)	(0.0138)	(0.0125)	(0.0104)	(0.0086)	(0.0128)
ln(*Cost*)^2	−0.0079***	−0.0076***	−0.0075***	−0.0071***	−0.0081***	−0.0069***
	(0.0009)	(0.0008)	(0.0008)	(0.0006)	(0.0011)	(0.0012)
ln(*Savings*)	−0.0048	−0.0110	0.0469	0.0160	0.0161	−0.0167
	(0.0226)	(0.0255)	(0.0284)	(0.0257)	(0.0144)	(0.0157)
ln(*Savings*)^2	0.0014	0.0051	0.0055	−0.0007	0.0023	0.0054
	(0.0027)	(0.0027)	(0.0028)	(0.0027)	(0.0016)	(0.0016)
Serial Position	−0.1820***	−0.1690***	−0.0911	−0.0536	−0.1504***	−0.4171***
	(0.0337)	(0.0386)	(0.0476)	(0.0449)	(0.0220)	(0.0452)
Coefficient of Variation of Payback					−0.0113	−0.6619***
					(0.0065)	(0.0766)
Serial Position × *Coefficient of Variation of Payback*						0.1450***
						(0.0173)
Number of Recommendations	0.7908***	0.6412***	0.3071	0.1825	0.5695***	0.7308***
	(0.1423)	(0.1656)	(0.1949)	(0.1860)	(0.0903)	(0.0769)
Variance of Payback	−0.0489***	−0.0784***	−0.0952***	−0.1178***	−0.0756***	−0.0847***
	(0.0125)	(0.0141)	(0.0155)	(0.0129)	(0.0074)	(0.0069)
Sales	0.0364	−0.0065	−0.0050	0.0282	0.0238	0.0367**
	(0.0262)	(0.0260)	(0.0264)	(0.0261)	(0.0135)	(0.0126)
Energy Costs	0.0143	0.0206	0.0201	0.0017	0.0162**	0.0179***
	(0.0098)	(0.0154)	(0.0136)	(0.0085)	(0.0057)	(0.0053)
Employees	0.0000	0.0003*	0.0001	−0.0002	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Assessment in 1st Quarter	0.0894*	0.0241	−0.0053	0.0101	0.0402*	0.0435*
	(0.0401)	(0.0404)	(0.0397)	(0.0395)	(0.0205)	(0.0189)
Assessment in 2nd Quarter	0.0162	0.0794	−0.0109	−0.0169	0.0242	0.0222
	(0.0386)	(0.0393)	(0.0382)	(0.0384)	(0.0197)	(0.0183)
Assessment in 3rd Quarter	0.0348	0.0217	0.0115	−0.0110	0.0231	0.0228
	(0.0398)	(0.0407)	(0.0405)	(0.0414)	(0.0207)	(0.0188)
Constant	0.4121	−0.3921	−3.4948	−1.1903	0.3890	1.2123***
	(0.5523)	(0.7161)	(2.4512)	(0.7951)	(0.3050)	(0.2921)
Other controls						
Recommendation type	Yes	Yes	Yes	Yes	Yes	Yes
IAC centers	Yes	Yes	Yes	Yes	Yes	Yes
Years	Yes	Yes	Yes	Yes	Yes	Yes
SIC code	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,464	22,236	22,158	22,239	76,052	76,052
Log-pseudolikelihood	−63,189***	−64,699***	−65,546	−65,076***	−220,976***	−191,107***
Exogeneity Wald statistic	16.73	11.83	1.65	0.63	27.95	38.22

Notes. Data pertain to recommendations made by IAC centers from 1981 to 2006. The estimation method is maximum likelihood. Standard errors are clustered at the assessment level and reported using a robust clustered variance–covariance matrix. The assessments are divided into four groups based on *Coefficient of Variation of Payback*. Model (1) represents assessments with lowest *Coefficient of Variation of Payback*, and model (4) represents assessments with highest *Coefficient of Variation of Payback*. Model (5) represents the main effects model, whereas model (6) represents the interaction effect model. The models use instrumental variables to instrument the serial position of a recommendation (using a sequence generated based on the ARC manual and the propensity with which each IAC makes recommendations). Standard errors are in parentheses.

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Table 8 Probit Estimates of Adoption of Recommendations, Grouped by Serial Position of Recommendations

Dependent variable: *Adopted* (equals 1 if recommendation is implemented, 0 otherwise)

	Probit for groups with serial positions 1, 3, 5, 7, 9, and 11					
	1	3	5	7	9	11
ln(*Cost*)	−0.1725***	−0.1914***	−0.2033***	−0.1880***	−0.1349***	−0.1398***
	(0.015)	(0.021)	(0.024)	(0.026)	(0.036)	(0.052)
ln(*Cost*)^2	−0.0116***	−0.0122***	−0.0155***	−0.0137***	−0.0025	0.0101
	(0.003)	(0.003)	(0.004)	(0.004)	(0.006)	(0.008)
ln(*Savings*)	0.0710***	0.1144***	0.1277***	0.0817**	0.0348	0.0181
	(0.017)	(0.025)	(0.025)	(0.027)	(0.034)	(0.049)
ln(*Savings*)^2	−0.0021	0.0018	0.0080	0.0007	−0.0073	−0.0197*
	(0.004)	(0.005)	(0.005)	(0.005)	(0.006)	(0.010)
Number of Recommendations	0.0196	0.0606	0.1011	0.0316	−0.0329	0.0135
	(0.047)	(0.046)	(0.053)	(0.070)	(0.108)	(0.187)
Variance of Payback	−0.0904***	−0.0738***	−0.0722***	−0.0397	−0.1192**	−0.1462*
	(0.020)	(0.020)	(0.023)	(0.024)	(0.039)	(0.061)
Sales	0.0157	0.0065	0.0246	0.0310	0.0544	0.1492
	(0.026)	(0.025)	(0.027)	(0.032)	(0.048)	(0.082)
Energy Cost	0.0283**	0.0032	−0.0114	−0.0020	−0.0155	−0.0641
	(0.011)	(0.010)	(0.011)	(0.013)	(0.017)	(0.033)
Employees	0.0001	−0.0001	0.0000	0.0000	−0.0001	−0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Assessment in 1st Quarter	0.0072	0.0581	0.0013	0.0642	0.0397	0.2517
	(0.039)	(0.039)	(0.042)	(0.052)	(0.076)	(0.126)
Assessment in 2nd Quarter	0.0143	0.0236	−0.0069	0.0649	0.0051	0.0816
	(0.038)	(0.038)	(0.041)	(0.050)	(0.074)	(0.119)
Assessment in 3rd Quarter	−0.0031	0.0326	−0.0444	0.0468	0.1249	−0.0412
	(0.039)	(0.039)	(0.042)	(0.053)	(0.078)	(0.131)
Constant	0.5300	0.2870	−0.1260	0.4910	14.9760	6.7290
	(0.731)	(0.701)	(0.630)	(1.374)	(11.801)	(21.535)
Controls						
Recommendation type	Yes	Yes	Yes	Yes	Yes	Yes
IAC centers	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,442	10,418	9,014	5,940	2,870	1,158
Log-pseudolikelihood	−6,468.9***	−6,743.5***	−5,880.4***	−3,869.6***	−1,819.5***	−698.59***

Notes. Data pertain to recommendations made by IAC centers from 1981 to 2006. The estimation method is maximum likelihood. Standard errors are clustered at the assessment level and reported using a robust clustered variance–covariance matrix. Standard errors are in parentheses.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

low managerial attention listed first. Our interviews with the eight IACs also suggested that the sequencing rules we examine are in line with some of the present practices. We also revisit the choice overload issue at the portfolio level. We examine the impact of the five sequencing rules and of the total number of recommendations on overall adoption at the portfolio level using the following model:

$$K_j = \alpha + \mathbf{P}_j\boldsymbol{\beta} + N_j\eta + GH_j\gamma + GL_j\kappa + \mathbf{CV}_j\boldsymbol{\rho} + \mathbf{L}_j\boldsymbol{\chi} + \xi_j. \tag{6}$$

Here K_j is the "knapsack gap," which measures adoption at the assessment level. It represents the deviation of total savings chosen from the maximum savings that could have been achieved without incurring higher implementation costs. The "knapsack gap" for assessment j is defined as $K_j = (\textit{Knapsack-Optimal Savings}_j - \textit{Total Savings Implemented}_j)/\textit{Total Savings Recommended}_j$, where $\textit{Knapsack-Optimal Savings}_j$ is the optimal value of the following knapsack problem:

$$\text{Maximize } \sum_{i=1}^{n} b_i x_i$$

$$\text{subject to } \sum_{i=1}^{n} c_i x_i \le \sum_{k=1}^{n} c_k IM_k; \quad x_i \epsilon [0, 1];$$

$$IM_k = \begin{cases} 1 & \text{if recommendation } k \text{ was implemented,} \\ 0 & \text{otherwise,} \end{cases}$$

where b_i and c_i represent the savings and costs for recommendation i; $\textit{Total Savings Implemented}_j = \sum_{k=1}^{n} b_k IM_k$, and $\textit{Total Savings Recommended}_j = \sum_{k=1}^{n} b_k$. We do not claim that the knapsack-optimal solution is necessarily the firm's best decision, but we believe that the "knapsack gap" can serve as a proxy for how efficiently firms utilize their limited resources. The average knapsack gap in our data is 11%.

In model (6), $\mathbf{P}_j$ represents the vector of sequencing rules obeyed by assessment j. We use five indicator variables to identify the sequencing rules each assessment obeys. The indicator variable *Assessment with Increasing Payback* identifies assessments with recommendations listed in order of increasing payback, where the recommendation with the lowest (fastest) payback is listed first, and the one with the highest (slowest) payback is listed last. *Assessment with Increasing Costs* and *Assessments with Decreasing Savings* are defined analogously. There are 207 assessments with recommendations listed by increasing payback, 99 by increasing costs, and 1,567 by decreasing savings. The indicator variable *High Managerial Attention on Top* identifies assessments where all recommendations that require high managerial attention are listed first, as occurs for 670 assessments, and *Low Managerial Attention on Top* is defined analogously, as occurs for 1,950 assessments.

The other variables in model (6) include N_j, which represents the number of recommendations; GH_j and GL_j, which identify whether assessment j has recommendations that need high managerial attention and low managerial attention, respectively; $\mathbf{CV}_j$, which is a vector with the coefficients of variation of costs and of savings to control for the variation in the financial characteristics of recommendations; the matrix $\mathbf{L}_j$, which includes controls for the sales, energy costs, number of employees, and the plant area; and ξ_j, which represents the error terms. We evaluated the impact of the five sequencing patterns individually and in various logical combinations using model (6). We present only a selection of these combinations, but our results are consistent across the various analyses. The first uses the variable *Assessments with Increasing Payback* for the vector $\mathbf{P}_j$; the second uses *Assessments with Increasing Costs* and *Assessments with Decreasing Savings* for $\mathbf{P}_j$; the third combines the variables used in the first and second models; the fourth uses *High Managerial Attention on Top* and *Low Managerial Attention on Top* for $\mathbf{P}_j$; and the fifth includes all five sequencing rules. The results are shown in Table 9.

In models (1), (3), and (5) of Table 9, the coefficient of *Assessment with Increasing Payback* is negative and significant at $p < 0.01$. Firms that receive assessments with recommendations listed in order of increasing payback choose options that are closer to the maximum energy savings achievable without incurring higher costs than other firms do. For the average firm in model (1), the "knapsack gap" is 8.09% when recommendations are listed by increasing payback, which increases to 11.47% when recommendations are not listed by increasing payback. This suggests that listing recommendations with increasing payback can enhance savings realized by 3.38%. In 2005 the IAC program assessed 551 facilities and recommended total annual savings of \$89.9 million. Implementing an additional 3.38% of recommended savings would constitute an increase in annual savings of \$3.04 million, without any additional investments. Although one should of course be cautious about such extrapolations from our models, this does at least suggest that listing recommendations by increasing payback might significantly enhance adoption of energy efficiency initiatives.

The coefficient of *High Managerial Attention on Top* is positive and significant at $p < 0.001$ in models (4) and (5) of Table 9, which indicates that overall adoption is lower when recommendations needing higher managerial attention are listed up-front. The other sequencing rules have no effect on overall adoption rates. Overall, these analyses at the portfolio level suggest that how recommendations are sequenced can have a significant impact on the overall adoption of energy efficiency initiatives.

The coefficient of *Number of Recommendations* is positive and significant at $p < 0.001$ in all models of Table 9. If more recommendations are listed, firms will deviate further from the knapsack-optimal choices and adopt a lower proportion of savings identified. In model (1), adding another average recommendation to an assessment reduces the recommended annual savings adopted by 1.11%. The latter does not mean that firms adopt less absolute savings, but, combined with our earlier observation that an individual recommendation is not less likely to be adopted if it is part of a longer list, it appears that firms faced with a longer list may gravitate more toward smaller recommendations than the most profitable ones. These results are consistent with choice overload while excluding the impact of budget constraint. However, the magnitude of the coefficient suggests that the overall effect is limited.

6. Discussion and Limitations

In this paper, we investigate whether the sequence in which recommendations are presented and their total number influence the adoption of energy efficiency initiatives made to small and medium-sized manufacturing firms. We find that recommendations that appear earlier in a list have higher adoption rates than recommendations that appear later. Taking a portfolio perspective, we find that overall adoption rates are higher for assessments where recommendations are listed by increasing payback, and are lower when recommendations that require higher managerial attention are listed first. The total number of recommendations has a mixed effect: we do not find any evidence

Table 9 **Estimation Results for OLS Models at Assessment Level for** *Knapsack Gap*

	Dependent variable: *Knapsack Gap*				
	(1)	(2)	(3)	(4)	(5)
Assessment with Increasing Payback	−0.0338**		−0.0334**		−0.0336**
	(0.0116)		(0.0117)		(0.0116)
Assessments with Increasing Costs		−0.0184	−0.0018		−0.0057
		(0.0179)	(0.0185)		(0.0183)
Assessments with Decreasing Savings		−0.0009	−0.0009		−0.0008
		(0.0049)	(0.0049)		(0.0049)
High Managerial Attention on Top				0.0284***	0.0287***
				(0.0085)	(0.0085)
Low Managerial Attention on Top				0.0084	0.0086
				(0.0048)	(0.0048)
Assessment Has High Managerial Attention Recommendations				0.0086*	0.0085*
				(0.0039)	(0.0039)
Assessment Has Low Managerial Attention Recommendations				−0.0097	−0.0126
				(0.0207)	(0.0207)
Number of Recommendations	0.0111***	0.0112***	0.0111***	0.0116***	0.0113***
	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)
Coefficient of Variation by Costs	−0.0405***	−0.0401***	−0.0405***	−0.0408***	−0.0412***
	(0.0046)	(0.0046)	(0.0046)	(0.0046)	(0.0046)
Coefficient of Variation by Savings	−0.0135**	−0.0137**	−0.0135**	−0.0141**	−0.0141**
	(0.0047)	(0.0047)	(0.0047)	(0.0047)	(0.0047)
Constant	0.0285	0.0266	0.0287	0.0244	0.0306
	(0.0295)	(0.0295)	(0.0295)	(0.0348)	(0.0350)
Other controls					
Sales	Yes	Yes	Yes	Yes	Yes
Energy costs	Yes	Yes	Yes	Yes	Yes
Employees	Yes	Yes	Yes	Yes	Yes
Plant area	Yes	Yes	Yes	Yes	Yes
R^2	0.0344***	0.0339***	0.0344***	0.0365***	0.0371***
Adjusted R^2	0.0337	0.0330	0.0335	0.0354	0.0358
Number	10,137	10,137	10,137	10,137	10,137

Notes. Data pertain to recommendations made by IAC centers from 1981 to 2006. The estimation method is ordinary least squares. Standard errors reported are using a robust clustered variance–covariance matrix. Data for 2,017 assessments are not used for these analyses because no recommendations were implemented. All of the analyses have been done by including these data, and the results of the study are still valid. The variables *Assessments with Increasing Payback*, *Assessments with Increasing Costs*, *Assessments with Decreasing Savings*, *High Managerial Attention on Top*, and *Low Managerial Attention on Top* take values of 1 for 297, 99, 1,567, 670, and 1,950 assessments respectively. Standard errors are in parentheses.

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

for choice overload on adoption of individual recommendations, but increasing the total number of recommendation has a modest adverse impact at the assessment level.

The magnitude and consistency with which we find the order effects are especially interesting. Moving a recommendation one position down the list in an average assessment has an impact similar to increasing its costs by at least 17% (as seen in our exogenous models, and more in the endogenous models). Given the large impact of the sequence effect on adoption of individual recommendations and of the entire portfolio, our results suggest that IACs should pay particular attention to how recommendations are sequenced in an assessment.

Our research contributes to the OM and energy efficiency literatures by identifying a significant new behavioral factor that influences how managers respond when given a list of recommendations. Organizations such as the IAC and other government agencies could leverage our results to structure recommendations to increase uptake. Our research could also be relevant for firms with complex supply chains, who may use external audit firms to drive operational changes in their supply chains. For instance, Philips (2011) deals with over 10,000 component suppliers and uses external audit firms to evaluate the sustainability performance of their supply chain partners once every three years. The audit firms provide specific recommendations on how supply chain partners can improve their sustainability performance. Philips could directly induce supply chain partners through monetary incentives or other economic means to implement key recommendations. However, given the complexity of the supply chain, the audit process yields numerous recommendations, and the implementation of many recommendations often gets left to the supplier's discretion. Managers in such firms could leverage our findings to induce suppliers to achieve higher sustainability performance. Similarly, many nonprofit organizations, such as the Fair Labor Association, Electronic Industry Citizenship Coalition, etc., often assess operational practices and provide reports with lists of recommendations to firms (e.g., see Fair Labor Association 2012 for a report on the assessment of the Nestlé cocoa supply chain in Ivory Coast). Our research can also be relevant for such nonprofit organizations. Though our results may have implications that extend beyond the specific context of energy efficiency, we believe additional testing is essential before extrapolating our results to other contexts.

Using field data rather than controlled experiments has many benefits, but also some inevitable limitations. We could not include firm-level profits and cash availability, because the identity of the firms in the IAC database is confidential. We conducted several additional robustness checks. First, we used each firm's four-digit SIC code to identify the average industry-level profitability and operating cash availability for the year in which the assessment was done, and used these as controls. Our results did not change. Another alternative explanation is that firms might plan to adopt all recommendations but decide to do so in the sequence in which they are presented. Hence, when the IAC contacts them within two years to check on the implementation status, they would have only implemented those recommendations that appeared earlier in the assessment. This does not invalidate our findings because it would still imply that firms use the sequencing of recommendations to guide their decisions rather than just their merits. Despite our robustness checks, no single study, especially with field data, can rule out all alternative explanations. We hope our work will stimulate further work on the role of sequencing using both field and lab experiments. For example, following the rich tradition of experiments based on the newsvendor model, one can envision an equally rich set of experiments on variations of the knapsack problem.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/msom.2013.0439.

Acknowledgments

The authors thank Richard Newell for sharing the analysis related to his paper. They thank Vered Blass, Magali Delmas, Craig Fox, Vishal Gaur, Srinagesh Gavirneni, Steven Lippman, Rudolph Marloth, Kumar Rajaram, Subramaniam Ramanarayanan, Lawrence Robinson, Rakesh Sarin, and Christopher Tang for valuable comments. They also thank Michael Muller and John Smegal from the IAC program for valuable inputs. The first author is grateful to the Harold Price Center for Entrepreneurial Studies at the University of California, Los Angeles Anderson School of Management for financial support. The authors are grateful for feedback from the participants at seminars given at the University of California, Berkeley; the University of California, San Diego; Pennsylvania State University; and Carlson School of Management, and from participants at various conference presentations. They thank Stephen Graves, the associate editor, and the reviewers for helpful feedback that improved this paper substantially.

References

Aflaki S, Kleindorfer PR, de Miera Polvorinos VS (2012) Finding and implementing energy efficiency projects in industrial facilities. *Production Oper. Management* 22(3):503–517.

Allcott H, Mullainathan S (2010) Behavior and energy policy. *Science* 327(5970):1204–1205.

Anderson NH (1971) Integration theory and attitude change. *Psych. Rev.* 78(3):171–206.

Anderson ST, Newell RG (2004) Information programs for technology adoption: The case of energy-efficiency audits. *Resource Energy Econom.* 26(1):27–50.

Benartzi S, Thaler RH (2007) Heuristics and biases in retirement savings behavior. *J. Econom. Perspect.* 21(3):81–104.

Bernstein L, Roy J, Delhotal KC, Harnisch J, Matsuhashi R, Price L, Tanaka K, Worrell E, Yamba F, Fengqi Z (2007) Contribution of Working Group III to the fourth assessment report of the Intergovernmental Panel on Climate Change. Metz B, Davidson OR, Bosch PR, Dave R, Meyer LA, eds. *Climate Change 2007: Mitigation of Climate Change* (Cambridge University Press, Cambridge, UK), 456–460, 475–477.

Bierman H Jr, Smidt S (2007) *The Capital Budgeting Decision* (Macmillan, New York).

Bureau of Labor Statistics (2008) Producer price index highlights—Finished goods (WPUSOP3000). Back data. Accessed June 2, 2010, http://www.bls.gov/xg_shells/ro4xgppihi.htm.

Cachon GP, Olivares M (2010) Drivers of finished-goods inventory in the U.S. automobile industry. *Management Sci.* 56(1):202–216.

Carney DR, Banaji MR (2012) First is best. *PLoS ONE* 7(6):e35088.

Charles D (2009) Leaping the efficiency gap. *Science* 325(5942): 804–811.

Chernev A (2003a) When more is less and less is more: The role of ideal point availability and assortment in consumer choice. *J. Consumer Res.* 30(2):170–183.

Chernev A (2003b) Product assortment and individual decision processes. *J. Personality Soc. Psych.* 85(1):151–162.

DeCanio SJ (1998) The efficiency paradox: Bureaucratic and organizational barriers to profitable energy-saving investments. *Energy Policy* 26(5):441–454.

Deshpande V, Cohen M, Donohue K (2003) An empirical study of service differentiation for weapon system service parts. *Oper. Res.* 51(4):518–530.

De Reyck B, Grushka-Cockayne Y, Lockett M, Calderini SR, Moura M, Slopper A (2005) The impact of project portfolio management on information technology projects. *Internat. J. Project Management* 23(7):424–537.

Dierderen P, Tongeren FV, Der Veen HV (2003) Returns on investments in energy-saving technologies under energy price uncertainty in Dutch greenhouse horticulture. *Environ. Resource Econom.* 24(4):379–394.

Evans WN, Schwab RM (1995) Finishing high school and starting college: Do Catholic schools make a difference? *Quart. J. Econom.* 110(4):941–974.

Expert Group on Energy Efficiency (2007) Realizing the potential of energy efficiency: Targets, policies, and measures for G8 countries. Expert report, United Nations Foundation, Washington, DC, 9–19.

Fair Labor Association (2012) Sustainable management of Nestlé's cocoa supply chain in the Ivory Coast—Focus on labor standards. Accessed December 14, 2012, http://www.fairlabor.org/sites/default/files/documents/reports/cocoa-report-final_0.pdf.

Gino F, Pisano G (2008) Toward a theory of behavioral operations. *Manufacturing Service Oper. Management* 10(4):676–691.

Gourville JT, Soman D (2005) Overchoice and assortment type: When and why variety backfires. *Marketing Sci.* 24(3):382–395.

Greene WH (2008) *Econometric Analysis*, 6th ed. (Prentice Hall, Upper Saddle River, NJ).

Hassett KA, Metcalf GE (1995) Energy tax credit and residential conservation investment: Evidence from panel data. *J. Public Econom.* 57(2):201–217.

Iyengar SS, Lepper MR (2000) When choice is demotivating: Can one desire too much of a good thing? *J. Personality Soc. Psych.* 79(6):995–1006.

Jaffe AB, Stavins RN (1994a) The energy-efficiency gap: What does it mean? *Energy Policy* 22(10):804–810.

Jaffe AB, Stavins RN (1994b) The energy paradox and the diffusion of conservation technology. *Resource Energy Econom.* 16(2):91–122.

Jaffe AB, Stavins RN (1995) Dynamic incentives of environmental regulations: The effect of alternative policy instruments on technology diffusion. *J. Environ. Econom. Management* 29(3):S43–S63.

Keizers JM, Bertrand JWM, Wessels J (2003) Diagnosing order planning performance at a navy maintenance and repair organization using logistic regression. *Production Oper. Management* 12(4):445–464.

Kleindorfer PR, Singhal K, Van Wassenhove LN (2005) Sustainable operations management. *Production Oper. Management* 14(4):482–492.

Landis JR, Koch GG (1977) The measurement of observer agreement of categorical data. *Biometrics* 33:159–174.

Li Y, Epley N (2009) When the best appears to be saved for last: Serial position effects on choice. *J. Behavioral Decision Making* 22(4):378–389.

Maddala GS (2003) *Limited Dependent and Qualitative Variables in Econometrics* (Cambridge University Press, Cambridge, UK), 22–27.

Mantonakis A, Rodero P, Lesschaeve I, Hastie R (2009) Order in choice: Effects of serial position on preferences. *Psych. Sci.* 20(11):1309–1312.

Martin M, Tonn B, Schmoyer R, Overly J, Schexnayder S, Johnson D (1999) Industrial Assessment Center program impact evaluation. ORNL/CON-473. Oak Ridge National Laboratory, Oak Ridge, TN.

Mueller S (2006) Missing the spark: An investigation into the low adoption paradox of combined heat and power technologies. *Energy Policy* 34(17):3153–3164.

Mulder P, de Groot HLF, Hofkes MW (2003) Explaining slow diffusion of energy-saving technologies; A vintage model with returns to diversity and learning by-using. *Resource Energy Econom.* 25(1):105–126.

Muller MR, Muller MB, Glaeser FW (2004) *The DOE Industrial Assessment Database Manual: User Information Version 8.2*. Accessed June 3, 2010, http://iac.rutgers.edu/manual_database.php.

Murray MP (2006) Avoiding invalid instruments and coping with weak instruments. *J. Econom. Perspect.* 20(4):111–132.

Olivares M, Cachon GP (2009) Competing retailers and inventory: An empirical investigation of General Motors' dealerships in isolated U.S. markets. *Management Sci.* 55(9):1586–1604.

Philips (2011) Supplier sustainability involvement program. Accessed January 12, 2013, http://www.philips.com/about/sustainability/oursustainabilityfocus/suppliersustainability.page.

Stern PC, Aronson E (1984) *Energy Use: The Human Dimension* (W. H. Freeman & Co., New York).

Tversky A, Sattath S, Slovic P (1988) Contingent weighting in judgment and choice. *Psych. Rev.* 95(3):371–384.

U.S. Department of Energy (2011) Department of Energy FY 2011 Congressional budget request. Accessed June 17, 2011, http://www.cfo.doe.gov/budget/11budget/Content/Volume3.pdf.

U.S. Department of Energy Industrial Technologies Program (2004). Energy use, loss, and opportunities analysis, U.S. manufacturing and mining. Accessed September 18, 2012, http://www1.eere.energy.gov/manufacturing/intensiveprocesses/pdfs/energy_use_loss_opportunities_analysis.pdf.

U.S. Department of Energy Industrial Technologies Program (2009) Impacts—Method of calculating results for the IAC program. Accessed January 12, 2012, http://www1.eere.energy.gov/industry/about/pdfs/impacts2007_appendix4.pdf.

U.S. Energy Information Administration (2012) Annual energy outlook, early release overview. Accessed October 9, 2012, http://www.eia.gov/forecasts/aeo/er/pdf/0383er%282012%29.pdf.

Wooldridge JM (2002) *Econometric Analysis of Cross Section and Panel Data* (MIT Press, Cambridge, MA), 83–113.