© *Academy of Management Discoveries*
2016, Vol. 2, No. 1, 7–32.
Online only
http://dx.doi.org/10.5465/amd.2015.0006

HOW TO SAVE A LEAKY SHIP: CAPABILITY TRAPS AND THE FAILURE OF WIN-WIN INVESTMENTS IN SUSTAINABILITY AND SOCIAL RESPONSIBILITY

JOHN LYNEIS
PA Consulting Group

JOHN STERMAN
MIT Sloan School of Management

Can managers enhance social responsibility while also improving profitability? Research demonstrates that there are "win-win" investments that improve both socially desirable outcomes and the bottom line, from energy and the environment to wages and workplace safety. Yet many such opportunities are not taken—money is left on the table. Here we explore this puzzle using the case of energy efficiency in a large research university, a setting that should favor implementation of win-win actions. However, despite a long time horizon, large endowment, and pro-social mission, the university failed to implement many programs offering both large environmental and financial benefits. Using ethnographic field study and panel regression, we develop a novel simulation model integrating energy use, maintenance, and facilities renewal. We find that the organization inadvertently fell into a *capability trap* in which poor performance prevented investments in win-win opportunities and the capabilities needed to realize them, perpetuating poor performance. Escaping the trap requires investments large enough and sustained long enough to cross tipping thresholds that convert the vicious cycle into a virtuous cycle of better performance, greater investment, and still better performance. We discuss how the organization is escaping from the trap and whether the results are applicable in other contexts.

Editor's Comment

The discovery in this paper is an empirically grounded model that explains why win-win opportunities—investments that provide both private gains to organizations and public benefits for society—often go unrealized. The model simulates the levels of investment and commitment that are needed to sustain a university's building infrastructure and realize large reductions in energy use that not only have positive net present value but

also cut greenhouse gas emissions. Along the way, we learn that sustaining infrastructure entails a nonlinear dynamic capability trap in which working harder can crowd out working smarter, creating more pressure to work harder. Escaping the trap involves crossing a tipping threshold that converts the vicious cycle to a virtuous cycle, and a J-curve effect of worse-before-better performance. The authors provide a simulation for readers to examine the nonlinear dynamics of the capability trap for themselves.

Andrew Van de Ven, Action Editor

http://bit.ly/LeakyShip provides an interactive simulator illustrating the concepts in the study.

INTRODUCTION

Many scholars and managers argue that organizations can take socially responsible actions that also improve the bottom line, creating both private gain and public goods in domains from the environment to wages, working conditions, safety, and public health (e.g., Christmann, 2000; Gunningham, Kagan, & Thornton, 2003; Levine, Toffel, & Johnson, 2012; Porter & van der Lind, 1995; Ton & Huckman, 2008; Ton, 2014). Such “win-win” actions include many investments that cut energy use, greenhouse gas (GHG) emissions, and air pollution and yield positive net present value (NPV) and short payback times (e.g., Creyts, Derkach, Nyquist, Ostrowski, & Stephenson, 2007; Eichholtz, Kok, & Quigley, 2010; Fuerst & McAllister, 2011; Hawken, Lovins, & Lovins, 1999; Lovins, 2012; Mills, 2011; Moser, Liu, Wang, & Zhang, 2012; Sullivan, Pugh, Melendez, & Hunt, 2010). Prosocial investments can also enhance competitive advantage by anticipating future regulatory requirements (Gunningham et al., 2003; Hart, 1995), building relationships with important stakeholders, and strengthening reputation (Barnett, 2007; Coglianese & Nash, 2001; Estlund, 2010; Locke, 2013; Parker, 2002).

Despite widespread evidence documenting win-win investments, organizations often fail to make them. Consider the case of buildings and physical infrastructure. Simple actions can generate substantial energy savings with positive NPV and short payback times (e.g., Eichholtz et al., 2010; Fuerst & McAllister, 2011; Mills, 2011; Moser et al., 2012; Sullivan et al., 2010), yet organizations often underinvest, leaving win-win opportunities on the table (e.g., Charles, 2009; DeCanio, 1998; Effinger, Friedman, & Moser, 2009; TIAX, 2005; Pérez-Lombard et al., 2008). Amin (2011) estimates that upgrading the U.S. electrical grid would more than pay for itself through reduced outage costs and improved reliability, and McKinsey (2010) found 12 $GtCO_2e$ per year of GHG emissions—nearly one-third of the global total—can be abated with existing technologies at negative cost. Industrial accidents frequently destroy firm value, harm the environment, and cost lives, yet are often easily avoided. For example, the 2010 Deepwater Horizon blowout, 2008 Imperial Sugar refinery explosion, and 2010 Upper Big Branch mine disaster—collectively causing 54 deaths—were all preventable through inexpensive investments in equipment, maintenance, and safety (National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling, 2011; U.S. Chemical Safety & Hazard Investigation Board, 2009; U.S. Department of Labor, 2010).

Why are win-win opportunities so often left on the table? Organizational theory on socially responsible action is largely silent on the question. Social responsibility is often examined in the context of debates regarding the theory of the firm as it relates to shareholders and other stakeholders (Donaldson & Preston, 1995; Freeman, 1999; Margolis & Walsh, 2003). Empirical studies often associate socially responsible actions with firm performance, industry or geography (e.g., McGuire, Sundgren, & Schneeweis, 1988; McWilliams & Siegel, 2001); the institutional and legal environment (Campbell, 2007; Short & Toffel, 2010); or individual agency (Howard-Grenville, 2007). Although such issues are critically important, particularly where the benefits of socially responsible actions are contested or conflict with private gain,

Author’s voice:
What motivated you personally to undertake this research?

We thank the Project on Innovation in Markets and Organizations at the MIT Sloan School of Management for financial support and the many members of the MIT community who participated in interviews, provided data, or reviewed the results and paper. We received helpful comments from seminar participants at the 2012 ARCS Conference (Alliance for Research in Corporate Sustainability), the 2015 International System Dynamics Conference, and the editors and anonymous referees of *Academy of Management Discoveries*. Particular thanks to Andy Van de Ven, Susan Zaid, and Curtis LeBaron. The results and views expressed here are those of the authors and do not necessarily reflect the views of these individuals or the policies of MIT.

existing theories are less useful in explaining the paradox of profitable pro-social investments that are not adopted.

A large literature addresses the issue in the context of energy efficiency. Some economists argue that win-win opportunities cannot exist since rational managers would already have implemented them; alleged win-wins must therefore reflect overoptimistic assessment of costs and benefits (Gillingham, Newell, & Palmer, 2009; Jaffe, Newell, & Stavins, 2004; Sutherland, 1991). Others acknowledge the existence of win-win investments and attribute underinvestment to market failures. Organizations may lack access to the capital necessary to finance up-front investments. Asymmetric information problems can arise when technology providers cannot credibly communicate the benefits to customers (Howarth & Sanstad, 1995). Principal–agent interactions can cause underinvestment, as in the famous "landlord–tenant" problem: when tenants pay the utility bill, landlords will underinvest in efficiency because they would bear the costs while the tenants reap the benefits (Jaffe & Stavins, 1994).

Behavioral biases can also lead to underinvestment. Managers may evaluate projects from the parochial perspective of their organizational function rather than what's best for the organization as a whole, choose investments with lower initial costs despite higher NPV life-cycle costs, underweight low-probability/high-consequence risks, and bow to competitive and capital market pressures for short-term financial results (Bazerman, 2009; Frederick, Lowenstein, & O'Donoghue, 2002; Rahmandad, 2012; Yates & Aronson, 1983).

Certainly, costs can be underestimated, and market failures, principal–agent problems, behavioral biases, and short-termism affect investment decisions. Yet, these explanations are only partly satisfactory. Win-win investments are well documented and many organizations have benefitted (e.g., Creyts, Derkach, Nyquist, Ostrowski, & Stephenson, 2007; Eichholtz et al., 2010; Fuerst & McAllister, 2011; Lovins, 2012). Large firms often have access to capital and strong incentives to overcome market failures and biases that limit profitability. Yet, many attempt to implement profitable pro-social investments only to see performance fall short of potential (e.g., Coglianese & Nash, 2001).

Failure to implement profitable opportunities afflicts improvement programs generally, not only pro-social opportunities (Keating, Oliva, Repenning, Rockart, & Sterman, 1999; Repenning & Sterman, 2002). From airline kitchens to health care, similar firms in the same industry and even different floors of the same hospital exhibit persistent performance differences despite financial incentives, market forces, and the availability of improvement methods that should lead to broad diffusion of best practices (Gibbons & Henderson, 2013; Wennberg, 2010). For example, total factor productivity varies by about a factor of two between the 10th and 90th percentile firms in the same 4-digit SIC industries in the United States, and by more than a factor of 5 in China and India (Syverson, 2011).

Author's voice:
How did your research paper evolve?

We argue that understanding the paradox of unexploited win-win investments requires us to consider not only the market failures, incentives, and behavioral biases that condition investment decisions, but also the dynamics of program implementation. To do so, here we report a longitudinal study of energy use and facility maintenance at a large research university, the Massachusetts Institute of Technology (MIT). Like many universities, MIT is well positioned to exploit win-win opportunities. MIT has an explicit pro-social mission, long time horizon, substantial endowment, AAA credit rating, bears no shareholder pressure for short-term results, and as the owner-operator of its facilities, does not face landlord–tenant agency problems. Nevertheless, in the past, the Institute failed to exploit many win-win opportunities to improve its facilities and energy efficiency.

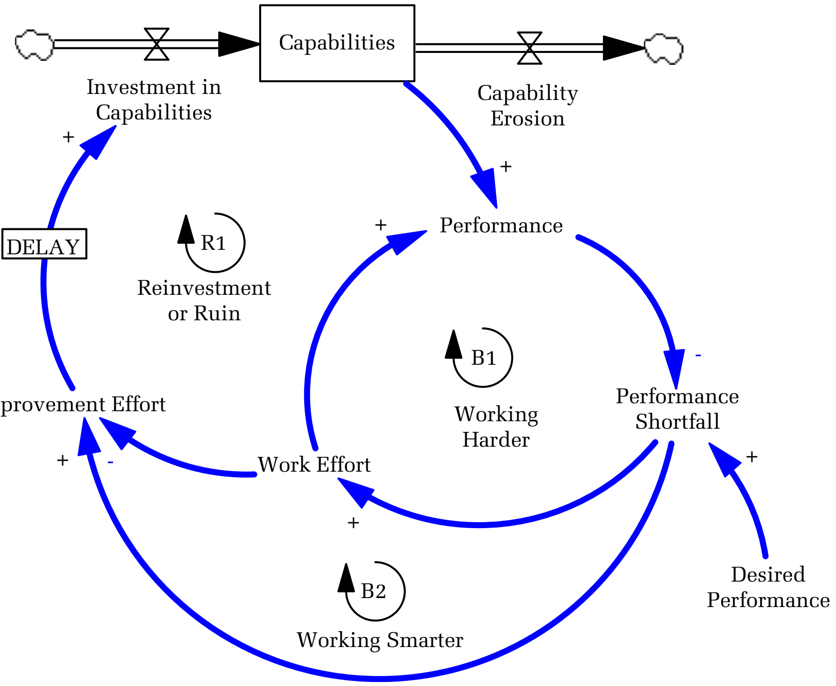
We develop a novel simulation model of the Institute's operations grounded in ethnographic interviews, archival records, and quantitative data including maintenance, energy use and the condition and renewal costs of every system in every building on campus. We use panel regression to estimate important physical and behavioral relationships, such as the rate at which energy efficiency deteriorates as buildings and systems age and how maintenance personnel allocate time to reactive versus proactive maintenance. We then use the model to explore why win-win opportunities were not taken. We conclude that the Institute inadvertently became stuck in a *capability trap* (Repenning & Sterman, 2001, 2002): a vicious cycle in which unreliable, inefficient facilities lead to high costs and a firefighting focus that prevent an organization from investing in the capabilities and programs needed to improve, thus perpetuating high costs and firefighting. Escaping the trap is often difficult: the first response to increased investment in process improvement is higher costs and/or fewer resources for urgent repairs, a Worse-Before-Better (WBB) dynamic. For example, in the short run, increasing proactive maintenance not only raises costs but requires reassigning technicians from repair to prevention and, often, taking operable equipment offline, cutting uptime.

To begin, we extend the theory of the capability trap to the context of facilities management, maintenance, and energy use. Next, we describe the research setting, the simulation model and the data used to develop and test it. We then use the model to explore policies to improve performance, including outcomes (do investments offer win-win benefits such as positive NPV and lower energy use?) and implementation dynamics (how long and deep is any WBB tradeoff?). We show how MIT, in part due to the work reported here, is overcoming the capability trap by investing in sustainable improvement. Finally, we note limitations of the study and consider the generality of the results. We close with implications for research and practice.

THE CAPABILITY TRAP

Repenning and Sterman (2001, 2002) developed the theory of the capability trap to explain the failure of many process improvement programs; Sterman (2015) applies the theory to sustainability and prosocial investments. Figure 1 shows the structure of the theory in the form of a causal diagram (Sterman, 2000). The managers of any process, whether production, product development, maintenance, human resources, or environmental quality, are responsible for the performance of that process against target or desired performance. If performance falls short of the target, managers have two basic options to close the gap: working harder or working smarter. Working harder includes adding resources (hiring, capacity expansion), increasing resource utilization (overtime, shorter breaks, speeding up), and boosting output per person-hour by cutting corners (skipping steps, cutting testing, deferring maintenance, failing to follow safety procedures). These activities form the balancing (i.e., negative) *Work Harder* feedback, B1: a performance shortfall leads to longer hours, corner cutting, deferring maintenance, and other shortcuts that improve performance. Alternatively, managers can interpret the performance gap as a sign that the organization's capabilities are inadequate. They can increase improvement activity designed to eliminate the root causes of poor performance and invest in the

FIGURE 1
The capability trap: Structure



Signs ("+" or "−") at arrowheads indicate the polarity of causal relationships: a "+" denotes that an increase in the independent variable causes the dependent variable to increase, ceteris paribus (and a decrease causes a decrease); formally, $X \rightarrow^{+} Y \Leftrightarrow \partial Y / \partial X > 0$. Similarly, a "−" indicates that an increase in the independent variable causes the dependent variable to decrease; that is, $X \rightarrow^{-} Y \Leftrightarrow \partial Y / \partial X < 0$. Boxes represent stocks; arrows with valves represent flows. A stock accumulates the difference between its inflows and outflows, e.g., $\text{Capabilities}(t) = \int[\text{Investment in Capabilities}(s) - \text{Capability Erosion}(s)]ds + \text{Capabilities}(t_0)$. See Sterman 2000. http://bit.ly/LeakyShip provides an interactive simulator of the capability trap.

capabilities that make improvement effort effective, including investments that enhance people's skills, knowledge of and adherence to best practices, and build cooperation and trust across organizational boundaries. Investing in capability improvement forms the balancing *Work Smarter* feedback, B2.

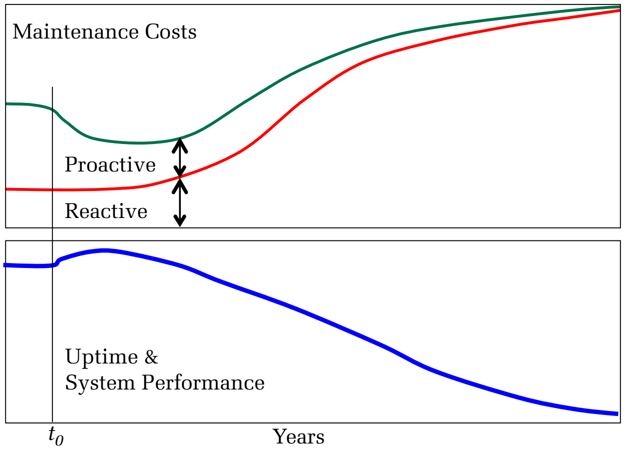
Consistent with the resource-based view of the firm and theories of dynamic capabilities, the organization's capabilities are shown as a stock. Capabilities—including productive, well-maintained equipment, skilled workers, effective improvement methods, organizational routines, and trust between workers and management and across organizational boundaries—are assets that build up as the result of investment and erode over time as equipment ages, employees leave, and changes in the environment render skills, knowledge, routines, and relationships obsolete.

Working harder and working smarter interact because time and resources are limited. When organizations are heavily loaded and resources constrained, greater work effort necessarily comes at the expense of maintenance, improvement, learning, training, coordination, and other activities needed to preserve and enhance capabilities. The result is the reinforcing (i.e., positive) feedback denoted *Reinvestment or Ruin* (R1). As the name suggests, the reinforcing feedback can operate as a virtuous cycle that builds capabilities and performance or as a vicious cycle that degrades both. An organization that increases the time and resources devoted to improvement sufficiently will, after a lag, build capabilities that boost performance, easing the performance gap and yielding still more time and resources for improvement. In contrast, if managers respond to a performance gap with greater pressure to boost output, improvement effort falls, the organization's capabilities erode, and the throughput gap grows still larger, forcing ever-greater reliance on working harder. The vicious cycle drives out improvement activity, leading to low capabilities and poor performance, and, all too often, environmental damage, accidents, or organizational failure.

How could an organization allow itself to fall into the capability trap? Consider managers and workers facing a performance gap. Working harder—including overtime, corner cutting, and deferring maintenance—will quickly boost output. Effort and outcome are closely related in time and space, observable and quite certain: a 10 percent increase in work hours quickly yields about 10 percent more throughput. In contrast, working smarter takes time, and both the length of the lag and the payoff are uncertain: improvement experiments take time and often fail, and it takes time to train people in improvement, develop routines and norms that prevent corner cutting, and build commitment, relationships, and trust. These features bias many organizations toward working harder even when the payoff to working smarter is higher.

Figure 2 illustrates using the example of maintenance in a manufacturing plant (Carroll, Sterman, & Marcus, 1998; Repenning & Sterman 2001, 2002). Initially, the plant is performing well, with high uptime, product quality, and safety. Maintenance spending is largely devoted to proactive maintenance and improvement. Now, imagine a company-wide budget cut (due to recession, competition, or other cause). The maintenance manager must cut expenses. Reactive maintenance cannot be cut: when equipment fails it must be fixed, lest plant uptime falls and customer commitments go unmet. Instead, process improvement and proactive maintenance (defined here to include scheduled, preventive, and predictive maintenance) suffer, along with part quality, equipment upgrades, training, and, all too often, adherence to safety protocols. The first impact? Maintenance costs fall, closing the budget gap, and plant uptime *rises*, because operable equipment is no longer taken down for proactive maintenance. Soon, however, breakdowns and failures grow, increasing reactive maintenance and costs, further cutting proactive maintenance and improvement. Worse, falling uptime and output erode revenue, and budgets may be cut further, increasing pressure to cut proactive maintenance and process improvement. The plant becomes trapped in a vicious cycle of increased

FIGURE 2
The capability trap: Dynamics



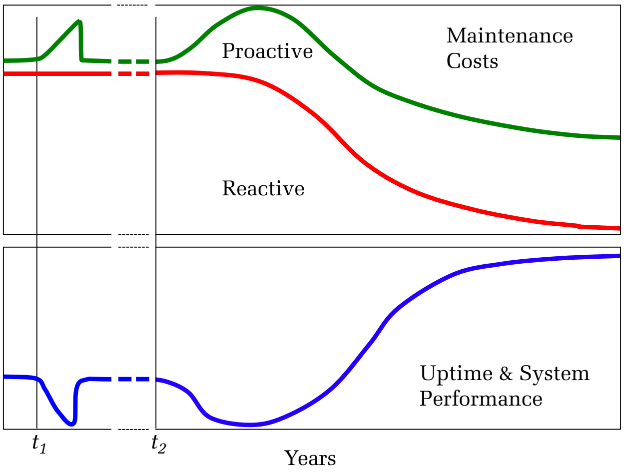
Budget cuts at time t_0 force the organization to cut proactive maintenance and improvement effort. As organizational capabilities fall, defects increase, increasing reactive maintenance, and forcing further reductions in proactive maintenance and process improvement. The self-reinforcing *Reinvestment or Ruin* feedback in Figure 1 operates as a vicious cycle, driving the organization to a state of high costs and low performance, reliability, and safety.

breakdowns, higher repair costs, lower uptime, greater production and financial pressure, less improvement effort, and still more breakdowns. Soon, the organization finds itself in a paradox: it spends more on maintenance than the industry average, yet gets less for it. Risks to the health and safety of employees and the community rise as equipment deteriorates.

Now, consider what happens when the organization seeks to escape the trap. Figure 3 shows a plant with high costs and low uptime, reliability, safety, and quality. At time t_1, managers initiate a major improvement program. The first impact? Costs rise while uptime and output fall. Costs rise because the organization must increase proactive maintenance and improvement activity, while still carrying out reactive repair work at the same rate. Uptime and output fall because operable equipment must be taken off line to perform proactive maintenance and test improvement ideas. In many organizations, the next impact is the abandonment of the improvement initiative.

What happens, however, if the organization does not give up when costs rise and uptime falls? After a new improvement program is started (at t_2 in Figure 3) the gradual growth in capabilities eventually begins to boost performance. Breakdowns begin to fall, uptime and output rise, and the burden of reactive maintenance eases, freeing up resources that can be reinvested in further capability development: the Reinvestment or Ruin feedback now operates as a virtuous cycle, bootstrapping the plant to low costs and high performance. Note, however, that the system exhibits WBB behavior.[1]

FIGURE 3
Escaping the capability trap: Worse-before-better



Improvement effort is given priority at time t_1, but the increase in costs and drop in uptime causes the organization to abandon the effort. If a new effort begins (at time t_2) and is not abandoned, the initial cost increase and performance drop eventually reverse, leading to lower costs and higher uptime, output, quality, reliability, and safety, in a worse-before-better pattern.

The theory of the capability trap yields three main insights: first, sustainable improvement requires transforming the vicious cycle into a virtuous cycle of improved performance, lower costs, greater investment in capabilities, and still better performance. Second, doing so creates WBB behavior because of the lag between investment in capabilities and results. Third, the system exhibits a tipping point because capabilities are stocks. To escape the trap, the investment in capabilities must be large enough so that capabilities are built faster than they decay. Managers may be willing to increase resources for improvement, despite the short-run costs, but unless those investments are large enough and sustained long enough to build capabilities faster than they erode, performance will still gradually deteriorate. Visit http://bit.ly/LeakyShip to download an interactive management flight simulator illustrating how the capability trap works. The simulator allows users to try different policies and see how a hypothetical organization responds. The link also provides instructions for the simulator and a presentation about this study.

As a metaphor, consider the organization as a leaky ship. To save the ship, the captain may order the crew to bail, accepting that doing so will slow progress, at least temporarily, as those who bail cannot also sail. But to avoid sinking, it is not enough for the crew to bail. They must bail faster than the water leaks in. Similarly, managers may boost resources for improvement, knowing that performance will suffer in the short run and may do so by what they consider to be substantial amounts relative to the organization's past or peers. But no matter how large the increase in resources for improvement, if it fails to build capabilities faster than they erode, capabilities will still fall. When the program fails to reverse the decline in performance, managers—or their successors—are likely to abandon it, leading organizations to give up too soon.

The original capability trap work (Repenning & Sterman 2001, 2002), however, does not provide a means for managers to determine *how* to escape

---

[1] Worse-Before-Better behavior also arises in economics and investment theory, where it is typically known as the "J curve." For example, the cash flow of a (successful) investment typically begins at zero, becomes negative, and only later becomes positive, tracing a "J" shape; a currency devaluation initially worsens the balance of payments, but later may improve it as imports fall and exports rise. See e.g. Giovannetti (2008).

the trap, including how much and how long to invest in different capabilities, how to coordinate those investments, how to evaluate their likely operational and financial benefits, or how to estimate the depth and duration of the WBB dynamic. Here we advance the theory by (a) showing that escaping the trap requires coordination of multiple capabilities and their interactions, and by (b) developing a formal simulation model, grounded in qualitative and quantitative data, which allows us to design policies for improvement and assess their resource requirements and likely outcomes. Our theory and model explicitly disaggregate capabilities and the determinants of performance. We model the number and productivity of staff, the routines and decision rules used to allocate those resources to reactive or proactive maintenance, and how these interact to affect performance. Performance is disaggregated to include defects in equipment, their root causes, the condition of the buildings and systems, and their energy efficiency. Each of these elements of the organization's capabilities is a separate stock in the model, responding to maintenance and improvement effort with different costs and lags. Interactions among these stocks play a major role in the dynamics and the response of the system to policies.

Our work has implications for theorists and practitioners. For theorists, we demonstrate the importance of explicitly modeling the determinants and interactions of different organizational capabilities, ranging from those embodied in physical infrastructure (e.g., defects in equipment, building condition) to those embodied in human capital, and organizational routines (e.g., maintenance technician skills and attitudes, routines for carrying out maintenance work, how people respond to financial and schedule pressure). Unfortunately, scholars have all too often treated these issues separately: theory and models in the operations management literature tend to focus on the physics of the system, while the organizational behavior literature tends to focus on decision making, norms, and group dynamics. For practitioners, the model we develop provides a tool they can use to determine how much to invest in each critical capability, how to coordinate those investments, and to assess the costs, benefits, and WBB dynamics likely to result. The model, though calibrated to MIT, is fully documented and can be modified to represent other organizations.

RESEARCH SETTING AND METHODS

We study the case of facilities maintenance and energy usage at the MIT, a large research university. For several reasons, universities are an excellent context in which to study the paradox of win-win investments that are not taken.

First, win-win opportunities in facilities and maintenance are well documented. Maintenance professionals have long known that inadequate maintenance is costly and inefficient (e.g., Levitt, 2009; Moubray, 1997) and, through, e.g., poor heating, ventilating, and air conditioning (HVAC) performance and equipment failures, increases energy use, GHG emissions and operating costs, cuts occupant comfort, and creates safety hazards. "Green" buildings consume significantly less energy than existing facilities or facilities built to current building codes, and many of these offer positive NPV and short payback times (e.g., Charles, 2009; Heo, Choudhary, & Augenbroe, 2012; Martani, Lee, Robinson, Britter, & Ratti, 2012; Perez-Lombard et al., 2008; TIAX 2005). Second, universities have a pro-social mission and do not experience shareholder pressure for short-term results. Some, including MIT, have large endowments that generate funds and buffer them from variations in cash flow, and AAA credit ratings providing access to low-cost capital. Third, as owner-operator of its facilities, MIT does not face landlord–tenant agency problems.

We employ a mixed methods research approach including ethnographic interviews with members of MIT's maintenance organization and administration, statistical analysis of building performance data, and simulation modeling. Qualitative data were collected starting in 2007. Semi-structured interviews were conducted with 30 individuals spanning the repair and maintenance (R&M) organization, the department of facilities, and university administrators. Interviews lasted between 45 and 120 minutes and were recorded. We interviewed hourly maintenance mechanics, supervisors, and managers. Individuals were asked to describe their history with the organization, their daily activities, challenges they faced in their work, and their views regarding department policies and priorities. We also interviewed representatives of academic departments and MIT engineers, analysts, and administrators charged with facilities administration, capital renewal projects, utilities, and finance. Interviewees were asked to explain department policies, investment priorities, and their views regarding opportunities for improved performance and energy efficiency.

The model simulates the condition of building stocks and systems, their energy use, building aging and renewal, and maintenance activity from 2005 through 2030. We use the system dynamics method (Sterman, 2000), which is widely used to model the dynamics of complex organizations (e.g., Freeman, Larsen, & Lomi, 2012; Morecroft, 2007; Repenning, 2002; Repenning & Sterman, 2002; Rudolph, Morrison, & Carroll, 2009; Sastry, 1997; Sterman, Repenning, & Kofman, 1997; Walrave, van Oorschot, & Romme, 2011). We build on existing system dynamics studies of service delivery and maintenance operations (Carroll, Sterman, & Marcus, 1998; Ledet, 1999; Oliva & Sterman, 2001; Sterman, 2000: 66–79).

Maintenance and Energy Use at MIT

MIT operates a large, diverse campus. As in many organizations, the R&M organization is part of a larger facilities group that is also responsible for new construction, renewal (i.e., renovation) of existing facilities, custodial work, utilities, security, and other operations. The R&M group has approximately 100 employees organized into teams. General maintenance groups are organized by zones of the campus and are supported by centralized groups of specialists including plumbers, HVAC mechanics, electricians, carpenters, and others. The daily activities of the R&M department are organized around an SAP work order system. Work orders arise from two sources: "reactive" maintenance works to resolve breakdowns and reported problems, and "proactive" maintenance includes scheduled, preventive, and predictive maintenance, including inspections of equipment and scheduled replacements. Many reactive work orders are directly initiated by members of the MIT community who report problems with temperature, plumbing, lighting, or other systems. Academic departments also have facilities liaisons who work closely with R&M.

Our interviews with maintenance personnel and MIT administrators revealed a central theme: maintenance operations are strongly driven by the large backlog of deferred maintenance—the work required to bring aging buildings and systems into conformance with current standards. In 2007, MIT commissioned an engineering firm to perform a detailed assessment of the state of its buildings. The report identified more than $1.4 billion (2007 dollars) in deferred maintenance. Examples include inoperable HVAC systems and controls; inefficient steam, electrical and plumbing systems original to the buildings; leaky single-glazed windows and cracks in building exteriors. A facilities engineer with extensive experience in other organizations explained:

> "Mechanically, [the buildings] are a mess. A mess. Here we have turn of the [20th] century buildings [and] a lot of those systems are still operational here. It just amazes me. Some of the systems I've seen here I've never seen anywhere else . . . To keep that equipment two to three times its life span and still be operational is kind of amazing. But it can't go on forever. We've gone on 2 to 3 times the normal life span of some of the systems, and we're still band-aiding them together . . ."

Author's voice:
How did you gain access to collect this data?

The backlog of deferred maintenance forced the R&M group to become highly reactive. Between 2005 and 2007, more than 85 percent of maintenance hours were spent responding to customer calls reporting problems or breakdowns. Approximately 40 percent were spent on urgent problems—those requiring a response within 2–3 days or sooner—compared to best practice benchmarks of 10 percent or less (Sullivan et al., 2010). R&M personnel identified many ways in which the reactive, customer focus undermined effectiveness. First, managers were unable to plan and schedule work, leading to inefficient time allocation and costly, expedited part procurement. Second, work quality suffered: because mechanics had to attend to the next emergency, they couldn't take the time to identify and resolve the root causes of problems. As one mechanic explained,

> "It's a fire drill . . . it's who's screaming right now. So your priorities change on an hourly basis, probably a half-hourly basis during the day, and it's kind of—it's basically a constant fire drill. [You] kind of have a tendency to leave it once you get to a point where no one is complaining."

Reactive maintenance also crowded out the "behind the scenes" proactive maintenance and improvement necessary to prevent future breakdowns, as a supervisor explained:

> "You know, we're a customer service organization. It's almost like we're afraid to commit completely to the behind the scenes stuff, because we want to get to the visible stuff so quickly. That's not spoken, but I think that's—having the resources available—the customer doesn't care if a belt is flapping in a fan. It might not matter for a year down the road, but to us it might be in January in the middle of the night that the fan shuts down—to us it's important, but to the customer it's not, so our resources go to what the customer wants, for the most part."

Deferred maintenance also increases energy use. Cracks in walls and roofs lead to energy loss, along with costly water damage, and, in winter, burst pipes. Dampers designed to reduce energy use by optimally mixing outside and conditioned air were rusted shut. Failed HVAC controls and steam traps increased energy use and caused some buildings to overheat; occupants would then open their windows to cool off, even in winter. The resulting breakdowns and customer calls kept the R&M staff from addressing the root causes of these problems, as a front-line mechanic described:

> "If you maintain [systems] correctly, the amount of heat calls, cold calls, failures, breakdowns is a minimum. We're at the high level of breakdowns and heat and cold calls. We're doing as much PM [preventive maintenance] as we can within that timeframe, [but] something's got to give. So we tweak something until no one is complaining, and then . . . walk away. I find systems that are heating and cooling at the same time because that makes the customer satisfied."

How did the Institute fall into the capability trap? Crises such as the Great Depression, the 1987 stock market crash, the implosion of the tech bubble in 2001, and the Great Recession of 2008 create temporary pressure to cut costs. Yet interviewees with long tenure described continual budget pressure, as a mechanic with more than 30 years of experience observed:

> "I can't think of a year that went by that we didn't have a budget crunch. . . . There were so many years of flat budgets or very minimal increases. In all my years here, this [maintenance] is one of the first places to get cut."

More important than the occasional fiscal crisis are the structural features that make it all too easy to begin a gradual slide into the capability trap. Despite the conditions that should favor high performance noted above, universities such as MIT experience continual competition for scarce resources: labs seek new equipment, departments seek to expand their programs and hire more faculty, admissions wants to increase student financial aid, etc. Long delays between cuts in maintenance and the resulting decline in capabilities, uncertainty over how long and large those impacts will be, and, especially, the failure to recognize the tipping point, however, mean pressure to maintain and improve facilities is weak. Consider a ship under pressure to cut costs. The captain can cut the crew without immediate consequence: by working harder, cutting corners, and deferring maintenance, the remaining hands can still pump out the bilge as fast as the water leaks in. Since corner cutting and capability erosion are difficult to observe the captain may interpret the result as a productivity gain, reinforcing the wisdom of further cuts. But as soon as bailing falls behind leakage, the water starts to rise. The problem is initially small, may not be noticed, and doesn't require urgent action. But the "Reinforcement or Ruin" feedbacks begin to operate as vicious cycles. As the water rises, it becomes gradually more dangerous to ignore the problem—and more costly and disruptive to address it. In fact, the deferred maintenance problem is neither recent nor attributable to any one leadership team. To the contrary, the problem grew gradually over the past century:

> In his annual report for 1917, MIT president Richard C. Maclaurin wrote of the importance of "clear[ing] ourselves of temporary embarrassments" by "learn[ing] from actual experience over a period of time, and not merely from estimates, what the cost of the maintenance of our plant actually is." (Plotkin, 2011)

MODELING MAINTENANCE AND ENERGY USE

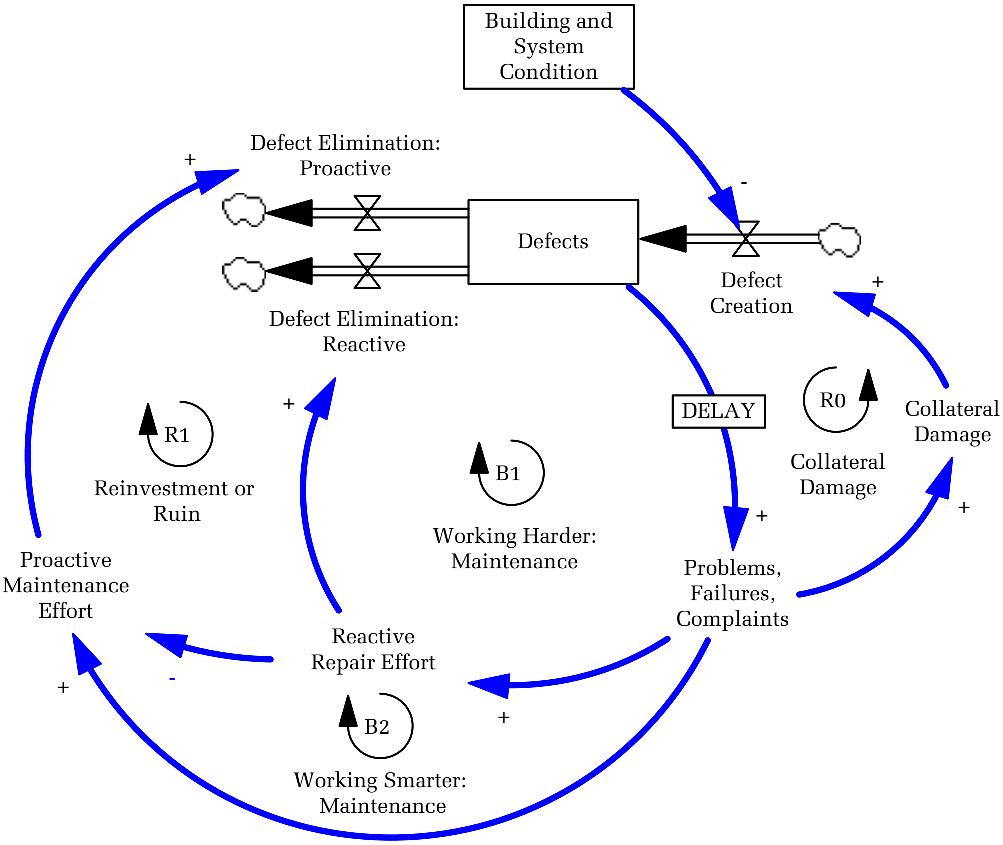
The simulation model captures both the physics of the facilities—e.g., building condition, energy use, equipment failures—and the behavior of organizational actors—e.g., the generation of maintenance work orders, their resolution by maintenance staff, and policies governing resource allocation to proactive and reactive work. The model represents the approximately 130 individual campus buildings and the systems within them individually. We model the current inventory of buildings and do not portray new facilities that may be added in the future. We use multiple methods to estimate parameters including panel regression, partial model estimation (Homer, 2012), interviews, archival data, expert judgment, and prior literature. The model represents the expected lifetime, scheduled year for renewal, and estimated renewal cost for all systems, by building, from 2007 through 2030 (approximately 6700 items), using a detailed engineering assessment MIT commissioned (Supplement Figure S1 illustrates the level of detail and provides full documentation; here we describe the model structure and several key formulations).

Feedback Structure Governing Defects in Building Systems

The concept of defects in building systems lies at the core of the model (Figure 4). Examples include worn

Author's voice:
What surprised you about this research project?

FIGURE 4
Feedback structure governing defect creation and elimination



The model represents every building separately. Building systems and the stocks of defects associated with them are disaggregated into six categories: exterior structures, interior structures, plumbing, electrical, HVAC, and other.

fan belts, HVAC systems out of calibration, cracked windows, or corroded steam line expansion joints. Defects are either known or latent, undiscovered problems. Defects are typically latent: a pump with a leaky bearing seal can continue to operate for some time. However, if not detected and corrected, the bearing will eventually seize and the pump will fail. The stock of defects increases through defect creation and decreases through defect elimination, which can result from reactive or proactive maintenance.

New defects arise from equipment aging and wear, at a rate determined by the condition of buildings and systems. Defects are also created by collateral damage arising from breakdowns: a pump that fails under load can break piping and throw shrapnel that causes other damage. These new defects can then cause other failures, shown as the reinforcing *Collateral Damage* feedback, R0.

Breakdowns and poor system performance generate work orders for reactive maintenance; such work eliminates the defects that caused the failure or complaint, closing the balancing *Working Harder: Maintenance* feedback, B1. Failures and complaints could also be interpreted as a signal that more proactive maintenance is needed. Doing so would eliminate defects before they can cause failures, closing the *Working Smarter: Maintenance* feedback, B2.

Although many R&M employees understood that it would be better to “work smarter” than “work harder,” constant pressure to resolve complaints and failures crowded out proactive work, as, for example, when mechanics “tweak something until no one is complaining.” As mechanics increasingly respond to customer calls they have less time to perform preventive maintenance or search for and correct the root causes of problems, so proactive work declines. Consequently, the stock of defects continues to grow, causing more failures, and further constraining the resources available for proactive work, shown as the reinforcing *Reinvestment or Ruin* feedback R1. Interviewees explicitly described the tradeoff between reactive and proactive work they constantly made:

> "This [equipment] is supposed to be looked at every week—well, I can't get to that every week because of our limitations, so we'll try to do all those weekly tasks every month or two months—a lot of stuff is only getting looked at once or twice a year because we don't have the resources to do it."—Maintenance Manager A

> "We have a lot of what's called deferred maintenance around here—basically, equipment that if you look at the recommendations, are well beyond their useful life . . . We could do 10 roofs this year if we had a lot of money, but we don't, so we do two roofs, that kind of thing. We look at the worst ones, we look at the ones that give us the most trouble or maybe cost us the most money on an operating [basis], and we pick those to try to get ourselves out of trouble." —Maintenance Manager B

Expanding the boundary of the model. Our model expands the boundary of the original capability trap theory to capture, endogenously, the determinants of the condition and energy efficiency of campus buildings and systems (Figure 5). Like any assets, buildings and systems gradually deteriorate and must eventually be renewed or replaced. The condition of the buildings and systems within them are stocks: if renewal (including retrofits and replacement) falls below deterioration, the condition of the buildings and systems declines, increasing defect creation.

The energy efficiency of buildings and systems, shown above the stock of building and system condition in Figure 5, also degrades over time as windows crack, insulation settles, gaps open in walls and roofs, pipes corrode, etc. Maintenance can partially compensate—windows can be repaired, cracks patched, ducts cleaned, etc. Building renewal also improves energy efficiency somewhat because building codes have generally tightened over time. However, efficiency can be improved further—at some additional cost—by installing windows, HVAC equipment, insulation, lighting, and other systems that are more efficient than code requires.

The physical processes, routines, and decision rules governing the evolution of building conditions and energy efficiency interact to create additional capability trap feedbacks.

First, just as collateral damage from equipment failure can create new defects, failures and breakdowns can degrade buildings and systems and compromise energy efficiency. A crack in a wall not only wastes energy, but on a cold night might cause sprinkler pipes to freeze and burst. The resulting flooding can damage structures, electrical and mechanical systems, and lab equipment, creating additional reinforcing *Collateral Damage* feedbacks, shown in Figure 5 as R0b, in addition to the original reinforcing loop (now labeled R0a).

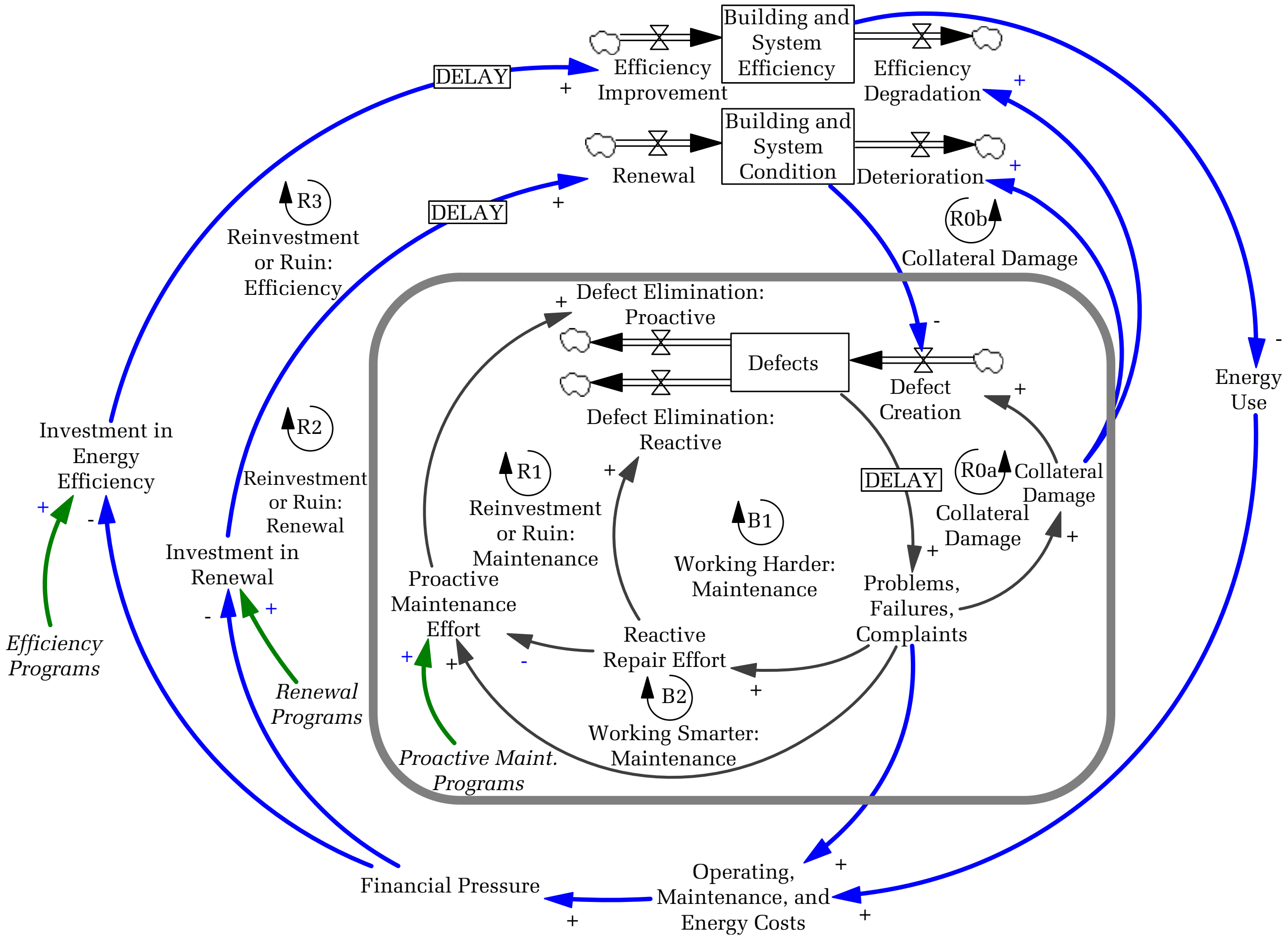
Second, endogenously modeling building condition creates new reinforcing feedbacks around facilities renewal. As the condition of buildings and systems deteriorates, the rate of defect creation increases. Eventually, breakdowns, failures and complaints increase, raising O&M costs. Higher O&M costs create financial pressure, reducing funds available for building and system renewal. With inadequate renewal, the condition of buildings and systems deteriorates further, leading to still more defects and still higher costs, creating the reinforcing *Reinvestment or Ruin: Renewal* feedback, R2.

Third, as energy efficiency degrades through aging, wear, and collateral damage from failures and breakdowns, energy use increases, raising operating costs, intensifying financial pressure, reducing the funds available for efficiency upgrades and leading to still higher energy use, forming the reinforcing *Reinvestment or Ruin: Efficiency* feedback, R3.

Fourth (not shown in Figure 5), the *effectiveness* of investments in buildings, systems, and energy efficiency depends on the policies, routines, and other capabilities of the organization. Effective renewal and efficiency investments require a holistic, systems approach, often called an integrated design process (Kinsley & DeLeon, 2009; Moser et al., 2012; Parrish & Regnier, 2013). It is generally more cost effective to plan lighting and office layout early in a project along with decisions about building orientation and window size, so natural light can be maximized, and to coordinate renewal of all systems in a building rather than renovating piecemeal. Doing so, however, requires more up-front planning and greater coordination among design and engineering specialties, and among the facilities department, building occupants and senior leadership. The stronger these intangible capabilities, the more likely the Reinvestment or Ruin feedbacks will operate as virtuous cycles and the stronger they will be, creating another layer of reinforcing feedbacks.

Modeling defects, work orders, and corner cutting. In addition to the stocks of defects shown in Figure 5, the model explicitly represents the backlogs of maintenance work orders for reactive and proactive work, the flows of work order creation and resolution that alter them, and the routines and decision rules governing the allocation of R&M resources between reactive and proactive work. We disaggregate defects and the work orders they generate into six categories using an industry standard

FIGURE 5
Interacting capabilities: Expanding the boundary of the model



Endogenously accounting for building and system condition, energy efficiency, operating costs and financial pressure creates new capability trap feedbacks. The three main policies for improvement are shown in italics, including programs to improve maintenance, to renew buildings and systems, and to improve energy efficiency.

classification scheme: exterior structures, interior structures, plumbing, electrical, HVAC, and other. For each category $i \in \{1, \ldots 6\}$, we represent two types of work orders: reactive and proactive ($j \in \{R, P\}$). New work orders of each type accumulate in a backlog until they are closed. New reactive work orders, $\mathrm{O}_{i,R}$, arise as latent defects lead to breakdowns, failures or customer complaints:

$$\mathrm{O}_{i,R} = \mathrm{h}_i \mathrm{D}_i, \tag{1}$$

where D_i is the stock of defects in category i and h_i is the hazard rate that a defect that category will lead to a failure or complaint that generates a reactive work order. We estimated the hazard rates from time series data on work orders and the stock of deferred maintenance in each category and from expert judgment of R&M and facilities personnel; as intuition suggests, defect generation rates are higher for mechanical systems, with their rotating equipment (motors, pumps, etc.) than for the structural elements of buildings (the supplement shows the formulation for proactive work orders).

The rate at which work orders for each category and type are closed, $\mathrm{C}_{i,j}$, is given by the total effort of the R&M department, determined by the number of maintenance personnel, N, average work hours per week, H, the share of that effort allocated to each category and type of work, and worker productivity for each:

$$\mathrm{C}_{i,j} = NHf_i \mathrm{s}_j \mathrm{p}_{i,j} \tag{2}$$

where $f_i, i \in \{1, \ldots 6\}$ is the fraction of total maintenance effort allocated to each of the six work categories, s_j is the share of total available work hours done reactively or proactively, and $p_{i,j}$ is the productivity of maintenance effort for each category and type of work.

Interviews and archival records showed that the R&M headcount, N, was relatively constant over the estimation period. Given budget constraints, the R&M organization did not have the ability to expand.

Both work hours and productivity may vary with the load on the R&M organization (Oliva & Sterman, 2001; Repenning & Sterman, 2002). When pressure to complete work is high, mechanics may put in longer hours and may also close work orders more quickly by cutting corners, including spending less effort searching for and eliminating the root causes of the problem. We estimated these effects using work order data from the Institute's SAP system including hours worked, productivity, and the allocation of time to different types of maintenance work, specifying:

$$H = H^{*} w^{\gamma_H} \tag{3}$$

$$\mathrm{p}_{i,j} = \mathrm{p}_{i,j}^{*} w^{\gamma_p} \tag{4}$$

Where, H^{*} is the standard work week, the $\mathrm{p}_{i,j}^{*}$ are the base productivities of each category and type of work, w is work pressure, and γ_p and γ_H are the sensitivities of productivity and work hours to work pressure, respectively. Work pressure, w, is the ratio of the total work hours needed to complete all work on schedule to the work hours available given the headcount and standard workweek.

Regression results yield a highly statistically significant response of productivity to work pressure, with $\gamma_p = 0.14$, ($t = 2.71$, $p < 0.0001$). The results provide evidence of corner cutting (Oliva & Sterman, 2001) and are consistent with the interviews, in which technicians described how high work pressure forced them to "leave it once you get to a point where no one is complaining" and "tweak something until no one is complaining, and then . . . walk away." The model captures these impacts of corner cutting: when less time is spent on each work order, the number of defects found and eliminated falls, and the number of defects created from poor quality work increases.

The estimate for the sensitivity of the workweek to work pressure, γ_H, although positive, was not statistically significantly different from zero. The low sensitivity of work hours to workload is consistent with both the interviews and work hour data: overtime was rare and more often used for scheduled shutdowns than to catch up when work pressure was high.

We assume maintenance effort across the six categories (both reactive and proactive), f_i, is allocated in proportion to the total desired rate of work completion in each category. We use a logit choice model to determine the share of time allocated to reactive and proactive work (see the Supplement). Consistent with the interviews, estimation results show that urgent, reactive work orders squeeze out proactive work. To illustrate, the large volume of reactive work orders yields a simulated allocation of approximately 91 percent reactive and 9 percent proactive work for 2005, close to the actual split. Further, if the volume of scheduled proactive work doubled, while the reactive workload remained the same, the proactive fraction of work would rise only to about 10 percent.

Modeling energy use. We estimated the relationship between energy use and building and system condition using data on energy consumption per gross square foot (gsf) for each building and for each of the three main energy carriers (steam, chilled water, and electricity) between 2000 and 2006. We ran panel regressions for each energy carrier and each building, with time (a proxy for building age) as an independent variable, fixed effects for buildings, and controls for annual heating and cooling degree-days. Results show highly statistically significant ($p < 0.0001$) time trends for all three energy carriers. Electricity use is rising fastest as it includes both the impact of aging and rising plug loads as the density of electronics has risen (a process co-linear with building aging). Rising steam and chilled water use, however, are predominantly driven by building and equipment deterioration.

Changes in energy efficiency are not likely to be linear over longer time horizons. We posit that energy efficiency decays exponentially over time, consistent with Toole and Claridge (2011). We estimated the exponential decay model using two methods (detailed in the Supplement). We then use the results to estimate the potential energy savings available from building renewal. We find that if every building and system were fully renewed, energy use per gsf would fall by 55 percent for chilled water, 41 percent for steam, and 23 percent for electricity relative to 2005 levels. For comparison, the new management school building, completed in 2010, uses less than half the energy per gsf than the ASHRAE 90.1 energy standards for comparable low-rise commercial buildings. Energy use for lighting is 55 percent less than the standard, and heating and cooling loads are 52 percent and 53 percent less than values for comparable MIT office/classroom buildings, respectively. The investments to achieve these reductions increased the capital cost of the building by less than 1 percent compared to a standard, code-compliant building, yielding an NPV of about $10 million (Sterman et al., 2015).

The interviews, quantitative data and estimation results support the feedback structure shown in Figure 5. That structure maps clearly onto the capability trap framework, but at multiple, interacting scales. R&M personnel felt strong pressure to resolve failures and customer complaints quickly, but resources were inadequate to do both. Consequently, proactive work suffered and the organization gradually sank into the trap through the self-reinforcing buildup of deferred maintenance, more frequent breakdowns and higher costs. At the same time, insufficient renewal investment caused buildings, systems, and energy efficiency to deteriorate, increasing defect creation and pushing the maintenance organization farther into reactive work while also increasing operating costs. The resulting financial pressure further limited maintenance building renewal. The descent into the capability trap was gradual, over many decades. By the 2000s the situation had become acute. The result was the "fire drill" atmosphere in which R&M personnel responded to "who's screaming right now," which forced them to defer proper maintenance still further.

To escape the capability trap an organization must invest in capabilities faster than they erode, just as saving a leaky ship requires the crew to bail at least as fast as water leaks in. However, that insight alone, although poorly understood (Booth Sweeney & Sterman, 2000; Cronin, Gonzalez, & Sterman, 2009), does not provide managers with sufficient guidance to select effective programs or allocate resources among maintenance, building and system renewal, and efficiency investments. Doing so requires explicit consideration of the different stocks that constitute the organization's capabilities and the interactions among them.

Consider again the leaky ship: to save the vessel the crew must bail faster than the water flows in. But bailing is exhausting, and the more sailors tasked to bail the fewer are available to sail. Thus the crew should also reduce the bailing required by plugging leaks faster than new ones spring. To do so, they must replace old boards and caulk faster than they rot and leak, but that further increases the workload or cuts the crew available to sail or bail. These tradeoffs could be eased if the crew's repair capabilities improved, but doing so requires building their skills faster than they decay, requiring, in the short run, still more time.

In the same way, escaping the maintenance-building condition-energy efficiency capability trap requires crossing multiple tipping points: to reduce the stock of reactive work orders, the rate they are closed must exceed the rate new ones are opened; to reduce the stock of defects, defect elimination must exceed defect creation; to improve the condition of buildings and systems, building renewal must exceed deterioration; to improve energy efficiency, upgrades must exceed efficiency degradation. Finally, to increase the organization's ability to carry out these investments effectively, intangible capabilities—skills, routines, cooperation, and trust across disciplinary and organizational boundaries—must be built faster than they erode.

The costs and characteristic time delays of these activities differ substantially. The feedbacks involving defects and the allocation of maintenance effort between reactive and proactive work are fastest: latent defects (e.g., worn fan belts, leaky bearing seals, drifting thermostats) can create breakdowns and complaints with delays on the order of days to months, and reactive work is typically done within a few days. In contrast, buildings and systems (e.g., walls, roofs, foundations, windows; HVAC, steam, and chilled water systems; water and sewer lines) have lifetimes on the order of many decades, while repairs are more costly and can take months to years. As seen below, the differential delays, costs, resources, and cross impacts of these different stocks strongly condition the dynamics, including the duration and depth of the WBB dynamic and the NPV and payback times of different policies.

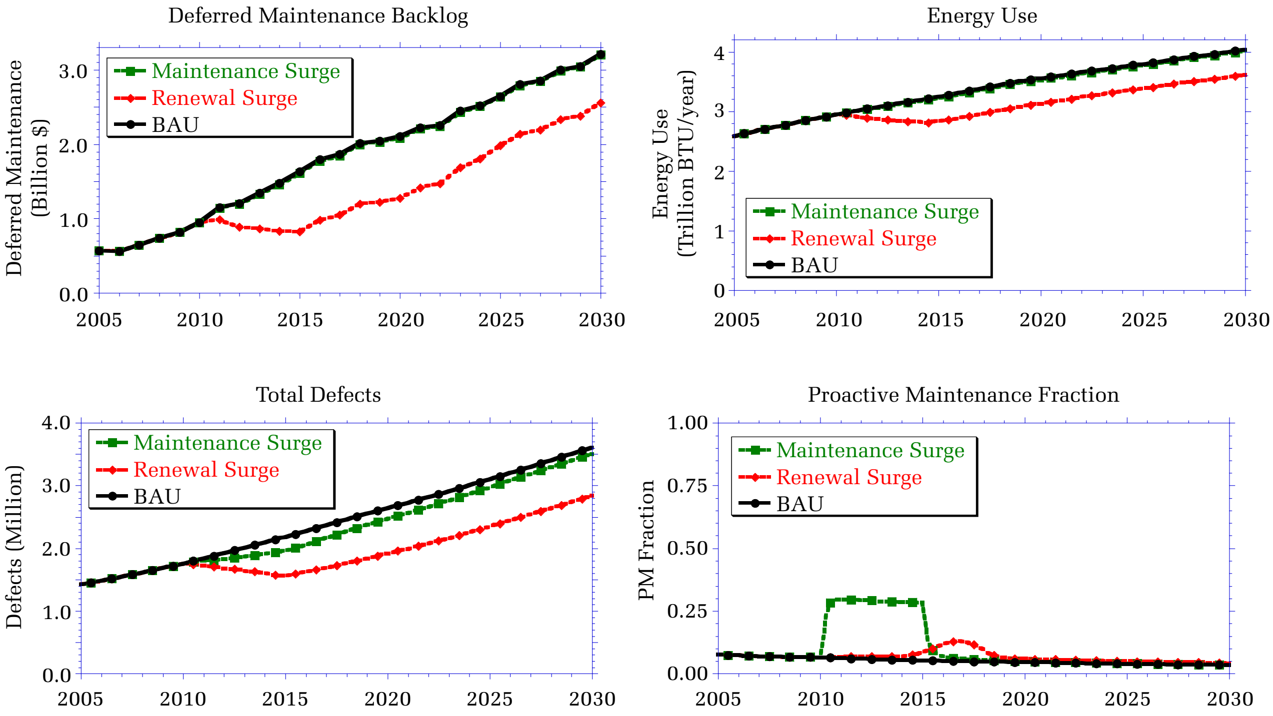
RESULTS

We start by simulating business as usual (BAU). The BAU case assumes capital renewal spending of \$19 million/year, the average rate between 1999 and 2010. Maintenance spending begins at approximately \$15 million/year but varies thereafter with the volume of work to be done. We assume energy prices remain constant in real terms (Table 2 presents sensitivity analysis). The result (Figure 6) is continued deterioration of campus conditions and capabilities. By 2030, the backlog of deferred maintenance is 5.6 times larger than the 2005 level as buildings and systems continue to deteriorate faster than they are renewed. The rising stock of deferred maintenance boosts the rate of defect creation well beyond the capacity of the R&M organization to eliminate defects, raising the stock of defects to 2.54 times the 2005 level despite a doubling in maintenance staff forced by the growing number of breakdowns. Energy use rises to about 1.6 times the 2005 level, while the proactive fraction of maintenance work sinks to 3.4 percent. The organization remains caught in the vicious "ruin" feedbacks described in Figure 5 even as maintenance spending grows.

Failing to Escape the Trap: Surge Funding for Maintenance and Campus Renewal

Figure 6 also shows two policies intended to reverse the deterioration. The "Maintenance Surge" consists of a temporary \$5 million/year increase in the R&M budget, roughly a third, from 2010 to 2015. The surge is intended to jump start the virtuous

FIGURE 6
Results for surges in maintenance and campus renewal



BAU: business as usual; Maintenance Surge: a $5 million/year increase in maintenance budget from 2010 to 2015 (over and above the BAU maintenance budget); Renewal Surge: a surge raising the campus renewal budget to $150 million/year from 2010 to 2015.

cycles of improvement by allowing more proactive work to be done. Indeed, during the surge, proactive R&M work rises from about 8 percent to nearly 30 percent, nearly stabilizing the stock of defects. However, without additional capital renewal the deferred maintenance backlog and energy use continue to climb as under BAU. Defects begin to grow again, even before the surge ends, forcing the R&M team to cut back on proactive work. When the surge ends proactive maintenance quickly falls back toward the BAU level. The surge fails to lift the organization above the tipping point. Using the leaky ship metaphor, the surge allows the organization to bail faster and even plug some leaks, but absent investments in campus renewal or energy efficiency, new leaks still spring faster than old ones are patched. Water flows in faster than even the expanded R&M organization can bail. The ship soon begins to sink once more.[2]

The outcome is similar for a surge in campus renewal ("Renewal Surge" in Figure 6). Here capital renewal jumps from $19 million/year to $150 million/year for 5 years (2010–2015), after which renewal spending returns to prior rates. The renewal surge is similar in magnitude to the actual increase that began roughly at that time (except that, as described below, MIT plans to continue renewal efforts beyond 2015). During the surge, the backlog of deferred maintenance, energy use, and the stock of defects all fall. The drop in defects allows the proactive fraction of maintenance work to rise slightly, to a peak of 13 percent. However, when the surge ends, the backlog of deferred maintenance, energy use, and defects all resume their rise, and the proactive maintenance fraction decays back toward the BAU level. Despite investing $750 million in renewal, the organization does not escape the capability trap. During the surge, capital renewal reduces the number of leaks, slowing the flow of water into the boat. For a time, the rate of bailing slightly exceeds the rate at which water flows in, causing the water (the stock of defects) to fall gradually. But when the surge ends the number of leaks grows. Water soon rushes in in faster than the crew can bail. The ship again starts to sink.

In both cases, substantial investment in a single activity is not enough to escape the capability trap. Significantly expanding the resources of the R&M

---

[2] A large enough surge can push the R&M organization over the tipping point, but without capital renewal, at least $15 million/year in additional R&M spending through 2030 is required (a total of $300 million in additional funds).

organization increases the amount of proactive maintenance, but the poor condition of buildings and systems means defect creation continues to exceed defect resolution. Similarly, a large surge in capital renewal removes some sources of defects, but the stock of deferred maintenance is so large that most maintenance work continues to be reactive, preventing the R&M organization from improving equipment reliability and efficiency or correcting latent defects before they cause breakdowns.

Escaping the Capability Trap

We next simulate coordinated policies for capital renewal, proactive maintenance, and energy efficiency. Figure 7 contrasts simulation results for four policies against the BAU simulation (Table 1 summarizes the policies and results).

Policy 1—Sustained Renewal. In the "Sustained Renewal" case (Policy 1) campus renewal investment increases to $150 million per year beginning in 2010 and remains at that rate thereafter (a total of $3 billion by 2030). Maintenance policies remain as in the BAU case, and any savings from lower energy consumption are harvested, that is, used to support overall Institute programs. The backlog of deferred maintenance falls steadily through 2030, to 29 percent of the 2010 level, $2.9 billion lower than the BAU case. Renewal also stabilizes energy use slightly below 2010 levels (a drop of 29 percent from BAU by 2030). Energy use does not fall as much as the stock of deferred maintenance: renewing buildings and systems upgrades their efficiency to current code, but then equipment and structures begin to deteriorate again. At first, the R&M organization remains stuck in the reactive, firefighting mode. But sustained renewal also gradually lowers the stock of defects, and by 2016, the proactive fraction of R&M work starts to increase. The maintenance organization slowly escapes the capability trap, with the proactive fraction of work reaching 42 percent by 2030. Sustained investment in renewal replaces rotting wood in the hull of ship with new oak, slowly reducing the rate at which new leaks spring. Bailing eventually outpaces the flow of water into the ship.

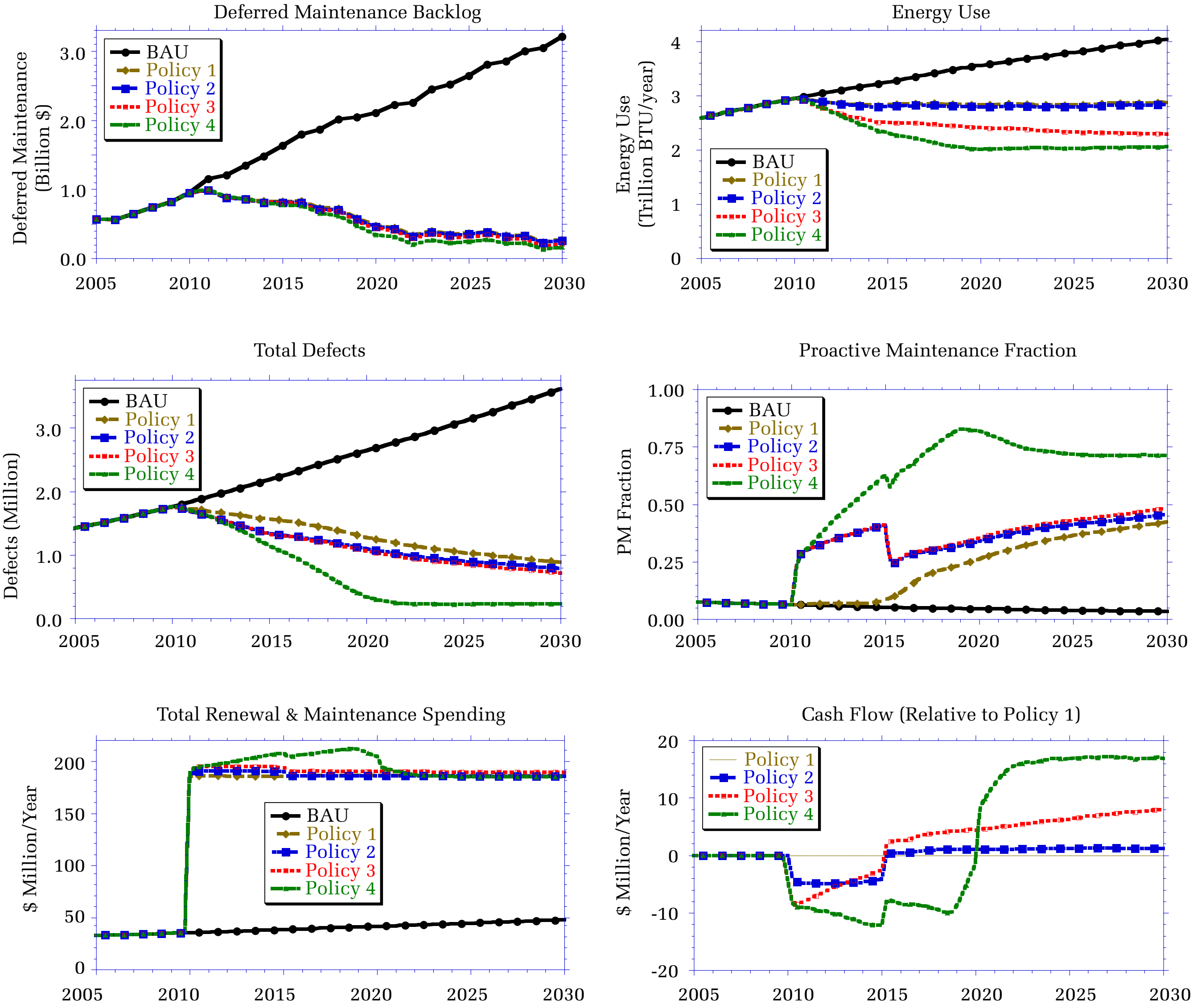
Policy 2—Sustained Renewal with Maintenance Surge. To speed improvement, Policy 2 augments Policy 1 with a surge in the maintenance budget of $5 million/year from 2010 to 2015 (Figure 7; Table 1). The backlog of deferred maintenance and energy use change only slightly (none of the surge goes to renewal or energy efficiency). However, the surge immediately increases the proactive maintenance fraction to 22 percent, cutting the stock of defects below Policy 1. With fewer defects, still more time is available for proactive effort, which rises to 41 percent by 2015. However, when the surge ends the proactive fraction immediately falls, ending only slightly higher than in Policy 1. Compared to sustained renewal alone, the surge causes negative cash flow of approximately $4.8 million/year through 2015, slightly less than the $5 million/year surge because higher proactive work cuts breakdowns and collateral damage. After the surge ends, these savings yield a small positive cash flow compared to Policy 1, but the savings do not outweigh the costs: the NPV of the maintenance surge relative to Policy 1 is $\$-2.9$ million.[3]

Policy 3—Investing in Energy Efficiency. Policies 1 and 2 assume that new buildings and systems are built to code. However, additional investment in energy efficiency can lower the energy consumption of renovated plant and equipment beyond the requirements of building code. Policy 3 builds on Policy 2 by specifying that all renewal projects include additional investment in energy efficient structures and systems beyond the levels code requires. Such investments include additional insulation, vapor and air barriers to eliminate air infiltration/exfiltration, high performance windows, energy recovery units in HVAC systems, variable speed lab hoods, LED lighting, occupancy sensors, and many others—and the use of an integrated design process that coordinates building and system design to optimize the performance of the buildings and systems as a whole. We conservatively assume that the extra investment yields half the potential efficiency improvement, at a cost of 2.5 percent of the base capital cost. As before, any savings from lower energy are harvested, that is, used to support general Institute programs. The additional investment totals $85 million by 2030. By 2030, campus energy use falls 22 percent below the 2010 level, far lower than the level achieved by Policy 1. Higher energy efficiency also yields spillovers to maintenance: as documented above, much of the load on the R&M organization arises from occupant complaints that spaces are too hot or too cold. Better windows and insulation, lower outside air infiltration, and better HVAC systems not only lower energy use but improve occupant comfort. With fewer urgent complaints about temperature, the R&M organization finds itself with slightly more resources for proactive work, allowing the stock of defects to fall somewhat compared to Policy 2. In addition, lower defect levels reduce collateral

---

[3] We use a discount rate of 5 percent/year, based on MIT's actual cost of capital: since 2010 the Institute has issued $1.3 billion in bonds to fund campus renewal at an average rate of approximately 5 percent/year.

FIGURE 7
Results for sustained renewal policies



BAU: business as usual; Policy 1: sustained renewal of $150 million/year 2010–2030; Policy 2: Policy 1 + $10 million/year maintenance surge 2010–2015; Policy 3: Policy 2 + additional investment in energy efficiency; Policy 4: Policy 3 + reinvestment of energy savings. See Table 1 for details. The supplement (Figure S23) shows total maintenance spending in the BAU scenario and Policies 1–4.

damage from breakdowns (through the reinforcing Collateral Damage feedbacks R0 in Figure 5). The cash flow of Policy 3 is initially worse than Policies 1 and 2, but the energy savings outweigh the cost of the extra investment in energy efficiency after only 3 years. Cash flow becomes positive in 2015, when the maintenance surge ends, and the savings continue to grow through 2030. The additional investment in energy efficiency yields a positive NPV of $58.5 million relative to Policy 1, a discounted benefit/cost ratio of 1.70, while reducing cumulative campus energy use (and associated GHG emissions) by 7.9 trillion British thermal units (GBTU) by 2030, 14 percent of cumulative consumption from 2010 to 2030: a substantial win-win.

Policy 4—Reinvesting Energy Savings. In Policy 4, we aim to speed the shift of the Reinvestment or Ruin feedbacks in Figure 5 from vicious to virtuous cycles by reinvesting the savings from lower energy use in further improvement. We allocate 25 percent of the energy savings to further efficiency programs and 75 percent to the maintenance budget. Reinvestment continues until 2020, when diminishing returns reduce the opportunities for productive use

TABLE 1
Comparing Investments in Capital Renewal, Maintenance, and Energy Efficiency

Policy:	**P1 Sustained Renewal**	**P2 P1 + Maintenance Surge**	**P3 P2 + Additional Energy Efficiency Investment**	**P4 P3 + Reinvestment of Energy Savings**
Cumulative Investment in Renewal, 2010–2030	$3 Billion	$3 Billion	$3 Billion	$3 Billion
Cumulative Investment in Maintenance, 2010–2015	0	$25 Million	$25 Million	$25 Million
Cumulative Additional Energy Efficiency Investment	0	0	$84.6 Million	$74.1 Million
Cumulative Energy Savings Reinvested	0	0	0	$173 Million
NPV of investment (relative to Policy 1)	0	$-2.9 Million	$58.5 Million	$98.5 Million
Discounted Benefit/ Cost Ratio (relative to Policy 1)	—	0.83	1.70	1.73
Payback time (relative to Policy 1)	—	After 2030	12 Years	16 Years
Cumulative energy savings relative to Policy 1 (GBTU)	—	0.6	7.9	12.9

of these resources; after 2020 the savings add to general Institute revenue. Reinvesting the energy savings strengthens the reinforcing feedbacks R1-R3 in Figure 5 by generating still more funds for efficiency; reinvesting in maintenance cuts new defect generation, generating still more resources for proactive maintenance. The policy generates $173 million in additional investment by 2030, yielding substantial benefits. By 2017, the proactive fraction of R&M effort exceeds 70 percent, further accelerating improvement; by 2020, defects are 68 percent lower than Policy 3, reducing collateral damage, which further improves the condition of plant and equipment. Cumulative energy savings through 2030 are 12.9 GBTU, a gain of 63 percent over Policy 3. The policy enables the Institute to escape the capability trap and enjoy higher quality, more efficient facilities and a safer campus—all for the same initial investment. The reinvestment policy intensifies the WBB tradeoff: program cash flow falls farther and remains negative longer than under Policy 3, but the NPV of the program rises from $58.5 to $98.5 million, a 68 percent gain. Reinvestment substantially increases the win-win benefits.

Sensitivity Analysis. Important assumptions in the model are uncertain. Table 2 summarizes results of sensitivity analysis across three critical uncertainties: the discount rate, energy prices, and the potential for energy efficiency improvement. MIT's actual cost of capital for investments in campus renewal is about 5 percent per year. Under that base case assumption, Policy 4 yields an NPV of $98.5 million, a discounted benefit/cost ratio of 1.73. At a discount rate of 3 percent/year the NPV rises to $284 million, a discounted benefit/cost ratio of 2.7. Under a discount rate as high as 9 percent per year, far higher than MIT's actual cost of capital, the program still yields a positive NPV of $7.5 million and a discounted benefit/cost ratio of 1.08. Future energy prices are highly uncertain. Many argue that prices are likely to rise as economic growth, particularly in developing nations, increases energy demand, and as policy responses to the risks of climate change increase fossil fuel prices. Similarly, innovation is increasing the potential for energy savings. Alternatively, petroleum prices fell dramatically in 2014 and may remain low for some time, and potential efficiency gains may be lower than we assume. Hence we varied assumed future energy prices from 20 percent below the base case to 50 percent above them, and the potential for efficiency improvements from 25 percent below the base case to 10 percent above it. The coordinated program of energy efficiency with reinvestment of savings remains the superior policy even under the pessimistic assumptions, with NPVs of $77 million under low improvement potential and $61 million under low energy prices.

IMPACT: CAMPUS RENEWAL AT MIT

We began this study in 2007. Since then, MIT has implemented substantial changes to its maintenance, energy, and campus renewal policies following the recommendations above. Since 2010, the Institute has issued $1.3 billion in bonds to fund major investments in facilities renewal and energy efficiency to reduce the stock of deferred maintenance. The current leadership team is committed to eliminating the backlog of deferred maintenance and is boosting total spending for campus renewal (including new construction, which we do not include in our model) to $200 million per year. Efficiency and sustainability programs are central to the effort:

TABLE 2
Sensitivity Analysis

	Base (Policy 1)	Discount Rate (%/year)		Energy Efficiency Potential		Energy Prices	
		3%/year	9%/year	+10%	−25%	−20%	+50%
NPV (relative to P1) (Million)	$98.5	$284.0	$7.5	$106.9	$77.2	$60.8	$196.9
Discounted Benefit/Cost ratio	1.73	2.69	1.08	1.79	1.58	1.46	2.39
Payback time (years)	16	16	16	15	18	18	12
Cumulative energy savings relative to Policy 1 (GBTU)	12.9	12.9	12.9	13.5	11.3	12.7	13.1

Values compare Policy 4 (continuous renewal + maintenance surge + additional investment in energy efficiency + reinvestment of energy savings) to Policy 1 (continuous renewal only). Discount rate and future real energy prices assumed to be constant at the indicated ratio to base case values.

both new construction and retrofits are designed to meet the LEED Silver standard as a minimum, and many projects since 2010 have achieved LEED Gold, with documented energy and other savings generating large financial benefits (Sterman et al., 2015). The R&M organization now emphasizes proactive maintenance and improvement, and is carrying out efficiency programs such as lighting and plumbing upgrades as a routine component of ongoing maintenance work. Under the Comprehensive Stewardship Program, dedicated maintenance teams carry out proactive maintenance of key zones of the campus.[4] The savings have been substantial. Consider the biology building, built in the early 1990s. Defects had crept in to the equipment after years of mostly reactive maintenance. Sensors and controls had drifted so that the building was heating and cooling itself simultaneously. Eliminating that waste, along with basic HVAC system cleaning and repairs, yielded immediate energy savings worth about $360,000/year. The total cost of the program was about $150,000 (Halber, 2010). The savings were so large and immediate that there was essentially no WBB behavior. Similar results have been realized in other buildings by carrying out long-deferred basic maintenance, such as cleaning steam traps. MIT is working with other organizations to build on these results. In 2010, the Institute partnered with the local electric utility to reduce campus electricity consumption. The $13 million program targeted a 15 percent reduction in electricity use, totaling 34 million kWh over the 3-year program. Actual reductions exceeded the targets every year, generating $4.4 million per year in operating cost savings, a projected total of $50 million over the life of the improvements, while reducing GHG emissions by 20,000 tCO_2 per year. The program has been renewed and expanded to include natural gas.[5] The Institute is reinvesting a portion of the savings in further improvement.

Finally, the Institute is building the intangible capabilities needed to enhance the effectiveness of these investments, including appointment of a campus sustainability director, reporting to the Executive Vice President (who oversees all campus operations including the management of the endowment and finances), coordination of previously disparate sustainability initiatives, commitment to use of the Integrated Design Process on all capital projects, and training in proactive best practices for R&M and facilities department staff.

DISCUSSION AND CONCLUSION: HOW TO SAVE A LEAKY SHIP

Profitable opportunities to improve organizational performance while benefiting society are well documented, yet such "win-win" investments are often not implemented. The prevalence of unexploited win-win opportunities is not fully explained by existing theories emphasizing shareholder pressure for quick returns, behavioral biases, market failures, and principal–agent problems such as the landlord–tenant problem. The case of energy efficiency and facilities maintenance at MIT

Author's voice:
What has been the impact of your research so far?

---

[4] See http://web.mit.edu/mit2030/themes/renovation-renewal-stewardship/csg-program.html.

[5] See http://newsoffice.mit.edu/2013/mit-nstar-extend-energy-efficiency-program-0702 and http://newsoffice.mit.edu/2013/energy-savings-add-up-to-success-for-efficiency-forward.

yields new insights into this puzzle. MIT has a large endowment, AAA credit rating and low cost of capital, an explicit pro-social mission, and as owner-operator of its facilities, does not face landlord–tenant problems. Nevertheless, over many decades the Institute gradually fell into the capability trap, accumulating a large backlog of deferred maintenance that raised energy, maintenance, and other operating costs, forcing the maintenance organization into a reactive, firefighting mode and preventing the investments needed to improve. We extended the theory of the capability trap to account explicitly for multiple capabilities and how they interact, including maintenance, capital investment to renew buildings and systems, and energy efficiency. To do so, we developed a formal simulation model, grounded in ethnographic study and detailed quantitative data on maintenance and facilities management, energy use, and the condition and renewal costs of every system in every building on campus. We use these data to estimate important physical relationships, such as the rate at which energy efficiency deteriorates as buildings and systems age, and critical behavioral decision rules, such as how maintenance personnel allocate time to reactive versus proactive maintenance. The results illustrate five reasons it is difficult to escape the capability trap, and how to overcome them: how to save a leaky ship.

1. To survive we must reassign hands from sailing to bailing, which will temporarily slow our progress

More generally, lags in capability development mean organizations experience "Worse-Before-Better" behavior: when programs to build capabilities are launched, the first response is a drop in organizational performance as resources are added or reassigned from firefighting to improvement, and as operable equipment is taken offline so that improvement work can be done. Escaping the trap and implementing win-win investment opportunities require all relevant stakeholders understand and be prepared for the WBB dynamic. If not, people, from senior leaders to front-line workers, may react to the initial drop in performance and/or rise in costs as evidence that the new policies do not work, abandon the program, and become cynical about the possibility of improvement (Repenning & Sterman, 2002; Keating et al., 1999). For MIT (or for-profit firms), these stakeholders range from the senior leadership, who set governance policies, goals and budgets and evaluate the performance of departmental managers, to the managers in those departments who experience those goals, budgets and evaluations, to the front-line workers who choose every day whether to work harder or smarter. Additional stakeholders at MIT include donors, alumni, students, and faculty; among for-profit firms, stakeholders would include investors and analysts, supply chain partners, customers, regulators, and members of affected communities.

Methods to assess the depth and duration of WBB behavior and set realistic goals include (a) estimating the "improvement half-life" of a process by assessing its technical and organizational complexity (Sterman et al., 1997; Sterman, 2015); (b) finding small, quick wins (Weick, 1984) to moderate the WBB dynamic; and (c) seeking synergies that yield increasing returns to investment. For example, a small amount of insulation in a building will save a little on heating bills, but more insulation, better windows, and reducing air infiltration may lower energy use enough to downsize the heating system or eliminate it altogether, resulting in far greater savings with higher NPV and ROI than smaller investments (Lovins, 2012; Sterman et al., 2015).

2. Bailing is not enough: To stay afloat, we must bail faster than water leaks in

Since capabilities are stocks, the system exhibits a tipping point. To escape the trap, investment must be large enough to build capabilities faster than they decay. Managers may be willing to boost spending on improvement by what they believe to be large amounts relative to their past or peers, yet will still fail if those investments are not large enough and sustained long enough to build capabilities faster than they erode. Many organizations do not measure or report their capabilities or fail to do so frequently enough to recognize whether they have crossed the tipping point (Rahmandad & Repenning, 2015). Incentives, performance evaluations and the tendency to "shoot the messenger" create pressure to avoid reporting problems or, all too often, deliberately covering them up, as in the "Liar's Club" observed in some product development processes (Ford & Sterman, 2003). But it is possible to assess capabilities and note changes in them. Capabilities embodied in physical infrastructure are comparatively easy to measure. Useful metrics include backlogs of deferred maintenance, defect and rework rates in products, processes, product returns, or warranty claims. Intangible capabilities such as organizational routines, employee skills, and trust within and across organizational functions and boundaries are harder to assess, but can be measured through benchmarking, customer and employee satisfaction surveys, and testing and recertification programs as is routinely done in high-hazard settings such as aviation and the military.

3. Bailing is not enough: We must plug leaks faster than they spring, rebuild the hull faster than it decays, and improve our design and carpentry skills

Even if the crew, through heroic efforts, can bail faster than the water flows in, bailing is exhausting and diverts the crew from sailing and other tasks necessary to survive: it cannot be sustained indefinitely. To save the ship, the crew must not only bail faster than the water leaks in, but also patch leaks faster than they spring, reducing the flow of water into the bilge; deploy new, better materials faster than old ones decay, reducing the rate at which new leaks spring; and, crucially, build the crew's skills in these activities. Doing so will divert even more hands from sailing in the short run, intensifying the WBB dynamic (see point 1). Generally, because there are multiple capabilities and multiple stocks of defects in any organization, there are multiple tipping points. These range from defects in equipment and backlogs of deferred maintenance to inefficient systems to intangible organizational routines, skills, and attitudes. Sustainable improvement requires understanding how these capabilities interact and implementing coordinated policies for improvement. To do so, organizations should avoid the widely used strategy of organizational decomposition, with individual sites, divisions, or functions working their own improvement programs. Organizations should approach the work of improvement as a system, explicitly considering the interactions among resources and capabilities through integrated design processes, cross-functional teams, and other methods to enhance collaboration across organizational silos.

4. To improve our effectiveness we must develop our ability to coordinate bailing, repair and design improvement

When the ship is sinking, it is tempting to put all effort into bailing and patching. But that may not be sufficient—the crew may soon become exhausted, threatening morale, or even mutiny. The officers and crew should also invest in learning how to organize more effectively so that the limited time available builds the crews' skills and productivity—and their will to carry out all these activities—faster than fatigue and failure sap energy and erode morale. Generally, the ability to coordinate improvement in the ensemble of capabilities is itself a critical capability, one likely to be weak when stuck deep in the capability trap. After years of downsizing and cost cutting, few organizations today have any slack to invest in capabilities: front-line workers are continually pressured to work harder; managers are told "do more with less" and face 24/7 work pressure through email and texts. Although a venerable concept in organization theory, managers think of "slack" as "waste"—hence, it is better to characterize it as "a strategic margin of reserve capacity." Organizations without significant slack cannot increase bailing (reactive maintenance), leak repair (proactive maintenance), and hull redesign (process improvement) without improving productivity by learning to coordinate these activities better. As part of its program of campus renewal, MIT not only increased investment in proactive maintenance and energy efficiency, but reorganized to coordinate these investments tightly with maintenance, facilities, operations and finance, and to invest in training to build the skills needed to sustain improvement.

5. As the need to bail eases, reinvest the savings to plug more leaks and strengthen the hull before sailing on

After stabilizing the water level in the bilge, it will be tempting to get underway immediately instead of using the crew freed up from bailing to plug and prevent new leaks so future bailing can be avoided. The urge to make up for lost progress is powerful, but likely to cause another crisis when new leaks spring. In general, savings from initial investment should be reinvested in further improvement. Escaping the capability trap requires shifting the positive "Reinvestment or Ruin" feedbacks from vicious to virtuous cycles. In the vicious cycle regime, managers often feel compelled to defer or cut investments in proactive maintenance, facilities renewal, and efficiency upgrades as they find it increasingly difficult to attain cost and throughput targets. If, despite these pressures, managers do the right thing and invest in improvement, they will then come under pressure to use the initial savings to close budget gaps or fund new programmatic activity. Doing so is tempting: initial gains are immediately evident and would likely be rewarded, while the opportunity costs of harvesting are not directly observable. But harvesting initial gains weakens or defeats the reinvestment feedbacks, preventing the organization from realizing many win-win opportunities, and possibly preventing it from escaping from the capability trap at all. Mechanisms to promote reinvestment within organizations include revolving loan funds financed by program savings, gain-sharing programs in which a unit retains a share of any savings for its own use, and performance evaluations and balanced scorecards that reward capability improvements. External stakeholders (e.g., donors, funders, regulators) can encourage reinvestment through matching fund programs, tax credits, and gain-sharing contracts (e.g., power purchase agreements to promote renewable energy; stronger building codes combined with subsidies to fund retrofits for low and medium income households).

These five principles show how difficult it can be to escape the capability trap, but also point to policies for success. To illustrate, consider the USS Constitution, a wooden-hulled frigate launched in 1797 in Boston, MA. Known as "Old Ironsides" for her resilience under close cannon fire in the War of 1812, she saw active duty until 1881, and was designated a museum ship in 1907. Congress, however, did not appropriate sufficient funds for maintenance, so by the 1920s

> she was starting to show the effects of her age and extended use. Faced with an ever-diminishing budget and the veritable loss of a generation of skilled wooden ship builders, no large scale repair effort could be considered; rather, all repairs at the time were minor in scope, and primarily cosmetic. . . . The stern had decayed to the point of nearly falling off, and cement was being used to patch rotted areas in the ship's decks and hull. Perhaps the most distressing news, however, was the rate at which Constitution was shipping water—over two feet a day, necessitating a daily visit by a tugboat to pump her out.[6]

The Navy responded by carrying out a full renovation and the Constitution became one of the most popular tourist attractions in Boston. Since then, the Navy has rigorously maintained and restored the ship, though doing so has meant periodic, multi-year layups in dry dock, during which she is not available to tourists. Today, Old Ironsides is the oldest commissioned naval vessel in the world. She sailed under her own power in 2012 to commemorate her victories in the War of 1812. And in 2015, she entered dry dock again for another major restoration, projected to last 3 years.

Similarly, the MIT experience suggests how large organizations can escape the capability trap, yielding substantial win-win benefits. We note, however, that although many actions have already paid back the initial investments, the ultimate test is whether organizational performance improves over the long term. Follow-up is needed to resolve uncertainties, learn from experience and continue to develop the capabilities needed for sustainable success.

The dynamics described here are likely to apply to a wide variety of win-win investments. Although some investments in safety or environmental performance are simple technological upgrades, quickly implemented, many others require substantial organizational changes, with long lags. New technologies may cause disruptive and unpredictable changes to work routines and intergroup relationships (e.g., Barley, 1986; Orlikowski, 1992). New organizational structures that support adherence to regulations, such as compliance offices and safety and environmental management systems, can produce similar effects (Huising & Silbey, 2011; Kelly & Dobbin, 2007). To make these technologies and management systems effective, organizations often must devote sizable resources to learning how to operate new technologies, resolve disagreements and interpretation challenges that emerge, coordinate across functions, and build new cultures. Underinvestment in these less tangible capabilities can cause resistance and conflict that limit the effectiveness of technical and administrative innovations (Lyneis, 2012).

Author's voice:
What was most challenging about this project?

Our results have implications for theories of self-regulation and corporate social responsibility. Scholars have long recognized that some organizations outperform others with regard to socially beneficial outcomes, even within the same industry and regulatory environment (e.g., Gunningham et al., 2003). Understanding such variation, however, has proven to be more difficult. Scholars point to differences in technical competency (Christmann, 2000), local institutional pressures, or legal environments (Bansal, 2005; Marquis, Glynn, & Davis, 2007; Short & Toffel, 2010) as important factors. The literature on self-regulation also attributes performance variation to differences in the commitment of managers and workers who must identify, advocate for, and implement improvements (Gunningham et al., 2003; Henriques & Sadorsky, 1999; Parker, 2002; Roome, 1992). The theory developed and tested here helps explain how commitment co-evolves endogenously with competence, capabilities, resource allocation, and performance.

In addition to commitment, technical competency, adequate capital and insulation from short-term pressures, we suggest managers must also possess an understanding of the complex dynamics of organizational improvement. Such understanding is often lacking (Repenning & Sterman, 2001). Actors in complex systems routinely misperceive the effects of accumulations, time delays, and feedback relationships (Cronin et al., 2009; Moxnes, 1998; Paich & Sterman, 1993; Sterman, 1989). Managers often implement policies that are thwarted by unanticipated consequences, a phenomenon known as "policy resistance" (Sterman, 2000) and that can lead to cynicism about the possibility of improvement (Keating et al., 1999; Sterman, Repenning, & Kofman, 1997). Commitment, self-efficacy, and belief in the power of

---

[6] USS Constitution Museum, National Cruise Scrapbooks, 1931-34, https://www.ussconstitutionmuseum.org/proddir/prod/495/39/.

individual agency should be seen as endogenous and coevolving with the physical and institutional structure of the complex systems in which we are all embedded (Repenning & Sterman, 2002; Sterman, 2000). Specifically, managers who fail to understand the five principles above may underinvest, so capabilities continue to erode, albeit perhaps more slowly. They may then interpret the continued slide in performance as evidence that the program is failing and abandon it. Even if they invest enough to get over the tipping point and are willing to endure the WBB behavior, they may then harvest initial savings, weakening the reinforcing feedbacks needed to escape the trap.

Like all studies, ours has limitations. The capability trap and unexploited win-win opportunities are common in many contexts, but our study is specific to a particular organization. First, follow-up research should explore whether and how the dynamics we describe apply in other settings, including for-profit firms and government, and in contexts beyond facilities, maintenance and energy use such as working conditions (e.g., Locke, 2013). Second, although the model is grounded in a wide range of data, the conclusions are robust to major uncertainties and we sought to make conservative assumptions throughout, the boundary of the analysis can be expanded further. For example, we assume no technical progress that could increase energy efficiency potential or lower its costs. We consider only energy and do not treat potential win-win opportunities from reducing water, toxic materials, and other forms of waste. We modeled the existing campus and did not seek to capture growth in programs and facilities. We omit a range of impacts, from improved safety to the impact of better facilities on research productivity, student success, organizational reputation and occupant comfort that can improve morale, health, productivity, and the recruitment and retention of students, staff and faculty. Although these benefits are difficult to quantify, research suggests they are much larger than the direct energy and maintenance savings (Heerwagen, 2010; Miller et al., 2009; World Green Building Council, 2013).

Finally, while many investments offer win-win benefits, other pro-social policies that improve environmental quality and human welfare may not. Society has justly banned slavery, child labor and many unsafe materials and working conditions despite the fact that these practices were highly profitable. Here, we sought to understand why well-documented win-win opportunities are so often unexploited. Our results should not be misconstrued to suggest that prosocial actions are justified only if they are profitable.

Win-win investments enable organizations to become more socially responsible while improving their own performance. Eliminating the barriers to implementation presents an important opportunity for both scholarly research and practical action.

Author's voice:
What would you differently, if you could do it over again?

REFERENCES

Amin, S. 2011. US grid gets less reliable [The Data]. ***IEEE Spectrum***, 48: 80.

Bansal, P. 2005. Evolving sustainably: A longitudinal study of corporate sustainable development. ***Strategic Management Journal***, 26: 197–218.

Barley, S. 1986. Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments. ***Administrative Science Quarterly***, 31: 78–108.

Barnett, M. L. 2007. Stakeholder influence capacity and the variability of financial returns to corporate social responsibility. ***Academy of Management Review***, 32: 794–816.

Bazerman, M. 2009. Barriers to acting in time on energy and strategies for overcoming them. In K. Gallagher (Ed.), ***Acting in time on energy policy***: 162–181. Washington, DC: Brookings Institution Press.

Booth Sweeney, L., & Sterman, J. 2000. Bathtub dynamics: Initial results of a systems thinking inventory. ***System Dynamics Review***, 16: 249–286.

Campbell, J. 2007. Why would corporations behave in socially responsible ways? An institutional theory of corporate social responsibility. ***Academy of Management Review***, 32: 946–967.

Carroll, J., Sterman, J., & Marcus, A. 1998. Playing the maintenance game: How mental models drive organizational decisions. In R. Stern and J. Halpern (Eds.), ***Debating rationality: Nonrational aspects of organizational decision making***. Ithaca, NY: Cornell University Press, 99–121.

Charles, D. 2009. Leaping the efficiency gap. ***Science***, 325: 804–811.

Christmann, P. 2000. Effects of "best practices" of environmental management on cost advantage: The role of complementary assets. ***Academy of Management Journal***, 43: 663–680.

Coglianese, C., & Nash, J. 2001. ***Regulating from the inside: Can environmental management systems achieve policy goals?*** Washington, DC: Resources for the Future.

Creyts, J. C., Derkach, A., Nyquist, S., Ostrowski, K., & Stephenson, J. 2007. ***Reducing US greenhouse gas emissions: How much at what cost?*** McKinsey & Company. http://www.mckinsey.com/client_service/sustainability/latest_thinking/reducing_us_greenhouse_gas_emissions. Accessed February 20, 2016.

Cronin, M., Gonzalez, C., & Sterman, J. 2009. Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. ***Organizational Behavior and Human Decision Processes***, 108: 116–130.

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

Donaldson, T., & Preston, L. 1995. The stakeholder theory of the corporation: Concepts, evidence, and implications. ***Academy of Management Review***, 20: 65–91.

Effinger, J., Friedman, H., & Moser, D. 2009. ***A study on energy savings and measure cost effectiveness of existing building commissioning***. Portland Energy Conservation, Inc. http://www.peci.org/sites/default/files/annex_report.pdf. Accessed June 14, 2012.

Eichholtz, P., Kok, N., & Quigley, J. 2010. Doing well by doing good? Green office buildings. ***The American Economic Review***, 100: 2492–2509.

Estlund, C. 2010. **Regoverning the workplace: From self-regulation to co-regulation**. New Haven, CT: Yale University Press.

Ford, D., & Sterman, J. 2003. The Liar's Club: Concealing Rework in Concurrent Development. ***Concurrent Engineering: Research and Applications***, 11(3): 211–220.

Frederick, S., Loewenstein, G., & O'Donoghue, T. 2002. Time discounting and time preference: A critical review. ***Journal of Economic Literature***, 40: 351–401.

Freeman, R. 1999. Divergent stakeholder theory. ***Academy of Management Review***, 24(2): 233–236.

Freeman, J., Larsen, E. R., & Lomi, A. 2012. Why is there no cannery in 'Cannery Row'? Exploring a behavioral simulation model of population extinction. ***Industrial and Corporate Change***, 21: 99–125.

Fuerst, F., & McAllister, P. 2011. Eco-labeling in commercial office markets: Do LEED and Energy Star offices obtain multiple premiums? ***Ecological Economics***, 70: 1220–1230.

Gibbons, R., & Henderson, R. 2013. What do managers do? In R. Gibbons & J. Roberts (Eds.), ***Handbook of organizational economics***: 680–731. Princeton, NJ: Princeton University Press.

Gillingham, K., Newell, R., & Palmer, K. 2009. ***Energy efficiency economics and policy***. Washington, DC: Resources for the Future. http://www.nber.org/papers/w15031.

Giovannetti, G. 2008. The J-Curve. In S. N. Durlauf & L. E. Blume (Eds.), ***The new Palgrave dictionary of economics*** (2nd ed.). Basingstoke, UK: Palgrave Macmillan. http://www.dictionaryofeconomics.com.

Gunningham, N., Kagan, R., & Thornton, D. 2003. ***Shades of green: Business, regulation, and environment***. Palo Alto, CA: Stanford University Press.

Halber, D. 2010. ***Gaining visibility into buildings' real-time energy performance***. http://mitei.mit.edu/news/gaining-visibility-buildings-real-time-energy-performance.

Hart, S. 1995. A natural-resource-based view of the firm. ***Academy of Management Review***, 20: 986–1014.

Hawken, P., Lovins, A., & Lovins, L. 1999. ***Natural capitalism: Creating the next industrial revolution***. Boston, MA: Little, Brown and Co.

Heerwagen, J. 2010. Green buildings, organizational success and occupant productivity. ***Building Research & Information***, 28: 353–367.

Henriques, I., & Sadorsky, P. 1999. The relationship between environmental commitment and managerial perceptions of stakeholder importance. ***Academy of Management Journal***, 42: 87–99.

Heo, Y., Choudhary, R., & Augenbroe, G. 2012. Calibration of building energy models for retrofit analysis under uncertainty. ***Energy and Building***, 47: 550–560.

Homer, J. 2012. Partial-model testing as a validation tool for system dynamics. ***System Dynamics Review***, 28: 281–294.

Howard-Grenville, J. 2007. Developing issue-selling effectiveness over time: Issue selling as resourcing. ***Organization Science***, 18: 560–577.

Howarth, R., & Sanstad, A. 1995. Discount rates and energy efficiency. ***Contemporary Economic Policy***, 13: 101–109.

Huising, R., & Silbey, S. 2011. Governing the gap: Forging safe science through relational regulation. ***Regulation & Governance***, 5: 14–42.

Jaffe, A., & Stavins, R. 1994. The energy-efficiency gap: What does it mean? ***Energy Policy***, 22: 804–810.

Jaffe, A., Newell, R., & Stavins, R. 2004. The economics of energy efficiency. In C. Cleveland (Ed.), ***Encyclopedia of Energy***: 79–90. Amsterdam: Elsevier.

Keating, E., Oliva, R., Repenning, N., Rockart, S., & Sterman, J. 1999. Overcoming the improvement paradox. ***European Management Journal***, 17: 120–134.

Kelly, E., & Dobbin, F. 2007. How to stop harassment: Professional construction of legal compliance in organizations. ***American Journal of Sociology***, 112: 1203–1243.

Kinsley, M., & DeLeon, S. 2009. ***Accelerating campus climate initiatives.*** Rocky Mountain Institute. http://www.rmi.org/Content/Files/RMI_AcceleratingCampusClimateInitiatives.pdf. Accessed February 26, 2016.

Ledet, W. 1999. Engaging the entire organization: Key to improving reliability. ***Oil & Gas Journal***, 97: 54–57.

Levine, D., Toffel, M., & Johnson, M. 2012. Randomized government safety inspections reduce worker injuries with no detectable job loss. ***Science***, 336: 907–911.

Levitt, J. 2009. ***Handbook of maintenance management***. New York, NY: Industrial Press.

Locke, R. 2013. ***The promise and limits of private power: Promoting labor standards in a global economy***. New York, NY: Cambridge University Press.

Lovins, A. 2012. ***Reinventing fire***. White River Junction, VT: Chelsea Green.

Lyneis, J. 2012. ***Resistance from top to bottom: The dynamics of risk management in complex crganizations.*** Doctoral Dissertation, MIT Sloan School of Management. http://hdl.handle.net/1721.1/79030.

Margolis, J., & Walsh, J. 2003. Misery loves companies: Rethinking social initiatives by business. ***Administrative Science Quarterly***, 48: 268–305.

Marquis, C., Glynn, M., & Davis, G. 2007. Community isomophism and corporate social action. ***Academy of Management Review***, 32: 925–945.

Martani, C., Lee, D., Robinson, P., Britter, R., & Ratti, C. 2012. ENERNET: Studying the dynamic relationship between building occupancy and energy consumption. ***Energy and Building***, 47: 584–591.

McGuire, J., Sundgren, A., & Schneeweis, T. 1988. Corporate social responsibility and firm financial performance. ***Academy of Management Journal***, 31: 854–872.

McKinsey. 2010. Impact of the financial crisis on carbon economics. Version 2.1 of the global greenhouse gas abatement cost curve. http://www.mckinsey.com/~/media/McKinsey/dotcom/client_service/Sustainability/cost%20curve%20PDFs/ImpactFinancialCrisisCarbonEconomicsGHGcostcurveV21.ashx. Accessed February 26, 2016.

McWilliams, A., & Siegel, D. 2001. Corporate social responsibility: A theory of the firm perspective. ***Academy of Management Review***, 26: 117–127.

Miller, N., Pogue, D., Gough, Q., & Davis, S. 2009. Green buildings and productivity. ***Journal of Sustainable Real Estate***, 1: 65–89.

Mills, E. 2011. Building commissioning: A golden opportunity for reducing energy costs and greenhouse-gas emissions. ***Energy Efficiency***, 4: 145–173.

Morecroft, J. 2007. ***Strategic modelling and business dynamics: A feedback systems approach***. West Sussex: John Wiley & Sons.

Moser, D., Liu, G., Wang, W., & Zhang, J. 2012. Achieving deep energy savings in existing buildings through integrated design. ***ASHRAE Transactions***, 118: 3–10.

Moubray, J. 1997. ***Reliability-centered maintenance***. New York, NY: Industrial Press Inc.

Moxnes, E. 1998. Not only the tragedy of the commons: Misperceptions of bioeconomics. ***Management Science***, 44: 1234–1248.

National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling. 2011. ***Deep water: The gulf oil disaster and the future of offshore drilling.*** Report to the President. http://www.gpo.gov/fdsys/pkg/GPO-OILCOMMISSION/content-detail.html. Accessed February 26, 2016.

Oliva, R., & Sterman, J. 2001. Cutting corners and working overtime: Quality erosion in the service industry. ***Management Science***, 47: 894–914.

Orlikowski, W. 1992. The duality of technology: Rethinking the concept of technology in organizations. ***Organization Science***, 3: 398–427.

Paich, M., & Sterman, J. 1993. Boom, bust, and failures to learn in experimental markets. ***Management Science***, 39: 1439–1458.

Parker, C. 2002. ***The open corporation: Effective self-regulation and democracy***. Cambridge, UK: Cambridge University Press.

Parrish, K., & Regnier, C. 2013. Proposed design process for deep energy savings in commercial building retrofit projects. ***Journal of Architectural Engineering***, 19: 71–80.

Pérez-Lombard, L., Ortiz, J., & Pout, C. 2008. A review on buildings energy consumption information. ***Energy and Building***, 40: 394–398.

Plotkin, N. 2011. ***To-do maintenance: $2.1 billion required. The Tech*** (MIT) online edition, January 19. http://tech.mit.edu/V130/N63/deferredmaintenance.html.

Porter, M., & van der Linde, C. 1995. Toward a new conception of the environment-competitiveness relationship. ***Journal of Economic Perspectives***, 9: 97–118.

Rahmandad, H., & Repenning, N. 2015. Capability erosion dynamics. ***Strategic Management Journal***, 10.1002/smj.2354.

Rahmandad, H. 2012. Impact of growth opportunities and competition on firm-level capability development trade-offs. ***Organization Science***, 23: 138–154.

Repenning, N. 2002. A simulation-based approach to understanding the dynamics of innovation implementation. ***Organization Science***, 13: 109–127.

Repenning, N., & Sterman, J. 2002. Capability traps and self-confirming attribution errors in the dynamics of process improvement. ***Administrative Science Quarterly***, 47: 265–295.

Repenning, N., & Sterman, J. 2001. Nobody ever gets credit for fixing problems that never happened. ***California Management Review***, 43: 64–88.

Roome, N. 1992. Developing environmental management strategies. ***Business Strategy and the Environment***, 1: 11–24.

Rudolph, J., Morrison, J., & Carroll, J. 2009. The dynamics of action-oriented problem solving: Linking interpretation and choice. ***Academy of Management Review***, 34: 733–756.

Sastry, M. 1997. Problems and paradoxes in a model of punctuated organizational change. ***Administrative Science Quarterly***, 42: 237–275.

Short, J., & Toffel, M. 2010. Making self-regulation more than merely symbolic: The critical role of the legal environment. ***Administrative Science Quarterly***, 55: 361–369.

Sterman, J. 2015. Stumbling towards sustainability: Why organizational learning and radical innovation are necessary to build a more sustainable world—but not sufficient. In R . Henderson, M. Tushman, & R. Gulati (Eds.), ***Organizational & strategic change and the challenge of sustainability***, 51–80. New York, NY: Oxford University Press.

Sterman, J. 2000. ***Business dynamics: Systems thinking and modeling for a complex world***. Chicago, IL: Irwin/McGraw-Hill.

Sterman, J. 1989. Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. ***Management Science***, 35: 321–339,.

Sterman, J., Repenning, N., & Kofman, F. 1997. Unanticipated side effects of successful quality programs: Exploring a paradox of organizational improvement. ***Management Science***, 43: 503–521.

Sterman, J., Cooper, P., Henry, W., Potes, A., & Mendoza, T. 2016. Assessing the business base for green buildings: The MIT Sloan School of Management. Working Paper, MIT Sloan School of Management.

Sullivan, G., Pugh, R., Melendez, A., & Hunt, W. 2010. ***Operations & maintenance best practices: A guide to achieving operational efficiency.*** Release 3.0. http://www1.eere.energy.gov/femp/pdfs/omguide_complete.pdf.

Sutherland, R. 1991. Market barriers to energy-efficiency investments. ***Energy Journal (Cambridge, Mass.)***, 12: 15–34.

Syverson, C. 2011. What determines productivity? ***Journal of Economic Literature***, 49: 326–365.

TIAX. 2005. ***Energy impact of commercial building controls and performance diagnostics: Market characterization, energy impact of building faults and energy savings potential.*** http://s3.amazonaws.com/zanran_storage/www.tiaxllc.com/ContentPages/42428345.pdf. Accessed February 26, 2016.

Ton, Z. 2014. ***The good jobs strategy: How the smartest companies invest in employees to lower costs and boost profits***. Boston, MA: New Harvest.

Ton, Z., & Huckman, R. 2008. Managing the impact of employee turnover on performance: The role of process conformance. ***Organization Science***, 19: 56–68.

Toole, C., & Claridge, D. 2011. ***The persistence of retro-commissioning savings in ten university buildings***. http://oaktrust.library.tamu.edu/bitstream/handle/1969.1/128798/ESL-IC-11-10-69.pdf. Accessed February 26, 2016.

U.S. Chemical Safety and Hazard Investigation Board 2009. ***Investigation report: Sugar dust explosion and fire (14 Killed, 36 Injured), imperial sugar company.*** Report 2008-05-I-GA. http://www.csb.gov/assets/1/19/Imperial_Sugar_Report_Final_updated.pdf. Accessed February 20, 2016.

U.S. Department of Labor 2010. ***Report of investigation: Fatal under ground mine explosion,*** April 5, 2010, Upper Big Branch Mine-South, Performance Coal Company. http://arlweb.msha.gov/Fatals/2010/UBB/PerformanceCoalUBB.asp. Accessed February 20, 2016.

Walrave, B., van Oorschot, K., & Romme, A. 2011. Getting trapped in the suppression of exploration: A simulation model. ***Journal of Management Studies***, 48: 1727–1751.

Wennberg, J. 2010. ***Tracking medicine***. New York, NY: Oxford University Press.

Weick, K. 1984. Small wins: Redefining the scale of social problems. ***American Psychologist***, 39: 40–49.

World Green Building Council 2013. ***The business case for green building***. http://www.worldgbc.org/activities/business-case/. Accessed February 20, 2016.

Yates, S., & Aronson, E. 1983. A social psychological perspective on energy conservation in residential buildings. ***American Psychologist***, 38: 435–444.

John Lyneis (john.lyneis@sloan.mit.edu) is a principal consultant at PA Consulting Group based in Cambridge MA, where he applies system dynamics to management and public sector challenges. He has worked with several organizations to model and understand large scale improvement initiatives. He received a PhD from the MIT Sloan School of Management in 2012.

John Sterman (jsterman@mit.edu) is the Jay W. Forrester Professor of Management at the MIT Sloan School of Management, where he directs the MIT System Dynamics Group and is faculty director of the Sloan School's Sustainability Initiative.