# KNOWLEDGE SOURCES AND OPERATIONAL PROBLEMS: LESS NOW, MORE LATER

Luca Berchicci
Rotterdam School of Management (RSM), Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands
lberchicci@rsm.nl

Nilanjana Dutt
Bocconi University, Department of Management and Technology, Via Roentgen 1-2, 20136, Milan, Italy
nilanjana.dutt@unibocconi.it

Will Mitchell
University of Toronto, Rotman School of Management, 105 St. George Street, Toronto, ON M5S 3E6, Canada
William.Mitchell@rotman.utoronto.ca

**ABSTRACT:** Unlike problems requiring *new-to-the-world* solutions that combine knowledge from multiple sources, operational problems can often be solved by repurposing existing knowledge from other contexts into *new-to-the-firm* solutions. Firms that seek new-to-the-firm solutions to operational problems face a cost-benefit tradeoff when deciding how many knowledge sources to use. With less need for knowledge recombination than for new-to-the-world solutions, greater knowledge breadth incurs greater screening and implementation costs without concomitant benefits. We study how U.S. manufacturing facilities, from 1991 to 2005, improve operational performance by reducing their rate of annual output of toxic chemical waste (i.e., improvements to operational effectiveness). Results show that search involving fewer knowledge sources in a given year is associated with greater improvements in operational performance (greater waste reduction). At the same time, though, using multiple knowledge sources over time helps improve operational performance, suggesting that avoiding satiation from a single source and learning across sources play temporal roles in toxic chemical waste reduction. Overall, the results suggest that the greatest improvements in operational performance arise with a focused search for new-to-the-firm solutions within periods, while also exploring multiple sources over time.

**Keywords: Knowledge sources, problem solving, search, solutions, operational performance**

Firms commonly need to solve problems (Felin and Zenger 2014, Katila and Ahuja 2002, Laursen 2012, MacDuffie 1997), often by searching for knowledge needed to create *new-to-the-world* solutions (Scherer 1982, Schumpeter 1942) or, instead, to adopt *new-to-the-firm* solutions used in other contexts (Mortensen and Bloch 2005). A key choice concerning the search for solutions is knowledge breadth—i.e., the number of information sources that firms utilize (Leiponen & Helfat, 2010a: 224), such as knowledge from customers, suppliers, or employees. Research typically emphasizing solutions involving new-to-the-world innovations, such as patents and new products, finds benefits from at least moderate breadth of search, enough to allow firms to recombine knowledge from multiple internal and/or external sources (Golovko and Valentini 2011, Leiponen and Helfat 2010a, Love et al. 2014). By contrast, little work has examined whether and under what conditions greater knowledge breadth is beneficial for new-to-the-firm solutions; because new-to-the-firm solutions apply pre-existing solutions to new contexts and therefore require less knowledge recombination, it is unclear whether the net gains might increase similarly with broader knowledge search. This paper delves into the link between knowledge breadth and operational performance, arguing that organizations seeking new-to-the-firm solutions benefit from narrow knowledge search within any single period and from broader search over time. Empirically, we examine how knowledge breadth is associated with the adoption of new-to-the-firm-solutions that improve operational performance in U.S. manufacturing facilities engaged in toxic chemical waste reduction.

Many problem-solving activities focus on implementing new-to-the-firm solutions by identifying and adopting pre-existing products, processes, and practices (Henderson and Cockburn 1994, McElheran 2015, Mortensen and Bloch 2005); this is sometimes referred to as imitation (e.g., Mansfield, 1961). For instance, to reduce product defect rates, a manufacturing firm may adopt an existing waste treatment filter for its production lines. New-to-the firm solutions are common across multiple settings (Rosenberg 1982), comprising the lion's share of problem-solving activities (Cohen et al. 2000, Reichstein and Salter 2006). To help clarify our lack of understanding of this widespread class of problems involving the search for new-to-the-firm solutions, we focus on the role that knowledge search plays in solving operational problems and, in turn, improving operational performance.

Operational problems are challenges that arise in meeting an organization's objectives for its operating activities, such as the need for lower costs or higher quality (Coelli et al. 2005, Rosenberg 1982). Operational problems may reflect internal goals or arise from external regulations. While some operational problems require new-to-the-world solutions, many involve new-to-the-firm solutions that repurpose existing knowledge to improve operational performance (Stevenson 2017). Operational performance can include operational efficiency in terms of cost per unit produced, and/or operational effectiveness in terms of quality per unit, such as the level of waste produced for a given volume.

To solve operational problems, managers search for potential solutions across one or more knowledge sources, where a greater number of knowledge sources indicates greater knowledge breadth (Katila and Ahuja 2002, Leiponen and Helfat 2010a). In the product defect rate example above, although knowledge about the potential benefits of a filter was pre-existing, it is only by speaking with employees, other firms, and/or their suppliers, that a focal firm is able to gain the knowledge needed to integrate the filter with its ongoing production line, and improve operational performance (in this example, operational effectiveness). Research emphasizing new-to-the-world solutions has highlighted the benefits of recombining complementary knowledge from multiple sources; it remains to be seen, however, whether operational problems that do not need the generation of new-to-the-world knowledge benefit similarly from broader knowledge search and recombination. The few empirical studies about the impact of knowledge on practice adoption offer mixed implications. Christmann's (2000) study of environmental best practice implementation suggests that operational performance may benefit by recombining knowledge involving multiple complementary assets. In contrast, an implication of Pil and Cohen's (2006) arguments concerning modular design is that improved environmental practices often arise from adopting pre-existing activities. These studies highlight our limited understanding of whether and how the breadth of knowledge search influences organizations' ability to solve operational problems.

To reconcile these arguments, we examine the short- and long-term relationship between knowledge breadth and problems related to operational performance, such as waste reduction, that primarily rely on new-to-the-firm solutions. In our framework, organizations have opportunities to solve operational problems by making regular changes to their production systems within relevant time periods. In settings where ongoing improvement is common, such as settings with annual targets, we expect that broader search (i.e., using a larger number of knowledge sources) in any given period will have a negative association with improvements in operational performance due to the use of less familiar sources, search time, and over-commitment to solutions. Instead, the benefits of using multiple knowledge sources are more likely to emerge over multiple periods as organizations exhaust the benefits from any one knowledge source, spread costs, and/or learn about improving operational performance. Thus, in contrast to the search for new-to-the-world solutions—where benefits arise from recombining ideas across multiple knowledge sources—ongoing operational improvements are more likely to arise from imitating existing product, process and practice changes, that can arise from narrow knowledge search within a period. Thus, in such cases the benefits of broader search and subsequent recombination are quickly outweighed by the higher screening and implementation costs associated with a continual search for solutions. Hence, rather than occurring within a single period, benefits tend to offset search costs when exploiting increasing knowledge breadth over multiple periods.

We test these ideas on a sample of U.S. manufacturing establishments drawn from the Environmental Protection Agency's (EPA) Toxics Release Inventory (TRI) database (Berchicci, Dowell, & King, 2017; Doshi, Dowell, & Toffel, 2013; King & Lenox, 2000; Li & Zhou, 2017). Because disposing of toxic chemical waste (hereafter waste) is financially expensive, and draws attention from regulatory and insurance watchdogs as well as local communities, manufacturing facilities strive to reduce it. To aid in waste reduction, facilities search different knowledge sources to identify solutions to improve operational performance annually. They must report all waste reduction activities, including the number of knowledge sources used in identifying solutions, the number of solutions adopted, and waste levels achieved, along with facility and chemical information to the EPA. We examine how knowledge breadth, both in a single period and across multiple periods, is associated with operating effectiveness in the form of annual waste reduction, while assessing the impact of two additional factors that clarify the baseline results: (1) facility and manager experience, and (2) the types of knowledge sources used. We find that the most effective strategy for improving waste reduction rates tends to be using a single knowledge source in a given year while using multiple sources across periods, with some differences based on whether we assess all knowledge sources or combine knowledge sources within four distinct classes.

Our study contributes to theory at the interface of literatures concerning problem solving

(Baer et al. 2013, Foss et al. 2016) and knowledge search (Argyres and Silverman 2004, Grant 1996, Helfat 1994, Katila and Ahuja 2002, Laursen 2012, Macher 2006, Puranam et al. 2006, Rothaermel and Deeds 2004, Valentini 2012). We use the problem-solving lens to complement existing studies of knowledge search, which emphasize search processes for uncovering new-to-the-world solutions, by considering search processes that solve problems that use new-to-the-firm solutions.

We do so by focusing on a widespread and important but understudied set of challenges: operational problems (Reichstein and Salter 2006, Rosenberg 1982). Our findings suggest that new-to-the-firm solutions such as those that are common for operational problems, rather than tending to arise from knowledge recombination across multiple complementary sources, arise most effectively from redeploying knowledge from one source at a time; however, searching more broadly *across multiple time periods* is associated with improved operational performance. Thus, our study reveals a different set of mechanisms in the relationship between knowledge breadth and problem solving. In parallel, the work enriches the literature on organizational problem-solving by linking knowledge search mechanisms with firms' problem-solving activities, thereby answering a call to extend this perspective (Baer et al., 2013; Nickerson, Yen, & Mahoney, 2012). We do so by connecting the implementation of problem-solving activities—through wide or narrow knowledge search—to the efficacy of the outcome. Our results illustrate ways in which the process of seeking solutions varies both based on the nature of the problem and over time.

The empirical context of this research offers its own contributions. We examine an operational problem that many firms face and address on an ongoing basis: waste reduction. We are able to measure how knowledge breadth at the level of each toxic chemical within the facility is being processed while accounting for the volume of the same chemical; hence, we can conduct a refined analysis of how knowledge breadth in any given period and across periods is associated with operational performance improvements within facilities, at the chemical level, over time. Because the final outcome—operational performance—is comparable across facilities, differences in performance by each facility (in our case, differences in waste reduction) indicate superior problem-solving strategies. Overall, our findings address solutions to ongoing operational problems by highlighting the benefits of using fewer knowledge sources at one time while experimenting with a broader set of knowledge sources over multiple periods.

**THEORY & HYPOTHESIS**

A significant body of research highlights the importance of knowledge search in driving improvements in firm performance by identifying and implementing new solutions (Ahuja and Katila 2001, Henderson and Cockburn 1996, Stuart 2000, Tushman and Rosenkopf 1996, Vivero 2002). A critical aspect of firms' success with new solutions is formulating the problems to be solved (Newell and Simon 1972) and then searching for appropriate knowledge that will allow them to devise effective solutions (Felin and Zenger 2014, Foss et al. 2016). Broadly, the knowledge that firms seek falls into two categories: knowledge that will lead to new-to-the-world solutions, often embodied in new products and patents (Arora and Gambardella 1994, Cohen and Klepper 1996, Cohen and Malerba 2001, Rosenberg 1982, Utterback and Abernathy 1975); and knowledge that will provide new-to-the-firm solutions, commonly embodied in operating processes (Dosi 1982, Pisano 1996, Reichstein and Salter 2006, Terjesen and Patel 2015, Utterback and Abernathy 1975) and practices (Bloom et al. 2012, Lawrence 2018). Despite the prevalence and importance of new-to-the-firm solutions, the drivers are less understood than those of new-to-the-world solutions, for two reasons. First, ongoing improvements to operational activities are less likely to be observed because they are not as likely to be embodied in new products or formally encoded in patents (Cohen and Klepper 1996). Second, seemingly minor changes to operational activities, such as adding an employee to the end of the production line to observe waste output, can have a major impact on improving operational performance (Dutt and King 2014), but are less likely to be explicitly recorded as new solutions.

The lack of understanding about new-to-the-firm solutions is particularly relevant to theories about the relationship between knowledge search and problem-solving (Felin and Zenger 2014, Helfat 1994, Katila and Ahuja 2002). For research on knowledge search and innovative activity, as shown by Laursen's (2012) review of thirty papers, we see that fourteen papers examine patent data, eleven examine new product introductions, and five examine R&D expenses as their key search variable—mostly focusing on new-to-the-world knowledge (Afuah and Tucci 2012, Laursen and Salter 2006, Leiponen and Helfat 2010a, Rosenkopf and Nerkar 2001). On the other hand, while the problem-solving literature has clarified how the nature of the problem and the governance structure affect problem solving (Felin & Zenger, 2014; Nickerson & Zenger, 2004), there is limited discussion of the link between solutions and knowledge search.

Our goal is to bridge these streams of knowledge search and problem-solving studies by examining how the breadth of knowledge search influences the efficacy of problem solving in the context of operational problems. We do so by considering costs and benefits that arise when firms make choices about how broadly to search to generate new-to-the-firm solutions. We start with the baseline expectation that in order to solve operational problems firms must either possess the knowledge needed to develop solutions or search one or more knowledge sources until they are able to identify solutions (Kogut & Zander, 1992; Nickerson & Zenger, 2004; Nonaka, 1994). As our primary independent variable, we use the number of knowledge sources as the measure of knowledge breadth (we considered also assessing knowledge depth but faced empirical limits that we discuss in the limitations section).

As part of the means of solving problems, we assume that a firm can identify solutions in three ways: invention, recombination, or imitation. First, it can invent something new. In practice, pure invention is rare (Ahuja and Lampert 2001). Second, a firm can combine multiple existing ideas, potentially with some new ideas, and so develop new solutions in the form of recombination. Recombination, highlighted by Schumpeter (1942), is more common than invention and has inspired a stream of work in strategy and economics (Galunic and Rodan 1998, Sorenson and Fleming 2004). Third, a firm can imitate practices that are already being used in other settings, whether inside or outside the organization (Mansfield 1961). While imitation is less likely to produce differentiation advantages than invention or recombination, such activities can help firms meet challenges that arise in the market and/or in the face of regulatory demands (Scherer 1982). In the remainder of the discussion, we will focus on problem-solving activities in the form of recombination and imitation, highlighting the role of knowledge sources.

Recombination yields benefits by integrating complementary ideas from multiple knowledge sources to create new and improved ideas (Milgrom and Roberts 1995). Although benefits may vary based on multiple factors including the relatedness of the knowledge sources (Tanriverdi and Venkatraman 2005), experience of the innovator (Katila and Ahuja 2002), and access to complementary knowledge (Cassiman and Veugelers 2006), typically, recombination is considered beneficial when using complementary knowledge sources.

In addition to considering the benefits of recombining multiple knowledge sources, it is important to recognize that using additional knowledge sources raises costs that accrue in two ways. First, firms expend resources such as employee hours and expertise in identifying which and how many knowledge sources to use: these are screening costs. Second, firms face costs associated with using knowledge that they have identified: these are implementation costs. Examples of implementation costs include training or hiring new personnel and modifying existing operations to utilize the new products and process (e.g., adding a new filter to a production line in order to implement a new waste management technique).

Screening and implementation costs can affect the performance of a solution both financially and in terms of quality. Greater financial costs reduce the efficiency of a solution. Quality costs, meanwhile, arise when the search process interferes with the ability to identify or adopt a solution, reducing the effectiveness of the solution, as we will discuss in more detail when we address negative margin returns in the hypothesis section. Collectively, the extent of screening and implementation costs will determine whether the benefits that arise from using multiple knowledge sources yield net gains in efficiency and/or effectiveness.

Let us consider how benefits and costs accrue for new-to-the-world and new-to-the-firm solutions. For new-to-the-world solutions, research suggests that viable new knowledge can arise from combining ideas from multiple complementary knowledge streams (Katila and Ahuja 2002, March 1991, Rosenkopf and Nerkar 2001, Stettner and Lavie 2014). While the number of viable combinations between different knowledge sources may vary, benefits of complementarity across knowledge sources have been shown to increase the quantity of viable solutions such as new products and patents. Although screening and implementation costs will tend to increase with greater knowledge breadth, the benefits of using more knowledge sources for new-to-the-world solutions will often at least initially out-weigh the growth in costs. These findings suggest that the greatest net gains for new-to-the-world solutions tend to arise from recombining knowledge from multiple sources (Laursen 2012, Leiponen and Helfat 2010a).

Now consider whether the key assumption that recombination results in the most benefits holds for new-to-the-firm solutions. If new-to-the-firm solutions typically arise by combining a variety of ideas, potentially including newly generated ideas, then using multiple knowledge sources, (as in the case of new-to-the-world solutions) should produce superior outcomes. However, if new-to-the-firm solutions commonly arise from opportunities to repurpose existing knowledge (Mansfield 1961, Scherer 1982, Szulanski 1996) that is already being used in other settings (Reichstein and Salter 2006, Rosenberg 1982, Rouvinen 2002, Utterback and Abernathy 1975), then a firm may not need to create new solutions and may benefit instead from adopting existing solutions. Moreover, even boundedly rational firms will often have reasonable insights about which sources can provide viable solutions; a solution of this type may not be "optimal" in the sense of being the best possible solution but it will prove, at least initially, to provide a satisfactory outcome by improving performance.

In such cases, when firms can ex-ante identify knowledge sources that should yield viable solutions, the costs and benefits of searching additional knowledge sources will diverge. On the one hand, increased search activities mean that screening and implementation costs will increase with increasing knowledge breadth. On the other hand, if broader search would yield better more effective solutions, the costs and benefits should balance. However, because there is no clear reason to expect a systematic increase in the effectiveness of solutions arising from broader search (as the firm does not need to create new knowledge and can adopt existing solution), this relationship between broader search for new-to-the-firm-solutions is likely to yield negative returns. While there undoubtedly are cases in which new-to-the-firm solutions involve substantial recombination, we will proceed with the assumption that they are *more likely* than new-to-the-world solutions to lie toward the end of the spectrum where benefits arise from single-source adoption and/or imitation.

**H1: Knowledge Breadth within a Period and Improvements in Operational Performance**

We apply the logic above to contexts where firms can search for knowledge needed to adopt new-to-the-firm solutions that will improve operational performance. We utilize two assumptions that are consistent with prior work and underlie the mechanisms by which firms engage in knowledge search. First, as we noted above, we assume that firms are boundedly rational and will satisfice when problem solving by using a subset of knowledge sources (Helfat 1994, Jacobides and Winter 2005, Knudsen and Levinthal 2007, Martin and Mitchell 1998) and focusing on viable (not necessarily optimal) solutions as long as they improve operational performance (Sanders and Carpenter 2003, Simon 1982). Second, consistent with the idea of the variety paradox (Dahlander et al. 2016), we assume that there are multiple viable solutions to improving operational performance (Afuah and Tucci 2012, Jeppesen and Lakhani 2010, Knudsen et al. 2014). While a technically focused firm might utilize technical knowledge sources, for instance, an employee-focused firm might utilize employee-based sources; because both firms have different competencies, they may find different paths to improving performance. Building on these two assumptions, we consider the link between knowledge breadth and operational performance.

When firms are choosing how many knowledge sources to use, the logic of local search suggests that they will start by using sources closest to their existing knowledge base (Dosi 1982, Helfat 1994, Martin and Mitchell 1998, Piezunka and Dahlander 2015, Rosenkopf and Nerkar 2001). Thus, the first knowledge source used is likely to be one that the firm knows well and that the firm expects will provide an opportunity for discovering a viable solution. With a single source, screening and implementation costs will tend to be limited. By contrast, when a firm searches more broadly, we expect it to engender diminishing and possibly even negative marginal returns as screening and implementation costs increase.

Costs will tend to increase for three reasons: less-familiar sources, search time, and over-commitment. The first two reasons involve screening costs, while the third reason involves implementation costs. First, additional knowledge sources are likely to include less-familiar sources, which require more extensive screening expenditure. For instance, if a firm has frequently used a government assistance program, it is likely to utilize this source first. If it decides to search more broadly, it might also, for example, seek information from an employee feedback program, which will require higher screening costs.

Second, screening costs will increase because firms that search multiple knowledge sources before selecting a solution will tend to spend more time searching than those that investigate a single source. Helfat and Leiponen (Leiponen and Helfat 2010b) and Gavetti and Levinthal (Gavetti and Levinthal 2000) highlight operational and cognitive challenges that arise in conducting broader search, including the need for greater time. In turn, firms that search multiple sources will have less time within a given period to achieve the productivity gains that arise from their solution. This temporal difference imposes screening costs on firms that use multiple sources, in the form of opportunity costs of time foregone for achieving the benefits of search.

Third, when firms search more broadly, implementation costs can increase if they become committed to using ideas from multiple sources (Kahneman and Lovallo 1993, O'Brien and Folta

2009) and attempt to combine across sources to identify solutions even though imitating a solution from a single source might be sufficient. In part, such attempts create screening costs arising from temporal challenges similar to the point above. In addition, attempts to create a recombined solution will tend to be more complex than single source solutions and may actually produce less-effective solutions than viable single-source solutions. Several scholars (e.g., (Argyres and Silverman 2004, Lieberman and Chacar 2003, Zhao and Islam 2017) have argued that search involving multiple actors (e.g., decentralized search) raises coordination challenges (Chen et al. 2012, Zhou 2013). Such challenges may reduce the effectiveness (e.g., quality) of an unnecessary multi-source solution compared to using a viable repurposed solution from a single source. This is a form of implementation cost that, rather than simply leading to declining marginal returns to knowledge breadth, can lead to negative marginal returns in the effectiveness (quality) of a chosen solution.

The assumption about the availability of multiple viable solutions is important to our argument. If there are few solutions with the potential for repurposing, a broader search process may be necessary to find even a single viable repurposed solution. However, in settings with multiple viable repurposed solutions for ongoing operational problems, narrower search with resources dedicated to implementation within a given period will often be a superior approach.

Moreover, empirical research on the adoption of environmental practices has not identified systematic benefits that would arise from complementarity in knowledge, again reinforcing our arguments. While both Pil and Cohen (2006) and Christmann (2006) highlight the benefits of adopting pre-existing practices to improve environmental performance, only Christmann finds additional benefits from the presence of complementary assets. Similar to Pil and Cohen, Lamin and Ramos (Lamin and Ramos 2016), and Bromiley and Rau (Bromiley and Rau 2016) highlight the benefits of access to knowledge when identifying and implementing best practices, but identify no specific need for knowledge recombination. The limited evidence for recombination benefits is also reflected in studies of implemented management practices, which show substantial variation in the types of practices adopted across different settings and highlight the contingent nature of practice adoption (Bloom et al. 2012) together with the high costs of changing established operational systems (Szulanski 1996).

These studies suggest two key points that are consistent with our arguments concerning screening and implementation costs. First, there is little evidence that imitation of existing solutions benefits from knowledge recombination. If recombination benefits exist, they are either small or do not exist uniformly across firms and, at least, are likely to be lower for new-to-the-firm than for new-to-the-world solutions. Second, research highlights high organizational costs of screening and implementing new processes (Szulanski 1996)(Szulanski 1996), which are exacerbated because firms must typically determine contingencies between the nature of the practice and the context in which they seek to adopt the practice. Together, these two conclusions suggest that screening and implementation costs will rise when firms search multiple knowledge sources for new-to-the-firm solutions that will be implemented in a given period, with no systematic evidence of an equivalent increase in benefits. Thus, for the average firm, greater knowledge breadth will tend to produce lower gains in operational performance compared to firms that consider fewer knowledge sources. Instead, firms will experience the greatest net gains by searching narrowly and imitating appropriate solutions.

These arguments provide the logic for the following hypothesis. We state the prediction in terms of the general concept of operational performance. In our empirical test, we focus on a problem for which solutions involve improving the operating effectiveness aspect of such performance (i.e., achieving higher quality by reducing toxic waste).

***H1****: The greater the knowledge breadth in a given period, the lower the improvements in operational performance.*

**H2: Cumulative Knowledge Breadth and Improvements in Operational Performance**

H1 highlights how and why the costs of using more knowledge sources will outweigh the benefits for solving operational problems *in any one period*. However, we do not expect this negative relationship to hold *across periods.* Instead, as we discuss next, there may be benefits to a broader knowledge search strategy over time.

There are three reasons why, across multiple periods, firms may benefit from increases in knowledge breadth when solving operational problems: knowledge satiation, cost spreading, and learning. First, narrower knowledge search may cease to yield viable solutions after firms exhaust the low-hanging fruit (Christmann 2000), i.e., initial solutions for operational improvement. Firms are likely to start by implementing solutions with relatively low costs; they may look to solutions that lie within their local search space and so have lower screening costs and/or seek those that require fewer interventions in the existing production system and so require lower implementation costs. Over time, these easily available solutions may cease yielding improvements or, at least, offer diminishing returns on improvements in operational performance. To continue improving as the benefits of initial sources reach satiation, firms may need to search more broadly by examining a greater number of knowledge sources and/or branch away from what they know by seeking more distant knowledge sources. In both cases, the cumulative number of knowledge sources will increase: firms will need to seek new knowledge by increasing knowledge breadth over multiple periods.

Beyond knowledge satiation, a second reason for searching more broadly arises if firms are able to spread screening and/or implementation costs over multiple periods. For instance, employees might identify multiple viable solutions in any one period but implement only a subset in the current period. Others could be implemented in a subsequent period, incurring low screening costs in this second period because improvements would come from previously discovered solutions. Alternately, employees working with a production system may discover techniques by which to implement changes to production processes more efficiently that may be reusable in future periods, thus reducing future implementation costs. In both cases, firms incur lower screening and implementation costs because they are drawing from multiple knowledge sources across multiple periods.

The third possibility is that firms learn over time, allowing them to engage in a greater number of problem solving activities, identify better solutions, and/or avoid corresponding increases in costs over multiple periods (Argote and Epple 1990, Eggers 2012, Pisano 1996). Although empirically equivalent to the argument above concerning cost spreading, the mechanism is one of learning. By accumulating experience in solving operational problems, employees will identify the most problem-prone areas in their production systems as well as better ways to implement changes (Rockart and Dutt 2015, Stan and Vermeulen 2012). Such expertise will allow employees to know which knowledge sources are best for solving which types of problems, when they can rely on their own skill sets versus when they should seek external help (King, 1999), what sequence to implement changes in, and how to stop looming problems before they disrupt the production system. Research on environmental management has highlighted how waste prevention is associated with learning that provides a springboard for the biggest portion of commercial benefits derived from process changes (King & Lenox, 2002).

In sum, even though H1 suggests that firms may be most likely to benefit from using fewer knowledge sources in any given period, they may also reap the benefits of multiple knowledge sources across periods. The benefits of a single source are likely to face diminishing marginal returns and, perhaps quickly, reach satiation. In addition, firms may be able to spread costs and/or learn to identify better ways to solve problems, which would allow them to reduce the costs of using more knowledge sources. Hence, in order to gain additional improvements in future periods, firms will often benefit by exploring more broadly over time. This idea underlies the second hypothesis.

***H2**: The greater the cumulative knowledge breadth across multiple periods, the greater the improvements in operational performance.*

## DATA & METHODS

### Data and Sample

The ideal setting to test our hypothesis would be one in which managers solve operational problems by investigating different knowledge sources to identify solutions whose impact on operational performance can be tracked over time. The U.S. manufacturing industry provides such a setting; we have access to detailed data collected by the EPA describing a specific operational problem (the need to reduce waste), along with relevant problem-solving activities such as the knowledge sources used to identify solutions and the resulting waste levels. The EPA's Toxics Release Inventory (TRI) database includes information about U.S. manufacturing establishments' waste management activities concerning toxic chemical waste generation and reduction.

The EPA's TRI differs from most other federal environmental programs. While federal government programs typically aim to achieve better environmental performance by setting standards and specifying how facilities must operate, the TRI is an information disclosure program. Part of a new approach to managing environmental protection, the TRI program makes information about industrial management of toxic chemicals available to the public. Thus, it "creates a strong incentive for companies to improve their environmental performance by sharing information about releases of toxic chemicals in their community" (EPA, 2017). The EPA has published TRI data annually, since 1987, in reports that account for 612 routinely processed chemicals and also include facility-level information about establishment location, size, industry affiliation, and managing personnel. Although TRI data are self-reported, there are two elements that give us confidence in the reliability of these data. First, the EPA can impose a fine of up to $25,000 per violation or misreporting. Second, prior studies have checked the accuracy of reporting in TRI data; De Marchi and Hamilton (de Marchi and Hamilton 2006) found that 95% of facilities reported information accurately.

These data have three characteristics that suit our study. First, they include all U.S. manufacturing facilities that face the ongoing challenge of managing toxic chemical waste (Berchicci, Dowell, & King, 2012; Doshi et al., 2013; Dutt & King, 2014; Li & Zhou, 2017). According to the TRI regulations, establishments that have more than ten employees and that manufacture, process, or use toxic chemicals above a specified quantity must report to the EPA. This means that the database includes about half the population of the U.S. manufacturing industry and more than two-thirds by production outcome. Second, these data capture annual facility-chemical level information for all establishments as long as they produce, process, or use chemical toxic waste and are annually scrutinized by the EPA. Chemical spills and negligence can result in fines and closure, making waste reduction a critical problem. Third, the Pollution Prevention Act, passed in 1990, encourages manufacturing facilities to reduce the amount of waste through cost-effective changes in production, operation, and raw materials use. A significant avenue for waste reduction is "source reduction," which refers to practices that increase efficiency in the use of materials and/or reduce the hazardous substances released into the environment prior to recycling, treatment, or disposal. The term includes equipment or technology modifications, process or procedure modifications, reformulation or redesign of products, substitution of raw materials, and improvements in housekeeping, maintenance, training, or inventory control. Thus, establishments can provide information on whether they reduce waste on site and, if so, how they reduce waste. In particular, they can specify newly adopted solutions to reduce waste and the methods by which solutions were identified for each chemical within the facility; the reported methods are the knowledge sources for this study. This detailed dataset allows us to analyze, for each chemical within a facility, the relationship between the number of knowledge sources and improvements in operational performance.

As one example, assume that a shoe manufacturer attempts to reduce ammonia use. Managers may first seek suggestions from their employees and thereafter implement a new production technique for reducing ammonia use. In such a case, managers would report to the TRI that the facility implemented a solution of type "improved application techniques", using the knowledge source “employee recommendation under a formal company program” for the specific chemical “ammonia”. The number of knowledge sources and the number of operational solutions both would be listed as one. These features make the TRI database suitable for our study.

We combine the TRI data with the National Establishment Time Series (NETS) Database, which contains information on the number of employees and sales, creating a comprehensive dataset describing production and waste activities of the population of U.S. manufacturing facilities operating from 1991 to 2005. Following other researchers, we excluded data from 1987 to 1990 due to a change in the TRI reporting guidelines in 1991 (Doshi et al., 2013; King & Lenox, 2000; King et al., 2005; King & Lenox, 2002). Similarly, we excluded data from 2006 on, when the EPA switched to reporting based on the North American Industry Classification System (NAICS) instead of Standard Industrial Classification (SIC) codes.

In contrast with most previous studies, which typically aggregate TRI data to the firm or facility level, we perform our analysis on the chemical level. This means that we examine changes in waste as the result of knowledge sources and solutions implemented for each chemical within each facility. Similar to Dutt and King (2014), our approach provides important empirical advantages, including a more precise understanding of changes in the waste reduction mechanisms for each chemical over time. By doing so, we are not forced to attach weights to chemicals of different types and from different production lines when aggregating chemical waste. Furthermore, we can directly control for changes in production output, as well as chemical- and location-specific regulatory shocks by incorporating facility-chemical fixed effects. Given a median of four chemicals per facility, our original sample includes 143,962 facility-chemical observations and 772,748 facility-chemical-year observations. Although all the facilities need to report their toxic chemical waste management, they might not necessarily report their source reduction or pollution prevention activities. Consequently, only 18.4% of the original sample engages in and reports formal source reduction activities. Thus, our sample declines to 45,541 facility-chemical cases and 142,833 facility-chemical-year cases. Given the nature of our dependent variable (waste change) and the empirical challenges we face, as we explain in the estimation section, our final sample will be smaller.

**Measures**

**Dependent variable and primary independent variable**

The dependent variable is operational performance, which we measure as a rate of change in waste (i.e., an operational effectiveness aspect of operational performance). Because facilities are engaged in waste reduction, increases in waste indicate lower operational performance while decreases in waste indicate higher operational performance. We measure operational performance $_{cit}$ as the logarithmic difference between the waste produced by chemical *c* in facility *i* in year *t* as compared to year *t-1*. We then multiply this value by 100 to facilitate interpretation of the coefficients. We can interpret the coefficients as percentage change for a unit change of the covariates. The calculation is as follows:

$$Operational\, performance_{\,cit} = ln\left(\frac{waste\ generation_{cit}}{waste\ generation_{cit-1}}\right) * 100$$

This measure of operational performance captures improvements in waste output, which is an outcome of the production process. This approach follows recent studies (e.g., Berchicci et al., 2017) that use the rate of change in waste as a dependent variable. More generally, multiple researchers have used the waste generation variable as well as related transformations to investigate links between green performance, corporate strategy, and operations management, making this approach highly reliable (Berchicci et al. 2012, Doshi et al. 2013, Dutt and King 2014, King and Lenox 2001, Terlaak and King 2006).

The primary independent variable is *knowledge breadth,* which captures the number of knowledge sources a firm looks to. This can potentially range from zero to nine; the observed maximum is eight. To measure this variable, we use the section of the TRI form in which managers specify the methods used to identify possible solutions to reduce waste. Appendix I contains a list of all knowledge sources and the relative frequencies. For instance, managers of firm Alpha may have used “Employee Recommendation (Under a Formal Company Program)” and “Employee Recommendation (Independent of a Formal Company Program)”. In this case knowledge breadth would count two knowledge sources. This is consistent with the measure of knowledge breadth used by Leiponen and Helfat (2010) and Laursen and Salter (2006).

We also operationalized an alternative measure of knowledge breadth, which combines similar knowledge source types. *Knowledge breadth (clustered)* counts only sources that are distinctive as independent knowledge sources. Appendix I depicts the four clusters of knowledge sources from which we build our *knowledge breadth (clustered)* measure: assistance programs (16.2%), audits (28.5%), employee programs (43.3%), and others (12.0%). Using the example of firm Alpha (which used both informal and formal employee recommendation methods with the “employee programs” cluster), *knowledge breadth (clustered)* would be equal to 1 as the two types of knowledge sources are similar in terms of screening and implementation costs.

Using *knowledge breadth (clustered)* allows us to further test our conjectures on the relationship between narrow search and performance. If searching multiple knowledge sources is associated with an increase of screening and implementation costs (with no systematic evidence of an equivalent increase in benefits) in a given period, we would expect a greater effect for searching across clusters of knowledge sources to determine a value for *knowledge breadth (clustered)* for that period. This is because the additional unit increase of this variable would capture greater impact in terms of screening and implementation costs.

Next, to test hypothesis 2, we examined whether there is a benefit to cumulative knowledge breadth over time. In H2 we argue that the benefit of using multiple diverse knowledge sources increases over time due to diminishing marginal returns of a single source and firms’ increased ability to identify better ways to solve problems. The *cumulative knowledge breadth* measure counts the total number of different knowledge sources used in the periods prior to the current year *t.* We again use an alternative measure, named *cumulative knowledge breadth* (*clustered),* which takes into account only the cumulative number of unique clusters over time.

**Control variables**

We control for two chemical-level factors that may influence operational performance. First, *number of operational solutions* assesses the number of operational solutions adopted, measured as the sum of techniques used to reduce waste for a given chemical in a given facility in a given year.[1] Although this can range from zero to 44, facilities in our sample in practice implemented up to 11 solutions to handle, process, transfer, dispose of, and reduce chemical waste. Appendix II shows the frequency of the most common solutions. While there is no dominant solution adopted, improvements in procedures, process modifications, new raw materials, and operating practices are the most common. Second, *production ratio* measures the rate of change in production output in the units directly related to the line where the waste chemical was used or produced within each facility in the year *t* relative to year *t-1;* positive values indicate increased production. In our setting, waste reduction is strongly predicted by production output; managers strive to reduce waste at a faster rate than production output. We transformed the production ratio to its logarithmic form and multiplied by 100 to make it consistent with our dependent variable.

We assess eight facility-level factors. First, *facility experience* captures the number of years of experience of a particular facility-chemical pair based on how long it has been active in waste reduction activities. We measure this by counting the number of years for which a specific facility has reported about a specific chemical to the EPA as part of the TRI program; even within facilities, there is variation in when specific chemicals were selected for waste reduction activities. Second, *technical officer experience* captures the number of years for which the main manager in the year *t-1* has been supervising that facility. Third, *facility size* is measured as the log of the number of employees and as log of sales in a facility. Fourth, *number of chemicals* counts the different chemicals a facility produces in year *t*. Fifth, the *number of inspections* captures the number of regulatory inspections that a facility has experienced in the year *t-1*. Sixth, the *inventory size* variable is the logarithmic amount in pounds of all the chemicals stored at the facility and addresses inventory levels. Next, to control for possible spillovers that go beyond the knowledge that the facility and technical office possess, we

[1] We assume that solutions adopted are highly correlated with solutions searched, but we are limited in that we cannot observe which solutions were searched but not implemented.

created two measures: *disposal spillovers* and *production spillovers*. The former is equal to 1 if multiple chemicals within a facility are disposed of (e.g., recycled, treated, or released by air) in the same way as the focal chemical being considered. The latter is equal to 1 if other chemicals are produced (rather than used) within the facility.

**Estimation**

Our main estimation model uses ordinary least squares (OLS) regressions for panel data with fixed effects (FE) at the facility-chemical level and chemical-year level. Facility-chemical FE controls for fixed facility-production attributes. Chemical-year FE controls for broader changes that could influence chemical-level changes (prices, disposal costs, regulation) and in aggregate reflects broader economic changes. Thus, in the FE regressions, we observe how changes in our primary independent variable are associated with rate of change in our dependent variable for each chemical within each facility while controlling for unobservable time invariant covariates at the chemical and facility level.

Before we examine the results, we need to consider two potential sources of bias: non-random choice of number of knowledge sources and selection bias in terms of the decision to conduct on-site search. We explored two methods to address the potentially non-random choice of the number of knowledge sources. First, we conducted a two-stage least squares (2SLS) analysis to test the main hypothesis. The first stage predicted the number of knowledge sources; the second stage predicted operational performance. We assessed several logical instrumental variables; in all cases, we used panel-specific regression models with chemical-year fixed effects and standard errors clustered by facility-chemical ID. Although the instrumental variables used in the first stage were theoretically sound, they were of poor statistical quality; we were not confident that they reliably improved our models. Second, we used Arellano-Bond estimation, which assumes that good instrumental variables can be found in lagged transformations of the main independent variables. We used several lag structures and found consistent estimators, but all post-estimation tests failed. Thus, we cannot eliminate concerns that there may be some bias due to non-random selection of the number of knowledge sources. Nonetheless, we conducted robustness tests to understand the degree to which this bias may or may not influence the results. Overall, as we discuss in greater depth below, even though we cannot claim causality, we believe these results are reliable.

We address the second challenge, selection bias, by recognizing that not all facility-chemical combinations engage in on-site waste reduction activities. To address possible bias from self-selection, we created a matched sample using coarsened exact matching (CEM) (Iacus et al. 2012). To create the matched sample, we generated a control subsample that has similar characteristics to a matched set of firms prior to their engagement in waste reduction. We then compared differences in changes in waste generated within facility-chemical pairs over time. In addition to the sample used in the analyses, we conducted multiple additional matches, as well as an inverse propensity treatment-weighting match technique that we report in the robustness section.

We created the matched sample by matching facilities that were formally searching knowledge sources to reduce waste when using chemicals with other facilities using the same chemicals that were not formally searching to reduce waste. The strength of any matching process depends on the quality of the covariates chosen for the matching itself. We decided to match facilities on exact values of industry affiliation (using a 2-digit SIC code), chemical identifier, and year. Furthermore, we matched on three continuous variables (transformed into standardized values) to automatically coarsen the data based on a binning algorithm provided by the CEM procedure. We used one variable for facility size (number of employees) and included trends in waste generation as well as trends in production volume. To calculate these trends, we used the relative average values in the two years prior to the treatment. We also used the option to produce a matching result that has the same number of treated and control observations *(k-to-k)*. In doing so, we created a sample where each chemical for which a facility was formally conducting waste reduction was matched with a control chemical with similar waste and production trends used by a facility of a similar size.

We then tested the strength of the matched sample vis-à-vis the whole dataset. Appendix III shows two descriptive statistics tables, the first of which refers to the full sample used for the matching. Given the lagged structure of the covariates, the sample size drops from 722,748 to 609,534 facility-chemical-year observations. As previously reported, only 18.4% of the full sample engaged in the waste reduction program (the mean of the waste reduction sample variable). The second table shows the descriptive statistics of the matched sample with 124,713 facility-chemical-year observations. Since the matching result has the same number of treated and control groups, the average for the waste reduction sample is now close to 50%, as we would expect after the matching.

Table 1 reports the results before and after matching. Model 1 contains a logistic regression testing the whole sample based on 609,534 facility-chemical-year observations. We see that two out of three main matching variable coefficients are significant in predicting whether a facility-chemical observation engages in waste reduction. Model 2, which contains the matched sample, then shows that the three coefficients are not significant, consistent with statistical requirements for matching. This suggests that the key variables created a balanced sample of facility-chemical pairs between those that conduct waste reduction and those that do not. Furthermore, because we analyze data at a fine-grained level of analysis (chemical rather than facility), we are confident that the subsample that served as the base for the rest of the analyses addresses the selection bias to the extent possible.

************ Table 1 about here ************

## RESULTS

Table 2 reports descriptive statistics of the treated sample, i.e., the sample in which we observe waste reduction activities. The final sample contains 61,253 facility-chemical-year observations. On average, facilities tend to have lower waste across years since the average of *operational performance* is -7.6%. The mean of knowledge breadth is 1.78 for the full sample. Although facilities can search up to nine knowledge sources, we found that the vast majority investigate a maximum of four (98% of the cases). This matches our assumptions about facilities being boundedly rational and searching only a subset of all possible knowledge sources. As such, the analysis focuses on facilities that examine up to four knowledge sources. In this case, as reported in Table 1, the mean of *knowledge breadth* is reduced to 1.74. The mean of *knowledge breadth (clustered)* is slightly lower, at 1.57; the correlation between the two knowledge-breadth variables is high (r=0.90), as one would expect.

************ Table 2 about here ************

### Test of H1: Knowledge Sources and Operational Performance

Table 3 contains the test of the hypothesis 1. The analysis uses ordinary least squares regression for panel data with fixed effects at facility-chemical and chemical-year, and standard errors clustered at the facility-chemical level. Models 1 and 2 report control variables only. *Facility experience* has a significant negative influence on waste change, i.e., the longer a facility has been reporting on a particular chemical, the greater the improvement in *operational performance*. One year of *facility experience* improves *operational performance* by 1.49%; this translates to a real waste reduction of about 7,300 pounds. This finding is consistent with existing research that links experience and problem-solving outcomes positively, while indicating that the facilities in the study have not reached diminishing marginal returns. The positive coefficient on *production ratio* suggests that firms producing more products also produce more waste, but the relationship is not 1 to 1; for one percentage increase of production volume there is only 0.83% increase of waste. Similarly, the positive influence of *number of chemicals* on *operational performance* suggests that facilities using more chemicals also produce more waste. The *number of operational solutions* (Model 2) does not significantly influence *operational performance*.

************ Table 3 about here ************

Model 3 introduces *knowledge breadth* to test hypothesis 1, which suggests that as *knowledge breadth* increases in a given period *operational performance* decreases. We find a significant positive influence of *knowledge breadth* on *operational performance* ($\beta$=2.39, p<5%). Hence, this result is consistent with the hypothesis 1.

Model 4 includes *knowledge breadth (clustered)*, which combines similar knowledge source types and counts only the number of knowledge sources that are distinctive. Its effect is larger and stronger ($\beta$=4.47, p<0.001%) than that of *knowledge breadth*. This result suggests that increasing the number of distinctive knowledge sources in a given period is even more detrimental than adding a similar one.

These results are materially significant. A unitary increase in *knowledge breadth* is associated with a 2.39% increase in waste output (recall that the dependent variable is logged). In real terms, this increase translates to almost 12,500 pounds of waste (based on average facility waste output). Even more significantly, a unitary increase in *knowledge breadth (clustered)* is associated with a 4.47% increase in waste output, which means an increase of 23,400 pounds of waste.

In line with learning studies, we consider both linear and quadratic effects of search. Models 4 and 5 include the quadratic terms for *knowledge breadth* and *number of operational solutions*, and for *knowledge breadth (clustered)* and *number of operational solutions*, respectively. In doing so, we test whether there is a curvilinear relationship between search and performance outcomes. No squared term shows a significant relationship with *operational performance*. Thus, the relationships are monotonic, significant only for the main effect of *knowledge breadth*.

The patterns suggest that variation in *operational performance* arises from the initial choice of how many knowledge sources to use rather than the number of solutions that are ultimately adopted. This conclusion strongly supports the hypothesis, highlighting the importance of the number of knowledge sources in driving *operational performance*. Furthermore, the more dissimilar the knowledge sources are, the more their increment is associated with poor operational performance in a given period.

**Test of H2: Cumulative Knowledge Sources and Operational Performance**

Table 4 contains the test for Hypothesis 2. Model 1 shows that there is no significant monotonic effect of *cumulative knowledge breadth*: counter to the prediction, examining more knowledge sources over time does not significantly influence operational performance. However, when we switch to the *knowledge breadth (clustered)* variable in Model 2, we find a negative and significant coefficient ($\beta$= -3.08, $p<0.1\%$). The result in Model 2 suggests that the cumulative effect of knowledge sources on *operational performance* is positive over time, as expected, when accounting for sources with distinct knowledge.

************ Table 4 about here ************

Models 3 and 4 consider both linear and quadratic effects. Model 3 reports a significant negative coefficient on *cumulative knowledge breadth* plus a significant positive squared term. Hence, the relationship between cumulative knowledge breadth and the dependent variable is negative up to a point, after that it become less negative.[2] This result is consistent with H1, even for the full set of knowledge sources, though with the benefits emerging only up to a moderate number of knowledge sources.

To ease the interpretation, we show the relationship graphically. Figures 1a and 1b show the margins of the *cumulative knowledge breadth* and its coefficient intervals, based on Model 3 from

[2] Following Haans et al. (Haans et al. 2016), we confirmed the inverted-U-shaped relationship ($p<0.01$) by using the command "utest" in Stata (Lind and Mehlum 2010).

Table 4. Figure 1a includes the full sample, while Figure 1b shows the same effect while right-censoring the *cumulative knowledge breadth* up to three times the standard deviations (this variable has mean of 8.12, standard deviation of 7.6, and maximum of 82); this limit excludes less than 2% of the full sample. The graphs show that the overall effect of *cumulative knowledge breadth* on the dependent variable is negative; i.e., waste initially declines as cumulative knowledge breadth increases, as Figure 1b highlights. Moreover, the positive effect captured by the increasing portion of the U-shaped curve in Figure 1a is driven by outliers: the increasing values arise for the few cases of knowledge breadth that are more than three standard deviations greater than the mean.

*********** Figure 1a & Figure 1b about here ***********

Model 4 does not show any curvilinear effect on the relationship between Cumulative Knowledge Breadth (clustered) and operational performance. Instead, there is a significant monotonic declining effect. Overall, the effects of both measures (cumulative knowledge breadth and its clustered version) on the dependent variable are similar and consistent, providing support to our H2.

To summarize, these findings suggest that there are benefits to using varied and distinctive knowledge sources over time, countering knowledge depletion from a single source or type of source and taking advantage of learning opportunities. When using complementary knowledge sources (as captured by our *knowledge breadth* variable), the accumulation of knowledge sources over time may limit the ability to reduce waste, possibly owing to diminishing returns on experimentation or to forgetting; in our case, though, outliers drive the inflection point to values of cumulative knowledge sources three times greater than its standard deviation. The key empirical result is that relying on greater knowledge sources over time is associated with an improvement in *operational performance*.

These results in Table 4 concerning cumulative knowledge sources identify an important extension to the within-period findings (Table 3). Within a period, drawing from more than one knowledge source is detrimental to improving operational performance. Over time, though, accumulating experience with different knowledge sources offers learning benefits and helps reduce waste. We believe this is an important insight: the most effective strategy for improving operational performance tends to be using one knowledge source in a given year, while exploring multiple sources over time.

**Robustness Tests**

We conducted four robustness tests. First, we estimated the main analyses with an alternate measure of number of operational solutions. Instead of counting the *number of operational solutions*, we counted the *number of different process activity types*, where activities are groups of related solutions; this variable ranged from 0 to 40 (instead of 0 to 44); the two measures of solutions are highly correlated. We found consistent results.

Second, we compared both random and fixed effects models and found consistent results. The reported analyses used restrictive fixed-effects models at the facility-chemical-year level with standard errors clustered at the facility-chemical level. The Hausman test suggested that either model was appropriate; to control for time-invariant factors, we report fixed-effects models.

Third, we explored CEM models with varying numbers of covariates to test whether our results were strongly dependent on the number of covariates. Increasing the number of covariates makes the matching more stringent, thereby reducing the number of matches and potentially reducing the reliability of the smaller sample. Although the number of matched observations changed substantially, these tests persistently showed the support for the main hypothesis. For instance, in one test, we included an additional facility size variable (facility sales). The matched sample decreased from 124,713 to 101,674; the results were robust.

Fourth, we conducted a complementary matching and regression adjustment using inverse probability of treatment weights (IPTW). IPTW typically improves upon propensity score matches by better accounting for extreme cases where the propensity score is almost 0 or 1. As discussed by Azoulay, Ding, and Stuart (Azoulay et al. 2009), the procedure better accommodates treatments that occur at different points in time. Additionally, as demonstrated by multiple scholars of epidemiology (Austin 2011, Funk et al. 2011, Robins et al. 2000), an advantage of IPTW is that it does not require both the treatment model and outcome model to be fully specified; the technique produces reliable estimates as long as either treatment model or the regression model is correctly specified. For instance, in cases where the covariates for matching might be specified incorrectly, IPTW can reduce bias in the coefficient estimates. Appendix IV reports models that first created matches using IPTW and then ran the regression adjustment specification. The results, using a more comprehensive set of covariates, are consistent with the CEM analysis.

**Additional Tests: Experience and Idiosyncratic Sources**

We conducted two sets of additional tests to explore alternative explanations to our hypotheses. First, we assessed three possible experience-based explanations for our results, which could be shaped by a possible experience-based selection effect. Research on efficiency improvements in manufacturing settings show that the biggest gains arise as organizations accumulate experience that allows them to reduce mistakes and increase productivity (Pisano 1996; Danneels 2002). As facilities gain experience in using knowledge sources, they might search fewer sources and reduce more waste. If so, greater experience rather than fewer annual knowledge sources would drive operational performance. At the same time, variation in the nature of experience can influence organizational learning in different ways (Hoang and Rothaermel 2010, Rockart and Dutt 2015, Salvato 2009). We examined two measures of experience based on how long a facility-chemical or manager-chemical pair has been active in waste reduction activities.

Figure 2a and 2b depict the facility experience results. Figure 2a for facility experience shows that, counter to the selection argument, the number of knowledge sources is slightly *more* likely to increase as facilities gain experience (one knowledge source is used 65% of the time in the first year, 52% in the second year, and around 45% in the following years). Figure 2b, for technical officer experience, shows that the number of knowledge sources remains largely constant as technical officers gain experience. Hence, these results in conjunction with the experience controls in the main models assuage concerns that selection based on facility or manager experience might drive the effect of knowledge sources on operational performance.

*********** Figure 2a & Figure 2b about here **********

Another possible experience-based explanation for the relationship between greater *knowledge breadth* and lower *operational performance* in a given period is that facilities might use up the benefits of initial choices of knowledge sources (i.e., use up the gains from low-hanging fruit) and then turn to broader search involving more knowledge sources to solve more complex problems. To help address the satiation possibility, we first checked whether the benefits of waste reduction occur only in the first year of operation. We reran our models in Table 3, excluding the first reporting year. The direction and significance of our main independent variable did not change, reinforcing support for the hypothesis.

Second, we tested whether knowledge source idiosyncrasies exist within our sample. In particular, we examined whether a given knowledge source, patterns of knowledge source types, or a particular sequence of knowledge sources drive our results. To assess this, we ran three analyses. Since many firms use a single source (48%), the first analysis examined the effect of every type of knowledge source on operational performance. The results of the regressions show that no specific knowledge source is superior at reducing waste; rather, for those using just one knowledge source, all sources are comparable.

The second analysis tested whether there are patterns in knowledge sources across periods across firms. If there were, they could indicate optimal knowledge source temporal combinations. Appendix I shows the frequency of use of different knowledge sources, while Appendix V depicts the trends across time. The most common is "participative team management", becoming somewhat more common over time as it grew from about 28% to 35% of the firms, while informal employee recommendations fell somewhat from about 21% to 15%. Nonetheless, there is no strong pattern of convergence or divergence over time. Hence, there is no evidence that firms are learning about a few optimal knowledge sources and converging on their use.

Third, we tested whether firms use similar sequences of knowledge sources across time. If common sequences of knowledge sources exist, it could be that firms tend to use a set of knowledge sources as best practice. We found that firms use unique sequences for 31% of all cases. We also found that only 15% of the total sample represents more than 500 observations per sequence (out of 35,999 observations), which depicts the same knowledge source type across 2, 3, or 4 years in a row. For example, there are 382 cases in which "participative team management" was used for four continuous years. These descriptive results suggest that there is no significant trend or popular sequence. Frequently, managers stick to the same type of knowledge source for a large number of years before experimenting with another knowledge source type.

**Assessing Assumptions**

We made three assumptions about the empirical setting of this study that we cannot fully test but, as we discuss here, we do not believe that these assumptions create extensive biases. First, we only observed operational solutions that were implemented. It is possible that in some cases a firm selects a group of solutions but does not adopt all solutions or adopts solutions randomly. However, because firms are searching, selecting, and adopting operational solutions and reporting waste activities annually, we expect selected and adopted solutions to be highly correlated. While it is highly unlikely that facilities would adopt solutions randomly, such randomness would create noise that should weaken but not systematically bias the results.

Second, a related concern stems from the fact that we observed the TRI forms annually. Because of the regulated nature of the industry, we expect facilities to search and then implement solutions, and document their activities in the sequence specified in the conceptual framework—that is, first investigate different knowledge sources and then adopt solutions—but it is difficult to verify the sequence. Nonetheless, even if the search and adoption processes were concurrent, the results show that the number of knowledge sources, rather than the number of operational solutions, has the greatest influence on operational performance.

Third, we are not able to exclude the possibility that an unobserved, time-varying factor is driving both changes in number of knowledge sources and operational performance. In sensitivity analyses, we conducted 2SLS and Arellano Bond models that predicted the number of knowledge sources in the first stage and found results that support the focal hypothesis. Nonetheless, because our instruments had limits, we did not focus on these results, even though these findings were consistent with our main results. Overall, we find stable and consistent results across multiple econometric specifications, using multiple models, and at an unusually granular level of data. We consider these findings to be reliable.

## DISCUSSION AND CONCLUSIONS

Research in the knowledge search and problem-solving literatures has identified the importance of knowledge breadth in driving search processes and outcomes (Felin and Zenger 2014, Katila and Ahuja 2002, Laursen 2012, Love et al. 2014, Reichstein and Salter 2006). This body of research has made great strides in clarifying the fundamental links between knowledge and problem solving (Felin & Zenger 2014; Macher, 2006; Rothaermel & Deeds, 2004). While the vast majority of empirical studies focus on the generation of *new-to-the-world* solutions (e.g., (Laursen and Salter 2006, Leiponen and Helfat 2010a, Schumpeter 1942, Sorenson and Fleming 2004), the more common type of problem-solving activity tends be the adoption of pre-existing knowledge into new contexts via *new-to-the-firm* solutions (Cohen et al. 2000, Rosenberg 1982).

Studies looking at the adoption of new-to-the-firm solutions, including when focusing on a context of environmental significance, find mixed benefits for knowledge recombination (Christmann 2000, Pil and Cohen 2006). To clarify the mechanisms and conditions under which knowledge breadth can be beneficial for identifying and implementing new-to-the-firm solutions, we examine how knowledge breadth influences performance outcomes in the context of firms solving operational problems. Because we can measure and distinguish between different knowledge sources used, solutions implemented, and performance improvement, our research is able to link antecedent search choices about knowledge breadth with relevant performance.

Our baseline results highlight two findings that are relevant for this literature as well as for managers who are seeking new-to-the-firm solutions. First, we find that firms tend to benefit from implementing a focused problem-solving strategy in a given period. Using one knowledge source in any one period—independent of the type of knowledge source or the point in the firm's learning trajectory—is associated with the greatest improvement in operational performance. We interpret the relationship to be a result of two factors: first, unlike the development of the novel ideas that are needed for generating new-to-the-world solutions, there is less systematic benefit to knowledge recombination for identifying new-to-the-firm solutions. Second, the costs of screening knowledge and implementing solutions in established production systems are high (Bloom et al. 2012, Christmann 2000, Laursen and Salter 2006), creating challenges that can undermine the efficiency and quality of new-to-the-firm solutions. In combination, the findings suggest that the lack of knowledge recombination benefits and increasing costs of screening and implementing solutions derived from increasing knowledge breadth result in decreasing net gains on knowledge breadth for new-to-the-firm solutions.

Second, because firms will tend to reach diminishing returns with any one knowledge source and learning tends to be a dynamic process, we considered the benefits of knowledge breadth over multiple periods. We find that the use of greater cumulative knowledge breadth is associated with significant improvements in operational performance. These results are relevant for managers who are organizing for problem solving—firms can often unlock the benefits of knowledge breadth by accessing a greater variety of knowledge over multiple periods, while focusing on fewer knowledge sources in any given period.

Three additional analyses assess mechanisms and assumptions. First, analysis that combines similar sources shows that costs increase as firms increase search across more different knowledge sources, providing support for the assumed mechanism. Second, we find firms become more likely to use more knowledge sources as they gain experience, reducing concerns that experience-based selection might drive the use of less knowledge breadth; this pattern is consistent with the assumption of bounded rationality. Third, consistent with the assumption of the variety paradox whereby different firms may reach similar levels of performance by pursuing different strategies, we find that there does not seem to be a single optimal source (e.g., employees or vendors) or optimal pattern of sources; instead, firms use a wide variety of knowledge sources with few repeated sequences.

Our results extend prior work that has highlighted the benefits of knowledge breadth. Our extension is to identify a class of problems and solutions for which knowledge breadth is not beneficial within a relevant decision period but instead provides net gains across multiple periods. In doing so, we build upon the limits to knowledge breadth identified in prior work (Leiponen and Helfat 2010b), clarifying the mechanisms by which knowledge recombination can be detrimental to operational performance. Because ongoing enhancements in operational activities can have ripple effects on improving the overall cost structure of the firm and its productivity, the benefits of a focused problem-solving strategy are relevant for both scholars and practitioners. Furthermore, a substantial proportion of organizational activity occurs in the form of process improvements within firms, which are often overlooked—in part because of measurement difficulties and in part because of their ubiquitous nature (Cohen et al. 2000, Cohen and Klepper 1996, McElheran 2015, Reichstein and Salter 2006, Rosenberg 1982); the results are highly relevant to such cases.

The empirical context is relevant in its own right. We examine ongoing operational performance in a manufacturing setting at the facility-chemical level—essentially inputs and outputs on the production line—such that the measures and concepts are closely linked. Furthermore, we can separate the influence of knowledge breadth on both number of solutions adopted and subsequent operational performance—at the fine-grained level of each chemical within each facility, rather than needing to rely on more general facility or firm-level measures—which illustrates the full set of problem-solving activities that firms undertake. We retest ideas proposed in prior research such as the direct effect of prior organizational experience, for which we find support, as being especially relevant for operational performance improvements. Because our study examines fine-grained data in manufacturing settings across many different industries, we believe our fine-grained analysis provides robust and reliable results.

The study has three limits that provide opportunities for future research. First, we assume a problem-solving model in which employees are making decisions about which knowledge sources to use in every period where they must report waste. It is possible that different problem-solving models are at work, although it is unclear how this may change our findings. For instance, some decisions might be made centrally rather than at the level of the chemical or some facilities may not prioritize waste reduction but simply go through the motions of implementing new changes. It is possible that not accounting for these different decision models is adding noise to our estimates. Second, we are unable to identify causality in our results. Although we implemented CEM to identify the most representative sample of facility-chemical pairs vis-à-vis the general population, this approach did not allow us to address the fact that firms choose how many knowledge sources to use: an unobserved factor may drive both their choice of the number of knowledge sources and operational performance improvements. However, even if we cannot unambiguously determine causality, the robust and material effects across multiple specifications allow us to interpret these findings as reliable and relevant correlations. Third, a natural corollary to knowledge breadth is knowledge depth. We are unable to measure this concept because the most relevant measure of depth in our data correlated highly with knowledge breadth, but we would expect increasing knowledge depth to contribute to operational performance improvements. Prior work (Ahuja and Katila 2001, Katila and Ahuja 2002) suggests that firms using knowledge more deeply will tend to uncover new avenues for exploiting improvement which, at least up to some point, is likely to be beneficial.

Overall, we extend research on knowledge sources and problem-solving to an important class of problems and solutions—operational problems and new-to-the-firm solutions. We show how and why the mechanisms of costs and benefits of knowledge breadth differ for such cases. The results highlight important implications for strategy and organizational theory scholars, as well as for managers who are allocating resources to solve operational problems.

**REFERENCES**

Afuah A, Tucci CL (2012) Crowdsourcing As a Solution to Distant Search. *Acad. Manage. Rev.* 37(3):355–375.

Ahuja G, Katila R (2001) Technological acquisitions and the innovation performance of acquiring firms: a longitudinal study. *Strateg. Manag. J.* 22:197–220.

Ahuja G, Lampert CM (2001) Entrepreneurship in the large corporation: a longitudinal study of how established firms create breakthrough inventions. *Strateg. Manag. J.* 22:521–543.

Argote L, Epple D (1990) Learning Curves in Manufacturing. *Science* 247:920–924.

Argyres NS, Silverman BS (2004) R&D, organization structure, and the development of corporate technological knowledge. *Strateg. Manag. J.* 25(8–9):929–958.

Arora A, Gambardella A (1994) Evaluating Technological Information and Utilizing It: Scientific Knowledge, Technological Capability, and External Linkages in Biotechnology. *J. Econ. Behav. Organ.* 24:91–114.

Austin PC (2011) An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivar. Behav. Res.* 46(3):399–424.

Azoulay P, Ding W, Stuart T (2009) The Impact of Academic Patenting on the Rate, Quality and Direction of (public) Research Output. *J. Ind. Econ.* 57(4):637–676.

Baer M, Dirks KT, Nickerson JA (2013) Microfoundations of strategic problem formulation. *Strateg. Manag. J.* 34(2):197–214.

Berchicci L, Dowell G, King AA (2012) Environmental capabilities and corporate strategy: exploring acquisitions among US manufacturing firms. *Strateg. Manag. J.* 33(9):1053–1071.

Berchicci L, Dowell G, King AA (2017) Environmental Performance and the Market for Corporate Assets. *Strateg. Manag. J.* 38(12):2444–2464.

Bloom N, Genakos C, Sadun R, Reenen JV (2012) Management Practices Across Firms and Countries. *Acad. Manag. Perspect.* 26(1):12–33.

Bromiley P, Rau D (2016) Operations management and the resource based view: Another view. *J. Oper. Manag.* 41:95–106.

Cassiman B, Veugelers R (2006) In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition. *Manag. Sci.* 52(1):68–82.

Chen PL, Williams C, Agarwal (2012) Growing pains: Pre-entry experience and the challenge of transition to incumbency. *Strateg. Manag. J.* 33(3):252–276.

Christmann P (2000) Effects of "Best Practices" of Environmental Management on Cost Advantage: The Role of Complementary Assets. *Acad. Manage. J.* 43(4):663–680.

Coelli TJ, Rao DSP, O'Donnell CJ, Battese GE (2005) *An Introduction to Efficiency and Productivity Analysis* (Springer Science & Business Media).

Cohen WM, Klepper S (1996) Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D. *Rev. Econ. Stat.* 78(2):232–243.

Cohen WM, Malerba F (2001) Is the Tendency to Variation a Chief Cause of Progress? *Ind. Corp. Change* 10(3):587–608.

Cohen WM, Nelson RR, Walsh JP (2000) *Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)* (National Bureau of Economic Research).

Dahlander L, O'Mahony S, Gann DM (2016) One foot in, one foot out: how does individuals' external search breadth affect innovation outcomes? *Strateg. Manag. J.* 37(2):280–302.

Doshi AR, Dowell GWS, Toffel MW (2013) How firms respond to mandatory information disclosure. *Strateg. Manag. J.* 34(10):1209–1231.

Dosi G (1982) Technological Paradigms and Technological Trajectories: A Suggested Interpretation of the Determinants and Directions of Technical Change. *Res. Policy* 11:147–162.

Dutt N, King AA (2014) The Judgment of Garbage: End-of-Pipe Treatment and Waste Reduction. *Manag. Sci.* 60(7):1812–1828.

Eggers JP (2012) All experience is not created equal: learning, adapting, and focusing in product portfolio management. *Strateg. Manag. J.* 33(3):315–335.

Felin T, Zenger TR (2014) Closed or open innovation? Problem solving and the governance choice. *Res. Policy* 43(5):914–925.

Foss N j, Frederiksen L, Rullani F (2016) Problem-formulation and problem-solving in self-organized communities: How modes of communication shape project behaviors in the free open-source software community. *Strateg. Manag. J.* 37(13):2589–2610.

Funk MJ, Westreich D, Wiesen C, Stürmer T, Brookhart MA, Davidian M (2011) Doubly Robust Estimation of Causal Effects. *Am. J. Epidemiol.* 173(7):761–767.

Galunic DC, Rodan S (1998) Resource Recombinations in the Firm: Knowledge Structures and the Potential for Schumpeterian Innovation. *Strateg. Manag. J.* 19(12):1193–1201.

Gavetti G, Levinthal D (2000) Looking forward and looking backward: Cognitive and experiential search. *Adm. Sci. Q.* 45:113–137.

Golovko E, Valentini G (2011) Exploring the complementarity between innovation and export for SMEs' growth. *J. Int. Bus. Stud.* 42(3):362–380.

Grant RM (1996) Toward a knowledge-based theory of the firm. *Strateg. Manag. J.* 17(S2):109–122.

Haans RFJ, Pieters C, He ZL (2016) Thinking about U: Theorizing and testing U- and inverted U-shaped relationships in strategy research. *Strateg. Manag. J.* 37(7):1177–1195.

Helfat CE (1994) Evolutionary Trajectories in Petroleum Firm R&D. *Manag. Sci.* 40:1720–1747.

Henderson R, Cockburn I (1994) Measuring Competence? Exploring Firm Effects in Pharmaceutical Research. *Strateg. Manag. J.* 15(S1):63–84.

Henderson R, Cockburn I (1996) Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery. *RAND J. Econ.* 27:32–59.

Hoang H, Rothaermel FT (2010) Leveraging internal and external experience: exploration, exploitation, and R&D project performance. *Strateg. Manag. J.* 31(7):734–758.

Iacus SM, King G, Porro G (2012) Causal Inference without Balance Checking: Coarsened Exact Matching. *Polit. Anal.* 20:1–24.

Jacobides MG, Winter SG (2005) The co-evolution of capabilities and transaction costs: explaining the institutional structure of production. *Strateg. Manag. J.* 26:395–413.

Jeppesen LB, Lakhani KR (2010) Marginality and Problem-Solving Effectiveness in Broadcast Search. *Organ. Sci.* 21(5):1016–1033.

Kahneman D, Lovallo D (1993) Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking. *Manag. Sci.* 39(1):17–31.

Katila R, Ahuja G (2002) Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction. *Acad. Manage. J.* 45:1183–1194.

King A (1999) Retrieving and transferring embodied data: Implications for the management of interdependence within organizations. *Manag. Sci.*:918–935.

King A, Lenox M (2002) Exploring the Locus of Profitable Pollution Reduction. *Manag. Sci.* 48:289–299.

King AA, Lenox MJ (2000) Industry Self-Regulation without Sanctions: The Chemical Industry's Responsible Care Program. *Acad. Manage. J.* 43:698–716.

King AA, Lenox MJ (2001) Does It Really Pay to Be Green? An Empirical Study of Firm Environmental and Financial Performance: An Empirical Study of Firm Environmental and Financial Performance. *J. Ind. Ecol.* 5:105–116.

King AA, Lenox MJ, Terlaak A (2005) The Strategic Use of Decentralized Institutions: Exploring Certification with the ISO 14001 Management Standard. *Acad. Manage. J.* 48:1091–1106.

Knudsen T, Levinthal DA (2007) Two Faces of Search: Alternative Generation and Alternative Evaluation. *Organ. Sci.* 18:39–54.

Knudsen T, Levinthal DA, Winter SG (2014) Hidden but in plain sight: The role of scale adjustment in industry dynamics. *Strateg. Manag. J.* 35(11):1569–1584.

Kogut B, Zander U (1992) Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organ. Sci.* 3(3):383–397.

Lamin A, Ramos MA (2016) R&D investment dynamics in agglomerations under weak appropriability regimes: Evidence from Indian R&D labs. *Strateg. Manag. J.* 37(3):604–621.

Laursen K (2012) Keep searching and you'll find: what do we know about variety creation through firms' search activities for innovation? *Ind. Corp. Change* 21(5):1181–1220.

Laursen K, Salter A (2006) Open for innovation: the role of openness in explaining innovation performance among U.K. manufacturing firms. *Strateg. Manag. J.* 27(2):131–150.

Lawrence M (2018) Taking Stock of the Ability to Change: The Effect of Prior Experience. *Organ. Sci.*
Leiponen A, Helfat CE (2010a) Innovation objectives, knowledge sources, and the benefits of breadth. *Strateg. Manag. J.* 31:224–236.
Leiponen A, Helfat CE (2010b) Location, Decentralization, and Knowledge Sources for Innovation. *Organ. Sci.* 22(3):641–658.
Levitt B, March JG (1988) Organizational Learning. *Annu. Rev. Sociol.* 14:319–340.
Li X, Zhou YM (2017) Offshoring Pollution while Offshoring Production? *Strateg. Manag. J.*:n/a-n/a.
Lieberman MB, Chacar AS (2003) Organizing for technological innovation in the u.s. pharmaceutical industry. *Geogr. Strategy*. Advances in Strategic Management. (Emerald Group Publishing Limited), 317–340.
Lind JT, Mehlum H (2010) With or Without U? The Appropriate Test for a U-Shaped Relationship*. *Oxf. Bull. Econ. Stat.* 72(1):109–118.
Love JH, Roper S, Vahter P (2014) Learning from openness: The dynamics of breadth in external innovation linkages. *Strateg. Manag. J.* 35(11):1703–1716.
MacDuffie JP (1997) The Road to "Root Cause": Shop-Floor Problem-Solving at Three Auto Assembly Plants. *Manag. Sci.* 43:479–502.
Macher JT (2006) Technological Development and the Boundaries of the Firm: A Knowledge-Based Examination in Semiconductor Manufacturing. *Manag. Sci.* 52(6):826–843.
Mansfield E (1961) Technical Change and the Rate of Imitation. *Econometrica* 29(4):741–766.
March JG (1991) Exploration and Exploitation in Organizational Learning. *Organ. Sci.* 2:71–87.
de Marchi S, Hamilton J (2006) Assessing the Accuracy of Self-Reported Data: an Evaluation of the Toxics Release Inventory. *J. Risk Uncertain.* 32:57–76.
Martin X, Mitchell W (1998) The influence of local search and performance heuristics on new design introduction in a new product market. *Res. Policy* 26:753–771.
McElheran K (2015) Do Market Leaders Lead in Business Process Innovation? The Case(s) of E-business Adoption. *Manag. Sci.* 61(6):1197–1216.
Milgrom P, Roberts J (1995) Complementarities and fit strategy, structure, and organizational change in manufacturing. *J. Account. Econ.* 19(2):179–208.
Mortensen PS, Bloch WC (2005) *Oslo Manual - Guidelines for Collecting and Interpreting Innovation Data, 3rd edition* Third. (Organisation for Economic Cooporation and Development, OECD).
Newell A, Simon HA (1972) *Human Problem Solving*, (CARNEGIE-MELLON UNIV PITTSBURGH PA).
Nickerson J, Yen CJ, Mahoney JT (2012) Exploring the Problem-Finding and Problem-Solving Approach for Designing Organizations. *Acad. Manag. Perspect.* 26(1):52–72.
Nickerson JA, Zenger TR (2004) A Knowledge-Based Theory of the Firm—The Problem-Solving Perspective. *Organ. Sci.* 15(6):617–632.
Nonaka I (1994) A Dynamic Theory of Organizational Knowledge Creation. *Organ. Sci.* 5(1):14–37.
O'Brien J, Folta T (2009) Sunk costs, uncertainty and market exit: A real options perspective. *Ind. Corp. Change* 18(5):807–833.
Piezunka H, Dahlander L (2015) Distant Search, Narrow Attention: How Crowding Alters Organizations' Filtering of Suggestions in Crowdsourcing. *Acad. Manage. J.* 58(3):856–880.
Pil FK, Cohen SK (2006) Modularity: Implications for Imitation, Innovation, and Sustained Advantage. *Acad. Manage. Rev.* 31(4):995–1011.
Pisano GP (1996) Learning-before-doing in the development of new process technology. *Res. Policy* 25(7):1097–1119.
Puranam P, Singh H, Zollo M (2006) Organizing for Innovation: Managing the Coordination-Autonomy Dilemma in Technology Acquisitions. *Acad. Manage. J.* 49(2):263–280.
Reichstein T, Salter A (2006) Investigating the sources of process innovation among UK manufacturing firms. *Ind. Corp. Change* 15(4):653–682.
Robins JM, Hernán MÁ, Brumback B (2000) Marginal Structural Models and Causal Inference in Epidemiology. *Epidemiology* 11(5):550–560.
Rockart SF, Dutt N (2015) The rate and potential of capability development trajectories. *Strateg. Manag. J.* 36(1):53–75.

Rosenberg N (1982) *Inside the Black Box: Technology and Economics* (Cambridge University Press).
Rosenkopf L, Nerkar A (2001) Beyond Local Search: Boundary-Spanning, Exploration, and Impact in the Optical Disk Industry. *Strateg. Manag. J.* 22:287–306.
Rothaermel FT, Deeds DL (2004) Exploration and exploitation alliances in biotechnology: a system of new product development. *Strateg. Manag. J.* 25(3):201–221.
Rouvinen P (2002) R&d-Productivity Dynamics: Causality, Lags, and "Dry Holes." *J. Appl. Econ.* 5(1):123.
Salvato C (2009) Capabilities Unveiled: The Role of Ordinary Activities in the Evolution of Product Development Processes. *Organ. Sci.* 20(2):384–409.
Sanders WG, Carpenter MA (2003) Strategic Satisficing? A Behavioral-Agency Theory Perspective on Stock Repurchase Program Announcements. *Acad. Manage. J.* 46(2):160–178.
Scherer FM (1982) Inter-Industry Technology Flows and Productivity Growth. *Rev. Econ. Stat.* 64(4):627–634.
Schumpeter JA (1942) *Capitalism, Socialism, and Democracy: Third Edition* (Harper Perennial Modern Classics).
Simon HA (1982) Models of Bounded Rationality: Behavioral Economics and Business Organization.
Sorenson O, Fleming L (2004) Science and the diffusion of knowledge. *Res. Policy* 33(10):1615–1634.
Stan M, Vermeulen F (2012) Selection at the Gate: Difficult Cases, Spillovers, and Organizational Learning. *Organ. Sci.* 24(3):796–812.
Stettner U, Lavie D (2014) Ambidexterity under scrutiny: Exploration and exploitation via internal organization, alliances, and acquisitions. *Strateg. Manag. J.* 35(13):1903–1929.
Stevenson WJ (2017) *Operations Management* 12 edition. (McGraw-Hill Higher Education).
Stuart TE (2000) Interorganizational alliances and the performance of firms: a study of growth and innovation rates in a high-technology industry. *Strateg. Manag. J.* 21(8):791–811.
Szulanski G (1996) Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strateg. Manag. J.* 17(S2):27–43.
Tanriverdi H, Venkatraman N (2005) Knowledge relatedness and the performance of multibusiness firms. *Strateg. Manag. J.* 26(2):97–119.
Terjesen S, Patel PC (2015) In Search of Process Innovations The Role of Search Depth, Search Breadth, and the Industry Environment. *J. Manag.*
Terlaak A, King AA (2006) The effect of certification with the ISO 9000 Quality Management Standard: A signaling approach. *J. Econ. Behav. Organ.* 60:579–602.
Tushman ML, Rosenkopf L (1996) Executive Succession, Strategic Reorientation and Performance Growth: A Longitudinal Study in the U.S. Cement Industry. *Manag. Sci.* 42(7):939–953.
Utterback JM, Abernathy WJ (1975) A dynamic model of process and product innovation. *Omega* 3(6):639–656.
Valentini G (2012) Measuring the effect of M&A on patenting quantity and quality. *Strateg. Manag. J.* 33(3):336–346.
Vivero RL (2002) The impact of process innovations on firm's productivity growth: the case of Spain. *Appl. Econ.* 34(8):1007–1016.
Zhao M, Islam M (2017) Cross-Regional R&D Collaboration and Local Knowledge Spillover. *Geogr. Locat. Strategy*. Advances in Strategic Management. (Emerald Publishing Limited), 343–385.
Zhou YM (2013) Designing for Complexity: Using Divisions and Hierarchy to Manage Complex Tasks. *Organ. Sci.* 24(2):339–355.

**Figure 1a: The Relationship Between Cumulative Knowledge Breadth and Operational Performance (based on Table 4, Model 3)**

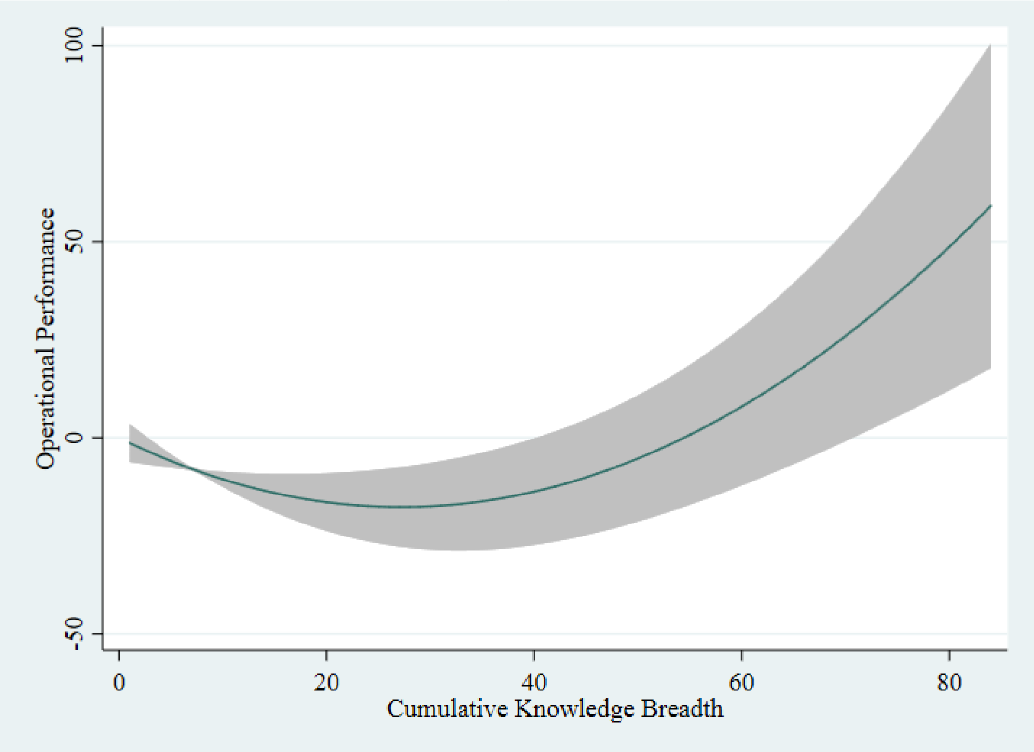


**Figure 1b: The Relationship Between Cumulative Knowledge Breadth and Operational Performance – Sample Limited to Three Times the Standard Deviation of the Cumulative Knowledge Breadth Variable (based on Table 4, Model 3)**

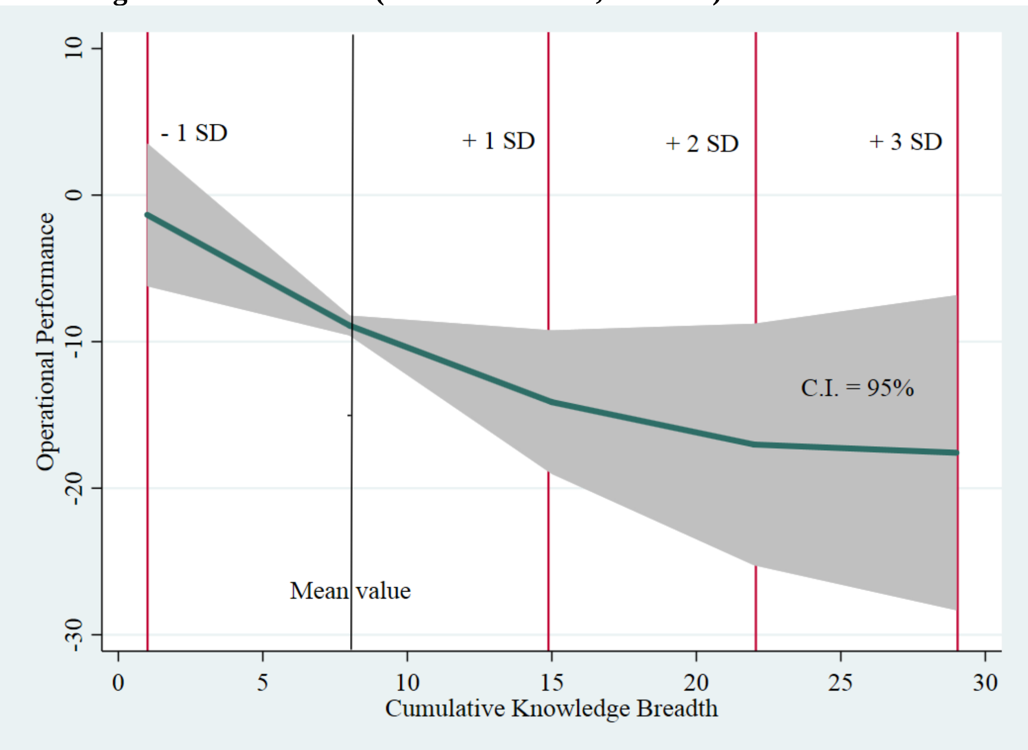

**Figure 2a: Facility Experience by No. of Knowledge Sources (KS)**

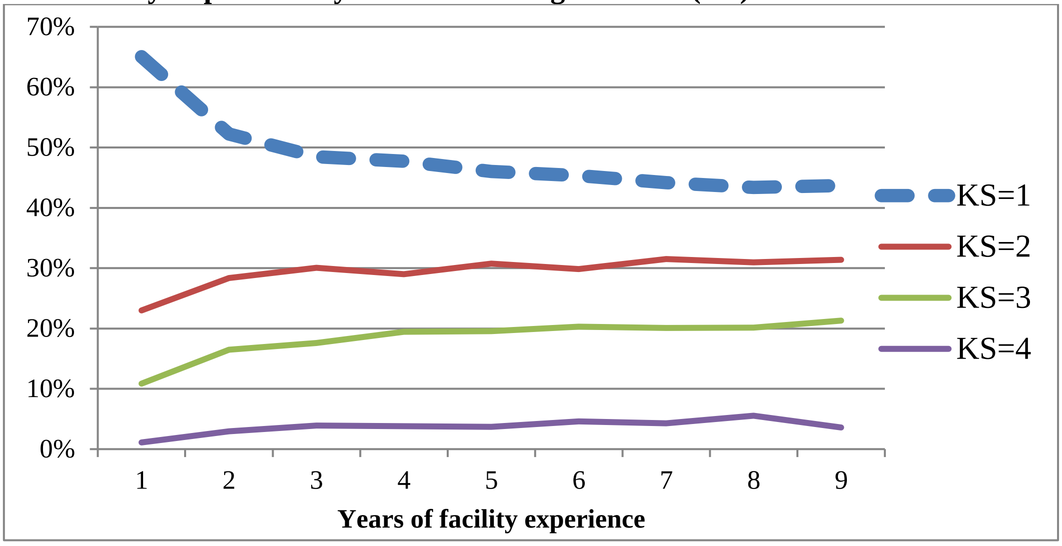


**Figure 2b: Technical Officer Experience by No. of Knowledge Sources**

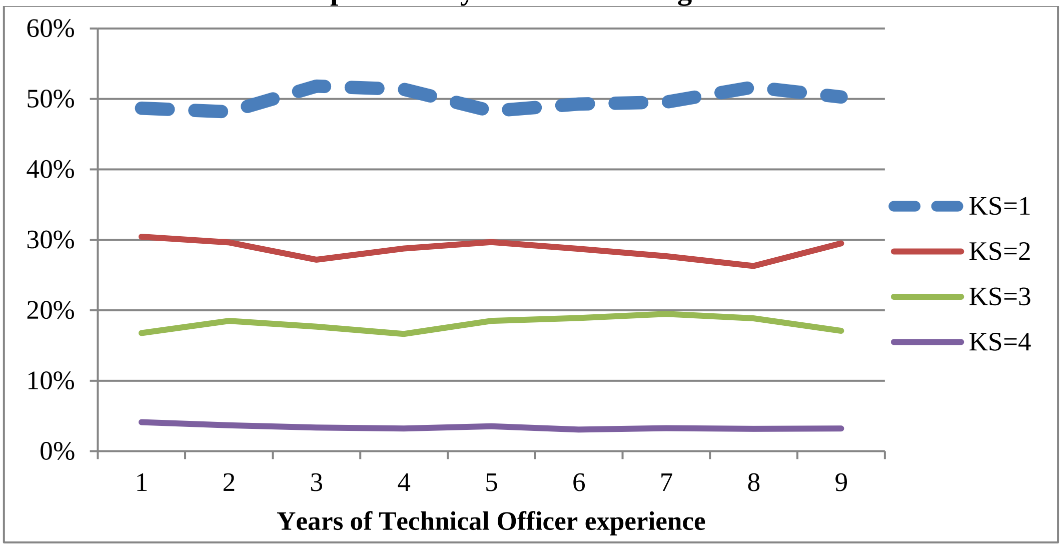


**Table 1: Matched Sample Test**

| DV: Waste Reduction Sample | (1) Full sample | (2) Matched sample |
|---|---|---|
| Plant size (employees) | -0.0437** | -0.00462 |
| | (0.0196) | (0.0269) |
| Trend in waste generation | 0.0866*** | 0.0104 |
| | (0.0157) | (0.0217) |
| Trend in production volume | 0.00513 | 0.0317 |
| | (0.00945) | (0.0269) |
| Within the same industry affiliation (2-digit SIC code) | | |
| For the same chemical | | Exact matching |
| For the same year | | |
| Constant | -1.337*** | 0.00135 |
| | (0.0590) | (0.0682) |
| Observations | 609,534 | 124,713 |
| Pseudo $R^2$ | 0.003 | 0.000 |

*** p<0.01, ** p<0.05 (robust standard errors in parentheses)

**Table 2: Main Sample Descriptive Statistics**

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Operational Performance | 1 | | | | | | | | | | | | | | | |
| 2 | Knowledge Breadth | 0.03 | 1 | | | | | | | | | | | | | | |
| 3 | Knowledge Breadth (clustered) | 0.03 | 0.90 | 1 | | | | | | | | | | | | | |
| 4 | Cumulative Knowledge Breadth | 0.02 | 0.45 | 0.41 | 1 | | | | | | | | | | | | |
| 5 | Cumulative Knowledge Breadth (clustered) | 0.00 | 0.11 | 0.11 | 0.52 | 1 | | | | | | | | | | | |
| 6 | Number of Process Solutions | 0.03 | 0.45 | 0.41 | 0.25 | 0.10 | 1 | | | | | | | | | | |
| 7 | Facility Experience | 0.01 | 0.10 | 0.10 | 0.83 | 0.53 | 0.11 | 1 | | | | | | | | | |
| 8 | Log Employees | -0.02 | -0.02 | -0.01 | -0.03 | 0.02 | -0.09 | -0.03 | 1 | | | | | | | | |
| 9 | Number of Chemicals | 0.02 | -0.04 | -0.02 | 0.05 | 0.04 | 0.03 | 0.09 | 0.29 | 1 | | | | | | | |
| 10 | Production Ratio | 0.16 | 0.02 | 0.02 | -0.02 | -0.01 | 0.02 | -0.04 | 0.00 | 0.00 | 1 | | | | | | |
| 11 | Technical Officer Experience | -0.02 | -0.03 | -0.02 | 0.16 | 0.07 | -0.03 | 0.22 | -0.01 | 0.02 | -0.03 | 1 | | | | | |
| 12 | Number of Inspections | 0.01 | -0.04 | -0.03 | 0.01 | 0.00 | -0.02 | 0.05 | 0.17 | 0.33 | 0.00 | 0.04 | 1 | | | | |
| 13 | Inventory Size | 0.02 | 0.01 | 0.03 | 0.06 | 0.04 | 0.03 | 0.07 | 0.00 | 0.24 | 0.01 | -0.04 | 0.13 | 1 | | | |
| 14 | Disposal Spillovers | 0.00 | 0.02 | 0.02 | 0.03 | 0.01 | 0.04 | 0.03 | -0.02 | 0.00 | 0.01 | 0.01 | -0.02 | -0.03 | 1 | | |
| 15 | Production Spillovers | 0.00 | 0.07 | 0.05 | 0.05 | 0.03 | 0.08 | 0.02 | -0.17 | -0.32 | -0.01 | -0.03 | -0.15 | -0.16 | 0.07 | 1 | |
| 16 | External knowledge source (KS) ratio | -0.01 | 0.14 | -0.01 | 0.02 | -0.01 | 0.00 | -0.04 | -0.01 | -0.13 | 0.00 | 0.02 | -0.06 | -0.16 | 0.02 | 0.08 | 1 |
| | Mean | -7.63 | 1.74 | 1.57 | 8.12 | 1.86 | 1.70 | 4.33 | 5.03 | 7.66 | 70.54 | 3.52 | 1.44 | 10.42 | 0.46 | 0.77 | 0.15 |
| | Std. Dev. | 101.15 | 0.87 | 0.71 | 7.60 | 1.18 | 0.96 | 3.01 | 1.31 | 8.40 | 20.65 | 3.49 | 3.74 | 1.78 | 0.50 | 0.42 | 0.29 |
| | Min | -1358.61 | 1 | 1 | 1 | 1 | 1 | 1 | 0.69 | 1 | 0 | 0 | 0 | 3.93 | 0 | 0 | 0 |
| | Max | 1358.61 | 4 | 4 | 82 | 11 | 9 | 15 | 10.31 | 91 | 904.79 | 18 | 181 | 19.70 | 1 | 1 | 1 |

**Table 3: Knowledge Breadth & Operational Performance (H1)**

| DV: Operational performance (negative coefficient = less waste) | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| Knowledge Breadth [H1 +] | | | 2.387* | | 4.145 | |
| | | | (1.036) | | (4.037) | |
| Knowledge Breadth Squared | | | | | -0.394 | |
| | | | | | (0.851) | |
| Knowledge Breadth (clustered) [H1 +] | | | | 4.471*** | | 3.890 |
| | | | | (1.213) | | (5.262) |
| Knowledge Breadth (clustered) Sq. | | | | | | 0.164 |
| | | | | | | (1.282) |
| No. Of Process Solutions | | 0.101 | -0.755 | -1.096 | -1.988 | -2.446 |
| | | (0.802) | (0.850) | (0.831) | (2.156) | (2.176) |
| No. Of Process Solutions Sq. | | | | | 0.257 | 0.282 |
| | | | | | (0.385) | (0.387) |
| Facility Experience | -1.491*** | -1.489*** | -1.448*** | -1.434*** | -1.448*** | -1.435*** |
| | (0.246) | (0.247) | (0.247) | (0.246) | (0.247) | (0.246) |
| Log Employees | -2.504+ | -2.505+ | -2.465+ | -2.423+ | -2.481+ | -2.430+ |
| | (1.446) | (1.446) | (1.447) | (1.447) | (1.448) | (1.448) |
| No. Chemicals | 0.873* | 0.873* | 0.887* | 0.899* | 0.890* | 0.900* |
| | (0.367) | (0.367) | (0.367) | (0.367) | (0.367) | (0.367) |
| Production Ratio | 0.832*** | 0.832*** | 0.832*** | 0.831*** | 0.831*** | 0.831*** |
| | (0.0495) | (0.0495) | (0.0495) | (0.0495) | (0.0495) | (0.0495) |
| Technical Officer Experience | -0.289 | -0.288 | -0.283 | -0.279 | -0.281 | -0.279 |
| | (0.184) | (0.183) | (0.183) | (0.184) | (0.183) | (0.183) |
| No. Inspections | 0.300 | 0.300 | 0.295 | 0.292 | 0.294 | 0.291 |
| | (0.223) | (0.223) | (0.223) | (0.223) | (0.223) | (0.223) |
| Inventory Size | 0.734 | 0.733 | 0.750 | 0.786 | 0.754 | 0.789 |
| | (0.772) | (0.772) | (0.772) | (0.773) | (0.772) | (0.773) |
| Disposal Spillovers | -2.930+ | -2.931+ | -2.893+ | -2.921+ | -2.887+ | -2.919+ |
| | (1.689) | (1.689) | (1.688) | (1.688) | (1.688) | (1.687) |
| Production Spillovers | 2.884 | 2.876 | 2.834 | 2.742 | 2.809 | 2.747 |
| | (2.533) | (2.538) | (2.536) | (2.538) | (2.537) | (2.538) |
| Constant | -57.77*** | -57.94*** | -61.15*** | -63.98*** | -61.55*** | -62.30*** |
| | (11.94) | (12.06) | (12.21) | (12.20) | (12.92) | (13.12) |
| Chemical-Year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Facility-Chemical Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 61,253 | 61,253 | 61,253 | 61,253 | 61,253 | 61,253 |
| R-squared | 0.024 | 0.024 | 0.024 | 0.025 | 0.024 | 0.025 |
| Number of facility-chemical IDs | 23,142 | 23,142 | 23,142 | 23,142 | 23,142 | 23,142 |

Results for CEM Sample; *** $p<0.001$, ** $p<0.01$, * $p<0.05$, + $p<0.10$ (robust standard errors in parentheses)

**Table 4: Cumulative Knowledge Breadth & Operational Performance (H2)**

| DV: Operational performance (negative coefficient = less waste) | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| Cumulative Knowledge Breadth [H2 -] | 0.0214 | | -1.294** | |
| | (0.198) | | (0.448) | |
| Cumulative Knowledge Breadth Squared | | | 0.0238*** | |
| | | | (0.00661) | |
| Cumulative Knowledge Breadth (clustered) [H2 -] | | -3.084*** | | -5.084** |
| | | (0.805) | | (1.930) |
| Cumulative Knowledge Breadth (clustered) Sq. | | | | 0.294 |
| | | | | (0.233) |
| No. Of Process Solutions | 0.0995 | -0.0577 | 0.447 | 0.00514 |
| | (0.803) | (0.804) | (0.810) | (0.802) |
| Facility Experience | -1.531** | -0.839** | -0.358 | -0.787** |
| | (0.472) | (0.286) | (0.594) | (0.290) |
| Log Employees | -2.507+ | -2.380+ | -2.531+ | -2.412+ |
| | (1.446) | (1.446) | (1.445) | (1.449) |
| No. Chemicals | 0.873* | 0.871* | 0.839* | 0.869* |
| | (0.368) | (0.367) | (0.367) | (0.367) |
| Production Ratio | 0.832*** | 0.830*** | 0.830*** | 0.829*** |
| | (0.0495) | (0.0495) | (0.0494) | (0.0495) |
| Technical Officer Experience | -0.288 | -0.287 | -0.288 | -0.289 |
| | (0.184) | (0.184) | (0.184) | (0.183) |
| No. Inspections | 0.301 | 0.280 | 0.303 | 0.279 |
| | (0.223) | (0.222) | (0.222) | (0.222) |
| Inventory Size | 0.734 | 0.782 | 0.762 | 0.760 |
| | (0.772) | (0.772) | (0.772) | (0.774) |
| Disposal Spillovers | -2.930+ | -2.992+ | -2.935+ | -2.987+ |
| | (1.689) | (1.691) | (1.688) | (1.692) |
| Production Spillovers | 2.881 | 3.015 | 2.861 | 3.013 |
| | (2.537) | (2.539) | (2.537) | (2.540) |
| Constant | -57.93*** | -55.84*** | -56.79*** | -53.61*** |
| | (12.07) | (12.06) | (12.08) | (12.38) |
| Chemical-Year Fixed Effects | Yes | Yes | Yes | Yes |
| Facility-Chemical Fixed Effects | Yes | Yes | Yes | Yes |
| Observations | 61,253 | 61,253 | 61,253 | 61,253 |
| R-squared | 0.024 | 0.025 | 0.025 | 0.025 |
| Number of facility-chemical IDs | 23,142 | 23,142 | 23,142 | 23,142 |

Results for CEM Sample; *** p<0.001, ** p<0.01, * p<0.05, + p<0.10 (robust standard errors in parentheses)

**Appendix I: Knowledge Sources–Methods Used to Identify Solutions Needed to Improve Operational Performance**

| Sources of knowledge for reducing waste | Frequency | Cluster |
|---|---|---|
| *Cluster A. Assistance programs (16.2%)* | | |
| Federal Government Technical Assistance Program | 0.5% | A |
| State Government Technical Assistance Program | 0.6% | A |
| Trade Association/Industry Technical Assistance Program | 2.5% | A |
| Vendor Assistance | 12.6% | A |
| *Cluster B. Audits (28.5%)* | | |
| Materials Balance Audits | 7.0% | B |
| External Pollution Prevention Opportunity Audit(s) | 2.3% | B |
| Internal Pollution Prevention Opportunity Audit(s) | 19.2% | B |
| *Cluster C. Employee programs (43.3%)* | | |
| Employee Recommendation (Under a Formal Company Program) | 4.6% | C |
| Employee Recommendation (Independent of a Formal Company Program) | 7.7% | C |
| Participative Team Management | 31.0% | C |
| *Cluster D. Others* | | |
| Other sources (not specified) | 12.0% | D |

**Appendix II: Most Common Solutions Adopted** (frequency ≥ 2%; total specified = 78%)

| | **Solutions adopted** | **Adoption frequency** |
|---|---|---|
| 1 | Improved maintenance scheduling, recordkeeping, or procedures | 14% |
| 2 | Other changes in operating practices | 12% |
| 3 | Other process modifications | 7% |
| 4 | Substituted raw materials | 7% |
| 5 | Modified equipment, layout, or piping | 7% |
| 6 | Implemented inspection or monitoring program of potential spill or leak sources | 5% |
| 7 | Changed production schedule to minimize equipment and feedstock changeovers | 5% |
| 8 | Other spill or leak prevention | 4% |
| 9 | Improved procedures for loading, unloading, and transfer operations | 3% |
| 10 | Instituted recirculation within a process | 3% |
| 11 | Other changes in inventory control | 3% |
| 12 | Modified design or composition of product | 2% |
| 13 | Substituted coating materials used | 2% |
| 14 | Instituted procedures to ensure that materials do not stay in inventory beyond | 2% |
| 15 | Other raw material modifications | 2% |

**Appendix III: CEM Descriptive Statistics**

**A. Full Sample (609,534 observations)**

| | | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 1 | Waste reduction sample | 1 | | | |
| 2 | Plant size (employees) | -0.012 | 1 | | |
| 3 | Trend in waste generation | 0.036 | -0.061 | 1 | |
| 4 | Trend in production volume | 0.009 | -0.021 | 0.002 | 1 |
| | Mean | 0.180 | 0.104 | 0.027 | 0.012 |
| | Standard deviation | 0.384 | 0.986 | 0.992 | 0.929 |
| | Minimum | 0 | -2.705 | -4.537 | -2.608 |
| | Maximum | 1 | 3.435 | 8.730 | 41.586 |

**B. Matched Sample (124,713 observations)**

| | | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 1 | Waste reduction sample | 1 | | | |
| 2 | Plant size (employees) | -0.0014 | 1 | | |
| 3 | Trend in waste generation | 0.0041 | 0.0214 | 1 | |
| 4 | Trend in production volume | 0.0058 | -0.0180 | 0.0218 | 1 |
| | Mean | 0.5041 | 0.057 | 0.073 | -0.030 |
| | Standard deviation | 0.5000 | 0.820 | 0.799 | 0.401 |
| | Minimum | 0 | -2.705 | -3.195 | -2.607 |
| | Maximum | 1 | 3.434 | 5.937 | 14.831 |

**Appendix IV: Analysis Via Inverse Probability of Treatment Weights (IPTW)**

| DV: Operational performance (negative coefficient = less waste) | Model 1a ATE | Model 1b Pop. Mean | Model 2a TM: KS=2 | Model 2b TM: KS=3 | Model 2c TM: KS=4 |
|---|---|---|---|---|---|
| Lagged 3year waste change | | | 0.0003** | 0.0005*** | 0.000004 |
| | | | (0.0001) | (0.0002) | (0.0002) |
| Log of employees | | | -0.0615*** | -0.003 | -0.081*** |
| | | | (0.0097) | (0.011) | (0.022) |
| SIC (2 Digit) | | | -0.018*** | -0.007*** | -0.003 |
| | | | (0.002) | (0.002) | (0.004) |
| Production Ratio | | | 0.0007* | 0.001** | 0.001 |
| | | | (0.0004) | (0.0004) | (0.0008) |
| Technical Officer Experience | | | -0.008*** | -0.0004 | -0.009 |
| | | | (0.002) | (0.003) | (0.006) |
| Number of Chemicals | | | -0.008*** | -0.016*** | -0.009*** |
| | | | (0.0009) | (0.001) | (0.002) |
| Inventory Size | | | 0.019*** | 0.037*** | 0.049*** |
| | | | (0.005) | (0.006) | (0.012) |
| Regulatory Intensity | | | -0.009 | 0.055*** | 0.122*** |
| | | | (0.010) | (0.011) | (0.023) |
| Total violation | | | 0.008 | 0.007 | -0.104*** |
| | | | (0.006) | (0.007) | (0.019) |
| Facility experience | | | 0.044*** | 0.0008 | 0.027* |
| | | | (0.006) | (0.006) | (0.015) |
| Environmentally Sensitive Industry | | | -0.133*** | -0.251*** | -0.106** |
| | | | (0.020) | (0.024) | (0.045) |
| 2 vs 1 knowledge source | 1.603** | | | | |
| | (0.814) | | | | |
| 3 vs 1 knowledge source | 4.09*** | | | | |
| | (0.987) | | | | |
| 4 vs 1 knowledge source | 5.677*** | | | | |
| | (1.695) | | | | |
| 1 knowledge source v. all others | | -8.169*** | | | |
| | | (0.484) | | | |
| Constant | | | -0.854*** | -1.676*** | -4.135*** |
| | | | (0.119) | (0.132) | (0.306) |
| Year Dummies | YES | YES | YES | YES | YES |
| Observations | 80,986 | 80,986 | 80,986 | 80,986 | 80,986 |

*** p<0.01, ** p<0.05, * p<0.10 (robust standard errors in parentheses); Note: ATE = Average Treatment Effect; TM = Treatment Models

Model 1a includes each treatment level. Similar to the main results, facility-chemical pairs using one knowledge source reduce more waste than those using two, three, or four sources. These results suggest that relative to facility-chemical pairs that use one knowledge source, those using two produce 1.6% more waste, those using three produce 4.1% more, and those using four produce 5.7% more waste. The order of magnitude is similar to that of the CEM-based results in Table 3**.** Model 1b then reports the average population effect of facilities using one knowledge source versus other levels of treatment as well as non-searching firms. The magnitude of 8.2% is a substantial reduction in waste for the facilities using a focused search strategy. Models 2a through 2c break down the analysis for each level of treatment versus facility-chemical pairs that use just one knowledge source. The results highlight the effect of the 11 different covariates and year dummies. We see similar trends for each level of treatment, with somewhat weaker results for the subset of facilities-chemical pairs that use four knowledge sources (Model 2c). These results, using a more comprehensive set of covariates, are consistent with the CEM analysis.

**Appendix V: Frequency of Knowledge Source Type by Year**