Effect of scenario planning on field experts' judgment of long-range investment decisions

Effect of Scenario Planning on Field Experts' Judgment of Long-range Investment Decisions

Shardul Phadnis
Postdoctoral Associate, Center for Transportation & Logistics
Massachusetts Institute of Technology
77 Massachusetts Ave E40-286, Cambridge, MA 02139
shardul@mit.edu
Phone: 617.258.8786; Fax: 617.253.4560

Chris Caplice
Executive Director, Center for Transportation & Logistics
Massachusetts Institute of Technology
77 Massachusetts Ave E40-275, Cambridge, MA 02139
caplice@mit.edu
Phone: 617.258.7975; Fax: 617.253.4560

Yossi Sheffi
Elisha Gray II Professor of Engineering Systems
Professor of Civil and Environmental Engineering
Massachusetts Institute of Technology
77 Massachusetts Ave E40-261A, Cambridge, MA 02139
sheffi@mit.edu
Phone: 617.253.5316; Fax: 617.253.4560

Mahender Singh
Rector
Malaysia Institute for Supply Chain Innovation
2A, Persiaran Tebar Layar, Seksyen U8, Bukit Jelutong, Shah Alam, 40150 Selangor, Malaysia
msingh@misi.edu.my
Phone: +603 7842.2177

Running head: Effect of Scenario Planning

(1)

Effect of Scenario Planning on Field Experts' Judgment of Long-range Investment Decisions

We present the results of three field experiments demonstrating the effect of scenario planning on field experts' judgment of several long-range investment decisions. Our results show, contrary to the past findings, that the use of multiple scenarios does not cause an aggregate increase or decrease in experts' confidence in their judgment. Rather, expert judgment changes in accordance with how an investment fares in a given scenario: it becomes more favorable if the investment is found to be useful for a particular scenario used by the expert, and vice versa. This scenario-induced change is moderated by the expert's confidence in his/her judgment before using the scenario. Finally, our results show that field experts prefer more flexible options to make specific long-range investments after using multiple scenarios.

INTRODUCTION

Scenario planning has been used for half a century by businesses and government agencies for strategic and long-range planning (Grant, 2003; Bradfield, *et al.*, 2005) and is viewed as a dynamic capability that can help today's firms remain adaptable in turbulent environments (Eisenhardt, 1999; Teece, 2007; Hodgkinson and Healey, 2011). Despite its use and promotion as a strategy process, scenario planning has undergone little empirical examination of its ability to influence planners' judgment when making ill-defined strategic decisions under uncertainty. The 'subjective and heuristic nature' of scenario planning is thought to leave 'many academics uncomfortable' (Schoemaker, 2004) deterring them from subjecting it to scholarly scrutiny.

In this paper, we elaborate how scenarios influence expert judgment based on three field studies in which scenario planning was used to evaluate several real-life long-range investments in the U.S. transportation infrastructure. Our results show that the use of multiple scenarios does

not categorically increase or decrease experts' confidence in judgment. Instead, experts update their judgment after using scenarios, either in favor of or against a specific investment, based on how that investment fares in the scenario used. The efficacy of scenarios to induce such changes is moderated by the expert's prior confidence in judgment. Finally, we find that experts show a greater preference for more flexible investment strategies after practicing scenario planning.

Scenario planning: Brief overview

Scenario planning is a decision-making process used for strategic and long-range planning. In this method, multiple views of future are used to envision and prepare for different environments that long-range plans could encounter. The practice of scenario planning originated in the 1950s when government agencies in the U.S. and France started using it for making policy decisions whose effects would last for decades (Bradfield, *et al*. 2005). The first reported use of scenario planning by a corporation is at Royal Dutch Shell ('Shell' hereafter) in late 1960s. The success of Shell in responding to a sudden rise in oil prices during the 1973 Oil Crisis brought scenario planning on corporate radar (Wack, 1985). By 1981, 38 percent of the Fortune-1000 companies reported using it, with greater use by firms with longer planning horizons (Linneman and Klein, 1983). Today, scenario planning is used by many public and private organizations for long-range planning (Cousens, *et al*., 2002; Royal Dutch Shell, 2005, 2008; Deutsche Post AG, 2012; etc.).

Scenario planning is often promoted as a cognitive aid to overcome limitations of human judgment in long-range planning (Wack, 1985; de Geus, 1988). However, evidence of the effect of scenario planning on managerial cognition is almost non-existent. Our extensive literature review unearthed only three experimental studies—all conducted with student subjects—that examined whether scenario planning influenced managerial judgment: two of them (Schoemaker, 1993; Kuhn and Sniezek, 1996) tested the effect of scenarios on subjects' confidence and reached

contrary conclusions; the third showed that scenario use reduced framing bias and improved self-reported decision quality (Meissner and Wulf, 2013). A field study correlated the use of scenario planning with faster growth, higher return on capital, and higher profitability (Phelps, Chan, and Kapsalis, 2001); however, the design of this study does not rule out the reverse causality that stronger firms may be more likely to practice scenario planning. Overall, none of these studies definitively answer whether the use of scenarios affects managerial judgment in the ill-defined long-range planning problems faced in real-world situations.

Scholars have acknowledged this gap in the literature. Paul Schoemaker, author of the three most cited peer-reviewed articles on scenario planning (Varum and Melo, 2010), calls the evidence of usefulness of scenario planning 'anecdotal' (2004: 288). Others admonish the lack of rigorous tests of scenario planning (Harries, 2003) and complain that even field studies of scenario exercises lack 'reliable accounts that render explicitly what has worked and what has not' (Wilkinson, 2009). Even the recent surveys of the scenario planning literature (Bradfield, *et al* 2005; Varum and Melo, 2010) conspicuously lack any mention of empirical assessments. At least three factors contribute to this gap in the literature. Scenario planning practices are highly personalized and hence difficult to compare (Wilkinson, 2009). Strategy practices used by firms are often not publicized, ruling out the use of panel analysis to test if the practice of scenario planning leads to superior firm performance. Scenario planning can also not be judged by the outcome of decisions made using the process, as numerous factors outside researchers' control influence the outcome and make it difficult to attribute the outcome to the decision process alone. One may test scenario planning by its effect on user judgment. The choice of test subjects is critical: since the method is used for making strategic decisions by senior executives, tests with student subjects may limit the results' external validity (Gordon, Slade, and Schmitt, 1987).

In the presence of these methodological challenges, it is unlikely that any one study will answer whether scenario planning works. A research program is necessary to examine the process using multiple research methods. In this paper, we make one attempt in this direction. We test if (and how) the use of one or more scenarios influences the judgment applied by experts to some real-life long-range investment decisions. We do not profess to evaluate the long-range performance of firms using scenario planning; our question is much more basic: does the process have any effect on expert judgment? The remainder of this paper reviews the pertinent literature, describes our study, presents its results, and concludes with suggestions for future research.

THEORETICAL BACKGROUND

Scenario planning is used for making strategic decisions, i.e. the ones that 'deal with the long-term allocation of existing resources and the development of new ones essential to assure the continued health and future growth of the enterprise' (Chandler, 1962). The planning horizon for such decisions is long, and it is generally difficult to predict the business environment over this period with reasonable accuracy. It can also be hard to estimate the effect of environmental forces on an organization's strategy. Therefore, decision-making in this domain is described as 'not the decision making under uncertainty of the textbook, where alternatives are given even if their consequences are not, but decision making under ambiguity, where almost nothing is given or easily determined' (Mintzberg, Raisinghani, and Theoret, 1976). Making such decisions requires going 'beyond the information given' and inferring what is missing 'from available information, especially the person's own experience and world knowledge' (Klayman and Schoemaker, 1993). In this process, decision-makers resort to cognitive simplification processes to make sense of the complex decision context (Schwenk, 1984; Russo and Schoemaker, 1989) using various mental heuristics (Kahneman, Slovic, and Tversky, 1982). While the heuristics

help managers arrive at decisions, they may also create a false sense of overconfidence (Camerer and Lovallo, 1999; Simon and Houghton, 2003). Overconfidence can engender 'a dangerous degree of hubris on the part of decision makers' (Levinthal, 2011) leading them to 'plunge in' to solve problems without defining them thoughtfully (Russo and Schoemaker, 1989) and to build 'forecasts of future outcomes anchored on plans and scenarios of success rather than on past results' (Kahneman and Lovallo, 1993). Field studies show that overconfidence is linked with poor executive decisions (Ben-David, Graham, and Harvey, 2006; Malmendier and Tate, 2005).

Scenario planning is argued to mitigate the negative effects of overconfidence. The two experimental studies of scenario planning we found, which measured whether scenario use affected confidence in judgment, reached contrary conclusions. Subjects (MBA students) in the Schoemaker (1993) study broadened confidence ranges around their point forecasts of parameters related to the strategic issues of their choice (e.g., profit margin, market share), indicating a *decrease* in confidence after scenario use. Conversely, subjects (undergraduate psychology students) in the Kuhn and Sniezek (1996) study expressed *higher* confidence in their five-decade point forecasts of various social issues (e.g. murder rate in Chicago, world population, etc.) after using one or more scenarios, compared to the control group. The discrepancy between these results may stem from the difference in experimental designs (use of a control group in the latter but not the former) or the subjects' familiarity with the experimental task (higher familiarity in the former). Regardless of the reason, the limited experimental evidence of the effect of multiple scenarios on confidence is inconclusive.

RESEARCH METHOD

We tested the effect of scenario planning on expert judgment in a series of field studies carried out as part of a larger project (Caplice and Phadnis, 2013), in which the method was used to

(6)

assess investments in the U.S. freight transportation infrastructure. Transportation infrastructure is an ideal setting for the use of scenario planning, as the process of making and implementing plans can take many years and the infrastructure remains in use for several years. Four scenarios were created for the project using the *scenario-axes technique*, a standard scenario creation method (Schwartz, 1991). Their application was tested in three scenario planning workshops (referred to as Studies I, II and III) conducted at U.S. transportation planning agencies (e.g., state departments of transportation). For each workshop, the host agency chose several transportation infrastructure investments (e.g. highways, rail lines, ports, etc.) to assess using the scenarios. The workshops were used to solicit insights from a diverse group of experts to help the agency prioritize the investments to meet the region's freight transportation needs for the next 30 years.

Field experts, as the study's subjects

The experts who participated in the scenario planning workshops served as the study's subjects. They were not self-selected, but handpicked and invited to participate in the workshop because their knowledge of the region was considered to provide valuable input to the agency's planning process. The experts came from business firms (shippers, carriers, logistics service providers), government agencies (federal and state planners, military), and other organizations (academia, community groups, transportation consultancies, etc.). The typical participants were CEOs and owners of carriers and consulting firms; transportation executives at large firms; and managers in government planning agencies. The participants were informed that their evaluations of the chosen investments would be used by the agency to inform its long-range planning process.

Design of field studies

The studies were conducted using a pretest-posttest design, with the scenario planning workshop as the experimental treatment. However, since the workshops were used to solicit expert insights

to be used in the agency's planning process, it was not possible to have a control group. Despite this limitation, our studies score well on all six criteria used to define field experiments (Harrison and List, 2004): (i) our subject pool consisted of field experts (ii) chosen for the 'information [they] bring to the task'; (iii) the study's task involved decision-making about the 'actual goods,' i.e. the infrastructure investments (iv) which was facilitated by applying rules appropriate to the domain of transportation infrastructure planning; (v) subjects were informed of the high 'stakes' of their decisions and (vi) the tasks were performed in an environment familiar to the subjects.

Each workshop used either three or four scenarios, depending on the number of experts, so that between 10 and 15 experts evaluated all investments in each scenario. The experts were assigned to scenarios using stratified random sampling, where the strata referred to the type of organization they represented (shipper, carrier, state and local planner, federal planner, third-party logistics provider, and other). Each scenario had a roughly equal number of experts of each type. The steps in the study are depicted in Figure 1. Before participating in the exercise, each expert completed the pretest using an online survey tool. After completing the pretest, each expert was sent a brochure of the scenario s/he was assigned to via email. The brochure included a detailed narrative of the scenario and several charts illustrating specific technological, economic, and demographic aspects of the scenario. All experts were asked to *read their scenario* before the workshop. At the start of the workshop, the head of the host agency informed the participants that their evaluations of the investments and insights would be considered in the agency's planning process. Following this, the participants were given an overview of scenario planning and then separated by the scenarios they were assigned to. The participants in each scenario met in separate groups and engaged in facilitated discussions. Each session began with *immersion of experts in the assigned scenario* to ensure that everyone understood the scenario

before judging the investments for that scenario. This involved asking the group to identify the key features of the scenario and showing a custom-designed video for the scenario. This was followed by a *discussion of the scenario's implications*, in which the experts shared their views on the region's transportation needs in that scenario. Based on the implications, the experts *evaluated the investments* individually and expressed them using the voting procedure described in the next section. Reading of the assigned scenario, immersion, discussion of its implications, and evaluation of investments constitute the experimental treatment 'single scenario evaluation.'

The single scenario evaluations were followed by a lunch break, after which the participants convened as a single group. They were first shown the videos of all scenarios used in the workshop. Following this, representatives from each scenario presented the evaluation of the investments in their scenario and described their rationale for it. The presentation was followed by a question-and-answer session. Typically, participants in the audience asked clarifying questions and described why they evaluated a particular investment in their scenario either differently from the presenting group or similarly as the presenting group but for a different reason. Following the presentations, the research team provided a one-slide graphical summary of evaluations from all scenarios used in the workshop. This slide was used to facilitate a discussion wherein the participants identified robust investments (i.e., ones found useful in all or most scenarios and wasteful in none) and those whose utility was contingent on the scenario (i.e., useful in some scenarios and wasteful in others). The viewing of videos of all scenarios, viewing and discussing investment evaluations in individual scenarios, and discussing evaluations across all scenarios constitute the experimental treatment 'multiple scenario evaluation.'

Insert Figure 1 about here

Voting mechanism to evaluate usefulness of an investment

Each participant was asked to assess the relative usefulness of the chosen investments for their scenario by allocating 100 points among them, such that the investments deemed more useful for the scenario were assigned more points. Each participant was also asked to veto between one and three investments s/he judged to be wasteful for the scenario. The participants assigned points and vetoes to the investments individually, and revealed their evaluations by placing poker chips on a group voting sheet. The facilitator tallied the vote and discussed it by asking the participants to share their rationale for the vote. The participants were allowed to change their vote; the final vote was used as the measure of usefulness of each investment for the particular scenario.

Pretest and posttest

The pretest was administered before the experts saw any scenario; the posttest was administered after the experts evaluated investments in a *single* scenario (Posttest-A) and after they observed evaluations in *multiple* scenarios (Posttest-B). The questionnaires used in the pretest and the posttests of a given study were identical, and were completed by the subjects individually. The subjects were asked to evaluate each investment for a 30-year planning horizon using a two-question format: if the planning agency should make the investment and the subject's confidence in that recommendation. Our design assumed that subjects will complete Pretest based on their mental image of the future 30 years, Posttest-A based on the perception of future influenced by one scenario, and Posttest-B based on the perception of future influenced by multiple scenarios. We assured the subjects anonymity of their responses to discourage any pro-social behavior.

RESULTS

We examined changes in expert judgment and confidence due to scenario use in studies I and II. Study I, conducted at a state Department of Transportation in the U.S., used scenario planning to

evaluate 16 infrastructure investments. The two-part question solicited investment recommendation (choices: Yes, No, and Do not know) and confidence in the recommendation on a four-point scale (ranging from 'Highly certain' to 'Highly uncertain'). Study II, conducted at a U.S. ocean port, used the scenarios to evaluate 15 infrastructure investments. To ensure that the results were not idiosyncratic to the instrument used, we used a different form of two-part question in Study II: the subjects were forced to provide a Yes or No recommendation, and their confidence was recorded using a finer seven-point ratio scale due to Windschitl and Wells (1996).

Result 1: Effect of *multiple*-scenario use on experts' confidence in decisions

Study I yielded 343 pretest-posttest pairs, where a subject evaluated a particular investment in both Pretest and Posttest-B, and answered either 'Yes' or 'No' to making that investment. In 189 of these (55.1%), a given expert had changed his/her judgment—either the investment advice and/or the confidence in it—from pretest to posttest. However, the aggregate distribution of votes among four confidence levels (Table 1) in the posttest was not different from that in the pretest ($\chi^2 = 5.65, df = 3, p = .13$). There were also no differences (at $p \leq .1$; z test of proportions) between the proportions of votes in the pretest and the posttest at any one confidence level.

<center>**Insert Table 1 about here**</center>

Study II yielded 285 pretest-posttest judgment pairs, where a subject had evaluated an investment in Pretest and Posttest-B. The Windschitl and Wells (1996) scale used in the study permits calculation of *average confidence*, as a sum of the products of proportion of votes and average value of the confidence interval. The pretest-posttest distribution of votes across the confidence levels is shown in Table 2. While a majority of judgments (198 out of 285; 69.5%) changed after multiple scenario evaluation in this study as well, the average confidence in the

<center>(11)</center>

pretest (0.811) was virtually identical to that in the posttest (0.814) ($p = .979$, two-tailed paired t test). The results of both these studies suggest the following:

Insert Table 2 about here

Proposition 1: The use of multiple scenarios to evaluate long-range investment decisions does not categorically increase or decrease the decision-makers' confidence in their evaluations.

Result 2: Change in expert judgment after using a *single* scenario

Next, we took a closer look at the change in judgment after single scenario evaluation. Any change in the expert judgment was denoted as to have become either *more favorable* or *less favorable* of the judged investment after single scenario evaluation. A judgment was designated to have become *more favorable* if the change in expert vote from Pretest to Posttest-A was one of the following three types: changed recommendation from opposing the investment to supporting it (voted 'No' in Pretest and 'Yes' in Posttest-A); remained opposed to making the investment but with lower confidence; or remained supportive of the investment and with higher confidence. The criterion for designating a judgment to have become *less favorable* was symmetric.

For each change in expert judgment, we checked how the corresponding investment was evaluated in the scenario used by that expert. Using the group vote from the particular scenario, we classified each investment as *Useful*, *Wasteful*, or *Neither* for that scenario. If *n* investments were assessed in a workshop, an investment was denoted *Useful for Scenario X* if it received a minimum of *1/n* points in Scenario X (Condition 1) and its proportion of points was at least three times higher than its proportion of vetoes in that scenario (Condition 2). Condition 1 ensures that the total points received is greater than what it is expected in a random assignment; Condition 2 ensures that only the investments receiving *sufficiently* high proportion of points compared to the proportion of vetoes are chosen as *Useful* investments. The criteria for *Wasteful* investments is symmetric: they receive a minimum of 1/n vetoes (Condition 3) and the proportion of vetoes is at

(12)

least three times that of points (Condition 4) (see the Online Appendix for an example). While this parameterization of the criteria of *Useful* and *Wasteful* is somewhat arbitrary, the sensitivity tests (in Online Appendix) show that the results hold over a wide range of parameter values.

Studies I and II yielded similar results; the results from Study II are presented here (see Figure 2). Study II yielded 404 pairs of judgment where an expert evaluated a given investment in both Pretest and Posttest-A. Of these, 288 judgments (71.3%) changed from the pretest to the posttest. Among the changed judgments, 62 were related to the investments denoted as *Useful* and 70 to those denoted as *Wasteful*. Out of 62 judgments of *Useful* investments, 39 (62.9%) became *more favorable* of the investment after single scenario evaluation. This likelihood of a changed judgment becoming *more favorable* is higher for the investments deemed *Useful* in the scenario used ($p < .001$; binomial test of proportions) than the average (136/288=47.2%). Similarly, among 70 judgments of *Wasteful* investments, 48 (68.6%) became *less favorable*. The likelihood of changed judgments becoming *less favorable* is also higher for investments found *Wasteful* in the scenario used ($p < .001$) than the average (152/288 =52.8%). The likelihood that investments are judged *more favorably* when deemed neither useful nor wasteful (75/156 = 48.1%) is not different from the average ($p = .291$). Overall, the results suggest:

Proposition 2: If a decision-maker's judgment of a long-term investment changes after evaluating it using one scenario, it will become (a) more favorable of the investment if it is found useful for the scenario and (b) less favorable if it is found wasteful for the scenario.

Insert Figure 2 about here

Result 3: Moderating effect of ex-ante confidence on scenario-induced change

In both Studies I and II, the change in expert judgment from single-scenario evaluation was moderated by the expert's confidence in the pretest (results in Figure 3). In Study I, 419 pairs of judgment were available where an expert evaluated a given investment in Pretest and Posttest-A;

(13)

234 of these (55.8%) changed after the single scenario evaluation. The proportion of changed judgments was lower (70/193=36.3%) than this average ($p < .001$; one proportion z test) for the investments judged with the highest level of confidence before using a scenario. For each of the remaining three levels of ex-ante confidence, almost 7 out of 10 judgments changed. Results of Study II are similar. 288 of 404 judgments (71.3%) changed after the single scenario evaluation; however, the proportion of changed judgments was lower (19/52=36.5%) than the average when experts were 'Certain' of their judgment in the pretest ($p < .001$). Overall, the results suggest:

Proposition 3: A one-time use of a scenario is less likely to cause a change in judgment if the decision-maker had the highest level of confidence in the judgment before using the scenario.

Insert Figure 3 about here

Result 4: Effect of multiple scenarios on type of strategies chosen

We tested the effect of multiple scenarios on experts' preference for flexibility in Study III. The study was conducted at a federal transportation agency, and used the scenarios to select strategies for investing in 13 types of transportation infrastructure segments (e.g., border crossings, ocean ports, highway corridors, etc.). Four generic strategies (described in the Online Appendix) were specified for each segment: they ranged from the least flexible Option-1 (implementation of specific projects) to the most flexible Option-4 (allocation of funds to the segment). The experts were asked in Pretest and Posttest-B to recommend an investment strategy for each type of segment. The questionnaire noted the variation in flexibility of the four strategies. 351 pretest-posttest pairs of recommendations by 27 experts were available (see Figure 4). A majority of the subjects (229/351=65.2%) recommended the least flexible strategy (Option-1) in the pretest. A majority of experts still favored Option-1 after multiple scenario evaluation; but, the proportion of experts choosing this option dropped to half (176/351). This drop was counteracted by almost doubling of support for a more flexible Option-3 (from 43/351=12.3% to 81/351=23.1%). The

changes in preferences for both these options are statistically significant ($p < .001$; z-test for two independent proportions). Support for the most flexible strategy (Option-4) also increased from 26/351 (7.4%) to 40/351 (11.4%) ($p = .056$). Overall, the results suggest:

Proposition 4: After evaluating a long-range investment using multiple scenarios, a decision-maker is more likely to prefer a solution with higher flexibility to implement that investment.

Insert Figure 4 about here

DISCUSSION

Practitioners and scholars have extolled scenario planning for its 'frame-breaking ability' (Eisenhardt, 1999) to change planners' 'assumptions about how the world works' (Wack, 1985) and influence their long-range investment decisions. However, management literature is devoid of rigorous studies that test effect of scenario planning on field experts' judgment (Schoemaker, 2004; Wilkinson, 2009). This study is our first attempt to address this gap in the literature.

Contradicting the previous two empirical studies of scenario planning, which themselves arrived at contrary conclusions (Schoemaker, 1993; Kuhn and Sniezek, 1996), our results show that the use of multiple scenarios does not categorically increase or decrease confidence (Result 1). The practice of multiple scenarios may not be a cognitive repair for *overconfidence* (Heath, Larrick, and Klayman, 1998) among expert decision-makers. Instead, the effect of scenarios may need to be assessed using some metric(s) other than aggregate confidence in judgment. One research strategy is to look for changes in the reasons used by the experts to justify their choices (Shafir, Simonson, and Tversky, 1993). The use of multiple scenarios may provide new pros and cons when evaluating long-term investments. This strategy can also detect the effect of scenarios when the reason for choosing a particular action changes, but the choice of action does not.

Although multiple scenario evaluation did not cause orderly changes in confidence, we observed systematic changes in expert judgment after the use of a single scenario. The judgment

(15)

change was contingent on suitability of the judged investment to the scenario used: a changed judgment was likely to become more favorable if the investment was found useful in the scenario and vice versa (Result 2). This may be a result of scenarios influencing the cognitive processes involved in assessing fitness of an investment for the future environment. Just as entrepreneurs use 'prior knowledge of markets to search for and think of opportunities for new technologies' (Gregoire, Barr, and Shepherd, 2010), scenario-based evaluation may allow experts to use their prior knowledge of an asset to think of new ways it may become a strength or a weakness in the environment envisioned after scenario use, and find new reasons (Shafir, *et al.,* 1993) to favor (disfavor) the investment. Judgment changes from the single scenario evaluation were moderated by the expert's ex-ante confidence: an expert with the highest level of confidence in judgment was less likely to change it after single scenario evaluation (Result 3). This relative inability of a scenario to change prior judgments of the highest confidence may be an artifact of the study's design (i.e., judgments could not become more favorable if the scenario matched the expert's ex-ante vision of the future) or a result of the disconfirmation bias, whereby a scenario inconsistent with the ex-ante vision is discounted by the expert (Edwards and Smith, 1996). Since a person's confidence depends on the 'amount and strength of the evidence' favoring the judgment (Koriat, Lichtenstein, and Fischhoff, 1980), scenarios may need to be highly elaborate—replete with cogent arguments for why and how the scenario may evolve—if they are to influence senior executives with extensive industry knowledge. The few publically available scenarios from large corporations are indeed extensive (Royal Dutch Shell, 2005, 2008; Deutsche Post, 2012; etc.)

Finally, our results show that the use of multiple scenarios can nudge experts towards more flexible strategies (Result 4). Due to this ability to help experts recognize the merits of flexible options, scenario planning could serve as a valuable process in a firm's arsenal of

dynamic capabilities (Teece, 2007; Eisenhardt, Furr, and Bingham, 2010). In our study, although the multiple scenario evaluation shifted expert preferences towards more flexible strategies, a majority of the experts still favored the least flexible option. There are at least two explanations for this effect. The experts were judging transportation infrastructure in the U.S., which was considered 'failing to keep pace with the current and expanding needs' and had been graded as either 'D/Poor: At risk' (roads, transit, etc.) or 'C/ Mediocre: Requires attention' (bridges, rail) by the American Society of Civil Engineers (2009). The expert participants may have felt that an immediate investment was necessary regardless of the scenario and chosen the least flexible option. Conversely, the efficacy of scenario planning in nudging experts towards more flexible options may be limited. We did not examine whether this ability to nudge is moderated by the urgency of making an investment. This could be easily examined in a laboratory experiment.

Limitations and directions for future empirical research

Although this work tested the effect of scenario planning on field experts' judgment of long-range decisions, our setting differs from the practice of scenario planning at firms like Shell in two ways. One, the subjects in our study were not accountable for their recommendations. While we cannot claim that our subjects were just as rigorous when making the recommendations as someone invested in them, we have no reason to believe that they took the exercise lightly either, especially since the importance of their input was emphasized to them. Two, our results are based on a one-time use of scenarios. Firms like Shell use scenario planning on a continual basis; this may have a different effect on executive judgment than a one-time use. Therefore, this study needs to be complemented by a longitudinal study of the on-going use of scenario planning.

Besides assessing if and how the effect of one-time use of scenarios differs from that of on-going use, we suggest three other directions—by no means exhaustive—for future research.

One area of interest is to identify the antecedents of scenario-induced change in expert judgment. Scenario planning is used at Shell for 'changing minds, not making plans' (de Geus, 1988). One could examine if scenario use changes 'minds' by recording the users' thinking using mental maps or brain imaging techniques (Reisberg, 2006). Temporal aspect of scenario effect is another interesting area. Knowing how long the effect of scenario use lasts, and whether the duration is moderated by scenario user's day-to-day responsibilities, would be of practical importance when planning scenario exercises for busy managers and planners. Finally, the intersection of scenario planning and decision analysis is a fertile area for scholarly exploration. Scenarios are thinking devices useful for structuring messy problems of long-range planning (van der Heijden, 2000), whereas decision analysis provides a rational process for making the 'right' decision for a given problem. Normative guidelines for combining the debiasing power of scenario planning with the structured decision analytic approach could enhance the quality of long-range decisions.

In summary, scenario planning is a rich area for management research. The process is practiced and recommended for making strategic decisions in unpredictable environments. In all our studies, a large portion of experts changed their judgment after using one or more scenarios. This ability of scenarios to influence even some expert judgment is noteworthy given the status-quo bias among decision-makers (Samuelson and Zeckhauser, 1988). A rigorous empirical scrutiny of scenario planning can help explore its true merits as a strategic thinking process.

(18)

REFERENCES

American Society of Civil Engineers. 2009. *Report Card for America's Infrastructure*. American Society of Civil Engineers: Washington, DC.

Bazerman MH, Moore DA. 2009. *Judgment in Managerial Decision Making*. John Wiley & Sons: Hoboken, NJ.

Ben-David I, Graham JR, Harvey, CR. 2006. Managerial overconfidence and corporate policies. Working Paper, Duke University.

Bradfield R, Wright G, Burt G, Cairns G, van der Heijden K. 2005. The origins and evolution of scenario techniques in long range business planning. *Futures* **37**(8): 795–812.

Camerer C, Lovallo D. 1999. Overconfidence and excess entry: An experimental approach. *American Economic Review* **89**(1): 306–318.

Caplice C, Phadnis S. 2013. *Strategic Issues Facing Transportation: Scenario Planning for Freight Transportation Infrastructure Investment*. Transportation Research Board: Washington, DC.

Chandler AD. 1962. *Strategy and Structure: Chapters in the History of the American Industrial Enterprise*. MIT Press: Cambridge, MA.

Cousens R, Steinberg T, White B, Walton S. 2002. Generic scenarios: A strategic futures paper. Strategic Futures.

de Geus AP. 1988. Planning as learning. *Harvard Business Review* **66**(2): 70–74.

Deutsche Post AG. 2012. *Delivering Tomorrow: Logistics 2050—A Scenario Study*. Deutsche Post AG: Bonn, Germany.

Edwards K, Smith EE. 1996. A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology* **71**(1): 5–24.

Eisenhardt KM. 1999. Strategy as strategic decision making. *Sloan Management Review* **40**(3): 65–72.

Eisenhardt KM, Furr NR, Bingham CB. 2010. Microfoundations of performance: Balancing efficiency and flexibility in dynamic environments. *Organization Science* **21**(6): 1263–1273.

Gordon ME, Slade LA, Schmitt N. 1987. Student guinea pigs: Porcine predictors and particularistic phenomena. *Academy of Management Review* **12**(1): 160–163.

Grant RM. 2003. Strategic planning in turbulent environment: Evidence from the oil majors. *Strategic Management Journal* **24**(6): 491–517.

Gregoire DA, Barr PS, Shepherd DA. 2010. Cognitive process of opportunity recognition: The role of structural alignment. *Organization Science* **21**(2): 413–431.

Harries C. 2003. Correspondence to what? Coherence to what? What is good scenario-based decision making? *Technological Forecasting and Social Change* **70**(8): 797–817.

Harrison G, List J. 2004. Field experiments. *Journal of Economic Literature* **42**(4): 1009–1055.

Heath C, Larrick RP, Klayman J. 1998. Cognitive repairs: How organizational practices can compensate for individual shortcomings. In *Research in Organizational Behavior* **20**, Staw BM, Cummings LL (eds). JAI Press: New York; 1–37.

Hodgkinson GP, Healey MP. 2011. Psychological foundations of dynamic capabilities: Reflexion and reflection in strategic management. *Strategic Management Journal* **32**(13): 1500–1516.

Kahneman D, Lovallo D. 1993. Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science* **39**(1): 17–31.

Kahneman D, Slovic P, Tversky A (eds). 1982. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge University Press: Cambridge, U.K.

Klayman J, Schoemaker PJ. 1993. Thinking about the future: A cognitive perspective. *Journal of Forecasting* **12**(2): 161–186.

Koriat A, Lichtenstein S, Fischhoff B. 1980. Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory* **6**(2): 107–118.

Kuhn KM, Sniezek JA. 1996. Confidence and uncertainty in judgmental forecasting: Differential effects of scenario presentation. *Journal of Behavioral Decision Making* **9**(4): 231–247.

Levinthal DA. 2011. A behavioral approach to strategy - what's the alternative? *Strategic Management Journal* **32**(13): 1517–1523.

Linneman RE, Klein HE. 1983. The use of multiple scenarios by U.S. industrial companies: A comparison study, 1977–1981. *Long Range Planning* **16**(6): 94–101.

Malmendier U, Tate G. 2005. CEO overconfidence and corporate investment. *Journal of Finance* **60**(6): 2661–2700.

Meissner P, Wulf T. 2013. Cognitive benefits of scenario planning: Its impact on biases and decision quality. *Technological Forecasting and Social Change* **80**(4): 801–814.

Mintzberg H, Raisinghani D, Theoret A. 1976. The structure of 'unstructured' decision processes. *Administrative Science Quarterly* **21**(2): 246–275.

Phelps R, Chan C, Kapsalis SC. 2001. Does scenario planning affect performance? Two exploratory studies. *Journal of Business Research* **51**(3): 223–232.

Reisberg D. 2006. *Cognition: Exploring the Science of the Mind*. WW Norton & Co.: New York.

Royal Dutch Shell. 2005. *The Shell Global Scenarios to 2025*. Shell International Limited.

Royal Dutch Shell. 2008. *Shell Energy Scenarios to 2050*. Shell International BV.

Russo JE, Schoemaker PJ. 1989. *Decision Traps: Ten Barriers to Brilliant Decision-making and How to Overcome Them*. Doubleday: New York.

Samuelson W, Zeckhauser R. 1988. Status quo bias in decision making. *Journal of Risk and Uncertainty* **1**(1): 7–59.

Schoemaker PJ. 1993. Multiple scenario development: Its conceptual and behavioral foundation. *Strategic Management Journal* **14**(3): 193–213.

Schoemaker PJ. 2004. Forecasting and scenario planning: The challenges of uncertainty and complexity. In *Blackwell Handbook of Judgment and Decision Making*, Koehler DJ, Harvey N (eds). Blackwell: Oxford, U.K.: 274–296.

Schwartz P. 1991. *The Art of the Long View: Planning for the Future in an Uncertain World*. Doubleday: New York.

Schwenk CR. 1984. Cognitive simplification processes in strategic decision-making. *Strategic Management Journal* **5**(2): 111–128.

Shafir E, Simonson I, Tversky A. 1993. Reason-based choice. *Cognition* **49**(1): 11–36.

Simon M, Houghton SM. 2003. The relationship between overconfidence and the introduction of risky products: Evidence from a field study. *Academy of Management Journal* **46**(2): 139–149.

Teece DJ. 2007. Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal* **28**(13): 1319–1350.

van der Heijden K. 2000. Scenarios and forecasting: Two perspectives. *Technological Forecasting and Social Change* **65**(1): 31–36.

Varum CA, Melo C. 2010. Directions in scenario planning literature: A review of the past decades. *Futures* **42**(4): 355–369.

Wack P. 1985. Scenarios: Uncharted waters ahead. *Harvard Business Review* **63**(5): 73–89.

Wilkinson A. 2009. Scenarios practices: In search of theory. *Journal of Futures Studies* **13**(3): 107–114.

Windschitl PD, Wells GL. 1996. Measuring psychological uncertainty: Verbal versus numeric methods. *Journal of Experimental Psychology: Applied* **2**(4): 343–364.
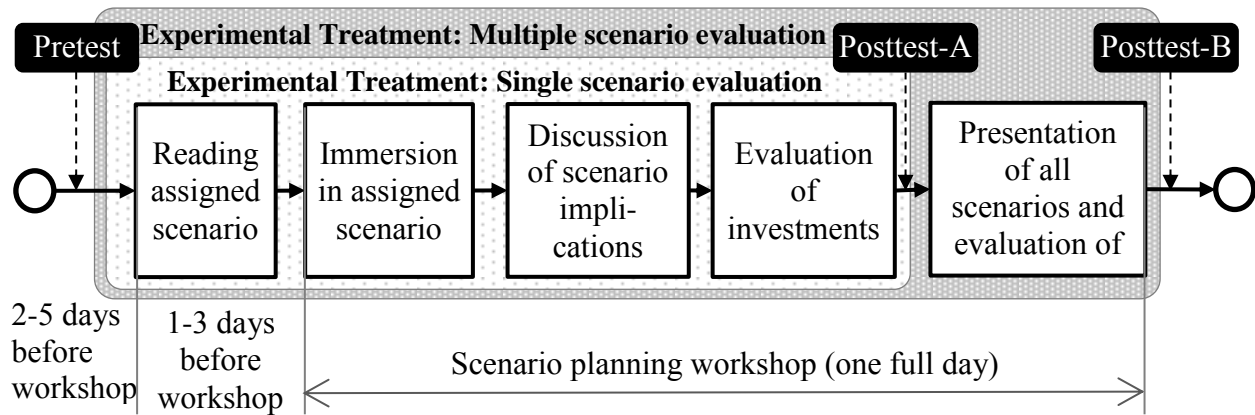
FIGURES AND TABLES

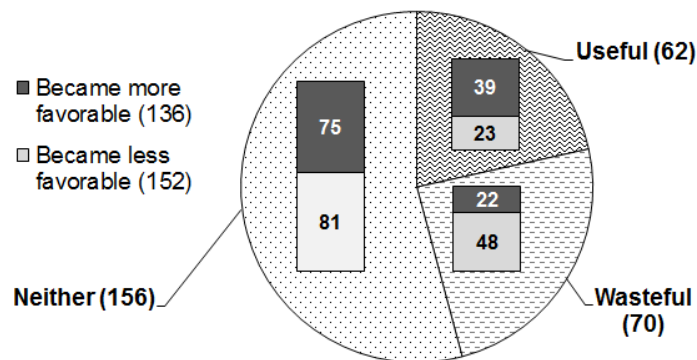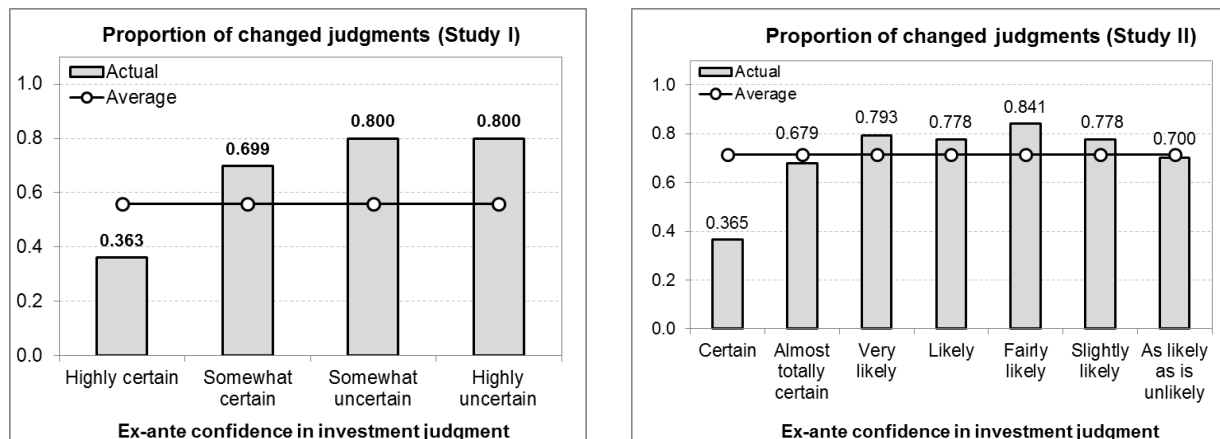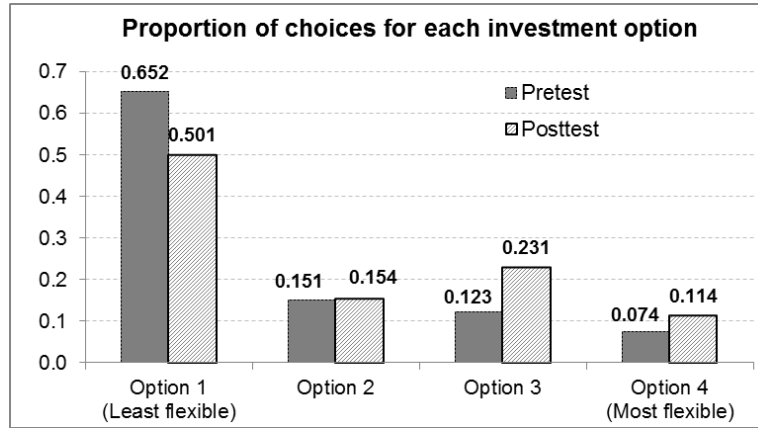
Figure 1: Design of field experiments

Figure 2: Type of change in expert judgment (based on investment assessments)

Figure 3: Prior confidence and change in judgment after using one scenario

Figure 4: Proportion of choices for four investment options in Study III

Table 1: Distribution of judgments in Pretest and Posttest-B (Study I)

Confidence level	Pretest			Posttest			*p* value* of differe-nce in proportions
	Yes	No	Total	Yes	No	Total	
Highly certain	129	29	158 (0.46)	123	34	157 (0.46)	0.469
Somewhat certain	107	26	133 (0.39)	82	38	120 (0.35)	0.152
Somewhat uncertain	34	13	47 (0.14)	41	17	58 (0.17)	0.122
Highly uncertain	4	1	5 (0.02)	3	5	8 (0.02)	0.200
Total assessments	274	69	343	249	94	343	*z test of proportions

Table 2: Distribution of judgments in Pretest and Posttest-B (Study II)

Confidence in judgment		Number of judgments		Confidence	
Level	Avg.	Pretest	Posttest	Pretest	Posttest
Certain (100%)	1.00	51	51	51.0	51.0
Almost totally certain (90–100%)	0.95	63	52	59.9	49.4
Very likely (80–90%)	0.85	45	71	38.3	60.4
Likely (70–80%)	0.75	46	40	34.5	30.0
Fairly likely (60–70%)	0.65	46	31	29.9	20.2
Slightly likely (50–60%)	0.55	14	22	7.7	12.1
As likely as is unlikely (50-50 chance)	0.50	20	18	10.0	9.0
Total		285	285	231.2	232.0
Average confidence				0.811	0.814

(22)

EFFECT OF SCENARIO PLANNING ON FIELD EXPERTS' JUDGMENT OF LONG-RANGE INVESTMENT DECISIONS

Scenarios

The scenarios were created using 'scenario axes technique,' which is a standard method for developing scenarios. The scenarios differ along two 'axes': nature of trade (global vs. regional) and resource availability (scarce vs. adequate). Scenario A[1] depicts a global economy where goods flow seamlessly between virtually all markets and the necessary resources can be acquired from wherever available. Scenario B paints the U.S. as a country with several local markets located in geographically-dispersed, highly self-reliant small cities, where self-sufficiency has been enabled by advances in small-scale manufacturing and adoption of renewable energy sources. Scenario C depicts a world fragmented into several regional blocs, with seamless trade within a bloc and minimal trade between different blocs. Scenario D describes a world with global trade dictated by a supranational organization that seeks to optimize demand and supply of the world's scarce resources.

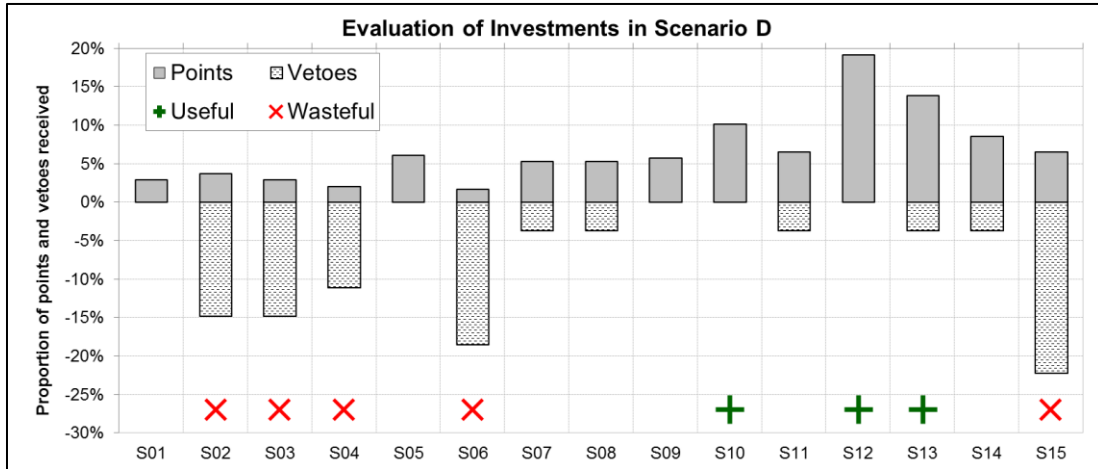**Example of 'Useful' and 'Wasteful' segments**

Figure 1 shows one example of the results of a voting session (Scenario D in Study II). The vertical axis shows the proportion of points (on positive side) and vetoes (on negative side) received by each investment in that scenario. Investments S10, S12, and S13 are denoted as *Useful for Scenario D*: they received 10 percent, 19 percent, and 14 percent of the points (i.e.,

---

[1] Scenario A, B, C, and D are the pseudonyms used during the review process; they to refer to scenarios 'Global Marketplace', 'Millions of Markets', 'Naftastique!', and 'One World Order', respectively in the Future Freight Flows project.

exceed the threshold of 1/15, satisfy Condition 1), and 0, 0, and 4 percent of the vetoes, respectively (i.e., exceed %vote-to-%veto-ratio threshold of 3, satisfy Condition 2). Investments S02, S03, S04, S06, and S15 are denoted as *Wasteful for Scenario D*: each received at least 1/15 of vetoes (Condition 3) and proportion of vetoes to was at least 3 times the proportion of points (Condition 4).

Figure 5: Evaluation of investments in Scenario D (Study II)

Sensitivity of Result 2 to criteria of 'useful' and 'wasteful'

An investment is termed *Useful* (*Wasteful*) if the proportion of votes (vetoes) it received is at

least $\frac{k}{\text{number of segments}}$ and the ratio of proportions of votes to vetoes (vetoes-to-votes) exceeds

p. In the results presented in the main body of the paper, $k = 1, p = 3$. To test robustness of the results to these somewhat arbitrary definitions, value of k was varied from 0.8 to 2 in increments of 0.2 (seven values), and p from 1.5 to 5 in increments of 0.5 (eight values). The proportions of expert judgments that changed according to Result 2, among all changed judgments for investments denoted *Useful* and *Wasteful*, are presented in Table 1.

(24)

Table 1: Proportion of changed judgments confirming Result 2 (Study II)

		Prop. of changed judgments for USEFUL assets that became MORE favorable							Prop. of changed judgments for WASTEFUL assets that became LESS favorable						
		Multiplier for prop. of points (k)							Multiplier for prop. of points (k)						
		0.8	1.0	1.2	1.4	1.6	1.8	2.0	0.8	1.0	1.2	1.4	1.6	1.8	2.0
Ratio of prop. of points to proportion of vetoes (p)	1.5	0.61	0.60	0.57	0.57	0.63	0.57	0.57	0.67	0.67	0.67	0.66	0.70	0.70	0.70
	2.0	0.58	0.58	0.57	0.57	0.63	0.57	0.57	0.67	0.67	0.67	0.66	0.70	0.70	0.70
	2.5	0.59	0.59	0.58	0.57	0.63	0.57	0.57	0.69	0.69	0.69	0.69	0.70	0.70	0.70
	3.0	0.62	0.63	0.63	0.63	0.63	0.57	0.57	0.69	0.69	0.69	0.69	0.70	0.70	0.70
	3.5	0.63	0.64	0.63	0.63	0.65	0.56	0.56	0.68	0.68	0.68	0.68	0.70	0.70	0.70
	4.0	0.62	0.63	0.63	0.63	0.65	0.54	0.54	0.67	0.67	0.67	0.67	0.68	0.67	0.67
	4.5	0.62	0.63	0.63	0.63	0.65	0.54	0.54	0.67	0.67	0.67	0.67	0.68	0.67	0.67
	5.0	0.62	0.63	0.63	0.63	0.65	0.54	0.54	0.67	0.67	0.67	0.67	0.68	0.67	0.67

The proportion of changed judgments for the investments deemed *Wasteful*, where the expert evaluated the investment *less favorably* after using one scenario, lies between 0.66 and 0.70; thus, the results are robust to the changes in the criteria used to define *Wasteful* investments. The proportion of changed judgments for investments deemed *Useful*, where the expert evaluated the investment *more favorably* after using one scenario, lies between 0.54 and 0.65. These proportions are lower than the ones for *Wasteful* investments, because the criteria used to denote a judgment as becoming *more favorable* are more stringent than those for used to denote a judgment as becoming *less favorable*. The definition of changed judgments becoming *more favorable* ignores all judgments where an investment was recommended with the highest level of confidence in both pretest and posttest, as there is no change in the judgment. For instance, 26 pretest-posttest judgment pairs were available for the investments denoted as *Useful* using the most stringent criteria: i.e., $k = 2, p = 5$. In 13 of these, the experts changed the judgment from the pretest to the posttest; in 7 of those instances the expert judgment became *more favorable* (7/13=0.54). However, the 13 unchanged judgments included 7 instances in which the experts had favored the investment with the highest confidence in both pretest and posttest. These

judgments could not become any *more favorable*. If these judgments are included in the analysis, the proportion of judgments for *Useful* investments that became *more favorable* would be (7+7)/(13+7) =0.70. Table 2 shows the proportion of judgments for *Useful* investments after including the unchanged judgments that recommended investment with the highest level of confidence in both pretest and posttest. These range from 0.64 to 0.75.

Table 2: Proportion of judgments confirming Result H2 for useful investments after including unchanged judgments supporting investment with highest confidence (Study II)

		0.8	1.0	1.2	1.4	1.6	1.8	2.0
	1.5	0.67	0.66	0.64	0.65	0.73	0.70	0.70
	2.0	0.65	0.65	0.64	0.65	0.73	0.70	0.70
Ratio of prop. of points to proportion of vetoes (p)	2.5	0.65	0.65	0.65	0.65	0.73	0.70	0.70
	3.0	0.68	0.70	0.70	0.70	0.73	0.70	0.70
	3.5	0.69	0.71	0.70	0.71	0.75	0.72	0.72
	4.0	0.68	0.70	0.70	0.70	0.74	0.70	0.70
	4.5	0.68	0.70	0.70	0.70	0.74	0.70	0.70
	5.0	0.68	0.70	0.70	0.70	0.74	0.70	0.70

Generic investment strategies in Study III

Study III specified four generic ways to invest in the chosen infrastructure segments. The flexibility of each strategy depended on its level of specificity. The most flexible strategy called for allocation of funds to an infrastructure segment (e.g., highways, ocean ports), but not to specific regions (e.g., highways in the Northeast, ocean ports in Florida) or specific projects (e.g., addition of a fourth lane in each direction of I-95 between exits 12 and 20 in Massachusetts, dredging of Port of Miami to allow larger ships). The least flexible strategy called for allocating funds to specific projects and starting their implementation. The four generic investment strategies are presented in Table 3.

Table 3: Four generic investment strategies in Study III

Investment options for each type of infrastructure segment	Allocate funds to segments	Allocate funds to regions	Allocate funds to projects	Start implementation
OPTION 1: Identify specific regions and projects where investments should be made, allocate funds, and start implementing those projects.	x	x	x	x
OPTION 2: Identify specific regions and projects where investments should be made, and allocate funds to the projects. BUT DO NOT START IMPLEMENTATION.	x	x	x	
OPTION 3: Identify specific regions where investments should be made, and allocate funds to those regions, BUT DO NOT ALLOCATE FUNDS TO PROJECTS.	x	x		
OPTION 4: Allocate funds to this segment, but do not allocate funds to individual regions within the segment.	x			