Mitigating Workpaper Default Effects that Impair Auditors' Risk Assessment Accuracy

Sarah Bonner
University of Southern California

Tracie Majors
University of Southern California

May 2025

This study was supported by a Center for Audit Quality and Auditing Section of the American Accounting Association Access to Audit Personnel Award. We thank Margot Cella and Lauren Tuite of the CAQ for facilitating access to participants. We thank Stacey Ritter and Meha Parekh for research assistance. We are grateful to workshop participants at Cornell University, as well as Lori Bhaskar, Nicole Donahoo, Cassandra Estep, Sean Hillison, Andrew Jones, Brett Patterson, Jeffrey Pickerd, Chad Proell, Mike Ricci, Sebastian Stirnkorb, Adam Vitalis, and Isaac Yamoah for helpful comments.

Mitigating Workpaper Default Effects that Impair Auditors' Risk Assessment Accuracy

Abstract

Regulators voice concerns that auditors miss changes in client risks, which can harm audit effectiveness and efficiency. In an experiment with audit staff, we replicate prior findings that auditors using prepopulated workpapers, which make prior year risk ratings *a default* for current year choices, are less accurate at identifying changes in risk, compared to auditors using non-prepopulated workpapers with no default. We test an intervention designed to target our posited underlying mechanisms: auditors inferring *implicit endorsement* that their firm finds it acceptable to stick with last year if uncertain, the default providing a path of inaction that reduces *anticipated regret* if auditors make incorrect choices, and the default providing a path to reduce *cognitive costs*. The intervention communicates the audit firm's priority of directionally accurate risk ratings and the related benefits. We find that the intervention fully attenuates these negative effects of prepopulation. Also important, side effects of the intervention that could have impaired accuracy do not emerge. In fact, the intervention also improves the accuracy of auditors using non-prepopulated workpapers for unchanging risks. Our findings provide a roadmap to mitigate undesired default effects, an important contribution given the strength and pervasiveness of defaults and consequent difficulty of identifying interventions.

JEL codes: G10, M40, M41, M42, D80, D91

Keywords: Audit risk assessment; default effects; prepopulation; audit effectiveness; audit workpapers; interventions

I. Introduction

Having voiced concerns about inaccurate risk assessment for several years (PCAOB [2015]), the Public Company Accounting Oversight Board (PCAOB) and Chief Accountant of the SEC recently re-emphasized the critical nature of risk assessment for high-quality audits, and the PCAOB prioritized it for inspections (PCAOB [2023a], [2023b]; SEC [2023]). Failure to increase risk ratings when necessary can lead to inadequate substantive testing and missed material misstatements (Allen et al. [2006]). Not appropriately decreasing risk ratings can lead to inefficiency, with more substantive testing than necessary for a high-quality audit. A key cause of auditors' rating inaccuracy for risks that have changed is that they are more likely to stick with prior year ratings when workpapers are *prepopulated* (Bonner, Majors, and Ritter [2018]). Prepopulated workpapers include the prior year's evidence and ratings in the spaces for current year choices. Non-prepopulated workpapers are blank; with these, auditors refer to prior year work in a separate file. Because prepopulation increases standardization, audit firms likely view it as beneficial for quality control (Boland, Galunic, and Sherwood [2024]). Thus, requiring use of non-prepopulated workpapers to improve risk rating accuracy likely is not desirable. In this study, we develop and test—in a controlled experiment—an intervention we hypothesize will reduce the negative effects of prepopulation on rating accuracy for risks that have changed.

Prepopulation effectively makes prior year choices *default options* for current year choices, that is, the options that will take effect if auditors do not actively select alternatives. By contrast, non-prepopulated workpapers introduce a required choice (Thaler and Sunstein [2021]). Decision-makers tend to stick with defaults, that is, adopt them as their personal choices, despite the freedom to choose otherwise (McKenzie, Liersch, and Finkelstein [2006]). Following, we

describe three mechanisms that are posited in economics and psychology to underlie default effects and how we expect they apply to auditors performing risk assessment.

The first mechanism—*implicit endorsement*—means that decision-makers infer that the "choice architect" deems it acceptable to stick with the default, at least if they are uncertain about the correct choice (Dinner et al. [2011]). Auditors likely are highly susceptible to this mechanism since the choice architect for the prepopulated format is an important authority figure (their firm). The second mechanism—*anticipated regret*—refers to decision-makers sticking because acceptance of the default, accomplished through *inaction*, leads to less future regret if their choice is wrong than does *actively* choosing an incorrect alternative (e.g., Ritov and Baron [1995]). Although auditors likely anticipate regret for incorrect choices of any sort since these choices can negatively affect performance evaluations, we expect they too will tilt toward anticipating greater regret for those due to inaction because inaction implies "same as last year" behavior, a mentality that likely feels natural (e.g., CAQ [2019]). The third mechanism—*cognitive costs—r*efers to decision-makers sticking with the default because it lowers their cognitive effort, and people typically prefer less effort to more (Kahneman [2011]). Sticking with the default and considering information related only to it is less effortful than also gathering information about and choosing among alternatives (Dinner et al. [2011]). This mechanism may be particularly important for risk assessment because standards require auditors to consider a large number of risk factors (PCAOB [2022]). Further, in choosing ratings for each risk factor, auditors often must deliberate between response categories that are similar, e.g., "medium-high" versus "high," a particularly effortful process (Thomson and Oppenheimer [2022]).

Because we hypothesize that all these mechanisms underlie auditors' stick-with-the-default behavior, our intervention targets them all. Consistent with more than one mechanism

driving default effects, Bonner et al. [2018] find that interventions designed to target only implicit endorsement are ineffective against prepopulation-created default effects. Further, because Bonner et al. [2018] and later research in the defaults literature (Van Gestal, Adriaanse, and De Ridder [2021]) find that *general* interventions are ineffective, ours is task-specific and detailed. It targets implicit endorsement by telling auditors that their firm would like them to focus on directional accuracy in their risk ratings and explains why the firm has this priority (i.e., the effectiveness and efficiency benefits), conveying that it is less acceptable to use uncertainty as a reason to stick. The intervention also removes one source of uncertainty. Specifically, it permits auditors to focus only on directional changes in risk, thus removing the uncertainty related to the exact rating to choose. The directional accuracy language also addresses anticipated regret by attempting to tilt this regret to be greater for rating errors that arise from *inaction via incorrectly sticking* than for those that arise from *action via incorrectly selecting an alternative*; it also removes selection of the wrong specific rating as a source of regret. Finally, the intervention addresses cognitive costs in that it limits any increases in cognitive effort by deleting one large source thereof—the deliberation among multiple ratings if more than one could be correct. Overall, we hypothesize an interaction in which the intervention reduces the negative effect of prepopulation on auditors' accuracy for changing risks.

We test our hypothesis in a 2 X 2 between-participants experiment, in which 111 staff auditors from multiple audit firms complete the risk assessment task from Bonner et al. [2018]. Consistent with Bonner et al. [2018], we manipulate whether auditors use prepopulated workpapers, where the spaces for the current year's risk ratings and evidence are filled in with the prior year's ratings and evidence, or non-prepopulated workpapers that are blank. All

auditors can refer to the prior year's ratings and evidence in a separate file. We also manipulate whether auditors receive the intervention, which includes the earlier mentioned elements.

To complete the task, auditors read current year client facts and provide ratings and evidence for 19 risk factors. Fourteen risk factors change in level from the prior year; for these, sticking with the default is incorrect. Our dependent variable is the number of times the auditor changes ratings for these risks in the correct direction. The other five factors are unchanged as to risk, such that sticking with the default is the correct choice; for these, Bonner et al. [2018] find that auditors using prepopulated (versus non-prepopulated) workpapers are *more accurate*. Inaccurate assessment of unchanging risks also can harm audit effectiveness and efficiency; thus, we include these factors to explore how our intervention affects accuracy for them.

We find the hypothesized interactive effect of prepopulation and the intervention on auditors' accuracy at changing risk ratings for risks that have changed. Absent the intervention, prepopulation has a significant negative effect on auditors' accuracy, replicating Bonner et al. [2018]. However, among auditors receiving the intervention, the effect of prepopulation is not significant. In fact, auditors using prepopulated workpapers who receive the intervention show accuracy similar to those using non-prepopulated workpapers. Auditors using non-prepopulated workpapers are not affected by the intervention, an important finding because it could have harmed them. That is. auditors using non-prepopulated (versus prepopulated) workpapers likely exert greater effort simply due to the workpaper format. Therefore, when receiving the intervention, these auditors might reach the point where the anticipated negative utility of effort outweighs the anticipated positive utility of accurate choices (Kool and Botvinick [2014, 2018]), causing them to seek ways to reduce effort, such as simply sticking for a subset of risk factors.

That the accuracy of auditors using non-prepopulated workpapers for changing risks is not harmed by the intervention rules out one unwanted plausible side effect of the intervention. Analyses of unchanging risks rules out another. Specifically, we find that the intervention does not harm the accuracy of auditors using prepopulated workpapers for these risks (there is no significant effect); this could have happened if auditors perceived an imperative from the intervention to almost always change their risk ratings. We also find that the intervention significantly improves the accuracy of auditors using non-prepopulated workpapers.

Finally, analyses of moderators—measures of auditors' beliefs that it is acceptable for them to stick when uncertain and their accuracy goals when performing their tasks—support that implicit endorsement and costliness, respectively, underlie default effects in risk assessment. The measures we use to examine the anticipated regret mechanism have no significant effects, however, possibly because they cannot cleanly capture inaction versus action. That is, given the nature of auditing, they also capture effectiveness and efficiency implications.

Our findings contribute to audit practice and research. Audit firms likely will find our intervention of interest, as auditors using prepopulated workpapers who also receive the intervention are invulnerable to the previously documented harmful effects of this workpaper format. In fact, we find that auditors using prepopulated workpapers under the guidance of the intervention show risk assessment accuracy that mirrors the accuracy of auditors using non-prepopulated workpapers. This aspect of our work is important because standardization, one reason firms likely prepopulate, has benefits for quality control. Indeed, the PCAOB emphasizes that standardized tools are a key intellectual resource for firms (PCAOB [2024]). However, the PCAOB also would find it concerning that prepopulated workpapers have this harmful effect, particularly given their concerns about risk assessment accuracy (PCAOB [2023a], [2023b]).

While it could be argued that the regulator could focus inspections more on this area, and PCAOB data suggests this is the case (PCAOB [2025]), inspectors could miss inaccuracies in ratings for risk factors that require knowledge of nuanced, client-specific information. Therefore, the regulator may view firms' use of our intervention as a reasonable solution for issues in this area. Another implication of our findings is that, if the intervention is applied across engagements, firm leadership could delegate the prepopulation decision to the lead audit partner for each engagement. Audit partners on engagements in which client risks change infrequently may prefer prepopulated workpapers because, while the intervention improves the accuracy of auditors using non-prepopulated workpapers for unchanging risks, those using prepopulated workpapers still are more accurate. Firms also could adapt our intervention for other tasks where prepopulation may have negative effects. For example, prepopulation could lead auditors not to consider new procedures that could detect fraud (Hoffman and Zimbelman [2009]).

For audit research, we contribute by building on Bonner et al. [2018] to outline the *multiple mechanisms* that underlie default effects stemming from prepopulation. We develop an intervention that addresses these mechanisms and mitigates the negative effect of prepopulation on auditors' accuracy for changing risks documented by Bonner et al. [2018]. Our intervention also increases the rating accuracy of auditors using non-prepopulated workpapers for unchanging risks, addressing this drawback of the non-prepopulated format identified by that study.

We also contribute to accounting research more broadly. Default effects go beyond audit settings. For example, financial reporting systems include defaults, such as recurring journal entries with which preparers work on a regular basis during interim periods. Financial statement errors could result if preparers do not adequately consider whether these defaults need adjustment. Our study provides a framework for future research to understand errors that may

result from defaults, along with a roadmap for developing interventions to address these effects. We also contribute to the defaults literature in economics and psychology by developing a task-specific intervention that mitigates undesired default effects. Few studies there have found effective interventions, and some that were successful would require significant resources. For example, one successful intervention for the negative effects of retirement plan defaults is detailed personal consultation (Blumenstock, Callen, and Ghani [2018]). Researchers in fields where devotion of such resources is unrealistic, e.g., medicine, may be inspired to explore interventions using our approach. We focus on the mechanisms that likely underlie default effects for a given task based on its characteristics and the context in which it is performed and create a practically feasible intervention that addresses them for the average decision-maker.

2. Background and Hypothesis Development

2.1 BACKGROUND

It is critical for auditors to make accurate risk assessment choices, as they directly affect audit effectiveness and efficiency (Allen et al. [2006]). Failing to increase ratings when risks have increased can harm audit effectiveness as auditors will not expand the scope of substantive testing and, thus, may miss material misstatements. Failing to decrease ratings when risks have decreased can harm efficiency as auditors will not scale back the scope of substantive testing. That is, firms could have conducted a high-quality audit with less testing. Efficiency can have economic implications outside the audit firm as well; recent research finds that efficient audit offices charge clients lower fees (Tan, Garg, Jeter, Kirubaharan, and Naiker [2025]).

Almost a decade ago, the PCAOB voiced concerns about harm to audit effectiveness from inaccurate risk assessment (PCAOB [2015]). Regulators recently revoiced such concerns, indicating that risk assessment inaccuracy may have persisted over time (SEC [2022]; PCAOB

[2023a]). The PCAOB classified risk assessment as a priority for future inspections (PCAOB [2023b]), and, indeed, a review of PCAOB data suggests a recent increase in deficiencies in this area (PCAOB [2025]). Risk assessment accuracy likely is challenging for inspectors to evaluate, however. For example, an audit team might miss information indicating that a particular risk rating should be changed from the prior year, but, unless the missed information is "top of mind" for inspectors because it affects many companies (e.g., macroeconomic changes), it might be difficult for them to detect the audit team's error. That is, to find missed client-specific risk information, inspectors might have to reperform most of the risk assessment, an impossibility given their limited time. Moreover, because of their mandate, external inspectors are focused on inaccuracies that harm audit effectiveness, while those that harm efficiency also are a matter of priority for audit firms and their clients. Thus, it is critical for audit firms to proactively identify the root cause of risk assessment inaccuracies and implement an effective intervention.

Recent research indicates that one cause of risk rating inaccuracy is that auditors tend to stick with prior year ratings when they occur as default ratings due to the use of prepopulated workpapers (Bonner et al. [2018]). While there are other factors that affect risk assessment accuracy (e.g., Allen et al. [2006]; Cassell, Dearden, Rosser, and Shipman [2022]), we believe that prepopulation is an especially critical factor because the defaults literature finds robust legitimating effects of defaults across a wide variety of decision contexts (e.g., Kesan and Shah [2006]). Moreover, as we discuss further below, we hypothesize that auditors may be particularly prone to default effects while performing the risk assessment task.

In this study, we examine the effects of an intervention that we design to mitigate the negative effects of prepopulation-created defaults. Because default effects are strong, it may seem that firms could "intervene" by requiring non-prepopulated workpapers. This solution is

undesirable, however, because the frequent use of prepopulated workpapers in practice, for risk assessment and other tasks such as planning substantive tests, suggests that firms reap benefits therefrom. Bonner et al. [2018] provide empirical evidence of one such benefit; for risk factors that *have not changed* from the prior year, auditors with prepopulated (versus non-prepopulated) workpapers are more accurate in their risk ratings, given sticking with last year is correct for these. Prepopulated workpapers also increase standardization, which helps firms achieve quality control. As the PCAOB notes, standardized documents promote consistency across engagements, and therefore are a key intellectual resource for firms (PCAOB [2024]). Related, Boland et al. [2024] conclude that audit firms are increasing standardization and automation in technology-based tools, particularly in audit planning, because they envision this increases both quality and efficiency.

2.2 HYPOTHESES DEVELOPMENT

In this section, we provide background on defaults and reiterate findings from Bonner et al. [2018]. We then consider theory illuminating *why* people tend to stick with defaults. Next, we explain how these theorized mechanisms particularly apply to auditors using prepopulated workpapers to perform risk assessment. We then develop an intervention, explaining *how* and *why* it addresses these mechanisms. We conclude by presenting our hypothesis that the intervention will reduce the negative effects of prepopulation (compared to non-prepopulation) on auditors' risk rating accuracy for risks that have changed.

2.2.1 Default Effects. Defaults are prevalent in society and the workplace, and literature in economics and psychology finds that decision-makers tend to stick with default options, despite the freedom to choose otherwise (Thaler and Sunstein [2021]). Defaults can materialize in dichotomous choices, where one option is the default and decision-makers must "opt out" to

choose the other option. Illustrating this, citizens are more likely to become organ donors when they must "opt out" (i.e., the default is to be a donor) versus "opt in" (i.e., the default is not to be) (Johnson and Goldstein [2003]). Defaults also can occur for *multiple alternative* choices, such as contribution rates in retirement plans. Studies in these settings show that an option is selected more than alternative options when it is the default, compared to when there is *no default designated* among the options (e.g., Madrian and Shea [2001]; Thaler and Benartzi [2004]).

Our setting maps to the latter scenario. For audit risk assessment, prepopulated workpapers have defaults of prior year ratings and supporting evidence filled into the spaces for the current year. Non-prepopulated workpapers have no defaults; they are blank, and auditors refer to prior year work in a separate file. Bonner et al. [2018] show that audit staff using prepopulated workpapers are less accurate in their risk ratings for risk factors that have changed from the prior year than are those using non-prepopulated workpapers. This occurs because auditors with prepopulated workpapers, consistent with other decision-makers, tend to stick with the default ratings. Consequently, auditors using prepopulated workpapers are less accurate at identifying changing risks than are auditors with non-prepopulated workpapers.

2.2.2 Theory Underlying Default Effects. In this section, we consider the mechanisms that the defaults literature posits underlie stick-with-the-default findings. The first is *implicit endorsement*, which refers to decision-makers inferring, perhaps nonconsciously, that a choice architect has selected the default because they endorse that choice or at least find it acceptable if the decision-maker is uncertain about the appropriate choice (e.g., Tannenbaum and Ditto [2011]); consequently, people who see the default tend to adopt it as their personal choice. For example, Madrian and Shea [2001], using 401(k) plan enrollment data from a large corporation, find that employees view the default as implicit advice vis-à-vis the "best" investment allocation.

As another example, McKenzie et al. [2006] conduct an experiment in which participants choose between reading "Summary A" and "Summary B." In one condition, Summary B is the *de facto* default because it is the only summary on participants' desks; participants can easily receive Summary A just by asking the experimenter, however. In the other condition, there is no default; both summaries are on the desk. Participants in the default condition choose Summary B more than those in the no-default condition and report greater agreement that they do so because the experimenter wanted them to read Summary B.

A second mechanism is *anticipated regret*, which can be reduced by a default because it allows for *inaction* via passive acceptance of the default. By contrast, switching from a default requires the decision-maker to take *action* to choose an alternative. Decision-makers likely experience less regret if their choice is wrong when they stick with the default than when they change. This occurs because sticking creates an "error of omission," which decision-makers generally prefer to the "error of commission" created by changing to an alternative (Ritov and Baron [1995]; Willis [2013]). To illustrate, Shevchenko et al. [2014] study two defaults. For the first one, participants continue to work on the same task, unless they "opt out" by indicating a desire to work on a new task; for the second, participants move to a different task, unless they "opt out" by indicating a desire to continue working on the same task. The authors view working on the same task as a form of inaction and working on a different task as action, and find strong sticking effects for both defaults. They conclude that, because even a default that involves action in the form of changing from the *given task* has strong effects, the desirability of the inaction inherent in *sticking with a default* likely is a strong mechanism mediating default effects.

A third mechanism underlying default effects is *cognitive costs*, which are higher when a decision-maker changes from the default versus sticks with it. This occurs because cognitive

effort is costly, and it requires more cognitive effort to consider information about *both the default and alternatives* than about only the default. Cognitive effort invokes costs in the form of negative utility due to the demands of "pushing [oneself]" (Inzlicht, Shenhav, and Olivola [2018, 337]; Kool and Botvinick [2018]), as well as negative emotions (Hockey [2013]). Perhaps unsurprisingly, people prefer less effort to more (Kahneman [2011]).[1] Madrian and Shea [2001] postulate an example of a cognitive cost; they suggest that employee contemplation of the high costs of considering the details of multiple alternative retirement plan options plays a role in employees sticking with the default plan. Blumenstock et al. [2018] report more direct evidence on cognitive costs from a field experiment related to retirement plan defaults. The study's design rules out implicit endorsement and finds that an intervention designed to reduce cognitive costs (detailed financial consultation) reduces choice of the default. Further supporting cognitive costs as a mechanism, employees tend to switch from the default most when they receive consultant assistance with effortful calculations.

2.2.3 Applying Theory to Auditing and Developing an Intervention. We posit that all three mechanisms underlie the effects of prepopulation-created defaults on auditors' risk rating accuracy given characteristics of the default, the audit environment, and the risk assessment task. Starting with *implicit endorsement*, auditors using prepopulated workpapers, like other decision-makers, may consciously or nonconsciously infer from the incorporated default ratings that sticking is acceptable to the "choice architect," at least if they are uncertain about current year ratings. Auditors could be uncertain as to whether a current year fact warrants a change in the rating from last year, such as contemplating whether the fact is of enough significance; they also could be uncertain about the specific rating to select if they contemplate a change. Giving

[1] The defaults literature also considers the costs of physical effort. We expect this is less applicable here, as auditors with non-prepopulated workpapers could copy and paste prior year information into current year workpapers.

particular gravity to the default for auditors is the fact that the choice architect is an authority figure, e.g., their firm, to whom they are accountable for performance of the task.

Second, *anticipated regret*, irrespective of whether it occurs through inaction or action, likely is at play in audit decisions because making mistakes can affect an auditor's performance evaluation and reputation. Despite this general concern about being wrong, however, we predict that auditors, like other decision makers, feel less anticipated regret about being wrong through the inaction of sticking with the default than being wrong through the action of changing from the default. This would occur because passive acceptance of the default provided by prepopulated workpapers means that they also are exhibiting "same as last year" behavior, a mentality that likely feels natural to auditors (CAQ [2019]; PCAOB [2023]).

Third, regarding *cognitive costs*, like other decision-makers, auditors using prepopulated workpapers might stick with the default because it allows for less cognitive effort, compared to choosing an alternative. Sticking with the default rating requires little input and processing effort because auditors can support this choice with a straightforward search for a few current year facts that indicate no change in risk. By contrast, if auditors consider whether an alternative is warranted, they must carefully read and combine a larger amount of current-year information and compare it to prior year information. We posit that auditors will especially welcome opportunities to reduce effort when performing risk assessment, for at least two reasons. First, the relevant standard prescribes a long list of individual risk factors to which auditors must apply the above effortful process (PCAOB [2022]). Second, if auditors are effortfully considering alternatives to the default and decide that a change is warranted, they then must exert further effort to determine the appropriate rating if more than one alternative to the default seems acceptable. The latter is especially effort-consuming if the alternatives are quite similar (see

Thomson and Oppenheimer [2022]). For instance, an auditor may have to decide between a rating of "4" versus "5" if she suspects that a risk has increased in level from last year, and the rating last year was a "3." Anticipating this extra effort might lead auditors to avoid making any decision to change, that is, to stick with the default (Thomson and Oppenheimer [2022]).

Based on these arguments, implicit endorsement, anticipated regret, and cognitive costs all likely underlie default effects in risk assessment that arise from the use of prepopulated workpapers. Limited previous research also supports the existence of multiple mechanisms. Although not the main focus of the study, Bonner et al. [2018] report that two interventions structured to target implicit endorsement, a key mediator of default effects (Jachimowicz, Duncan, Weber, and Johnson [2019]), are unsuccessful at reducing the negative risk rating accuracy effects of prepopulation.[2] As such, we structure an intervention to target *each of the three mechanisms*. Our intervention does so in a *detailed, context-specific* manner, which is important for at least two reasons. First, of the limited interventions examined in the defaults literature, including the ones designed and tested by Bonner et al. [2018], a substantial number have been ineffective. For example, monthly telephone surveys asking employees about their financial behaviors and understanding of their retirement plan program did not reduce default effects in that setting, likely because the intervention did not target the higher cognitive costs employees would incur if considering an alternative to the default (Blumenstock et al. [2018]). Second, interventions that are general now are known to be ineffective against defaults. For instance, the general intervention in Bonner et al. [2018] was ineffective, and subsequent research in the defaults literature finds that general instructions to deliberate over one's choices

[2] The first, *general*, intervention reminded auditors that "it is not sufficient for auditors to merely rely on their perceived knowledge of the industry or information obtained from *prior audits*…" (emphasis added); it also cautioned auditors to exercise professional skepticism. The second was similar but provided a *task-specific* example.

do not mitigate default effects (Van Gestal et al. [2021]). [3] Even task-specific interventions may be ineffective; the task-specific version of the intervention in Bonner et al. [2018] was no more effective than the general version. We note, however, this could have occurred because the intervention targeted only implicit endorsement or perhaps also needed to be even more detailed.

There are a few successful interventions in the defaults literature, but we cannot adapt them to auditing because they could impair effectiveness. For example, asking decision-makers if they would like to make a different decision than the default can reduce default effects (Sakaguchi et al. [2022]), but asking auditors if they would like to choose another risk rating if they initially stick could produce excessive changing that leads to inaccuracy for unchanged risks; effectiveness would be compromised if auditors decrease ratings for these factors, and efficiency would be compromised if auditors increase ratings for these factors. As another example, choice architects at some companies have set higher default contribution rates for employee retirement plans to guide employees to make more appropriate retirement savings choices for themselves (Thaler and Sunstein [2021]). Yet, in risk assessment, changing the default ratings from the prior year ratings to, say, all be the lowest or highest rating, is undesirable. The correct risk rating depends on current year facts for each risk factor. In the former scenario, audit effectiveness could be compromised for risk factors where current risks are not that low, and in the latter scenario, audit efficiency could be compromised for risk factors where current risks are not that high. We now turn to defining the elements of our task-specific intervention.

Our intervention targets implicit endorsement by counteracting the idea that sticking is acceptable to the choice architect—in our study, the audit firm—if the auditor is uncertain.

[3] Thus, broad interventions in auditing to increase skepticism, such as priming intrinsic motivation (Kadous and Zhou [2019]) or making professional identity salient (Bauer [2015]), likely would be ineffective for default effects.

Specifically, the intervention communicates the audit firm's priority of directional accuracy in risk ratings and explains the underlying reason, i.e., the effectiveness benefits of correctly increasing risks and efficiency benefits of correctly decreasing risks. This communication conveys that it is less acceptable for auditors to use uncertainty as a reason to stick with the default. By giving auditors permission to focus on accuracy in *directional changes*, the intervention also removes a potential source of uncertainty, that is, the specific risk rating to choose if more than one response category is consistent with a given directional change.

These instructions also address anticipated regret. That is, communicating a priority for directional accuracy and the benefits for the audit firm of the auditor taking action to correctly increase and decrease risk ratings is intended to change how they think about regret. Specifically, the communication is intended to tilt their anticipated regret about the possibility of being inaccurate to be relatively greater for the scenario of inaccuracy due to *inaction* (i.e., incorrectly sticking with the default rating) than inaccuracy due to *taking action* (i.e., incorrectly choosing an alternative rating). As above, the focus on directional accuracy also is intended to remove the element of regret that would derive from choosing the wrong numerical rating.

The directional accuracy element targets the cognitive costs mechanism as well. Auditors who are envisioning considering alternative ratings for each individual risk factor (of the nineteen in our task, which still is only a subset of those in the standard) likely are anticipating extensive effort. As mentioned, people tend to prefer less effort to more (e.g., Kahneman [2011]). Anticipation of this extensive effort could act as a hurdle to changing from the default then, in that it causes sticking with the default to appear attractive in comparison. By permitting a focus on directional changes only, we expect the intervention will cause auditors to anticipate *a more moderate amount* of effort to consider alternatives to the default than they would without

the intervention; that is, it removes the effort associated with choosing among specific ratings and reduces the aforementioned hurdle. Overall, by targeting the three mechanisms that we theorize underlie default effects in the risk assessment task, we pose the following hypothesis:

> *Hypothesis:* The negative effect of prepopulation on auditors' risk rating accuracy for changing risks will be weaker for auditors receiving the intervention, compared to auditors not receiving the intervention.

This prediction is not without tension. First, to the extent a "same as last year" mentality feels natural, the intervention may not shift auditors' anticipated regret to be stronger for *failing to take action* when necessary (i.e., to change the risk rating). Second, our hypothesis implies an ordinal interaction but one of disordinal form could result if the intervention *harms* the accuracy of auditors using *non-prepopulated* workpapers. While the intervention has one element to ease cognitive costs, auditors still must exert more effort to fully apply it for each risk factor. Non-prepopulated workpapers also require greater effort than the prepopulated format. This combination of effort requirements means that the accuracy of auditors using non-prepopulated workpapers could be *harmed* by the intervention because they find that the negative utility of the combined effort outweighs the positive utility of performing the task well. If so, they may resort to low- effort tactics, e.g., simply sticking with last year for a subset of risk factors.

3. Method

3.1 PARTICIPANTS

Participants are 111 audit staff (mean experience of 13.9 months) from multiple audit firms.[4] The Center for Audit Quality facilitated study administration, and participants completed the study online via Qualtrics. We employ the risk assessment task from Bonner et al. [2018], and note that staff perform risk assessments in practice (Saunders, Keune, and Hawkins [2023]).

[4] We obtained Institutional Review Board approval for the study.

3.2 DESIGN AND MANIPULATIONS

Our study includes two manipulated independent variables that are fully crossed. The first variable is the prepopulation manipulation used by Bonner et al. [2018]. We vary whether the current year workpapers are prepopulated with the prior year's conclusions. Auditors in the prepopulated condition complete a workpaper that contains last year's risk ratings and evidence in the spaces for the current year's risk ratings and evidence. If auditors with prepopulated workpapers take no action for a given risk factor, the prior year's rating and evidence become the current year's rating and evidence. Auditors in the non-prepopulated condition complete a workpaper that is blank. These auditors must actively indicate a risk rating and fill in evidence to the current year workpaper. All auditors receive the prior year workpaper in a separate file.[5]

The second manipulated variable is whether the auditor receives the intervention. Auditors receiving the intervention read the following statement in the task instructions:

> *Your audit firm offered the following additional guidance:*
> - *Firm leadership would like you to just focus on getting the direction of the risk change correct (i.e., whether it has increased, decreased, or not changed from the prior year), rather than worrying about the specific numerical rating. They highlighted the benefits of doing so:*
>   - ***Increasing** a risk rating (when the risk has actually increased) increases audit effectiveness because this will change the nature, timing, and/or extent of procedures performed (i.e., in a way that better responds to the risk), making it more likely that the audit team will detect any material misstatements that exist.*
>   - ***Decreasing** a risk rating (when the risk has actually decreased) increases audit efficiency because this will change the nature, timing, and/or extent of the procedures performed (i.e., in a way that uses fewer resources to respond to the risk), making it more likely that the audit team will finish the audit in fewer hours than last year.*

[5] Participants with prepopulated (non-prepopulated) workpapers are informed that the prior year's rating and evidence have (or have not) been copied over into the current year workpapers. As in Bonner et al. [2018], participants are informed that their audit firm set this policy because it strikes the right balance between effectiveness and efficiency – consistent with the earlier discussion of audit firms and regulators prioritizing standardization and consistency across engagements (PCAOB [2024]). This design choice also is intended to reduce noise; otherwise, for instance, participants might infer that the choice was made at the audit-team level because a *particular client's* risks tend to change little, which could bias participants toward a belief it is acceptable to stick.

3.3 TASK, DEPENDENT VARIABLE, AND OTHER MEASURES

3.3.1 Risk Assessment Task. Auditors are asked to envision that they are on an audit team and have been asked to complete a risk assessment task. They view a brief example of a completed risk assessment workpaper and read an excerpt (para. 0.03-0.07) from the relevant audit standard (PCAOB [2022]). Participants then are provided brief task instructions and asked to start the task. They read background information, which explains current year client circumstances that are relevant to the various risk factors. Then, participants are told that, consistent with auditing standards, they have access to the prior year's workpaper. The prior year workpaper includes the prior year ratings, along with the evidence that explains the facts that supported each rating. Following, they read the manipulated language about prepopulation and the intervention. Participants complete the task, which entails selecting a risk rating from Low "1" to High "5" (with "3" labeled as Medium) for 19 risk factors – eight that have increased in risk from the prior year, six that have decreased from the prior year, and five that are unchanged; they also are asked to document evidence supporting each risk rating.

3.3.2 Dependent Variable and Other Measures. Consistent with Bonner et al. [2018], we analyze changing and unchanging risks separately. We do so because sticking with the default produced by prepopulation is *incorrect for changing risks*, but *correct for unchanging risks*. Our hypothesis, and thus primary dependent variable, relates to auditors' accuracy for the *changing* risk factors. We measure the number of times, for the 14 changing risks, the auditor provides a directionally correct rating – that is, increases the rating from the prior year if the risk increased, or decreases the rating from the prior year if the risk decreased.

We include the five unchanging risk factors to test for a potential undesirable "side effect" of the intervention on audit effectiveness. Because of the intervention's strong emphasis

on correctly increasing and decreasing risks, auditors might perceive an imperative to almost always change ratings. If this occurs, and auditors receiving the intervention more frequently change risk ratings for the unchanging risk factors than they do without it, the intervention would impair accuracy for these risks. This "side effect" could occur for auditors irrespective of whether they are using prepopulated or non-prepopulated workpapers. Finally, we measure moderators such as auditors' accuracy goals to illuminate the mechanisms underlying default effects, and noise variables such as pre-existing task knowledge.

4. Results

4.1 MANIPULATION CHECKS

To evaluate attention to our *Prepopulation* manipulation, we ask auditors "to confirm your understanding of the structure of the workpapers, were the prior year's numerical risk ratings and evidence documentation automatically copied over into your current year workpapers?" 82 of 111 participants (73.9 percent) indicate the correct response; responses differ across *Prepopulation* conditions ($\chi^2 = 25.77$; $p < 0.001$). For the *Intervention* manipulation, we ask "to confirm your understanding of the experimental materials, were you explicitly told by your audit firm that you should just focus on getting the direction of the risk change from the prior year correct (i.e., whether it has increased, decreased, or not changed from the prior year)?" 83 of 111 participants (74.8 percent) indicate the correct response, and responses differ across *Intervention* conditions ($\chi^2 = 28.48$; $p < 0.001$). Results for hypothesis tests are robust to dropping the observations of participants who answered one or both questions incorrectly.

4.2 TEST OF HYPOTHESIS

Our hypothesis predicts that the negative effects of prepopulation on auditors' accuracy at changing ratings for risks that have changed from the prior year will be weaker among auditors

receiving the intervention, compared to auditors not receiving it. Table 1, Panel A reports descriptive statistics by experimental condition, that is, across the *Prepopulation* and *Intervention* manipulated variables. The pattern of means reflects our hypothesis (see Figure 1). We create an Analysis of Variance (ANOVA) with the independent variables of *Prepopulation,* indicating whether the participant's workpapers were prepopulated or non-prepopulated; *Intervention*, indicating whether the participant received the intervention; and the interaction term. *Accuracy for Changing Risks* is the dependent variable. Panel B displays the results.[6]

[Insert Table 1 and Figure 1 About Here]

Replicating Bonner et al. [2018], we find a significant main effect of *Prepopulation*, indicating that auditors using prepopulated workpapers are less accurate in their risk ratings for changing risks, versus auditors using non-prepopulated workpapers ($p = 0.001$). We find a significant main effect of the *Intervention* ($p = 0.019$), indicating that auditors receiving the intervention are more accurate than auditors not receiving it. Supporting our hypothesis, there is a marginally significant *Prepopulation* X *Intervention* interaction ($p = 0.063$) in the ordinal form. The nature of the interaction suggests that our intervention reduces the observed negative effect of prepopulation. Simple effects analyses (Panel C) show a negative effect of *Prepopulation* among auditors not receiving the *Intervention* ($p < 0.001$); among those receiving the intervention, the effect is not significant ($p = 0.124$). Importantly, the accuracy of auditors using prepopulated workpapers with the intervention is indistinguishable from the accuracy of auditors using non-prepopulated workpapers with no intervention (two-tailed $p = 0.386$; untabulated).

We now consider effects of the intervention separately among auditors using prepopulated versus non-prepopulated workpapers. Reflecting the above findings, auditors using

[6] Unless otherwise specified, we use one-tailed *p*-values to reflect directional predictions.

prepopulated workpapers who receive the *Intervention* perform significantly better than those with no *Intervention* ($p = 0.006$; Panel C). The ordinal form of our interaction provides preliminary evidence that the potential "side effect" of the intervention for auditors using non-prepopulated workpapers does not materialize. In support, we find that the *Intervention* has no significant effect on auditors using non-prepopulated workpapers ($p = 0.347$; Panel C). This result is important because firms may want to retain the option of using this format while also employing the intervention and, in general, clearly would prefer an intervention that would not impair audit effectiveness. In sum, our intervention not only reduces the negative effect of prepopulation on risk rating inaccuracy for changing risks, consistent with our hypothesis, but makes auditors invulnerable to it. The intervention also does not reduce the risk rating accuracy of auditors with non-prepopulated workpapers, leaving firms the option of choosing the non-prepopulated format for any benefits it may provide for a specific engagement.

4.3 ANALYSES OF AUDITORS' ACCURACY FOR UNCHANGING RISKS AND POSSIBLE SIDE EFFECTS

In this section, we test for the presence of another potential side effect of the intervention. Specifically, auditors might perceive an imperative from the intervention to almost always change their risk ratings; this could impair auditors' accuracy for unchanging risks because the prior year's rating is the accurate choice for the current year. To examine this issue, we create the variable *Accuracy for Unchanging Risks*, which is the number of times, for the five unchanging risks, the auditor selects the same rating as last year. Table 2, Panel A and Figure 2 show descriptive statistics for this variable by condition.

[Insert Table 2 and Figure 2 About Here]

We run the ANOVA used to test our hypothesis, but with *Accuracy for Unchanging Risks* as the dependent variable. Consistent with Bonner et al. [2018], we find a significant main effect of *Prepopulation* ($p < 0.001$), indicating that auditors' *Accuracy for Unchanging Risks* is greater

when they use prepopulated workpapers, versus non-prepopulated. We also find a marginally significant *Prepopulation* X *Intervention* interaction (two-tailed $p = 0.077$; Panel B); when interpreted in light of the visual pattern, the interaction suggests that the effect of *Prepopulation* is weaker among auditors receiving the *Intervention*. Following, we examine the simple effects of the *Intervention*, which shed insight on whether the side effect—of auditors perceiving an imperative from the intervention to almost always change their ratings—emerges.

Simple effects analyses (Panel C) show that the effect of the *Intervention* on accuracy for unchanging risks is not significant for auditors using prepopulated workpapers ($p > 0.500$), an important finding as it could have been negative. Among auditors using non-prepopulated workpapers, those receiving the intervention *perform better* than those not receiving it, and this effect is significant (two-tailed $p = 0.025$). This finding suggests that not only did the side effect not emerge for the latter group of auditors, but the intervention appears to help them, perhaps because it also mentions that auditors should not change ratings for risks that are unchanged.

We also examine the simple effects of *Prepopulation*. Unsurprisingly, in the no *Intervention* condition, the effect of *Prepopulation* is significant and positive ($p < 0.001$). The effect of *Prepopulation* also is positive and significant in the *Intervention* condition ($p = 0.041$). Therefore, while the intervention aids auditors using non-prepopulated workpapers in being more accurate for unchanging risk factors, the benefit of prepopulation for these types of risk factors persists with the intervention. This finding suggests that, if a client's risks tend to stay the same year over year, firms may consider prepopulated workpapers to be an advantageous format.

4.4 ANALYSES OF MODERATORS TO ILLUMINATE MECHANISMS

In this section, we examine the effects of three potential moderators, each theoretically allowing for illumination of a mechanism that we hypothesize underlies our results (see Asay et

al. [2022]).[7] We start with *implicit endorsement*, which we predict will be moderated by auditors' belief that it is acceptable to stick with last year when they are uncertain. We expect that auditors with higher intrinsic levels of this belief will more readily infer from prepopulation that their firm finds it acceptable to stick with last year, and in turn, be most aided by the intervention in counteracting this idea. *Acceptable to Stick* is auditors' agreement (Likert Scale from 1-7) that "When performing the risk assessment task, I felt like it was acceptable, if I was uncertain, to stick with the prior year risk rating." We augment the model used to test our hypothesis with the *Acceptable to Stick* independent variable and related interactions.

Table 3, Panel A reveals a marginally significant, positive *Prepopulation* X *Intervention* X *Acceptable to Stick* interaction on *Accuracy for Changing Risks* ($p = 0.088$). For auditors with higher levels of *Acceptable to Stick* (i.e., one standard deviation above the mean), there is a significant, positive *Prepopulation* X *Intervention* interaction ($p = 0.023$; Panel B); for those with a medium level of this belief (i.e., at the mean), there is a marginally significant, positive interaction ($p = 0.072$). For these two types of auditors, the intervention reduces the negative effect of prepopulation. Indeed, with no *Intervention*, there is a significant, negative effect of *Prepopulation* on *Accuracy Changing Risks* (larger $p = 0.002$; Panel C) for these auditors, but the effect is not significant among those who receive the *Intervention* (smaller $p = 0.166$). By contrast, and as expected, the interaction is not significant for auditors who have a lower level of the *Acceptable to Stick* belief (i.e., one standard deviation below the mean; $p > 0.500$).

[Insert Table 3 About Here]

We now consider *anticipated regret*. We expect that auditors more concerned with the ramifications of *incorrectly changing* risk ratings (versus *incorrectly sticking*) will be most

[7] We confirm that the moderator measures are not significantly correlated with our manipulations or their interaction.

susceptible to this mechanism. We measure concern about incorrectly *changing* using auditors' responses to two questions (Likert Scale 1-7). We elicit agreement that "I worried about what would happen to audit efficiency if I increased these risk ratings (e.g., that hours might go over budget)" (i.e., for risk factors they thought may have increased) and "I worried about what would happen to audit effectiveness if I decreased these risk ratings (e.g., that we might miss a misstatement)" (i.e., for risk factors they thought may have decreased). We measure concern about incorrectly *sticking* using auditors' response to two questions. We elicit agreement that "I worried about what would happen to audit effectiveness if I did not increase these risk ratings (e.g., that we might miss a misstatement)" (i.e., for risk factors they thought may have increased) and "I worried about what would happen to audit efficiency if I did not decrease these risk ratings (e.g., that hours might go over budget)" (i.e., for risk factors that they thought may have decreased). However, none of these individual measures is a moderator (three-way interaction p's > 0.500). To explore why, we conduct a factor analysis with the four questions. Two factors emerge; instead of the incorrect changing (i.e., action) and incorrect sticking (i.e., inaction) questions loading together, the efficiency and effectiveness questions load together.[8] Thus, our measures appear not to cleanly capture the construct of action versus inaction due to the necessary conflation of this construct with the efficiency/effectiveness construct.

Finally, we turn to *cognitive costs*, which we predict will be moderated by auditors' *Accuracy Goal*. We expect that auditors with a lower accuracy goal, because they tend to prioritize a low-effort path, will be most susceptible to this mechanism. Absent the intervention—when envisioning the extensive effort required to choose a rating (e.g., for a prior year rating of "2" that has likely increased, contemplating which of the three higher rating

[8] We confirm (untabulated) that neither combined measure is a moderator (smaller three-way interaction p of 0.231).

categories is suitable)—we expect auditors with a lower accuracy goal will experience aversive thoughts and emotions. As a result, they will stick with the default, a low-effort path, to neutralize the aversive thoughts and emotions. The intervention, however, establishes a path to completing the task that encourages focusing on directional accuracy, and not specific ratings. Because this path is less effortful than the above path of focusing on "correct ratings" for each risk factor, we expect that receiving the intervention will reduce the likelihood that they will experience the aversive thoughts and emotions. Altogether, we expect the intervention to aid these auditors in being less susceptible to the negative effects of prepopulation by affecting costliness. By contrast, we expect auditors with a higher accuracy goal will be less susceptible to the effects of prepopulation, and, therefore, have less need for the intervention. That is, consistent with psychology theory (Kunda [1990]), we expect they will place more weight on the positive utility of performing well on the task, making them less daunted by the thought of exerting extensive effort to achieve this desired outcome.

We measure auditors' accuracy goal using four measures. The first, professional identity, is auditors' selection from one of seven images of two overlapping circles that reflect the self and the profession, ranging from no overlap to nearly overlapping (Bauer [2015]).[9] Additional measures of the goal include auditors' agreement (Likert Scale from 1-7) that (1) "It was important to me to be accurate in my performance on the risk assessment task," and (2) I felt like my firm wanted me to be accurate on these risk ratings." We also ask about accuracy goals more generally using agreement (Likert Scale from 1-7) with the statement: "On my audit engagements, I try to make sure the client reports the most accurate numbers possible." The four

[9] Bonner et al. [2018] use professional identity as a proxy for accuracy goal and report that auditors with higher professional identity are less susceptible to the sticking response and its negative effect on accuracy. We measure *Accuracy Goal* using the professional identity proxy combined with more direct measures to obtain a more valid measure; however, we obtain inferentially similar results when using only the professional identity measure.

measures load on one factor (eigenvalue = 2.16; smallest loading = 0.58). Cronbach's alpha is 0.68, which is acceptable for newly developed measures (Nunnally and Bernstein [1994]).

The model reveals a significant, negative *Prepopulation* X *Intervention* X *Accuracy Goal* interaction ($p = 0.049$; Panel D). For auditors with a lower *Accuracy Goal*, there is a significant, positive *Prepopulation* X *Intervention* interaction ($p = 0.016$; Panel E); for auditors with a medium *Accuracy Goal*, the interaction is marginally significant and positive ($p = 0.070$). With no *Intervention*, there is a significant, negative effect of *Prepopulation* on *Accuracy Changing Risks* for these auditors (larger $p = 0.001$; Panel F), but the effect is not significant among those who receive the *Intervention* (smaller $p = 0.169$). Overall, the intervention reduces the negative effect of prepopulation for these two types of auditors. Finally, as expected, for auditors with a higher *Accuracy Goal*, the interaction is not significant ($p > 0.500$).

5. Discussion and Conclusions

In this study, we develop a detailed, task-specific intervention that targets the three mechanisms we posit underlie the negative effects of prepopulation-created defaults on auditors' accuracy at identifying changes in their clients' risks: implicit endorsement, anticipated regret, and cognitive costs. In an experiment with audit staff completing a risk assessment task, we find that our intervention reduces these negative effects. Encouragingly, the intervention causes auditors using prepopulated workpapers to show accuracy that mirrors the accuracy of auditors using non-prepopulated workpapers; further, it does not harm the accuracy of auditors using non-prepopulated workpapers, ruling out one possible concerning side effect.

Additional analyses—of auditors' accuracy at keeping the prior year's rating for risks that are unchanged in the current year—rule out other side effects that could have occurred if auditors perceived an imperative from the intervention to almost always change their risk ratings.

Specifically, for these unchanging risks, we find that the intervention does not decrease the accuracy of auditors using prepopulated workpapers, and it *improves* the accuracy of auditors using non-prepopulated workpapers. Finally, in additional analyses we find moderating effects of measures of auditors' intrinsic beliefs that it is acceptable for them to stick when uncertain and their accuracy goals when performing their tasks; these results support that implicit endorsement and costliness, respectively, are at work in default effects on risk assessment.

Our work contributes to audit practice and research. For audit practice, given the direct implications of risk assessment for audit effectiveness and efficiency (Allen et al. [2006]), and that risk assessment is a matter of priority for both the SEC and PCAOB (PCAOB [2023b]; SEC [2023]), we expect that firms and regulators will find our intervention of interest. Implementing it would allow firms to retain the prepopulated format and, thus, at least some of its perceived benefits such as standardization, while eliminating the troublesome outcome that auditors inaccurately stick with defaults for changing risks, which arises from this format choice. Because regulators both acknowledge the benefits of standardized audit support systems and likely have difficulty efficiently monitoring the quality of risk assessment, they may view firm use of our intervention as a reasonable solution for problems in this area. On the other hand, should firms wish to place the prepopulation decision at the team level, our intervention remains valuable in that it increases the risk rating accuracy of auditors using non-prepopulated workpapers for risks that have not changed. Also important, receiving the intervention does not decrease these auditors' accuracy for risks that have changed, and thus does not introduce side effects.

Our work contributes to audit research by developing an intervention to target the three mechanisms that we posit underlie undesired default effects from prepopulated workpapers. In doing so, we build on work by Bonner et al. [2018] who empirically demonstrated these effects,

and also found that two interventions targeted at implicit endorsement were ineffective against the strong default effects observed there. Our work also contributes to the broader accounting literature. As Bonner et al. [2018] discuss, stickiness is prevalent in other accounting settings, including budgeting, debt covenants, and executive compensation (*The Economist* [2009]; Kahan and Klausner [1997]; Indjejikian, Matejka, Merchant, and Van Der Stede [2014]). Prepopulated documents exist in those settings, and therefore, the mechanisms of implicit endorsement, anticipated regret, and cognitive costs of decision-making could be at work if stickiness is caused by the use of these documents. As another example, it seems conceivable that firms prepopulate current year disclosure content with the prior year's disclosures, and this could be a driver of the lengthy, boilerplate nature of disclosures that tend to persist (Cazier, McMullin, and Treu [2021]). Overall, future archival work that observes stickiness in outcomes can consider our theoretical mechanisms, and application of default theory to accounting settings, to consider whether defaults could be at work. Future research can directly test for the presence of default effects in such settings, by conducting controlled experiments in which decision-makers can be randomly assigned to receive prepopulated or non-prepopulated documents. Future research also can use our approach to structure and test the effectiveness of interventions that target the mechanisms hypothesized to underlie any default effects, with these hypotheses being based on a detailed understanding of the focal task, choice architect, and decision-makers.

Our work also contributes to the extensive psychology and economics literature on defaults. Only a limited number of studies have examined interventions, and of those, only a few have been successful. Our study may inspire researchers in other areas, for instance medicine and retirement plans, to use the roadmap we lay out to create task-specific interventions that reduce undesired default effects by targeting the appropriate mechanisms. Some of the successful

interventions studied previously, such as personal consultation for employees contemplating retirement plan choices (Blumenstock et al. [2018]), could be prohibitively costly in real-world settings. Altogether, our intervention may inspire researchers in those other fields to develop successful interventions that might be more practically implementable than prior solutions.

Limitations of our study provide opportunities for future research. First, because we use audit staff participants, our results may not generalize to more experienced auditors. Future research could examine prepopulation effects, as well as the effects of the intervention, for auditors with greater seniority. Interventions might need to be different, for instance, if these auditors have developed a stronger tendency to "SALY" with experience, e.g., as a result of developing habits (Bonner, Kadous, and Majors [2022]). Second, our measures of auditors' anticipated regret related to inaction versus action do not moderate our results as expected. Future research can consider other measures for this construct, ones that do not also invoke implications for efficiency and effectiveness, the factor we believe explains the lack of results. Third, we select the audit firm as the choice architect; while we expect our results to generalize to other choice architects who are similarly viewed as an authority figure, future research can explore this issue. Fourth, our intervention and audit task pertain to audit risk assessment. Exploring effects of prepopulation in different audit tasks, the underlying mechanisms involved, and developing interventions to target those mechanisms is an important avenue for future research.

REFERENCES

ALLEN, R.; D. HERMANSON; T. KOZLOSKI; and R. RAMSAY. "Auditor Risk Assessment: Insights from the Academic Literature." *Accounting Horizons* 20 (2006): 157-177.

ASAY, H. S.; R. GUGGENMOS; K. KADOUS; L. KOONCE; and R. LIBBY. "Theory Testing and Process Evidence in Accounting Experiments." *The Accounting Review* 97 (2022): 23-43.

BAUER, T. "The Effects of Client Identity Strength and Professional Identity Salience on Auditor Judgments." *The Accounting Review* 90 (2015): 95-114.

BLUMENSTOCK, J.; M. CALLEN; and T. GHANI. "Why Do Defaults Affect Behavior? Experimental Evidence from Afghanistan." *American Economic Review* 108 (2018): 2868-2901.

BOLAND, C.; N. GALUNIC; and M. SHERWOOD. "The Purposes and Functions of Audit Firms' Technology-Based Audit Tools." *Working Paper*, Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4258373

BONNER, S.; K. KADOUS; and T. MAJORS. "A Habit Strength-Based Explanation for Auditors' Use of Simple Cognitive Processes for Complex Tasks." *The Accounting Review* 97 (2022): 107-129.

BONNER, S.; T. MAJORS; and S. RITTER. "Prepopulating Audit Workpapers with Prior Year Assessments: Default Option Effects on Risk Rating Accuracy." *Journal of Accounting Research* 56 (2018):1453-1481.

CASSELL, C.; S. DEARDEN; D. ROSSER; and J. SHIPMAN. "Confirmation Bias and Auditor Risk Assessments: Archival Evidence." *AUDITING: A Journal of Practice & Theory* 41 (2022): 67-93.

CAZIER, R.; J. MCMULLIN; and J. TREU. "Are Lengthy and Boilerplate Risk Factor Disclosures Inadequate? An Examination of Judicial and Regulatory Assessments of Risk Factor Language." *The Accounting Review* 96 (2021): 131-155.

CENTER FOR AUDIT QUALITY (CAQ). "Professional Judgment Resource." Available at: https://www.thecaq.org/wp-content/uploads/2019/03/professional-judgment-resource.pdf

DINNER, I.; E. JOHNSON; D. GOLDSTEIN; and K. LIU. "Partitioning Default Effects: Why People Choose Not To Choose: Correction to Dinner et al. (2011)." *Journal of Experimental Psychology: Applied* 17 (2011): 432-432.

THE ECONOMIST. "Zero-Base Budgeting." 2009. Available at: https://www.economist.com/node/13005039.

HOCKEY, R. *The Psychology of Fatigue*. Cambridge, United Kingdom: Cambridge University Press, 2013.

HOFFMAN, V., and M. ZIMBELMAN. "Do Strategic Reasoning and Brainstorming Help Auditors Change Their Standard Audit Procedures in Response to Fraud Risk?" *The Accounting Review* 84 (2009): 811-837.

INDJEJIKIAN, R.; M. MATEJKA; K. MERCHANT; and W. VAN DER STEDE. "Earnings Targets and Annual Bonus Incentives." *The Accounting Review* 89 (2014): 1227-1258.

INZLICHT, M.; A. SHENHAV; and C. OLIVOLA. "The Effort Paradox: Effort is Both Costly and Valued." *Trends in Cognitive Sciences* 22 (2018): 337-349.

JACHIMOWICZ, J.; S. DUNCAN; E. WEBER; and E. JOHNSON. "When and Why Defaults Influence Decisions: A Meta-Analysis of Default Effects." *Behavioural Public Policy* 3 (2019): 159-186.

JOHNSON, E., and D. GOLDSTEIN. "Do Defaults Save Lives?" *Science* 302 (2003): 1338-1339.

KADOUS, K., and Y. ZHOU. "How Does Intrinsic Motivation Improve Auditor Judgment in Complex Audit Tasks?" *Contemporary Accounting Research* 36 (2019): 108-131.

KAHAN, M., and M. KLAUSNER. "Standardization and Innovation in Corporate Contracting." *Virginia Law Review* 83 (1997): 713-770.

KAHNEMAN, D. *Thinking Fast and Slow*. New York, NY: Farrar, Straus, and Giroux, 2011.

KESAN, J., and R. SHAH. "Setting Software Defaults: Perspectives from Law, Computer Science, and Behavioral Economics." *Notre Dame Law Review* 82 (2006): 583-634.

KOOL, W., and M. BOTVINICK. "A Labor/Leisure Tradeoff in Cognitive Control." *Motivation Science* 1 (2014): 3-18.

KOOL, W., and M. BOTVINICK. "Mental Labour." *Nature Human Behavior* 2 (2018): 899-908.

KUNDA, Z. "The Case for Motivated Reasoning." *Psychological Bulletin* 108 (1990): 480-498.

MADRIAN, B., and D. SHEA. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *The Quarterly Journal of Economics* 116 (2001):1149-1187.

MCKENZIE, C.; M. LIERSCH; and S. FINKELSTEIN. "Recommendations Implicit in Policy Defaults." *Psychological Science* 17 (2006): 414-420.

NUNNALLY, J., AND I. BERNSTEIN. *Psychometric Theory*. New York, NY: McGraw Hill, 1994.

PUBLIC COMPANY ACCOUNTING OVERSIGHT BOARD (PCAOB). *PCAOB Report Encourages Auditors to Take Action in Response to Risk Assessment Deficiencies Identified in Inspections.* Washington, DC: PCAOB, 2015.

PCAOB. *AS 2110: Identifying and Assessing Risks of Material Misstatement.* Washington, DC: PCAOB, 2022.

PCAOB. *Staff Update and Preview of 2022 Inspection Observations*. Washington, DC: PCAOB, 2023a.

PCAOB. *Inspections to Prioritize Audit Risks Related to Fraud, the Financial Services Sector, Crypto*. Washington, DC: PCAOB, 2023b.

PCAOB. *PCAOB Adopts New Quality Control Standard With a Risk-Based Approach Designed to Drive Continuous Improvement in Audit Quality*. Washington, DC: PCAOB, 2024.

PCAOB. *PCAOB Makes Available New Downloadable Datasets Featuring PCAOB Inspection Findings From Audit Firm Inspection Reports*. Washington, DC: PCAOB, 2025.

RITOV, I., and J. BARON. "Outcome Knowledge, Regret, and Omission Bias." *Organizational Behavior and Human Decision Processes* 64 (1995): 119-127.

SAKAGUCHI, H; N. STEWART; J. GATHERGOOD; P. ADAMS; B. GUTTMAN-KENNEY; L. HAYES; and S. HUNT. "Default Effects of Credit Card Minimum Payments." *Journal of Marketing Research* 59 (2022): 775-796.

SAUNDERS, K.; KEUNE, M.; and E. HAWKINS. "More than Making Copies: Survey Evidence on the Work of Novice Auditors." *AUDITING: A Journal of Practice & Theory* 42 (2023): 131-156.

SEC. *The Importance of a Comprehensive Risk Assessment by Auditors and Management.* Washington, DC: SEC, 2023.

SHEVCHENKO, Y; B. HELVERSEN; and B. SCHEIBEHENNE. "Change and Status Quo in Decisions with Defaults: The Effect of Incidental Emotions Depends on the Type of Default." *Judgment and Decision Making* 9 (2014): 287-296.

TAN, E.; M. GARG; D. JETER; D. KIRUBAHARAN; and V. NAIKER. "Audit Office Effectiveness and Efficiency." *AUDITING: A Journal of Practice & Theory* 44 (2025): 181-209.

TANNENBAUM, D., AND P. DITTO. "Information Asymmetries in Default Options." *Working Paper, University of Chicago*, 2011.

THALER, R., AND S. BENARTZI. "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy* 112 (2004): 164-187.

THALER, R., AND C. SUNSTEIN. *Nudge: The Final Edition*. New York, NY: Penguin Books, 2021.

THOMSON, K., and D. OPPENHEIMER. "The 'Effort Elephant' in the Room: What is Effort, Anyway?" *Perspectives on Psychological Science* 17 (2022): 1633-1652.

VAN GESTAL, L.; M. ADRIAANSE; AND D. DE RIDDER. "Do Nudges Make Use of Automatic Processing? Unraveling the Effects of a Default Nudge Under Type 1 and Type 2 Processing." *Comprehensive Results in Social Psychology* 5 (2021): 4-24.

WILLIS, L. "When Nudges Fail: Slippery Defaults." *University of Chicago Law Review* 80 (2013): 1155-1229.

FIGURE 1
Results for Test of Hypothesis

FIGURE 1—Results of our experiment reported by manipulated conditions. Auditors complete a risk assessment task. The dependent variable is *Accuracy for Changing Risks*, specifically, the number of times that auditors choose directionally correct risk ratings for the 14 risk factors that have changed from the prior year. Our first manipulated variable is *Prepopulation.* Auditors in the prepopulated condition complete a workpaper that contains last year's risk ratings and evidence in the spaces for the current year's risk ratings and evidence. Auditors in the non-prepopulated condition complete a workpaper that is blank. All auditors have access to the prior year workpaper in a separate file. Our second manipulated variable is *Intervention*. Auditors are randomly assigned to receive or not receive an intervention that we design to target the three mechanisms that we hypothesize underlie default effects in the risk assessment task: implicit endorsement, anticipated regret, and cognitive costs.

FIGURE 2
Results for Analyses of Auditors' Accuracy for Unchanging Risks

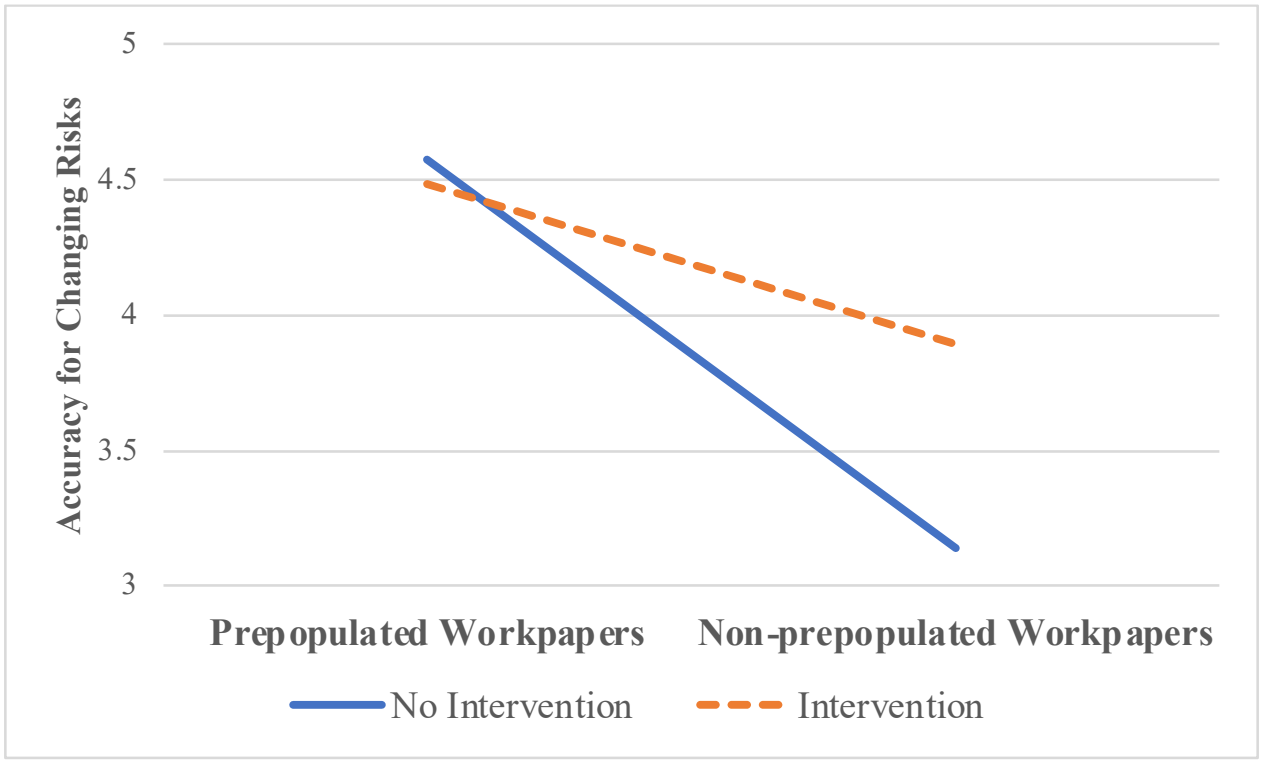

FIGURE 2—Additional analyses of auditors' accuracy for unchanging risks. Figure 1 defines the experimental setting and independent variables. The dependent variable is *Accuracy for Unchanging Risks*, specifically, the number of times that auditors choose the same risk rating as last year for the five risk factors that are unchanged from the prior year.

TABLE 1

Effects of Prepopulation and the Intervention on Auditors' Accuracy for Changing Risks

Panel A: Descriptive statistics for dependent variable by experimental condition

	No Intervention	Intervention
Prepopulated Workpapers	7.39 (3.56) n=28	9.26 (2.67) n=27
Non-prepopulated Workpapers	9.82 (2.07) n=28	10.11 (2.25) n=28

Panel B: Analysis of Variance to Test Hypothesis

Dependent Variable = *Accuracy for Changing Risks*	**df**	**SS**	**MS**	**F-statistic**	**p-value**
Prepopulation	1	74.46	74.46	10.21	0.001
Intervention	1	32.12	32.12	4.40	0.019
Prepopulation X Intervention	1	17.33	17.33	2.38	0.063
Error	107	780.65	7.30		

Panel C: Simple Effects

Effect of Prepopulation	**df**	**SS**	**MS**	**F-statistic**	**p-value**
Within No Intervention	1	82.57	82.57	11.32	< 0.001
Error	107	780.65	7.30		
Within Intervention	1	9.88	9.88	1.35	0.124
Error	107	780.65	7.30		

Effect of Intervention	**df**	**SS**	**MS**	**F-statistic**	**p-value**
Within Prepopulated Workpapers	1	47.88	47.88	6.56	0.006
Error	107	780.65	7.30		
Within Non-prepopulated Workpapers	1	1.14	1.14	0.16	0.347
Error	107	780.65	7.30		

We manipulate, between-participants, whether auditors use workpapers that are prepopulated with prior year risk ratings and evidence, compared to non-prepopulated (i.e., blank), and whether auditors receive an intervention that we design to target the three mechanisms that we posit underlie prepopulation-created default effects: implicit endorsement, anticipated regret, and cognitive costs. The dependent variable is *Accuracy for Changing Risks*, specifically, the number of times that auditors choose directionally correct risk ratings for the 14 risk factors that have changed from the prior year. Panel A reports descriptive statistics (mean, standard deviation, and number of audit staff participants). Panels B and C report results of the ANOVA that we create to test our hypothesis that the negative effect of prepopulation on auditors' risk rating accuracy for changing risks will be weaker for auditors receiving the intervention, compared to auditors not receiving the intervention. *P*-values are one-tailed for directional predictions.

TABLE 2

Analyses of Auditors' Accuracy for Unchanging Risks and Possible Side Effects

Panel A: Descriptive statistics for dependent variable by experimental condition

	No Intervention	Intervention
Prepopulated Workpapers	4.57 (1.00) n=28	4.48 (0.75) n=27
Non-prepopulated Workpapers	3.14 (1.60) n=28	3.89 (1.40) n=28

Panel B Analysis of Variance

Dependent Variable = *Accuracy for Unchanging Risks*	**df**	**SS**	**MS**	**F-statistic**	**p-value**
Prepopulation	1	28.22	28.22	18.45	< 0.001
Intervention	1	3.02	3.02	1.98	0.163*
Prepopulation X Intervention	1	4.89	4.89	3.20	0.077*
Error	107	163.71	1.53		

Panel C: Simple Effects

Effect of Prepopulation	**df**	**SS**	**MS**	**F-statistic**	**p-value**
Within No Intervention	1	28.57	28.57	18.68	< 0.001
Error	107	163.71	1.53		
Within Intervention	1	4.76	4.76	3.11	0.041
Error	107	163.71	1.53		

Effect of Intervention	**df**	**SS**	**MS**	**F-statistic**	**p-value**
Within Prepopulated Workpapers	1	0.11	0.11	0.07	> 0.500
Error	107	163.71	1.53		
Within Non-prepopulated Workpapers	1	7.88	7.88	5.15	0.025*
Error	107	163.71	1.53		

We examine effects of our manipulated independent variables on auditors' accuracy at risks that have not changed from the prior year, which allows for examining possible side effects of the intervention. Specifically, the intervention could cause auditors to perceive an imperative to almost always change their risk ratings, which could lead to (1) auditors using prepopulated workpapers (who Bonner et al. [2018] show perform better for these risks) being harmed by the intervention and/or (2) auditors using non-prepopulated workpapers not being assisted by the intervention. The independent variables are defined in the notes to Table 1. The dependent variable is *Accuracy for Unchanging Risks*, specifically, the number of times that auditors choose the same risk rating as last year for the five risk factors that are unchanged from the prior year. Panel A reports descriptive statistics (mean, standard deviation, and number of audit staff participants). Panels B and C report results of the ANOVA that we create for our analysis. *P*-values are one-tailed for directional predictions, except those with a *.

TABLE 3
Analyses of Moderators to Illuminate Mechanisms

Panel A: Regression Model – Auditors' Belief it is Acceptable to Stick as a Moderator

Dependent Variable =*Accuracy for Changing Risks*	df	Unstandardized Coefficient	Standard Error	t-statistic	p-value
Intercept	103	9.00	1.30	6.94	< 0.001
Prepopulation	103	0.49	2.13	0.23	> 0.500
Intervention	103	3.00	1.83	1.64	0.052
Acceptable to Stick	103	0.24	0.35	0.69	0.492*
Prepopulation X Intervention	103	-1.89	2.73	-0.69	0.492*
Prepopulation X Acceptable to Stick	103	-0.76	0.53	-1.44	0.077
Intervention X Acceptable to Stick	103	-0.81	0.50	-1.63	0.054
Prepopulation X Intervention X Acceptable to Stick	103	0.95	0.70	1.36	0.088

Panel B: Model Estimates of Prepopulation X Intervention Effect at Levels of Moderator

	Effect	df	F-statistic	p-value
Lower Acceptable to Stick	0.08	1,103	0.00	> 0.500
Medium Acceptable to Stick	1.52	1,103	2.16	0.072
Higher Acceptable to Stick	2.97	1,103	4.09	0.023

Panel C: Model Estimates of Effects of Prepopulation within Intervention Conditions at Levels of Moderator

Within No Intervention	Effect	Standard Error	df	t-statistic	p-value
Lower Acceptable to Stick	-1.07	1.16	103	-0.92	0.179
Medium Acceptable to Stick	-2.23	0.74	103	-3.03	0.002
Higher Acceptable to Stick	-3.39	1.02	103	-3.34	0.001

Within Intervention	Effect	Standard Error	df	t-statistic	p-value
Lower Acceptable to Stick	-1.00	0.95	103	-1.05	0.148
Medium Acceptable to Stick	-0.71	0.73	103	-0.98	0.166
Higher Acceptable to Stick	-0.42	1.06	103	-0.40	0.347

Panel D: Regression Model – Auditors' Accuracy Goal as a Moderator

Dependent Variable =*Accuracy at Changing Risks*	df	Unstandardized Coefficient	Standard Error	t-statistic	p-value
Intercept	103	9.83	0.46	21.43	< 0.001
Prepopulation	103	-2.03	0.66	-3.09	0.001
Intervention	103	0.03	0.66	0.04	0.485
Accuracy Goal	103	0.53	0.36	1.47	0.072
Prepopulation X Intervention	103	1.39	0.93	1.49	0.070
Prepopulation X Accuracy Goal	103	1.39	0.56	2.50	0.007
Intervention X Accuracy Goal	103	0.81	0.70	1.16	0.125
Prepopulation X Intervention X Accuracy Goal	103	-1.77	1.06	-1.67	0.049

Panel E: Model Estimates of Prepopulation X Intervention Effect at Levels of Moderator

	Effect	df	F-statistic	p-value
Lower Accuracy Goal	3.16	1,103	4.78	0.016
Medium Accuracy Goal	1.39	1,103	2.21	0.070
Higher Accuracy Goal	-0.39	1,103	0.08	> 0.500

Panel F: Model Estimates of Effects of Prepopulation within Intervention Conditions at Levels of Moderator

Within No Intervention	Effect	Standard Error	df	t-statistic	p-value
Lower Accuracy Goal	-3.42	0.81	103	-4.22	< 0.001
Medium Accuracy Goal	-2.03	0.66	103	-3.09	0.001
Higher Accuracy Goal	-0.64	0.91	103	-0.70	0.242

Within Intervention	Effect	Standard Error	df	t-statistic	p-value
Lower Accuracy Goal	-0.26	1.20	103	-0.22	0.415
Medium Accuracy Goal	-0.64	0.67	103	-0.96	0.169
Higher Accuracy Goal	-1.02	1.04	103	-0.99	0.163

We test moderators to illuminate the mechanism underlying our findings. The independent variables (except for the moderators) and dependent variable are defined in the notes to Table 1. The first moderator, *Acceptable to Stick* (targeting implicit endorsement), is auditors' agreement (Likert Scale from 1-7) that "When performing the risk assessment task, I felt like it was acceptable, if I was uncertain, to stick with the prior year risk rating." The other moderator, *Accuracy Goal* (targeting cognitive costs) is the factor score of auditors' responses to four post-experimental questions. One question, professional identity, is auditors' selection from one of seven images of two overlapping circles that reflect the self and the profession, ranging from no overlap to nearly overlapping (Bauer [2015]). Two questions elicit auditors' agreement (Likert Scale from 1-7) that (1) "It was important to me to be accurate in my performance on the risk assessment task," and (2) I felt like my firm wanted me to be accurate on these risk ratings." The final question elicits auditors' agreement (on a Likert Scale from 1-7) that: "On my audit engagements, I try to make sure the client reports the most accurate numbers possible." We also examine a moderator to target the anticipated regret mechanism, but the results are insignificant; to reduce clutter, we do not tabulate this model. *P*-values are one-tailed for directional predictions, except those that are accompanied by a *.