Accounting, Organizations and Society xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Accounting, Organizations and Society

journal homepage: www.elsevier.com/locate/aos

The effects of forecast type and performance-based incentives on the quality of management forecasts

Clara Xiaoling Chen [a,*], Kristina M. Rennekamp [b], Flora H. Zhou [c]

[a] University of Illinois at Urbana-Champaign, United States
[b] Cornell University, United States
[c] Georgia State University, United States

A B S T R A C T

Understanding forecasts is important because of their pervasiveness in business decisions such as budgeting, production, and financial reporting. In this study we use an abstract experiment to examine how the preparation of disaggregated forecasts interacts with performance-based incentives to influence the accuracy and optimism of forecasts. We manipulate two factors between subjects at two levels each: forecast type (*disaggregated* or *aggregated*) and performance-based incentives (*present* or *absent*). Consistent with our predictions, we find that (1) preparing disaggregated forecasts leads to greater improvements in forecast accuracy (compared to preparing aggregated forecasts) in the absence of performance-based incentives than in the presence of performance-based incentives, and (2) preparing disaggregated forecasts leads to greater increases in forecast optimism (compared to preparing aggregated forecasts) in the presence of performance-based incentives than in the absence of performance-based incentives. Our study contributes to our understanding of unintentional biases in the forecasting process. Our results have important practical implications for designers of management control systems who elicit internal forecasts from managers. Finally, our results also have important practical implications for those who either prepare or use external management forecasts.

Introduction

Understanding forecasts is important because of their pervasiveness in business decisions such as budgeting, compensation, and financial reporting. Inaccurate forecasts can reduce the effectiveness of the production planning process and negatively impact production efficiency, cost management, and ultimately firm performance (e.g., Bruggen, Grabner, & Sedatole, 2013). To increase the chance of obtaining accurate forecasts from an agent, a principal needs to be careful in designing the management control system that elicits such forecasts from the agent (e.g., Osband, 1989).

One such control system that is commonly used is the planning and budgeting system of a firm (Merchant & Van der Stede, 2012). Within the planning and budgeting system, an important design choice is the level of aggregation at which the principal elicits forecasts from the agent. In practice, firms vary considerably in the level of aggregation of the information elicited by the planning and budgeting system (Merchant & Van der Stede, 2012). For example, top management can request that divisional managers prepare either an aggregated forecast (e.g., forecast total sales for the division) or a disaggregated forecast (e.g., forecast sales for individual products within the division) (see Kahn, 1998 and Lapide, 2006). Although

* Corresponding author at: Department of Accountancy, University of Illinois at Urbana—Champaign, 389 Wohlers Hall, 1206 South Sixth Street, Champaign, IL 61820, United States. Tel.: +1 (217) 244 3953.

E-mail addresses: cxchen@illinois.edu (C.X. Chen), kmr52@cornell.edu (K.M. Rennekamp), hzhou17@gsu.edu (F.H. Zhou).

http://dx.doi.org/10.1016/j.aos.2015.03.002
0361-3682/

Please cite this article in press as: Chen, C. X., et al. The effects of forecast type and performance-based incentives on the quality of management forecasts. *Accounting, Organizations and Society* (2015), http://dx.doi.org/10.1016/j.aos.2015.03.002

managers are likely to prepare both disaggregated and aggregated forecasts for internal decision-making purposes, the level of forecast aggregation required by the budgeting system will determine which forecast is more salient to them. Further, research on the anchoring and adjustment bias suggests that managers likely anchor on the numbers in the forecast that are most salient to them (Bromiley, 1987; Tversky & Kahneman, 1974). Therefore, the level of aggregation at which the principal elicits forecasts from the agent should affect managers' forecasts even when both types of forecasts are prepared.

Although economic theory suggests that a rational agent will provide the same forecast of a summary performance measure regardless of the level of forecast aggregation (or forecast type), psychology theory suggests that forecast type will influence the quality of the agent's forecasts, where forecast quality can refer to both the accuracy and optimism (or bias) in a forecast. We investigate how a control system design choice—forecast type—interacts with incentives to affect two dependent measures of forecast quality: *forecast accuracy* and *forecast optimism*. *Forecast accuracy* refers to the degree of closeness between a forecast and the actual outcome. *Forecast optimism* refers to consistent differences between forecasts and actual outcomes; that is, the extent to which forecasts exhibit a general tendency to be too high relative to actual outcomes. Specifically, we examine how forecast type affects forecast accuracy and forecast optimism in the presence or absence of explicit performance-based incentives that are tied to the measure being forecasted.

Drawing on psychology, forecasting, and accounting literatures on forecasts, we generate the following predictions for forecast accuracy and forecast optimism, respectively. First, we predict that preparing disaggregated forecasts leads to greater improvements in forecast accuracy (compared to preparing aggregated forecasts) in the absence of performance-based incentives than in the presence of performance-based incentives. When performance-based incentives are absent, disaggregated forecasts involve more careful and objective consideration of forecast components, which should improve forecast accuracy compared to preparing aggregated forecasts. Second, we predict that preparing disaggregated forecasts leads to greater increases in forecast optimism (compared to preparing aggregated forecasts) in the presence of performance-based incentives than in the absence of performance-based incentives. When managers produce disaggregated forecasts but *do* have explicit incentives to achieve favorable performance on the forecasted measure, they have both the motivation and opportunity to produce optimistic forecasts. In other words, while the preparation of disaggregated forecasts involves more complete consideration of information, theory suggests that individuals with performance-based incentives are likely to consider that additional information in a biased way that helps them reach their desired conclusions (Hales, 2007).

To test our predictions we conduct an abstract laboratory experiment where participants complete a knowledge task with questions from four different categories (e.g., English, math, grammar, and logic) and prepare forecasts of their performance. Participants complete two rounds of the task. After the initial round, participants receive feedback on their performance. Before the second round begins, participants provide forecasts of their second-round performance. Participants then answer the second round of questions and learn their actual performance.

We use an abstract task in our study for two reasons. First, we are interested in examining a fundamental psychological bias rather than reactions to rich, institutional features. An abstract knowledge test allows us to test the fundamental processes that affect the characteristics of our two types of forecasts while avoiding noise in participants' responses that could arise from asking them to do an unfamiliar task like forecasting revenues and expenses. Second, using a task with rich institutional features could introduce other incentives that may lead to intentional biases in the forecasts. For example, in an internal budgeting setting, managers may intentionally provide lower forecasts to increase the probability of achieving targets or intentionally provide higher forecasts to increase resource allocations (Fisher, Maines, Peffer, & Sprinkle, 2002). Using an abstract task removes the institutional features that might drive managers to intentionally produce biased forecasts, allowing us to isolate the effects of unintentional bias.

We manipulate two factors between subjects at two levels each. First, to manipulate forecast type, participants in the disaggregated forecast condition forecast their scores in all four categories of the test (e.g., English, math, grammar and logic), while participants in the aggregated forecast condition forecast their total score.[1] Second, we manipulate whether explicit performance-based incentives are present or absent.[2] We hold average participant compensation constant across the two incentive conditions. We examine two dependent variables: (1) forecast accuracy, where overestimation of scores is treated as equivalent to underestimation of scores; and (2) forecast optimism, which captures systematic tendency to overestimate scores.

Consistent with our predictions, we find that: (1) preparing disaggregated forecasts leads to greater improvements in *forecast accuracy* (compared to preparing aggregated forecasts) in the *absence* of performance-based incentives than in the *presence* of performance-based incentives; and (2) preparing disaggregated forecasts leads to greater increases in *forecast optimism* (compared to preparing aggregated forecasts) in the *presence* of performance-based incentives than in the *absence* of performance-based incentives. Given that participants' pay would be higher in the absence of the forecast error and forecast optimism described above, our results show that participants' judgments conflict with their financial incentives and therefore suggest that the biases we observe are unintentional.

[1] Although we manipulate the level of disaggregation at two levels in our experiment, the level of disaggregation can vary in degrees in practice. We expect the directional effects we document in our study to hold with varying levels of disaggregation.

[2] We manipulate incentives at two levels in our experiment, but the absence of performance-based incentives versus presence of performance-based incentives conditions can also map into low-powered incentives versus high-powered incentives in the real world.

Our study contributes to our understanding of unintentional biases in the forecasting process. Since unintentional biases may be more difficult to discipline than intentional, incentive-driven biases, our study provides insights that are likely useful to both preparers and users of forecasts. First, our results contribute to the budgeting literature. Prior budgeting literature focuses heavily on the opportunistic behavior of agents in the budgeting process and the effectiveness of truth-inducing incentives (e.g., Chow, Cooper, & Waller, 1988; Church, Hannan, & Kuang, 2012; Shields & Young, 1993; Waller, 1988; Webb, 2002; Young, 1985). However, unintentional biases such as those documented in our paper are more difficult to mitigate. Specifically, our results show that a control system design choice that has so far been largely overlooked in management accounting research—the level of forecasts elicited—can have unintended consequences for potential bias and accuracy in management forecasts.

Second, by highlighting the potential effect of an internal planning and budgeting system design choice (i.e., forecast type) on externally reported management forecasts, our study complements the accounting literature on management forecasts as well as an emerging literature that examines the link between external disclosures and internal decision-making (e.g., Goodman, Neamtiu, Shroff, & White, 2014; Hemmer & Labro, 2008; McNichols & Stubben, 2008). Prior research on management forecasts has shown that disaggregated forecasts increase the market's perception of the informational value and credibility of management forecasts (Hirst, Koonce, & Venkataraman, 2007; Hutton, Miller, & Skinner, 2003; Lansford, Lev, & Tucker, 2013), reduce investors' fixation on announced earnings (Elliott, Hobson, & Jackson, 2011), and decrease auditors' tolerance for misstatement (Libby & Brown, 2013). Our study differs from these prior studies by: (1) taking the perspective of the preparer, rather than the users, of management forecasts; and (2) by focusing on the actual, rather than perceived, quality of disaggregated forecasts. Despite the documented perceived benefits of disaggregated forecasts, our results suggest that the actual quality of disaggregated management forecasts may depend on the incentives that managers face.

Finally, our study also adds to the forecasting literature by highlighting unintentional behavioral biases in the forecasting process. The literature on forecasting has largely ignored behavioral explanations for unintentional optimism. Our study also answers the call in the forecasting literature for more research that sheds light on the circumstances under which disaggregated forecasts are more likely to improve on aggregated forecasts (e.g., Henrion, Fischer, & Mullin, 1993; Kremer, Siemsen, & Thomas, 2012).

Theory and hypotheses

Prior research suggests that decomposition, or disaggregation, is a useful technique for reducing information processing demands on the estimator, which may lead to more complete consideration of available information and, ultimately, more accurate estimates (Raiffa, 1968; Ravinder, Kleinmuntz, & Dyer, 1988). However, prior research also suggests that disaggregation and more complete consideration of information are not necessarily always beneficial to judgment quality (e.g., Henrion et al., 1993). We argue that while disaggregation can improve the accuracy of estimates in the absence of performance-based incentives, it can give forecasters greater opportunity to inject optimistic bias into their forecasts in the presence of explicit performance-based incentives.

Forecast type and consideration of available information

We first consider how processing of information differs between preparing disaggregated forecasts and preparing aggregated forecasts, regardless of the incentives that a manager faces. Prior research suggests that preparing disaggregated forecasts can reduce information processing load for the forecaster, which may lead to more careful consideration of all available information than preparing aggregated forecasts (Henrion et al., 1993; Raiffa, 1968; Ravinder et al., 1988). This occurs because forecasting a number holistically is often considered a more complex task than decomposing the forecasting problem into multiple components first and then combining the components into an aggregated forecast (e.g., Henrion et al., 1993; Ravinder et al., 1988). As task complexity increases, individuals are more likely to choose strategies that lower total cognitive costs (Bonner, 2008). When individuals use these strategies, they do not search for all relevant information in making decisions and, as a consequence, decision quality is often reduced. In a management forecast setting, making an aggregated forecast requires the manager to attend to all relevant information at once, which can be mentally taxing. To make the forecasting task more manageable, the manager may choose to attend to the most salient pieces of information and ignore or underweight less important information. Consistent with this argument, in an audit setting, Zimbelman (1997) shows that auditors' attention to fraud-risk factors is higher when they separately assess the risk of intentional and unintentional misstatement than when they assess only the overall risk of misstatement.

Related work on support theory also suggests that disaggregation leads to more careful consideration of individual components of a given set of information. Research on support theory finds that unpacking an event into two or more of its components helps respondents recall more evidence from memory and/or makes existing evidence more salient, such that the rated likelihood of the event occurring increases (Tversky & Koehler, 1994). Support theory has primarily focused on the assessments of probability or frequency of alternative hypotheses, but the cognitive mechanism underlying the unpacking phenomenon is quite general. Van Boven and Epley (2003) confirm that unpacking also influences evaluations when events are simply described in greater detail as opposed to being unpacked into non-overlapping components. Specifically, Van Boven and Epley (2003) show that unpacking leads people to think about the details of a category or event, thereby making it easier to mentally generate evaluative evidence.

Please cite this article in press as: Chen, C. X., et al. The effects of forecast type and performance-based incentives on the quality of management forecasts. *Accounting, Organizations and Society* (2015), http://dx.doi.org/10.1016/j.aos.2015.03.002

Combined, the prior literature suggests that preparing disaggregated forecasts leads to lower information processing demands and more complete consideration of available information. However, greater consideration of information has the potential to either benefit or harm managers' forecasts, depending on whether the managers have explicit performance-based incentives.

Forecast accuracy in the absence of performance-based incentives

Drawing on prior literature on forecasting and accounting, we first consider the effects of preparing disaggregated forecasts in the absence of explicit performance-based incentives. More specifically, we predict that preparing disaggregated forecasts will lead to greater improvements in forecast accuracy (compared to preparing an aggregated forecast) in the absence of performance-based incentives than in the presence of performance-based incentives for at least two reasons. First, as discussed in the previous subsection, preparing disaggregated forecasts reduces the cognitive load of the forecaster. In the absence of performance-based incentives, a reduction of the cognitive load may lead to more complete consideration of all available information and improve the accuracy of forecasts. Second, in the absence of performance-based incentives, disaggregated component forecasts are likely to contain random errors, some of which overstate performance and some of which understate performance. These random errors will at least partially cancel each other out when they are combined to derive the top-level forecast, leading to less error and greater forecast accuracy for disaggregated forecasts compared to aggregated forecasts (e.g., Kleinmuntz, Fennema, & Peecher, 1996; Ravinder et al., 1988).

Based on the above discussion, we expect that disaggregated forecasts will result in both greater precision in the forecast of each component (due to greater consideration of information in forecasting each component) and greater reduction of random errors when component forecasts are combined (due to cancellation of error terms). In turn, we expect these effects to result in overall greater improvements in forecast accuracy in disaggregated forecasts than in aggregated forecasts when performance-based incentives are absent compared to when performance-based incentives are present.

We note, however, that greater consideration of information in forecasting each component may induce greater bias in the forecasts in the presence of explicit incentives tied to the forecasted measure. In addition, the error reduction effect discussed above could be undermined if the errors associated with the component forecasts are positively correlated, i.e., the errors are systematic rather than random (Ravinder et al., 1988). Prior research suggests that this is more likely to be the case in the presence of explicit performance-based incentives, which we consider next.

Forecast accuracy in the presence of performance-based incentives

Prior literature suggests that individuals are naturally, and often unintentionally, optimistic, and that performance incentives or directional goals can exacerbate this optimism (Hales, 2007). Although managers also have incentives for accurate forecasts because investors associate accurate forecasts with management credibility and reward accurate forecasts (Ajinkya & Gift, 1984; Graham, Harvey, & Rajgopal, 2005; Healy & Palepu, 2001; Jennings, 1987; Mercer, 2005), managers' incentives for accurate forecasts may be dominated by their incentives for favorable performance when they are provided with explicit performance-based incentives.

The discussion above suggests that when managers are rewarded for higher performance on the forecasted measure, forecasts are likely more optimistic, regardless of whether disaggregated or aggregated forecasts are prepared. Although most individuals have at least some intrinsic motivation for favorable performance, we expect the bias to be greater when performance-based incentives are explicit. However, these theories only predict a main effect of performance incentives on forecast optimism. Within a group of individuals that have explicit performance incentives, greater forecast optimism among those who prepare a disaggregated rather than an aggregated forecast would be consistent with motivated reasoning theory, which we discuss next.[3]

Motivated reasoning theory predicts that directional preferences will affect how people attend to and process information (Kunda, 1990). In an accounting setting, Hales (2007) shows that investors' forecasts of earnings are affected by the investment position they hold and by whether they are facing the prospect of a gain or loss on those investments. Specifically, investors' forecasts of earnings are biased in a direction consistent with their directional preferences, even if they are only provided with an incentive to be accurate. Building on Hales (2007), Thayer (2011) shows that investors seek additional information consistent with their desired conclusions about an investment.

As discussed at the beginning of this section, theory suggests that preparing a disaggregated forecast will make a manager more likely to attend to more detailed information to forecast the individual components, leading to more accurate forecasts in the absence of performance-based incentives. However, when managers are provided with explicit incentives that are tied to their performance on the forecasted measure, preparing disaggregated forecasts is less likely to lead to higher forecast accuracy than preparing aggregated forecasts for two reasons.

First, in the presence of performance-based incentives, a manager making a judgment about future performance has a preference for positive performance. Disaggregation will cause the manager to attend to more detailed information about how a favorable outcome is likely to be achieved, and hence, allow more opportunity for the manager to interpret information in the direction consistent with his or her preferences (Kunda, 1990). Thus, when performance-based incentives are present,

[3] Theory on optimism alone does not predict that the magnitude of optimism should vary by forecast type. However, motivated reasoning theory, which more explicitly incorporates biased *processing*, does help to make that prediction.

the potential benefit of considering more information might be partially or completely offset by the negative effect of attending to more *preference-consistent* information (Kunda, 1990). Motivated reasoning theory suggests that individuals will not consider a balanced set of reasons for a given outcome when making a judgment (Ditto & Lopez, 1992).

Second, the presence of performance-based incentives is likely to lead to an overall optimistic bias in component forecasts, i.e., systematic overestimation errors. Systematic overestimation errors in component forecasts will not cancel each other out when the component forecasts are combined into an overall forecast. Therefore, we expect that in the presence of performance-based incentives, preparing disaggregated forecasts will not lead to greater forecast accuracy.

Combined, our theory suggests an interaction between forecast type and performance-based incentives on forecast accuracy. Specifically, we expect that preparing disaggregated forecasts rather than aggregated forecasts increases forecast accuracy in the absence of performance-based incentives, but not necessarily in the presence of performance-based incentives. This discussion leads to our first hypothesis on forecast accuracy:

H1. Preparing disaggregated forecasts leads to greater improvements in forecast accuracy (compared to preparing aggregated forecasts) in the absence of performance-based incentives than in the presence of performance-based incentives.

Performance-based incentives and forecast optimism

The theory that we have outlined suggests that disaggregated forecasts may exhibit greater forecast optimism than aggregated forecasts when performance-based incentives are present compared to when performance-based incentives are absent. Even though managers have the motivation to produce optimistic forecasts when they prepare an aggregated forecast in the presence of performance-based incentives, they have less opportunity to inject optimistic bias into their forecasts. Thus, providing disaggregated forecasts should lead to greater forecast optimism than providing aggregated forecasts in the presence of performance-based incentives.

In the absence of performance-based incentives, however, managers have weaker motivation to make optimistically biased forecasts regardless of whether they prepare aggregated or disaggregated forecasts. Thus, we expect a smaller difference in forecast optimism between disaggregated forecasts and aggregated forecasts in the absence of performance-based incentives than in the presence of performance-based incentives. Our second hypothesis is therefore:

H2. Preparing disaggregated forecasts leads to a greater increase in forecast optimism (compared to preparing aggregated forecasts) in the presence of performance-based incentives than in the absence of performance-based incentives.

Method

Participants

We recruit ninety-two undergraduate business students from a large public university as participants. In the experiment, participants complete a knowledge test with questions from four categories (English, math, grammar and logic) and make associated forecasts of their performance. Because we examine a fundamental psychological bias rather than reactions to rich, institutional features, we believe students have sufficient knowledge for the task and can be used as participants (Libby, Bloomfield, & Nelson, 2002; Libby & Rennekamp, 2012). Further, undergraduates take knowledge tests (either the SAT or ACT) before entering the university that are similar to those we use in our study and have the ability to understand the incentives associated with our forecasting task.

Research design

To test our hypotheses, we use a 2 (Forecast Type) × 2 (Performance-Based Incentives) between-subjects experimental design. We manipulate forecast type at two levels: disaggregated forecast versus aggregated forecast. Participants complete a first round of the question-based task to get a sense of their skill in the four topic categories. In the disaggregated forecast condition, participants provide a separate forecast of their performance in each of the four categories for the second round. In the aggregated forecast condition, participants provide a forecast of their overall performance (or total score) in the second round.

We also manipulate performance-based incentives at two levels: absent or present. Following prior literature (Hales, 2007), we provide subjects with an incentive to make accurate forecasts in both conditions to reduce noise in the results. Specifically, in the condition with performance-based incentives, a participant's pay is based on two components: the participant's actual performance on the task and the accuracy of the participant's forecast. The formula is as follows:

$$\begin{aligned}\text{Total pay} = \text{£}4 * \text{Number of questions answered correctly}\\ + (\text{£}20 - \text{£}2 * \text{absolute value of the difference}\\ \text{between forecast and actual performance})\end{aligned}$$

For the performance-based component, participants receive £4.00 in experimental currency for each question answered correctly, up to a total of £112.00 if the participant answers all 28 questions correctly.[4] For the forecast accuracy component, participants receive a bonus that is £20.00 for a completely accurate forecast. The bonus is reduced by £2.00 per question if the forecast deviates from the actual performance and drops to zero if the forecast either over- or underestimates actual performance by ten

[4] All currency amounts described here are denoted in experimental laboratory currency unless stated otherwise. Laboratory earnings are converted to U.S. dollar earnings upon completion of the experiment. Participants do not know in advance the exchange rate between the two currencies, but do know that earning more laboratory currency will always translate to higher U.S. dollar-denominated earnings.

or more questions. The participant is always better off answering as many questions correctly as he/she can, regardless of the forecast he/she provides because the participant earns £4.00 for every correct answer but loses only £2.00 of the forecast bonus for each question by which the actual performance differs from the forecast.[5] Thus, for a given forecast, the participant will always receive higher compensation by performing to the best of his/her ability rather than by withholding effort after meeting his/her forecast. In the condition without a performance-based incentive, pay is based only on the accuracy of the participant's forecast.

Immediately after receiving information on how they will be paid, participants are asked to answer a manipulation check question on the same page to ensure that they understand the incentive scheme to which they are assigned. Specifically, we ask participants to indicate the components of their compensation by choosing between two options: (1) my compensation will increase if my forecast of my performance on the round of questions I answer is more accurate; and (2) my compensation will increase if I perform better on the round of questions I answer. Results reveal that 100% of participants correctly indicate whether their compensation is based on forecast accuracy only or on both forecast accuracy and actual performance in the second round of questions.

Participants are informed that once all participants have completed the task, their earnings in £ will be converted to real U.S. dollars at a positive but unspecified rate and that they are always better off trying to earn more £ in the study, since that translates to greater earnings in U.S. dollars. Participants are informed before the start of the experiment to expect payments approximately two weeks after all sessions are conducted. In addition, each participant receives a $5 show-up fee. On average, participants receive $20 in U.S. currency for their participation across all conditions.

Task and experimental procedures

We randomly assign each laboratory session to either the presence or absence of explicit performance-based incentives treatment to ensure participants are not aware of our manipulations. Upon arrival at the experiment, participants are randomly assigned to one of the two forecast type conditions. We ask participants to read the informed consent form and sign the form before they start the task. In the experiment, participants complete two rounds of a mini SAT-type test. We use an initial round of SAT-type questions to familiarize participants with the task and form expectations about their future performance. We intentionally choose relatively difficult questions for our task in order to increase variation in participants' forecasting performance. This allows us to better detect the effects of our independent variables on our dependent measures. To keep the total time required for the task to a minimum, the first round contains two questions from each of the four categories, while the second round contains 28 questions, with seven questions from each of the four categories.

After the first round, participants receive feedback on their performance. Before the second round begins, participants make a private forecast of their second-round performance. Participants then answer the second round of questions. Participants in the disaggregated forecast condition are asked to provide forecasts of their scores for each of the categories of SAT-type questions (English, math, grammar and logic). Participants in the aggregated forecast condition are asked to provide a forecast of their total score on the test. After participants complete the second round of questions, they answer a post-experimental questionnaire, which includes debriefing and demographic questions.

Dependent and control variables

Forecasts of the four components in the disaggregated forecast condition are summed to form total score forecasts, which are compared to the total scores forecasted in the aggregated forecast condition. We examine two aspects of forecast quality: accuracy and optimistic bias. Following prior literature (e.g., Duru & Reeb, 2002; Goodman et al., 2014; Henrion et al., 1993; Mikhail, Walther, & Willis, 1999), we capture forecast accuracy with the absolute forecast error, i.e., the absolute value of the difference between the forecast and the performance. A smaller absolute forecast error indicates greater forecast accuracy. To facilitate interpretation, we transform this reverse measure of forecast accuracy by calculating the difference between the maximum performance, 28, and the absolute forecast error and use this as our measure of forecast accuracy. A larger number for this transformed measure indicates greater forecast accuracy. We measure the optimistic bias in a forecast as the signed forecast error, i.e., the signed difference between the forecast and the actual performance. A larger positive forecast error indicates a higher level of optimism.

We control for participants' performance because prior research has shown that although skillful individuals often overestimate their performance relative to others, they also underestimate their own absolute performance (Klayman, Soll, González-Vallejo, & Barlas, 1999; Krueger & Mueller, 2002; Kruger & Dunning, 1999; Larrick, Burson, & Soll, 2007). Therefore, we expect a negative relationship between participants' performance in the actual round and their forecast optimism. We control for participants' performance in the first round because higher performance in the first round will lead to higher expected performance in the actual round, which is likely to lead to higher forecast errors.

[5] For example, if a participant forecasts that she can get 16 out of the 28 questions correct, she will get £78 in laboratory currency ($£4 * 15 + (£20 - £2 * 1)$) if she ends up answering 15 questions correctly, £84 ($£4 * 16 + £20$) if she ends up answering 16 questions correctly, and £92 ($£4 * 20 + (£20 - £2 * 4)$) if she ends up answering 20 questions correctly.

Results

Table 1 provides descriptive statistics of average performance and average forecasts of total score for the four conditions.[6,7] Consistent with performance-based incentives increasing effort and performance, we observe that participants' performance in the second round is significantly higher in the presence of performance-based incentives than in the absence of performance-based incentives (17.55 versus 15.15, $p < 0.01$, two-tailed).

Test of H1: Forecast type, performance-based incentives, and forecast accuracy

H1 predicts that disaggregated forecasts will lead to greater improvement in forecast accuracy in the absence of performance-based incentives than in the presence of performance-based incentives, compared to aggregated forecasts. Again, we measure forecast accuracy by calculating the difference between the maximum possible score of 28 and the absolute forecast error, where a larger difference corresponds to greater forecast accuracy. We test this interaction using contrast coding as well as follow-up simple effects tests using an analysis of covariance (ANCOVA) (Buckless & Ravenscroft, 1990). We include participants' performances in the first and second rounds of questions as covariates to control for variation in the data that is not the focus of our study. We control for participants' performance in the first round because higher performance in the first round will lead to higher expected performance in the second round, which, in turn, will lead to higher forecast error holding the actual performance constant. Consistent with this conjecture, a regression analysis shows that forecast errors are positively associated with actual performance in the first round. We also control for participants' performance in the second round to control for the negative correlation between performance and forecast optimism documented in prior literature (Klayman et al., 1999; Krueger & Mueller, 2002; Kruger & Dunning, 1999; Larrick et al., 2007). Consistent with prior research, a regression analysis shows that forecast optimism is negatively associated with actual performance in the second round.

Based on our first hypothesis, we use contrast weights of +3 in the absence of performance-based incentives/disaggregated forecast condition and −1 in the other three conditions. The results presented in Panel C, Table 2 show that the planned contrast is marginally significant, supporting our hypothesis ($p = 0.07$, one-tailed).[8] The follow-up simple contrasts confirm the interaction between performance-based incentives and forecast type on forecast accuracy. Specifically, when there are no performance-based incentives, preparing a disaggregated forecast leads to significantly greater forecast accuracy ($p = 0.02$, one-tailed). By contrast, when participants have performance-based incentives, there is no significant difference in forecast accuracy between the disaggregated and the aggregated forecast types ($p = 0.42$, two-tailed).

Table 1
Descriptive statistics.

	Forecast type			
	n	Aggregated forecast	*n*	Disaggregated forecast
Panel A: Mean (standard deviation) for performance				
Absence of Performance-Based Incentives	30	14.40 (5.88)	18	16.39 (2.91)
Presence of Performance-Based Incentives	18	17.89 (2.93)	26	17.31 (2.28)
Panel B: Mean (standard deviation) for forecasts				
Absence of Performance-Based Incentives	30	14.93 (8.03)	18	17.78 (3.78)
Presence of Performance-Based Incentives	18	18.72 (3.61)	26	20.15 (4.18)

The table presents descriptive statistics for participants' performance and forecasts for the four experimental conditions.
We manipulate *Performance-Based Incentives* at two levels: In the *absence of performance-based incentives* condition, participants are only compensated for the accuracy of their performance forecasts; in the *presence of performance-based incentives*, participants are compensated for both their performance in the second round of SAT-type questions and the accuracy of their performance forecasts.
We manipulate *Forecast Type* at two levels: In the *aggregated forecast condition*, participants provide a holistic forecast for their performance in the second round of SAT-type questions; in the *disaggregated forecast condition*, participants provide a separate forecast for their performance in each of the four components of the second round of SAT-type questions.
Performance = Participants' actual performance (number of correct answers to the questions) in the second round of SAT-type questions.
Forecast = Forecast of total score in the second round of SAT-type questions in the aggregated forecast condition; sum of forecasts of the four components in the second round of SAT-type questions in the disaggregated forecast condition.

Test of H2: Forecast type, performance-based incentives, and forecast optimism

H2 predicts an interaction between forecast type and performance-based incentives such that the disaggregated forecast condition leads to a greater increase in forecast optimism than the aggregated forecast condition in the presence of performance-based incentives than in the absence of performance-based incentives. To test H2, we estimate an analysis of covariance (ANCOVA) using signed forecast errors as the dependent variable and performance-based incentives and forecast type as the independent variables. A more positive signed forecast error indicates a higher level of forecast optimism. We also test

[6] The difference in cell sizes is due to the number of participants who showed up for a given session and imperfect randomization of the online survey software.

[7] Of the ninety-two participants, four participants forecasted zero in the aggregated forecast condition in the absence of performance-based incentives, indicating an intention to game the incentive system. Excluding the four participants who made forecasts of zero strengthens the results. Among the other eighty-eight participants, there is no significant difference in performance between disaggregated and aggregated forecast types when performance-based incentives are absent ($p = 0.57$) or present ($p = 0.55$), indicating other participants were not engaged in similar gaming behavior.

[8] The planned contrast is statistically significant ($p = 0.03$, one-tailed) when we exclude the four participants who made forecasts of zero.

Table 2
The effects of forecast type and performance-based incentives on forecast accuracy.

	Forecast type			
	n	Aggregated forecast	*n*	Disaggregated forecast
Panel A: Mean (standard deviation) for forecast accuracy				
Absence of Performance-Based Incentives	30	23.20 (4.26)	18	25.50 (1.54)
Presence of Performance-Based Incentives	18	24.94 (2.01)	26	24.15 (3.20)
Source	*df*	Mean square	*F*	*p*-value
Panel B: ANCOVA model of forecast accuracy				
Performance-Based Incentives	1	0.68	0.07	0.80
Forecast Type	1	8.68	0.87	0.35
Performance-Based Incentives × Forecast Type	1	42.36	4.23	**0.02**
Trial Performance	1	30.49	3.05	0.08
Performance	1	0.10	0.01	0.92
Error	82	10.00		

Source	*df*	*F*	*p*-value
Panel C: Planned contrast coding and follow-up simple effect tests			
Overall tests:			
Preparing disaggregated forecasts (compared to preparing aggregated forecasts) leads to a greater improvement in forecast accuracy in the absence of performance-based incentives than in the presence of performance-based incentives. Contrast weights (−1, 3, −1, −1)	1	2.19	**0.07**
Follow-up simple effect tests:			
Effect of forecast type in the absence of performance-based incentives	1	4.51	**0.02**
Effect of forecast type in the presence of performance-based incentives	1	0.64	0.42
Effect of performance-based incentives in the disaggregated forecast conditions	1	1.59	0.21
Effect of performance-based incentives in the aggregated forecast conditions	1	2.61	0.11

The table presents descriptive statistics, the ANCOVA model, and simple contrasts for forecast accuracy for the four treatments. See Table 1 for descriptions of the manipulations of *performance-based incentives* and *forecast type*. Reported *p*-values are two-tailed unless testing a one-tailed prediction, as signified by bold face.
Forecast Accuracy = The difference between the maximum possible score of 28 and the absolute difference between forecast and performance in the second round of SAT-type questions, where forecast is the forecast of total score in the second round of SAT-type questions in the aggregated forecast conditions and the sum of component forecasts in the second round of SAT-type questions in the disaggregated forecast conditions. Higher measures indicate greater forecast accuracy.
Trial Performance = Participants' performance in the first round of SAT-type questions.
Performance = Participants' actual performance in the second round of SAT-type questions.

this interaction using contrast coding as well as follow-up simple effects tests where, based on our second hypothesis, we use contrast weights of +3 in the presence of performance-based incentives/disaggregated forecast condition, +1 in the presence of performance-based incentives/aggregated forecast condition, and −2 in the absence of performance-based incentives/aggregated forecast and absence of performance-based incentives/disaggregated forecast conditions.[9]

The results presented in Panel C of Table 3 show that the planned contrast is statistically significant, supporting H2 ($p = 0.03$, one-tailed).[10] The follow-up simple contrasts (Table 3, Panel C) confirm the ordinal interaction between performance-based incentives and forecast type on forecast optimism. Specifically, in the absence of performance-based incentives, there is no significant difference in forecast optimism between disaggregated and aggregated forecasts ($p = 0.63$, two-tailed). By contrast, when participants receive performance-based incentives, forecast optimism is higher in the disaggregated forecast condition than in the aggregated forecast condition ($p = 0.08$, one-tailed).[11]

Overall, our results are consistent with H2. When participants' incentives are not tied to performance, producing disaggregated forecasts does not lead to more optimistically biased forecasts. However, when participants' incentives are tied to performance, they have a preference for favorable performance. As a result, producing disaggregated forecasts gives participants both the motivation and the opportunity to engage in biased processing of information and interpret it in a way that is consistent with their preferences, which leads to significantly more optimistically biased forecasts compared to producing aggregated forecasts.

Supplemental analyses

In this section we conduct additional analyses to support the theoretical arguments underlying our hypotheses.

[9] Since the simple effect of performance-based incentives in the aggregated forecast condition is insignificant, we verify that our results are robust to an alternative allocation of contrast weights (specifically, +3 in the presence of *performance-based incentives/disaggregated forecast* condition and −1 in the other three conditions). This set of contrast weights is more restrictive by not allowing for a simple effect of performance-based incentives in the aggregated forecast condition.

[10] The planned contrast is also statistically significant ($p = 0.02$, one-tailed) when we exclude the four participants who made forecasts of zero.

[11] These results are stronger when we exclude the four participants who made forecasts of zero: when participants receive performance-based incentives, forecast optimism is higher in the disaggregated forecast condition than in the aggregated forecast condition ($p = 0.07$, one-tailed).

Please cite this article in press as: Chen, C. X., et al. The effects of forecast type and performance-based incentives on the quality of management forecasts. *Accounting, Organizations and Society* (2015), http://dx.doi.org/10.1016/j.aos.2015.03.002

Table 3
The effects of forecast type and performance-based incentives on forecast optimism.

	Forecast type			
	n	Aggregated forecast	*n*	Disaggregated forecast
Panel A: Mean (standard deviation) for forecast optimism				
Absence of Performance-Based Incentives	30	0.53 (6.45)	18	1.39 (2.64)
Presence of Performance-Based Incentives	18	0.83 (3.63)	26	2.85 (4.14)
Source	*df*	Mean square	*F*	*p*-value
Panel B: ANCOVA model of forecast optimism				
Performance-Based Incentives	1	42.81	2.29	**0.07**
Forecast Type	1	33.98	1.82	0.18
Performance-Based Incentives × Forecast Type	1	8.29	0.44	**0.25**
Trial Performance	1	302.70	16.18	<0.01
Performance	1	131.66	7.04	<0.01
Error	82	18.71		

Source	*df*	*F*	*p*-value
Panel C: Planned contrast coding and follow-up simple effect tests			
Overall tests:			
Preparing disaggregated forecasts (compared to preparing aggregated forecasts) leads to a greater increase in forecast optimism in the presence of performance-based incentives than in the absence of performance-based incentives. Contrast weights (3, 1, −2, −2)	1	3.92	**0.03**
Follow-up simple effect tests:			
Effect of forecast type in the absence of performance-based incentives	1	0.23	0.63
Effect of forecast type in the presence of performance-based incentives	1	2.01	**0.08**
Effect of performance-based incentives in the disaggregated forecast conditions	1	2.42	0.12
Effect of performance-based incentives in the aggregated forecast conditions	1	0.37	0.54

The table presents descriptive statistics, the ANCOVA model, and contrast coding tests for forecast optimism for the four experimental conditions. See Table 1 for descriptions of the manipulations of *performance-based incentives* and *forecast type*. Reported *p*-values are two-tailed unless testing a one-tailed prediction, as signified by bold face.
Forecast optimism = Excess of forecast over the actual performance in the second round of SAT-type questions, where forecast is the forecast of total score in the second round of SAT-type questions in the aggregated forecast conditions and the sum of the four component forecasts in the second round of SAT-type questions in the disaggregated forecast conditions.
Trial Performance = Participants' performance in the first round of SAT-type questions.
Performance = Participants' actual performance in the second round of SAT-type questions.

Effect of disaggregation on forecast accuracy in the absence of performance-based incentives

To develop H1 we rely on arguments suggesting that in the absence of performance-based incentives the disaggregated forecast type results in: (1) greater attention to information for each component forecast and (2) random errors in the component forecasts that cancel each other out in the top-level forecast. These two effects should lead to greater forecast accuracy in the disaggregated forecasts than in the aggregated forecasts in the absence of performance-based incentives.

The first argument implies the forecast for each component is more precise and better calibrated under the disaggregated forecast type than under the aggregated forecast type (e.g., Henrion et al., 1993). To test this, we compare the standard deviation of the absolute forecast error in the absence of performance-based incentives/aggregated forecast condition and the standard deviation of the sum of the absolute forecast errors of the component forecasts in the disaggregated forecast condition. Since we expect disaggregated forecasts to be better calibrated than aggregated forecasts, we expect the standard deviation to be lower in the condition where participants prepare disaggregated forecasts. As shown in Table 4, Panel B, a Levene's test of equal variances confirms this conjecture (2.12 vs. 4.26, $p = 0.03$, one-tailed) (Levene, 1960).

The second argument implies that the sum of the absolute forecast errors of the four component forecasts in the disaggregated forecast condition is not necessarily lower than the absolute forecast error in the aggregated forecast condition. This is because the greater accuracy for the disaggregated condition is partially driven by the error cancellation in the aggregation process. Our additional analyses are consistent with this argument. Due to the unequal variances based on the Levene's test as discussed above, we use Welch's *t*-test to compare the means of the absolute forecast errors between the aggregated forecast and the disaggregated forecast conditions (Welch, 1951). As summarized in Table 4, Panel C, Welch's *t*-test indicates no significant difference between the mean absolute forecast error of aggregated forecasts and the sum of the absolute forecast errors of the components of disaggregated forecasts (4.80 for aggregated forecasts and 4.50 for the sum of the components of disaggregated forecasts, $p = 0.75$). Combined with our result for H1 that disaggregated forecasts are more accurate than aggregated forecasts in the absence of performance-based incentives, this result is consistent with our second argument that errors in the component forecasts cancel each other out.

Effect of performance-based incentives on forecast optimism in disaggregated forecasts

Both H1 and H2 rely on the argument that, in the presence of performance-based incentives, preparing disaggregated forecasts will cause the manager to attend to more detailed information about how a favorable outcome is

Table 4
Comparison between the absolute forecast error of aggregated forecasts and the sum of the absolute forecast errors of the four components of disaggregated forecasts in the absence of performance-based incentives.

	n	Aggregated forecast	*n*	Disaggregated forecast
Panel A: Mean (standard deviation) for absolute forecast errors in the absence of performance-based incentives condition				
Absolute Forecast Error	30	4.80 (4.26)	18	4.50 (2.12)
		df	*F*	*p*-value
Panel B: Levene's test of equal variances				
Absolute Forecast Error	Aggregated vs. Disaggregated	1	4.04	**0.03**
		df	*F*	*p*-value
Panel C: Welch's t-test of equal means				
Absolute Forecast Error	Aggregated vs. Disaggregated	1	0.11	0.75

The table presents descriptive statistics, Levene's test of equal variances, and Welch's *t*-test of equal means between the absolute forecast error in the aggregated forecast condition and the sum of absolute forecast errors of the four components in the disaggregated forecast condition in the absence of performance-based incentives. See Table 1 for descriptions of the manipulations of *performance-based incentives* and *forecast type*. Reported *p*-values are two-tailed unless testing a one-tailed prediction, as signified by bold face.
Absolute Forecast Error is defined as the absolute difference between the forecast and the actual performance in the second round of SAT-type questions in the aggregated forecast/absence of performance-based incentives condition and the sum of absolute forecast errors of the four component forecasts in the second round of SAT-type questions in the disaggregated forecast/absence of performance-based incentives condition.
Levene's test is a statistical test to assess the equality of variances between two groups (Levene, 1960).
Welch's *t*-test is a statistical test to assess the equality of means between two groups when the variances are unequal between the two groups (Welch, 1951).

likely to be achieved and, hence, allow the manager to inject more optimism into each component forecast of a disaggregated forecast. The argument suggests that when comparing disaggregated forecasts in the presence of performance-based incentives with those in the absence of performance-based incentives we should observe a significantly positive effect of performance-based incentives on forecast optimism (as measured by the signed forecast errors) in the component forecasts. To test this argument, we conduct a 2×4 repeated measures ANCOVA on the signed errors of the component forecasts, where the first factor is performance-based incentives (present vs. absent) and the second factor is the repeated measure. To be consistent with our tests of H1 and H2, we control for the component performance in both the first and second rounds of the task.

The repeated measures ANCOVA analysis is reported in Table 5. As shown in Table 5, Panel A, the least square mean of the signed errors in component forecasts in the presence of performance-based incentives is greater than that in the absence of performance-based incentives. Panel B further shows that the effect of performance-based incentive on component forecast optimism is significantly positive ($p = 0.01$, one-tailed). These results support our argument that, in the presence of performance-based incentives, preparing disaggregated forecasts induces greater optimism in each component forecast.

Overall, the above results are consistent with our theoretical arguments. When participants' incentives are not explicitly tied to performance, producing disaggregated forecasts leads to detailed information processing and cancellation of errors in the forecasting process. However, when participants' incentives are tied to performance, participants have a preference for favorable performance, which results in more optimistically biased forecasts for each component, mitigating the error cancellation effect for the disaggregated forecasts. Consequently, although producing disaggregated forecasts leads to greater forecast accuracy when explicit performance-based incentives are absent, this effect is reduced when explicit performance-based incentives are present.

Table 5
The effects of performance-based incentives on component forecast optimism in the disaggregated forecast conditions.

		n	Component forecast optimism
Panel A: Least square means (standard deviation) for component forecast optimism			
Absence of Performance-Based Incentives		72	0.21 (0.20)
Presence of Performance-Based Incentives		104	0.81 (0.17)
Source	*df*	*F*	*p*-value
Panel B: Repeated measures ANCOVA analysis of component forecast optimism			
Performance-Based Incentives	1	5.23	**0.01**
Component Forecasts	3	8.26	<0.01
Performance-Based Incentives × Component Forecasts	3	1.82	0.15
Component Trial Performance	1	148.68	<0.01
Component Performance	1	57.95	<0.01

The table presents descriptive statistics and the repeated-measures ANCOVA model for component forecast optimism in the disaggregated forecast conditions, with component forecasts being the repeated measure. See Table 1 for descriptions of the manipulations of *performance-based incentives*. Reported *p*-values are two-tailed unless testing a one-tailed prediction, as signified by bold face.
Component Forecast Optimism = Excess of component forecast over the actual component performance in each category of the second round of SAT-type questions, where component forecast is the forecast of each of the four categories in the second round of SAT-type questions in the disaggregated forecast conditions.
Component Trial Performance = Participants' performance in each category of the first round of SAT-type questions.
Component Performance = Participants' actual performance in each category of the second round of SAT-type questions.

Table 6
The effects of forecast type and performance-based incentives on perceived forecasting difficulty.

	Forecast type			
	n	Aggregated forecast	*n*	Disaggregated forecast
Panel A: Mean (standard deviation) for forecasting difficulty				
Absence of Performance-Based Incentives	30	7.37 (2.50)	18	8.00 (1.91)
Presence of Performance-Based Incentives	18	7.67 (1.85)	26	6.42 (2.14)
Source	*df*	Mean square	*F*	*p*-value
Panel B: ANCOVA model of forecasting difficulty				
Performance-Based Incentives	1	15.60	3.59	0.06
Forecast Type	1	1.58	0.36	0.55
Performance-Based Incentives × Forecast Type	1	17.78	4.09	**0.02**
Trial Performance	1	28.35	6.52	0.01
Performance	1	20.76	4.78	0.03
Error	82	4.35		

Source	*df*	*F*	*p*-value
Panel C: Planned contrast coding and follow-up simple effect tests			
Overall tests:			
Producing a disaggregated forecast in the presence of performance-based incentives is perceived as the easiest. Forecasting in the other three conditions is perceived as more difficult. Contrast weights (3, −1, −1, −1)	1	7.86	**<0.01**
Follow-up simple effect tests:			
Effect of forecast type in the absence of performance-based incentives	1	1.03	0.31
Effect of forecast type in the presence of performance-based incentives	1	3.44	**0.03**
Effect of performance-based incentives when disaggregated forecasts are prepared	1	7.76	**<0.01**
Effect of performance-based incentives when aggregated forecasts are prepared	1	0.00	0.95

The table presents descriptive statistics, the ANCOVA model, and contrast coding tests for the perceived forecasting difficulty for the four treatments. See Table 1 for descriptions of the manipulations of *performance-based incentives* and *forecast type*. Reported *p*-values are two-tailed unless testing a one-tailed prediction, as signified by bold face.
Forecasting Difficulty = Participants' response on an 11-point scale eliciting the extent to which they agree with the statement "I think that forecasting my performance was difficult". Higher scores indicate higher perceived forecasting difficulty.
Trial Performance = Participants' performance in the first round of SAT-type questions.
Performance = Participants' actual performance in the second round of SAT-type questions.

Effects of forecast type and performance-based incentives on forecasting difficulty

Our theoretical development argues that individuals with explicit performance-based incentives accept preference-consistent information more readily and attend less to preference-inconsistent information, particularly when disaggregated components are more accessible. Therefore, we expect that in the presence of explicit performance-based incentives it is easier for individuals to reach their desired conclusion when they make disaggregated forecasts. That is, preparing disaggregated forecasts should be perceived as easier than preparing aggregated forecasts when explicit performance-based incentives are present compared to when such incentives are absent.

To test this prediction, we measure perceived forecasting difficulty by asking participants to indicate on an 11-point scale to what extent they agree with the statement "I think that forecasting my performance was difficult". Higher scores indicate greater perceived forecasting difficulty. Table 6 shows that the results are consistent with our expectations. Specifically, Table 6, Panel B shows a significant interaction effect between performance-based incentives and forecast type on perceived forecasting difficulty ($p = 0.02$, one-tailed). The simple contrasts in Table 6, Panel C also support our prediction and show that when participants' compensation is linked to performance, they indicate that producing disaggregated forecasts is less difficult than producing aggregated forecasts (6.42 vs. 7.67, $p = 0.03$, one-tailed). In contrast, when participants' compensation is not tied to performance, the perceived forecasting difficulty is not significantly different between the two forecast types (8.00 vs. 7.37, $p = 0.31$). This result on perceived forecasting difficulty is consistent with our theoretical argument.

Conclusion

In this study, we use an abstract experiment to examine how forecast type interacts with performance-based incentives to influence both the accuracy and optimism of management forecasts. We find that: (1) preparing disaggregated forecasts leads to a greater improvement in forecast accuracy (compared to preparing aggregated forecasts) in the absence of performance-based incentives than in the presence of performance-based incentives; and (2) preparing disaggregated forecasts leads to a greater increase in forecast optimism (compared to preparing aggregated forecasts) in the presence of performance-based incentives than in the absence of performance-based incentives.

Although our study focuses on the disaggregation of information in managers' performance forecasts, our results have implications for a wide variety of disclosures prepared by managers. Both textual (e.g., MD&A) and verbal (e.g., conference call) disclosures can vary in the extent to which qualitative and quantitative information is disaggregated. Furthermore, the Financial Accounting Standards Board's (FASB's) emphasis on disaggregated financial reporting as part of their Financial Statement Presentation Project suggests that disaggregation may play an increasingly important role in mandatory disclosures, in addition to voluntary disclosures. Our study suggests that the quality of these disclosures may vary depending on the extent to which these disclosures are disaggregated and the forecasting approach managers use to arrive at these disclosures. We also expect that, to the extent that analysts work closely with the firms on which they provide forecasts and have incentives to provide optimistic forecasts (Koonce & Mercer, 2005; Libby, Hunton, Tan, & Seybert, 2008), they likely succumb to similar biases to those that we document in this study. Such biases are likely exacerbated when analysts forecast components of earnings before forecasting a bottom-line earnings number.

Results of our study should be interpreted with several caveats in mind. First, we use an abstract experimental task to maximize internal validity. In practice, a firm's operational environment contains much more information than the abstract task in our study. We expect that preparing disaggregated forecasts may have an even stronger effect on forecast optimism when performance-based incentives are present in the real world with a richer information environment.

Second, while our results are consistent with the predictions of motivated reasoning theory, we do not capture actual measures of participants' biased information processing. Given the nature of our task, it was not clear how objective measures of information processing could be captured without potentially influencing participants' attention to information and their resulting forecasts. Further, we do not manipulate the direction of performance incentives in our task because managers typically do not have an incentive to perform poorly. Future research could investigate whether disaggregated forecasts are more pessimistic than aggregated forecasts when managers have a negative directional goal (for example, if they have a preference to write-down a particular asset).

Third, we focus on unintentional biases and hold managers' other incentives constant in our experiment. In reality, managers likely have other incentives that may lead to intentional biases in their forecasts. For example, managers may have incentives to provide pessimistic management forecasts to guide down analysts' forecasts and increase the probability of achieving targets. Future studies can examine how forecast type interacts with other managerial incentives to influence the quality of management forecasts.

Acknowledgements

We thank the Editor (Lisa Koonce), the Discussant (Jeffrey Hales), and an anonymous reviewer for helpful guidance and suggestions. We also thank Nathalie Beckers, Chris Chapman, Kristina Demek, Bart Dierynck, Brooke Elliott, Cassandra Estep, Christoph Feichter, Brent Garza, Steph Grant, Lan Guo, Kevin Jackson, Qiwei Li, Bob Libby, Joan Luft, Mark Peecher, Steve Salterio, Karl Schuhmacher, Timothy Shields, Huaxiang Yin and workshop participants at Maastricht University, Nanyang Technological University, Tilburg University, Wilfrid Laurier University, the *Accounting, Organizations and Society* Conference on Accounting Estimates, the American Accounting Association Annual Meeting, and the American Accounting Association Management Accounting Section Midyear Meeting for their helpful comments. We gratefully acknowledge financial support from Deloitte & Touche LLP and the National Natural Science Foundation of China Funds (No. 71402086).

References

Bonner, S. E. (2008). *Judgment and decision making in accounting*. Upper Saddle River, NJ: Prentice Hall.

Bromiley, P. (1987). Do forecasts produced by organizations reflect anchoring and adjustment? *Journal of Forecasting, 6*, 201–210.

Bruggen, A., Grabner, I., & Sedatole, K. (2013). *Toward the unknown: The effect of forecast accuracy on production decisions*. Working paper, Maastricht University and Michigan State University.

Buckless, F. A., & Ravenscroft, S. P. (1990). Contrast coding: A refinement of ANOVA in behavioral analysis. *The Accounting Review, 65*(4), 933–945.

Chow, C. W., Cooper, J., & Waller, W. S. (1988). Participative budgeting: Effects of a truth-inducing pay scheme and information asymmetry on slack and performance. *The Accounting Review, 63*(1), 111–122.

Church, B. K., Hannan, R. L., & Kuang, X. (2012). Shared interest and honesty in budget reporting. *Accounting, Organizations and Society, 37*(3), 155–167.

Ditto, P. H., & Lopez, D. A. (1992). Motivated skepticism: Use of differential decision criteria for preferred and non-preferred conclusions. *Journal of Personality and Social Psychology, 63*, 568–584.

Duru, A., & Reeb, D. M. (2002). International diversification and analysts' forecast accuracy and bias. *The Accounting Review, 77*(2), 415–433.

Elliott, W. B., Hobson, J. L., & Jackson, K. E. (2011). Disaggregating management forecasts to reduce investors' susceptibility to earnings fixation. *The Accounting Review, 86*(1), 185–208.

Fisher, J. G., Maines, L. A., Peffer, S. P., & Sprinkle, G. B. (2002). Using budgets for performance evaluation: Effects of resources allocation and horizontal information asymmetry on budget proposals, budget slack, and performance. *The Accounting Review, 77*(4), 847–865.

Goodman, T. H., Neamtiu, M., Shroff, N., & White, H. D. (2014). Management forecast quality and capital investment decisions. *The Accounting Review, 89*(1), 331–365.

Graham, J., Harvey, C. R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics, 40*, 3–73.

Hales, J. (2007). Directional preferences, information processing, and investors' forecasts of earnings. *Journal of Accounting Research, 45*, 607–628.

Healy, P., & Palepu, K. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics, 31*, 405–440.

Hemmer, T., & Labro, E. (2008). On the optimal relation between the properties of managerial and financial reporting systems. *Journal of Accounting Research, 46*(5), 1209–1240.

Henrion, M., Fischer, G. W., & Mullin, T. (1993). Divide and conquer? Effects of decomposition on the accuracy and calibration of subjective probability distributions. *Organizational Behavior and Human Decision Processes, 55*, 207–227.

Hirst, D. E., Koonce, L., & Venkataraman, S. (2007). How disaggregation enhances the credibility of management earnings forecasts. *Journal of Accounting Research, 45*(4), 811–837.

Hutton, A. P., Miller, G. S., & Skinner, D. J. (2003). The role of supplementary statements with management earnings forecasts. *Journal of Accounting Research, 41*, 867–890.

Jennings, R. (1987). Unsystematic security price movements, management earnings forecasts, and revisions in consensus analyst earnings forecasts. *Journal of Accounting Research, 25*, 90–110.

Kahn, K. B. (1998). Revisiting top-down versus bottom-up forecasting. *Journal of Business Forecasting, 17*(2), 14–19.

Please cite this article in press as: Chen, C. X., et al. The effects of forecast type and performance-based incentives on the quality of management forecasts. *Accounting, Organizations and Society* (2015), http://dx.doi.org/10.1016/j.aos.2015.03.002

Klayman, J., Soll, J. B., González-Vallejo, C., & Barlas, S. (1999). Overconfidence: It depends on how, what, and whom you ask. *Organizational Behavior and Human Decision Processes, 79*, 216–247.

Kleinmuntz, D. N., Fennema, M. G., & Peecher, M. E. (1996). Conditional assessment of subjective probabilities: Identifying the benefits of decomposition. *Organizational Behavior and Human Decision Processes, 66*(1), 1–15.

Koonce, L., & Mercer, M. (2005). Using psychology theories in archival financial accounting research. *Journal of Accounting Literature, 24*, 175–214.

Kremer, M., Siemsen, E., & Thomas, D. (2012). *The sum and its parts: A behavioral investigation of top-down and bottom-up forecasting processes*. Working paper. Penn State University and the University of Minnesota.

Krueger, J., & Mueller, R. A. (2002). Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology, 82*, 180–188.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*, 1121–1134.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*, 480–498.

Lansford, B., Lev, B., & Tucker, J. W. (2013). Causes and consequences of disaggregating earnings guidance. *Journal of Business Finance and Accounting, 40*(1&2), 26–54.

Lapide, L. (2006). Top-down & bottom-up forecasting in S&OP. *Journal of Business Forecasting, 25*(2), 1–3.

Larrick, R. P., Burson, K. A., & Soll, J. B. (2007). Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes, 102*, 76–94.

Levene, H. (1960). *Contributions to probability and statistics*. Stanford, CA: Stanford University Press.

Libby, R., Bloomfield, R., & Nelson, M. (2002). Experimental research in financial accounting. *Accounting, Organizations and Society, 27*, 777–812.

Libby, R., & Brown, T. (2013). Financial statement disaggregation decisions and auditors' tolerance for misstatement. *The Accounting Review, 88*(2), 641–665.

Libby, R., Hunton, J., Tan, H.-T., & Seybert, N. (2008). Relationship incentives and the optimistic/pessimistic pattern in analysts' forecasts. *Journal of Accounting Research, 46*(1), 173–198.

Libby, R., & Rennekamp, K. (2012). Self-serving attribution bias, overconfidence, and the issuance of management forecasts. *Journal of Accounting Research, 50*, 197–231.

McNichols, M., & Stubben, S. R. (2008). Does earnings management affect firms' investment decisions? *The Accounting Review, 83*(6), 1571–1603.

Mercer, M. (2005). The fleeting effects of disclosure forthcomingness on management's reporting credibility. *The Accounting Review, 80*(2), 723–744.

Merchant, K., & Van der Stede, W. A. (2012). *Management control systems: Performance measurement, evaluation, and incentives* (3rd ed.). London: Prentice Hall.

Mikhail, M. B., Walther, B. R., & Willis, R. H. (1999). Does forecast accuracy matter to security analysts? *The Accounting Review, 74*(2), 185–200.

Ajinkya, B., & Gift, M. (1984). Corporate managers' earnings forecasts and symmetrical adjustments of market expectations. *Journal of Accounting Research, 22*, 425–444.

Osband, K. (1989). Optimal forecasting incentives. *Journal of Political Economy, 97*(5), 1091–1112.

Raiffa, H. (1968). *Decision analysis: Introductory lectures on choices under uncertainty*. Reading, MA: Addison-Wesley.

Ravinder, H. V., Kleinmuntz, D. N., & Dyer, J. S. (1988). The reliability of subjective probabilities obtained through decomposition. *Management Science, 34*(2), 186–199.

Shields, M., & Young, S. M. (1993). Antecedents and consequences of participative budgeting: Evidence on the effects of asymmetrical information. *Journal of Management Accounting Research, 5*, 265–280.

Thayer, J. (2011). Determinants of investors' information acquisition: Credibility and confirmation. *The Accounting Review, 86*(1), 1–22.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.

Tversky, A., & Koehler, D. J. (1994). Support theory: A non-extensional representation of subjective probability. *Psychological Review, 101*, 547–567.

Van Boven, L., & Epley, N. (2003). The unpacking effect in evaluative judgments: When the whole is less than the sum of its parts. *Journal of Experimental Social Psychology, 26*, 269–280.

Waller, W. (1988). Slack in participative budgeting: The joint effects of a truth-inducing pay scheme and risk preferences. *Accounting, Organizations, and Society, 13*, 87–98.

Webb, A. (2002). The impact of reputation and variance investigations on the creation of budget slack. *Accounting, Organizations, and Society, 27*(4–5), 361–378.

Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika, 38*, 330–336.

Young, S. M. (1985). Participative budgeting: The effects of risk aversion and asymmetric information on budgetary slack. *Journal of Accounting Research, 23*, 829–842.

Zimbelman, M. F. (1997). The effects of SAS No. 82 on auditors' attention to fraud risk factors and audit planning decisions. *Journal of Accounting Research, 35*, 75–97.