The Journal of Behavioral Finance
2007, Vol. 8, No. 3, 154–160

Reporting Frequency and Sample Size: Effects on Prediction, Confidence Levels, and Confidence Intervals

Terence J. Pitre

Very little research has examined the possible consequences of more frequent financial reporting. Using a between-subjects experiment, I examine one possible consequence—increased sample size of data—and its effect on non-professional investor uncertainty (as measured by confidence intervals and confidence levels) and predictions. I report three principal findings: 1) Confidence intervals increase with larger sample sizes, rather than decrease as statistical theory suggests, 2) confidence levels are unaffected by sample size when the investor does not view it as important for accuracy, and 3) estimates generated from larger sample sizes are nearer to the sample mean and significantly different from those from smaller samples, which also contradicts statistical theory.

keywords: *disclosure frequency, forecasts, accuracy, confidence levels.*

Introduction

The idea of more frequent financial reporting has become increasingly popular. Past Securities & Exchange Commission Chairman Harvey Pitt has argued that "quarterly filings produce an out-of-date snap shot rather than a real-time window" (Levitt [2002]). And as one Fortune 500 CFO states, "We need to reach beyond the quarterly reports to provide more timely information and more context about future prospects" (McNamee [2001]).

Firms such as Cisco, with the aid of in-house technology, have already adopted the concept of real-time internal reporting. The continued emergence and acceptance of technologies such as XBRL (eXtensible Business Reporting Language) is expected to make more frequent external reporting a possibility for more companies (Watson, McGuire, and Cohen [2000]).

More frequent reporting, however, can affect financial data in many ways that have not been fully examined in either the popular press or the academic literature. For some accounts, it will result in more disaggregated data (e.g., weekly rather than quarterly income). For others, the result may be a repetition of unchanged values (e.g., the par value of common stock or the book value of assets). Still other accounts may find that more frequent reporting will result in a larger sample of observations (e.g., accounts receivable). Each of these changes affects the statistical properties of accounting data, as well as how investors cognitively process it and use it to make important judgments and decisions.

The purpose of this study is to examine the cognitive effect of one change that could result from more frequent reporting—larger samples sizes—on non-professional investor judgments. In an accounting-based experimental setting, I analyze how larger sample size affects investor predictions of future account values, as well as two dimensions of investor uncertainty about their predictions: *confidence level,* and *confidence interval*. Using psychology theory, I formulate and test hypotheses relating to the accuracy and the uncertainty of predictions.

Confidence levels and confidence intervals have an inverse relationship: as confidence levels increase, confidence intervals should decrease. Levels and intervals are often used as alternative measures of a one-dimensional uncertainty construct. Statistical theory suggests this uncertainty should be lower when individuals base their predictions on a larger sample.

Following Peterson and Pitz [1986, 1988], I predict and show that sample size affects the two measures differently than conventional statistical theory. Although the small and large samples have identical standard deviations in the experimental setting, the larger sample *increases* subject uncertainty when measured by confidence intervals and volatility judgments. Confidence levels (the stated probability of being correct) are not affected by sample size differences.

Subjects believe smaller samples are less volatile. Thus, when an observation deviates moderately from the mean, subjects in the small-sample condition are more likely to view that observation as an outlier, and either underweight it or not use it at all. Underweighting or ignoring an observation from a small sample

Terence J. Pitre PhD, Assistant Professor, Graduate School of Business, Department of Accounting, TMH 343-A, 1000 LaSalle Street, Minneapolis, MN 55403, Phone: (651) 962-4710; Fax: (651) 962-4246. Email: tjpitre@stthomas.edu

(four quarterly observations) has a larger impact on the prediction than similar actions would have on a larger sample (fifty-two weekly observations). Consequently, predictions made from small samples are more likely to deviate systematically from the sample mean.

My results have several implications for increasing reporting frequency. On the one hand, I find that a larger sample increases individuals' use of an unbiased predictor (the sample mean). One the other hand, the effect of large samples on subjective judgments of uncertainty (and ultimately risk judgments) can be problematic. Individuals' uncertainty about predictions increases rather than decreases when they have a larger sample on which to base their predictions.

The remainder of the paper is organized as follows: Section 2 contains a literature review and develops my hypotheses. Section 3 contains the experimental design and statistical tests. Section 4 contains my empirical results, and section 5 presents conclusions.

Literature Review and Hypothesis Development

Statistical theory regarding the relationship among sample size, predictions, and uncertainty is clear, if not intuitive. Estimates based on small samples have larger standard errors than those based on larger samples. Thus the use of larger samples should result in more accurate predictions, and individuals should be more certain about their predictions when the sample size increases. In other words, a larger sample should increase confidence levels and decrease confidence intervals.

Psychology research has found that judgment is often insensitive to sample size. Kahneman and Tversky [1972] demonstrated that people typically assess the likelihood of a sample result by the similarity of the result to the corresponding parameter of the population, regardless of sample size. Consequently, the judged probability of a sample statistic (and the uncertainty about it) is independent of sample size.

Subsequent research has found mixed support for Kahneman and Tversky's [1972] assertions, however. Sedlmeier and Gigerenzer [1997] have summarized this research, and propose that the inconsistency results from whether the subjects' task was a frequency distribution task or a sampling distribution task. Subjects in the frequency distribution task were required to decide whether a single large-sample mean or a single small-sample mean was more likely to be closer to the population mean. Subjects in the sampling distribution task were required to decide whether sample means much larger than the population mean were more common in large samples or in small samples. Performance differed between these two tasks: Individuals were appropriately sensitive to sample size in the frequency distribution, but not in the sampling distribution.

Neither of these tasks required subjects to predict a future value and state how uncertain they were about their predictions based on sample sizes. Such predictions and uncertainty estimates are an important use of accounting information, and it is unclear from the previous literature how sample size will affect performance on these tasks. Previous studies have also not examined the impact of the small sample being a subset of the larger sample, a condition specific to the increased reporting setting.

To develop my hypotheses, I used research on the effects of information quantity in multiple-cue prediction tasks as a theoretical foundation. A classic study by Oskamp [1965] found that an increase in the amount of information increased confidence levels without increasing prediction accuracy. Conversely, Peterson and Pitz [1986] found that an increase in the amount of information increased both accuracy and uncertainty.

Peterson and Pitz [1986] offered two explanations for these differences. First, they found that accuracy increased because the additional information consisted of informative cues that were not related to the cues subjects already had. In prior studies (like Oskamp [1965]), the additional information tended to be redundant. Second, uncertainty increased because it was measured differently, using confidence intervals rather than confidence levels.

In a later related study, Peterson and Pitz [1988] found that both confidence levels and uncertainty increased with an increase in information (the number of cues). They measured uncertainty with a confidence level question, and found that confidence levels increased significantly when the amount of information was manipulated within-subjects. This suggests that the amount of information can lead to an increase in confidence levels, *if* it is important for accuracy. Taken together, these studies suggest that an increase in the amount of information (cues) will result in an increase in confidence intervals, an increase or no change in confidence levels, and an increase in accuracy.

Peterson and Pitz [1986, 1988] argue that uncertainty increases when measured with confidence intervals as the number of alternatives increases. In this view, uncertainty increases because additional information provides individuals with alternatives that were not considered when not explicitly given. In an accounting setting, a sample of fifty-two weekly observations will provide more alternative values than a sample of four quarterly observations. Thus, the larger sample should result in wider confidence intervals, not narrower ones as statistical theory suggests.

Moreover, in an accounting setting, the small sample of quarterly observations is a subset of the fifty-two weekly observations. Thus the larger sample will *always* have a range equal to or wider than the smaller sample. There is evidence that individuals base their judgments of uncertainty on the range or extreme

values of a distribution, and do not weight outcomes by probabilities in the way standard deviation calculations do (Shapira [1995]). Therefore, I predict that the uncertainty as measured by confidence intervals will be wider in the large-sample condition than in the small-sample condition.

H_1 Subjects in the large-sample condition will have wider confidence intervals than subjects in the small-sample condition.

Increases in information do not decrease all measures of certainty. An increase in the amount of information has been shown to increase confidence levels even if it does not increase accuracy (Oskamp [1965], Hoge [1970], Slovic [1973]). Peterson and Pitz [1988] demonstrate, however, that increases in confidence levels associated with increases in the amount of information are significant only in a within-subjects manipulation that renders the additional information salient because it is important for accuracy. I use a between-subjects design to avoid demand effects. Accordingly, I predict that confidence levels will not be significantly different across the large- and small-sample conditions.

H_2 There will be no significant difference in confidence levels between the large-sample condition and the small-sample condition.

The increase in accuracy exhibited in Peterson and Pitz [1986] was associated with subjects' willingness to enter extreme estimates on the basis of little information. As the amount of information (the number of cues) increased, the estimates tended to be closer to the mean of the criterion and closer to the actual outcomes.

Stated mathematically, subjects tended to overweight the predictors in the equation $y = \alpha + \beta_i X_i + \varepsilon$ when fewer predictors were used. But in this study I use a task in which the sample mean is the best predictor ($y = 1/n\ X_i$). It thus seems less likely that the individuals will overweight all the X_is in the small sample. It is possible that the weights on the X_is will not be identical across conditions, however. Therefore, I test the following hypothesis.

H_3 Predictions will be different across the large- and small-sample conditions.

Research Design

I use a between-subjects design experiment in which the task is to predict a future period value based on historical data. The independent variable is the sample size of the historical data. The two treatment conditions consist of a large-sample condition resulting from weekly reporting (fifty-two observations), and a small-sample condition resulting from quarterly reporting (four observations). The dependent variables are the predictions of the future period value, the confidence level, and confidence intervals about the predictions. Appendix A contains the actual questions given to participants.

In order to simulate the weekly historical data, I use a random number generator to produce fifty-two numbers. From those, I randomly chose four to represent the quarter-end balances. The four quarterly values were chosen so that both data sets have equal means and variances. This allows me to test for pure sample size effects, avoiding confounds between sample size and sample mean or variance.

Sixty-eight accounting majors served as subjects for this experiment. Subjects were asked to predict the value of a marketable securities account at the end of the next quarter. To obtain comparable results, both conditions required subjects to predict the next future quarter-end value. Subjects were further told that company management always has a target balance amount for the account, that the target has not changed in the past year, and that it is not expected to change in the next year. The participants were instructed that the balance sometimes fluctuates above or below the target amount due to changes in value of the securities and the cash needs of the firm. This scenario implied that the best predictor of the future balance is the mean of the historical observations. This portion of the instructions was intended to simplify the task and make it more suitable for the subject pool.[1]

In order to better measure the impact of more frequent reporting on confidence intervals and confidence levels, subjects answered questions in one of two formats: 1) predicted value and confidence interval, or 2) predicted value and confidence level. This reduces any demand effects that might arise if a subject received both confidence level and confidence interval questions and believed he or she was supposed to respond consistently to both.

In a post-experiment questionnaire, subjects provided demographic information and responded to several questions eliciting their judgments about the volatility of the data, their fundamental statistical knowledge, and the method they used to generate their predictions.

Results

Subjects had completed few academic courses in accounting and statistics, as indicated by the mean numbers of courses taken, 0.27 and 0.72, respectively. Average overall investing experience was 6.20 months. The post-experiment questionnaires revealed that subjects were risk-neutral. This was determined by asking students to indicate how much ($0–$20) they would invest if there were a fifty-fifty chance of winning or losing $20. The average amount invested by all subjects was $10.13.

Table 1. *Correlations*

	Sample Size	Confidence Level	Confidence Interval	Prediction	Volatility
Sample Size	1				
Confidence Level	0.171	1			
Confidence Interval	0.521**	a	1		
Prediction	0.326**	−0.069	0.517**	1	
Volatility	0.241*	0.113	0.398*	0.120	1

*Correlation is significant at the 0.01 level (two-tailed).
**Correlation is significant at the 0.05 level (two-tailed).
a = Cannot be computed because subjects did not have to respond to both types of questions.

The correlations in Table 1 indicate that the frequency of reporting (as measured by the sample size of the data) is positively and significantly correlated with predictions, confidence intervals, and perceived volatility (0.326, 0.521, and 0.241, respectively). The measure "sample size" was coded 1 for quarterly reporting (small sample), and 2 for weekly reporting (large sample). Thus the inference is that larger samples lead to higher predictions, wider intervals, and higher judgments of volatility. Frequency of reporting was negatively but not significantly correlated with confidence levels (−0.126).

Table 2 reports the dependent variable means and standard deviations and the results of hypothesis tests. H1 predicted that the confidence interval would be wider in the large-sample condition (more frequent reporting). The mean interval for the large-sample group was 29.0, while it was 17.66 for the small-sample group. The difference was significant ($t = -3.457$, $p < 0.002$), indicating subjects were more certain about predictions based on a smaller sample.

H2 predicted that confidence levels would not be significantly different across the two groups, and the data were consistent with this prediction ($t = 0.982$, $p < 0.33$). As in previous research, when the difference in the quantity of information is not important for accuracy, there is no effect on sample size. Unlike previous studies (Peterson and Pitz [1986, 1988]), however, the mean of the small-sample group (45.16) was higher than the mean of the larger-sample group (35.62). This might imply that subjects' confidence levels decreased with sample size, rather than increased as previous research indicated. I examine this issue further in the conclusion.

H3 predicted that subject predictions of future account values would differ across sample size conditions. The mean prediction for the small-sample group was 16.25, while it was 19.81 for the large-sample group ($t = -2.679$, $p < 0.011$[2]).

Supplemental Analysis

I conducted an additional test to determine whether larger samples led to predictions closer to the mean of the sample. Specifically, I test the null hypothesis that

Table 2. *Descriptive Statistics and Test Results of the Effects of Sample Size on Predictions, Confidence Levels, Confidence Intervals, and Perceived Volatility*

Variable	N	Mean Small Sample(Std. Dev.)	Mean Large Sample(Std. Dev.)	t-Statistic	df	Sig. p >
Panel A:						
Hypothesis Tests						
Confidence Interval	34	17.66 (10.11)	29.00 (8.85)	−3.457	32	0.002
Confidence Level	34	45.16 (27.10)	35.62 (29.57)	0.982	32	0.33
Prediction	68	16.25 (2.82)	19.81 (7.03)	−2.679*	39.77*	0.011*
Panel B:						
Supplemental Tests						
Difference Between Prediction and Sample Mean of 21						
Small Sample	36			−10.09	35	0.000
Larger Sample	32			−0.955	31	0.347
Volatility	68	3.88 (1.84)	4.84 (2.04)	−2.020	66	0.047

*Equal variances not assumed.
Confidence interval = the subjective confidence interval about the estimate given.
Confidence levels = the subjective probability that the prediction is correct.
Prediction = the estimated future ending balance for the account.
Volatility = the judged volatility of the data (1 = 10% and 10 = 100%).

each of the predicted means is equal to 21 (the mean value of both samples).

Table 2 reports the results of the t-test. I failed to reject the null hypothesis with the predicted mean of 19.81 from the large-sample group ($t = -0.955$, $p < 0.347$). However, the null hypothesis is clearly rejected for the small-sample group, with a mean of 16.25 ($t = -10.09$, $p < 0.000$). Jointly considered with the results of hypothesis 3, I find that sample size differences lead to predictions that are significantly different from each other, and that large sample sizes result in more frequent use of an unbiased predictor (the sample mean).

The difference in predictions can further be explained by examining the answers in the post-experiment questionnaire. One question concerned subjects' volatility judgments of their data sets. Because both samples had identical means and variances, I expected both groups to have similar beliefs about the volatility of the data. However, mean volatilities for the small- and large-sample groups, respectively, were 3.88 and 4.84 (on a scale of 1 to 10, where 1 = the lowest volatility and 10 = the highest).

Results of the t-test for differences due to sample size in Table 2 indicate that the means were different from each other ($t = -2.020$, $p< 0.047$). This result helps explain the difference in predictions. Those in the small-sample group clearly believed their data were less volatile. As a result, they believed there was an "outlier" or an "extreme" observation in the data, and either underweighted it or completely ignored it when making their predictions (as noted from the comments section of the questionnaire).

This may also help explain why confidence levels were higher in the small-sample condition, contradicting the findings of previous research. If subjects clearly believed one observation was an outlier, they probably believed that underweighting or ignoring it would enhance the quality of the sample and subsequently their prediction, thus leading to increased confidence in their prediction.

Alternatively, because subjects in the large-sample condition had more observations to consider, they were not affected by the extreme values in their predictions. But perhaps they were less confident due to the larger amount of data and the larger cognitive load. Future research could investigate these phenomena further.

Conclusions

This paper examines one possible effect of increased reporting—larger sample sizes—on non-professional investor decision-making. I find three principal results: First, inconsistent with statistical theory, confidence intervals increased (i.e., the uncertainty increased) with increases in sample size. Second, as hypothesized, an alternative measure of uncertainty, confidence level, was not significantly affected by sample size. Finally, predictions based on larger samples were closer to the sample mean and significantly different from predictions based on smaller samples.

This experiment provides evidence that subjects in the small-sample condition made less use of what they believed to be "extreme" values in the sample. These outcomes were consistent with prior literature and my hypotheses, but inconsistent with normative statistical theory, suggesting that confidence levels and confidence intervals are inversely related. The fact that the alternative measures of uncertainty do not behave according to statistical theory is important to researchers because individuals' subjective sense of uncertainty may have different dimensions that respond differently to changes in sample size. Consequently, individuals' risk and uncertainty judgments can be affected by how questions are asked.

Thus, despite strong enthusiasm for more frequent reporting, it appears that the consequences are not fully understood yet. Future research should include tasks that involve security valuation and risk estimates and should examine the relationship between more (less) frequent reporting of accounting information.

Additionally, because income numbers behave differently from balance sheet items when disaggregated, and because they are vital to securities valuation, research should examine how more (less) frequent reporting of income numbers affects investor judgments. If subsequent research confirms my results, one could conclude that increasing the frequency of reporting, which leads to larger amounts of information, could have both positive and negative effects on decision makers. Ironically, the end result could be *more* informed investors who fail to act or act improperly based on their newfound information.

Notes

1. In many actual financial prediction tasks, the sample mean would not be the optimal predictor, and a more sophisticated predictive model would be required. Because naïve investors may use a variety of suboptimal models to predict asset values or earnings, explaining their behavior would require disentangling the effects of the suboptimal valuation or earnings model from the pure sample size effects investigated here. This more complex task is left for future research.
2. The t-test is adjusted for unequal variances.

References

Byrnes, N. "The Downside of Disclosure. Too Much Data Can Be a Bad Thing. It's Quality of Information That Counts, Not Quantity." *BusinessWeek*, August 26, 2002.

Dimson, E., and A. Jackson. "Performance Evaluation: High Frequency Performance Monitoring." *Journal of Portfolio Management*, Fall, (2001), pp. 33–43.

Freudenthal, H., 1974. "The Empirical Law of Large Numbers or the Stability of Frequencies." *Educational Studies in Mathematics*, 4, (1974), pp. 484–490.

Hoge, R.D. "Confidence and Decision as an Index of Perceived Accuracy of Information Processing." *Psychonomic Science*, 18, (1970), pp. 351–353.

Kahneman, D., and A. Tversky. "Subjective Probability: A Judgment of Representativeness." *Cognitive Psychology,* 3, (1972), pp. 430–454.

Levitt, Arthur, Jr. *Take on the Street: What Wall Street and Corporate America Don't Want You to Know*. New York: Random House, 2002.

McNamee, M. "New Yardsticks for Investors." *BusinessWeek*, November 5, 2001.

Oskamp, S. "Overconfidence in Case Study Judgments." *Journal of Consulting Psychology*, 29, (1965), pp. 261–265.

Peterson, D., and G.F. Pitz. "The Effects of Amount of Information on Predictions of Uncertain Quantities." *Acta Psychologica,* 61, (1986), pp. 229–241.

Peterson, D., and G.F. Pitz. "Confidence, Uncertainty and the Use of Information." *Journal of Experimental Psychology: Learning Memory and Cognition,* 1, (1988), pp. 85–92.

Pitz, G.F. "Subjective Probability Distributions for Imperfectly Known Quantities." In L.W. Gregg, ed., *Knowledge and Cognition*. New York: Wiley, 1974, pp. 29–41.

Sedlmeier, P. "The Distribution Matters: Two Types of Sample-Size Tasks." *Journal of Behavioral Decision Making*, 11,)1998), pp. 281–301.

Sedlmeier, P., and G. Gigerenzer. "Intuitions about Sample Size: The Empirical Law of Large Numbers." *Journal of Behavioral Decision Making,* 10, (1997), pp. 33–51.

Shapira, Zur. *Risk Taking: A Managerial Perspective*. New York: Russell Sage Foundation, 1995.

Slovic, P. "Behavioral Problems of Adhering to a Decision Policy." Working paper, Oregon Research Institute, Eugene, Oregon, 1973.

Watson, L., L. McGuire, and E. Cohen. "Looking at Business Reports Through XBRL-Tinted Glasses." *Strategic Finance*, (2000), pp. 40–45.

Appendix A

Case Material for the Large-Sample (high-frequency) Condition

Assume the role of an investor in the DAS company. As an investor, it is important to monitor the firm's marketable securities account because it is a large portion of the asset base and will thus affect market valuation. DAS management has a target balance for this account, which has remained constant over the last year, but the actual account balance has gone up and down along with changes in securities values and firm cash requirements.

There have been no changes in the investment strategy in the prior year, and the firm does not foresee any changes in the future. The economy and other external forces are expected to be similar in the next period.

You have been provided with a history of the account balances for the past year, as well as weekly data. Your task is to estimate the next quarters' ending asset value to the best of your ability.

Your instructions are as follows:

Based on the following information, please provide your best prediction of the next quarters' account balance to the nearest (000).

Numbers are rounded to the nearest (000)

Data represent the prior fifty-two weeks

Week 1	23	**Week 14**	32	**Week 27**	13	**Week 40**	34
Week 2	25	**Week 15**	16	**Week 28**	20	**Week 41**	34
Week 3	24	**Week 16**	25	**Week 29**	30	**Week 42**	12
Week 4	12	**Week 17**	16	**Week 30**	22	**Week 43**	10
Week 5	7	**Week 18**	11	**Week 31**	31	**Week 44**	9
Week 6	27	**Week 19**	−3	**Week 32**	16	**Week 45**	8
Week 7	20	**Week 20**	43	**Week 33**	6	**Week 46**	34
Week 8	26	**Week 21**	25	**Week 34**	7	**Week 47**	26
Week 9	18	**Week 22**	21	**Week 35**	26	**Week 48**	32
Week 10	40	**Week 23**	33	**Week 36**	28	**Week 49**	20
Week 11	2	**Week 24**	15	**Week 37**	9	**Week 50**	22
Week 12	23	**Week 25**	10	**Week 38**	22	**Week 51**	15
Week 13	12	**Week 26**	15	**Week 39**	38	**Week 52**	17

1. Your prediction of the account balance at the end of the *next quarter* is: $_____ (Enter your prediction less the zeros; for example, if your estimate is 22,000 enter 22)
2. How confident are you that your prediction is correct? _____%. Rate on a scale from 0–100 (a confidence level of 0 = not confident at all; a confidence level of 100 = absolutely confident).

Alternatively, some subjects answered the following question regarding confidence interval in place of the confidence level question above.

2.

a. You are 95% certain that the balance in the account will be *more* than $_____ at the end of the next quarter.
b. You are 95% certain that the balance in the account will be *less* than $_____ at the end of the next quarter.

Assume the role of an investor in DAS company. As an investor, it is important to monitor the firm's marketable securities account because it makes up a large portion of the asset base and will thus affect market valuation. DAS management has a target balance for this account, which has remained constant over the last year, but the actual account balance has gone up and down with changes in securities values and in firm cash requirements.

There have been no changes in the investment strategy in the prior year and the firm does not foresee any changes in the future. The economy and other external forces are expected to be similar in the next period.

You have been provided with a history of the account balances for the past year, as well as with

quarterly data. Your task is to estimate the next quarters' ending asset value.

Your instructions are as follows:

Based on the following information, please provide your best prediction of the next quarters' account balance to the nearest (000).

Numbers are rounded to the nearest (000)

Data represents the prior fifty-two weeks

Quarter 1	12
Quarter 2	15
Quarter 3	38
Quarter 4	17

Response questions were identical to the large-sample condition.