Journal of Accounting and Economics 51 (2011) 204–218

Contents lists available at ScienceDirect

Journal of Accounting and Economics

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

Challenges and opportunities in disclosure research—A discussion of 'the financial reporting environment: Review of the recent literature' ☆

Philip G. Berger *

Booth School of Business, University of Chicago, Chicago, IL 60637, USA

ARTICLE INFO

Article history:
Received 14 October 2010
Received in revised form
3 January 2011
Accepted 3 January 2011
Available online 12 January 2011

JEL classification:
G14
M41

Keywords:
Financial reporting
Disclosure
Regulation

ABSTRACT

Beyer et al. (2010) review the financial reporting literature related to voluntary and mandatory firm disclosures, and sell-side analyst reports. The discussion summarizes their approach, highlights some of their main conclusions, and presents alternative ideas about promising avenues for future research.

1. Introduction

The review article by Beyer, Cohen, Lys, and Walther (BCLW, 2010), summarizes a large body of literature related to disclosure choices, disclosure regulation, and disclosure intermediation by sell-side analysts. The authors provide a useful framework for thinking about research questions in these areas, point out some of the more important contributions and limitations of papers written during the past decade, and briefly suggest some general ideas for promising areas of future research. This discussion summarizes the portions of BCLW's review related to voluntary and mandatory disclosure, discusses some related literature, and offers thoughts about promising areas for future work.

Section 2 addresses the empirical work in the BCLW review and discusses alternative empirical approaches being introduced to study related research issues. Section 3 covers voluntary disclosure models with empirical voluntary disclosure research discussed in Section 4. Section 5 reviews some of the research on models of mandatory disclosure and Section 6 explores issues in the empirical research on mandated disclosure. Concluding remarks are offered in Section 7.

☆ This discussion was prepared for the *Journal of Accounting and Economics* Conference held October 2–3, 2009. I thank Ross Watts for inviting me to be the discussant and for sharing with me his comments on an earlier version of BCLW's paper. I appreciate discussions with Ryan Ball about regressions of low frequency data on high frequency data, helpful input from Stan Markov and Shiva Shivakumar, and helpful comments from Anne Beyer, Daniel Cohen, John Core, Thomas Lys, Jonathan Rogers, Haresh Sapra, Beverly Walther, and Jerry Zimmerman (editor).

* Tel.: +17738348687; fax: +17737530817.
E-mail address: pberger@chicagobooth.edu

0165-4101/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jacceco.2011.01.001

2. Empirical analysis by BCLW

The BCLW survey begins with a brief empirical assessment of the relative contribution to the information reflected in security prices of five accounting disclosures: management forecasts, analyst forecasts, SEC filings, earnings guidance, and earnings announcements.[1] The assessment is performed with a decomposition of the quarterly stock return variance. The decomposition shows that about 28% of quarterly stock return variance occurs on days when accounting disclosures are made and that the majority (about 55%) of the accounting-based information is provided by management forecasts. The remaining 45% of the accounting-based information is provided about equally by analyst forecasts versus the combination of earnings pre-announcements, earnings announcements, and SEC filings.

I have two main concerns with this analysis. First, it provides evidence similar to that already in Ball and Shivakumar (2008), without fully acknowledging its debt to their paper. Second, Ball and Shivakumar note (p. 1010) that the higher informativeness of management forecasts over earnings announcements is likely due to managers self-selecting when they forecast. BCLW do not address or discuss the impact of the greater selectivity of management forecasts versus earnings announcements. If managers choose to issue forecasts when they think the forecast will be more informative, the less frequent management forecasts will be more informative than earnings announcements for selection reasons.

There are some minor measurement and econometric issues with the BCLW regression. These issues are of little concern as long as one realizes that the point of the return variance decomposition approach is to give a simple and descriptive illustration of well-known points, such as earnings explaining relatively little of a period's total return variance.

The minor econometric issues are as follows. First, there is likely to be measurement error in some of the short window CARs. For example, without a control for the expected earnings announcement date, the CARs for the earnings announcements are influenced by deviations from the expected report date. Measurement error in the announcement dates results in a biased estimate of the explained return variance. Second, in a variance decomposition, the covariance between items must be grouped into the portion of the variance explained by one of the items, which can be done only arbitrarily. Third, the attempt to decompose quarterly return variance is based on coding each information item as zero whenever the specific event did not occur during the quarter. Thus, none of the return variance being decomposed can be assigned to an information item for the firm-quarters in which the item is coded as zero.

The econometric issues are of little concern given the descriptive goal of the variance decomposition approach, and the approach has merit. It is not, however, the only possible approach for use in studying questions such as how various information events contribute to the total information conveyed over a period of time. I therefore discuss an alternate approach that can be used to address questions like those in Ball and Shivakumar (2008) and BCLW, but that can also allow potentially more interesting and complex questions to be examined.

Decomposing stock return variance from a long window (such as a quarter or a year) to assess the relative contribution of various three-day-window information items is an example of studying how a low frequency item of interest (quarterly return) is related to higher frequency variables (the information events such as management forecasts and earnings announcements). Ball and Easton (2010) develop a method for regressing low frequency data on high frequency data. Their focus is on the timeliness of the incorporation of the value changes captured by daily stock returns into earnings. The method they develop permits the possibility that the earnings–return relation changes over the earnings measurement period. In other words, instead of regressing annual earnings on annual returns and estimating a single regression coefficient to capture the average relation, Ball and Easton (2010) regress annual earnings on 251 daily returns and estimate a separate coefficient for each trading day of the fiscal year.

With regard to the findings of Ball and Shivakumar (2008) and the related Table 1 findings in BCLW, Ball and Easton (2010) use their approach to show that a much higher proportion of the stock price change on earnings announcement days is incorporated into the year's earnings compared to the proportion of stock price change incorporated into the year's earnings from non-earnings-announcement days. These findings suggest earnings announcements may be more important information events than the descriptive variance decomposition approach indicates.

With regard to expanding the set of researchable questions with their method, the main focus of Ball and Easton (2010) is on examining the portion of stock price change that is recognized in the earnings of the period, which they refer to as earnings recognition timeliness (ERT). By allowing the earnings–return relation to change over the earnings measurement period, Ball and Easton are able to separate two elements of ERT that have been aggregated together in previous empirical approaches. The two elements are the recognition in the current period's earnings of the effects of (1) news regarding sales and matched expenses for the current period and (2) news that changes expectations about future periods' earnings. Ball and Easton (2010) provide evidence that separating these two elements can provide additional insights related to the many studies that examine differences in ERT across data partitions, with the best known example being partitions based on positive return (good news) versus negative return (bad news) samples following Basu (1997).

A useful related technique is the Mixed Data Sampling (MIDAS) regression approach used in Ball (2010). MIDAS regression models were recently developed by Ghysels et al. (2006, 2007) and provide a parsimonious and flexible class of regression models that allow the dependent variable to be sampled at a low frequency (e.g., earnings announcements) while explanatory variables are sampled at a high frequency (e.g., daily returns). MIDAS regressions resolve severe

[1] I appreciate comments from John Core that helped considerably in improving the writing of this section.

multicollinearity by restricting the individual coefficient estimates on the high frequency explanatory variables to be a function of a small vector of parameters. This reduces the sampling variability of the coefficient estimates and counteracts increased variability from multicollinearity.

3. Voluntary disclosure models

BCLW's review of voluntary disclosure research raises several interesting observations. With regard to recommendations for future voluntary disclosure models, I agree that assessing the impact of incentive systems on management disclosure decisions may be misleading if it is assumed that the main goal of these systems is to influence disclosure choices. BCLW note that incentive systems are instead likely to be at least as concerned with influencing investment, effort, and competitive behavior choices (and, I would add, risk-taking). BCLW rightly argue that disclosure models would benefit from incorporating these real choices.

I think it is worthwhile to push further on this theme, which relates to the fact that most of the models of disclosure choice focus on a pure exchange economy (i.e., one with no real effects on production choices). It is not clear that further modeling of pure exchange economies is likely to provide significant additional insights for our understanding of discretionary disclosure decisions. Instead, the literature seems to have reached a point where incorporating real effects on production and investment choices needs to occur to provide substantial new insights into the causes and consequences of managers' disclosure choices.

As discussed by BCLW (see their footnote 30), until recently theoretical literature on the real effects of voluntary disclosure has focused on the effects on product market competitors. Recent working papers that instead consider the impact of voluntary disclosures on firms' investment decisions include Gox and Wagenhofer (2009) and Beyer and Guttman (2010). These papers are promising, but incorporating investment choices into voluntary disclosure models adds challenges because of the extra complexity required of the models.

Beyer and Guttman (2010) provides a nice illustration of the promise of incorporating investment decisions within a model of voluntary disclosure. The paper adds voluntary disclosure to a Myers and Majluf (1984) setting in which a firm with assets in place is considering issuing equity to raise capital for a new investment. The firm's current owners have private information about the value of assets in place. Myers and Majluf show that (in a setting without the ability to make disclosures about, or signal the value of, the assets in place) the firm will give up the new investment opportunity to avoid issuing equity if the value of the assets in place is sufficiently large.

Beyer and Guttman (2010) adds the possibilities that the firm can make credible disclosures about the assets in place or signal their value and then proceeds to analyze the equilibrium disclosure and investment policies that result. The authors find that the cost of credibly communicating the value of the assets in place has a nonlinear relation to the value of the assets, with the maximum cost applying for intermediate values and smaller costs applying for both small and large values of the assets in place. Thus, in contrast to what one might expect based on the Myers and Majluf (1984) paper, firms would forsake the investment opportunity and the associated opportunity to disclose when they have assets in place with intermediate values rather than high values.

These results are interesting, but they depend on the cost to the manager of the firm's strategy for reporting the assets in place. The assumption made is a common one in the literature—namely, that the personal cost to the manager of reporting a value of assets in place different from his privately informed valuation is represented by a quadratic-like function of the difference between the reported and true values. Such cost functions are often justified by appealing to concerns by the manager about his legal liability or personal reputation. Whether such justifications are compelling in the setting studied by Beyer and Guttman is worth considering carefully given the centrality of this assumption to the reported results. Nevertheless, despite any concerns about specific assumptions, papers like Beyer and Guttman (2010) that incorporate investment effects of disclosure choices are likely to raise important doubts about the conventional wisdom on the costs and benefits of discretionary disclosure decisions.

4. Empirical voluntary disclosure research

Two large areas within the empirical voluntary disclosure research domain are studies that explore the impact of proprietary information costs on reducing discretionary disclosure and papers that examine the benefits of voluntary disclosures for liquidity and the cost of capital. Despite considerable bodies of published work in these research areas, each continues to have limitations.

4.1. Proprietary information costs

As BCLW note, proprietary information costs of discretionary disclosure have generally been proxied by some measure of product market competition. Many studies using such measures conclude that product market competition affects disclosure. These conclusions are, however, subject to several limitations. These include conflicting evidence, measurement error in the product market competition proxies commonly used, a general failure to control for the potential benefits of discretionary disclosure (even though the benefits are likely correlated with the proprietary costs of disclosure), and identification problems related to disentangling the proprietary cost hypothesis from an alternative cost of disclosure

that can be labeled an entrenchment or agency cost. I discuss ways of addressing most of these issues, but hasten to add that no general solution currently exists for simultaneously addressing all of them.

Studies examining the association between discretionary disclosure and product market competition produce conflicting results using very similar measures of product market competition, but different measures of disclosure. Using industry concentration measures constructed with Compustat data, Bamber and Cheon (1998) find that management earnings forecasts are less likely in more concentrated industries. Conversely, Verrecchia and Weber (2006) find a negative relation between a Compustat-based concentration measure and the likelihood a firm asks the SEC to redact proprietary information from its filings. Thus, as BCLW recognize, Verrecchia and Weber find more informative disclosure in more concentrated industries whereas Bamber and Cheon find less.

The appropriate measure of product market competition is also open to question. Most papers use measures calculated from Compustat, which includes firms with publicly traded equity or debt, but generally excludes privately held companies. BCLW note that Ali et al. (2009) suggest that using Compustat data generates poor proxies for product market competition. It is worthwhile to elaborate in some detail on Ali et al.'s concerns. First, they show that using U.S. Census industry concentration measures (i.e., including private and public firms) often reverses prior results based on Compustat concentration metrics. Second, they show that the correlation of Compustat and Census measures of product market concentration is only 13%. Moreover, Compustat measures of concentration proxy for declining industries while Census measures do not. An important limitation of the Census measures, however, is that they are available only for firms with manufacturing operations located in the United States.

Given the differences between the Compustat and Census measures, it is not surprising that results documented in the literature using Compustat are often reversed if Census measures of concentration are used instead. Ali et al. demonstrate this for Harris's (1998) finding that segment disclosure is less likely in more concentrated industries, as measured using Compustat data. Ali et al. first replicate Harris's finding within the Census sample of manufacturing firms and then show that Harris's result no longer holds when using Census-based concentration measures. Numerous other studies of the relation between disclosure and product market competition draw inferences based on Compustat-based measures of concentration, including Bamber and Cheon (1998), Botosan and Harris (2000), Botosan and Stanford (2005), Rogers and Stocken (2005) and Verrecchia and Weber (2006). Inferences in these studies may be sensitive to the measure of product market competition that was used. While the measure that incorporates private firms need not necessarily be better (depending on the underlying notion of product market competition in the particular study and the importance of including non-manufacturing firms in the sample), the apparent fragility of inferences to the specific measure used is disconcerting.

A recent paper that attempts to address the concerns related to the measurement of product market competition and the mixed results on the relation between such competition and voluntary disclosure is Karuna (2010). He uses product substitutability, market size, and entry costs to capture multiple dimensions of competition derived from product market fundamentals. In addition, his voluntary disclosure measures are both forward-looking (and thus of greater proprietary concern because rivals have more time to react before the disclosed information will affect the product market) and measured at the industry segment level (allowing greater accuracy in identifying the industry in which competition is measured). Interestingly, his results show a nonlinear relation between product market competition and voluntary disclosure. More specifically, he finds support for both a negative competition-disclosure relation (the proprietary cost hypothesis motive) and a positive competition-disclosure relation (the strategic deterrence motive hypothesis).

At a more conceptual level, empirical studies of the role of proprietary information costs on discretionary disclosure decisions often ignore the potential benefits from disclosure and other potential costs of disclosure. Turning first to the benefits, models often assume away any correlation between proprietary costs of disclosure and the capital market benefits of voluntary information releases. Verrecchia (1983), for example, assumes that proprietary information costs are fixed and exogenous whereas the capital market benefits of disclosure are increasing in how good the news is that is being considered for release. Thus, only sufficiently good news gets disclosed. In contrast, empirical research that examines the impact of proprietary cost concerns on disclosure must examine samples in which proprietary costs vary. The dimensions on which empiricists typically claim proprietary costs are varying are the extent of product market competition and how positive the news being considered for release is.

Proprietary costs are usually hypothesized to be higher when product market competition is lower and when the news in question is better. These papers typically do not, however, control for variation across observations in the capital market benefits of disclosure. If the benefits of disclosure are larger when the disclosing firm faces less competition and has better news to disclose, the proxies in these papers that are interpreted as measuring proprietary information costs could also be capturing capital market benefits. Obviously, what the proxies really are capturing affects the interpretations drawn in these studies.

Are disclosure benefits likely to be larger when competition is reduced or when the news being disclosed is better? The answer is almost certainly yes. Consider a firm developing a new drug aimed at treating a widespread health problem. News about the drug having great efficacy or being approved for patent protection is likely much more beneficial to the disclosing firm than if the news were instead that the drug has modest efficacy or will be immediately subject to competition with similar products. Yet it is also likely true that disclosing news about development of the drug too early would have serious proprietary costs even though doing so would create capital market benefits.

A somewhat different concern is that measures used to capture proprietary costs may be capturing other costs of disclosure and that other costs of disclosure are not usually studied in conjunction with proprietary costs. Berger and Hann (2007)

raise this issue by arguing that much of the prior evidence interpreted as consistent with the proprietary cost hypothesis is also consistent with an agency (or entrenchment) cost hypothesis that posits disclosures are withheld as a result of conflicts of interest between managers and shareholders.

Bens et al. (2011) expand on this line of reasoning as follows. First, the plausibility of the proprietary cost motive is necessary for the agency/entrenchment cost motive to potentially exist. If agency costs were the only plausible motive for non-disclosure, an unraveling result would occur in which full disclosure would obtain. The research implications of this line of reasoning go even further—the agency and proprietary cost motives are reasonable only if outsiders cannot use *publicly available* disclosures to fully unravel these motives for aggregation. Taken together, these observations imply that the agency and proprietary cost motives should be studied simultaneously and that data from publicly disseminated financial statements have limitations for developing unambiguous empirical inferences about either of these motives for non-disclosure.

Using confidential U.S. Census data at the manufacturing plant level, Bens et al. investigate aggregation in external reporting. They find segment aggregation is more likely when the agency and proprietary costs of separately reporting the pseudo-segment are higher. Both motives affect the segment aggregation decisions of firms that report multiple external segments, but only the proprietary cost motive appears to affect the decision to group business units into a single externally reported segment.

Bens et al. is the first accounting study to exploit the Census Bureau's confidential data. As argued by Ali et al. (2009), these data have potential advantages for measuring product market competition. More generally, Bens et al. make use of several industry-level and industry-adjusted measures in addition to concentration ratios and all are potentially more accurately measured by including the private as well as public firms in the industry.

In addition, Bens et al. (2011) construct a new proprietary cost measure based on the proportion of industry sales made by private firms. The fraction of industry activity represented by private firms is a potentially more meaningful dimension of how product market competition affects proprietary information costs than simple measures of industry concentration. The reason is that, in the U.S., private firms are not required to make public their financial statements.

Finally, it is worth noting that while most of the variables constructed by Bens et al. can be measured only with access to the confidential Census data, the Census Bureau does make publicly available some industry-level variables in the manufacturing sector, such as the four-firm and Herfindahl concentration indexes.[2] While these measures are limited, recent studies such as Ali et al. (2009) and Tang (2009) demonstrate that the incorporation of private firms into industry-level measures can alter inferences about disclosure choices.

4.2. Disclosure and the cost of capital

Another large empirical literature examines the relation between various measures of disclosure and the cost of capital. BCLW's review of the theory literature in this area very reasonably leads to the advice that empirical researchers should admit that whether or not information risk is priced is an open issue. I think an additional warning is in order for empiricists seeking to apply insights from "direct effects" papers such as Lambert et al. (2007). My concern is that empiricists may seek theoretical justification for predictions relating information quality to the cost of capital without fully understanding the necessary conditions for the prediction to hold.

In addition to the theoretical challenge of why the cost of capital should be affected by disclosure choices, measuring the cost of capital presents empirical challenges. Easton and Monahan (2005) use variance decomposition to assess widely used implied cost of capital metrics. They show that the absolute magnitude of noise in each measure is large, and that significant improvement in the measures is not readily obtained either from using portfolio approaches to attempt to average out measurement error across firms or from using instrumental variables. Moreover, it is difficult to deal with the pervasive estimation problems for cost of capital because the alternatives to implied cost of capital measures have their own failings. Realized returns are likely biased and noisy measures of expected returns. Realized returns are (ex-post) biased if information surprises do not cancel out across sample observations (Elton, 1999; Fama and French, 2002). As investigated by Naranjo (2010), mispricing will also result in realized returns being biased measures of true expected returns. Target prices from analysts would, even if accurate, proxy for the market's expected return only for Hold recommendations. In addition, analysts' target prices are not accurate (Bradshaw and Brown, 2006).

Despite the challenges of measuring cost of capital and examining its association with voluntary or mandated disclosures, accounting researchers are likely to continue pursuing these questions and some consensus on best practices in such studies appears within reach. An example of a paper that I think tackles these problems well is Hail and Leuz (2006). A key lesson from this study is that best practices cannot be described by a simple template for all research situations. Rather, context is critical both in identifying the main dimensions of the measurement and research design problems and in deciding how to address them.

Hail and Leuz (2006) is a cross-country study. Thus, one concern is that analysts' long-term growth assumptions could be less accurate for some countries and that the accuracy of analysts' expectations might be affected by variation in the

[2] The most recent online data from the *Census of Manufacturing* are at http://www.census.gov/econ/census07, where links are also available to censuses from 2002, 1997 and 1992.

accounting systems in the different countries. Hail and Leuz address this concern by assessing the influence of long-term growth differences across countries on their results. In addition, they take four steps to address general concerns about their use of implied cost of capital measures. First, as in Gode and Mohanram (2003) and Botosan and Plumlee (2005), Hail and Leuz show that their implied cost of capital estimates are systematically related to traditional risk and country factors. Second, Hail and Leuz use robustness checks aimed at addressing specific deficiencies in analyst forecasting behavior, including lagging stock price by three months relative to the forecast measurement date as suggested by Guay et al. (2005). Third, Hail and Leuz (2006) compute accuracy-weighted country-year means of their cost of capital estimates to give more weight to observations with higher forecast accuracy. Finally, they perform sensitivity analyses using dividend yields and expected returns derived from country credit-risk ratings and country-index returns as alternatives to analyst-based cost of capital proxies (that also avoid some of the problems with realized returns and target prices).

Even if a researcher applies the best possible practices given the research setting to the issue of proxying for the cost of capital, she must then confront the issue of how best to measure disclosure or reporting quality. BCLW's main points in this regard are that the measurement issues are difficult and the commonly used proxies may not measure what they aim to capture.

4.3. *Measuring reporting "quality"*

Given the challenges in measuring reporting quality, BCLW note that Li (2008, 2010) has introduced natural language processing techniques to accounting. Thus, accounting research is beginning to heed Core's (2001) call to use such techniques to create meaningful disclosure quality measures for large samples.

Before offering some thoughts about recent and potential future work using natural language processing, I discuss some exciting recent research in other areas that I think offers hope for moving forward with the construction of better measures of reporting "quality." BCLW's remarks on this issue quite properly point out that a reporting quality measure developed in one context should not be used too hastily by researchers in other contexts. The example that BCLW focus on is the use by other researchers of the Dechow and Dichev (2002) measure of accruals quality, but BCLW's critique here is limited to the Wysocki (2009) point that the Dechow and Dichev measure is dominated by the negative contemporaneous correlation between cash flows and accruals.

A bigger issue with measures of accruals quality is identified by Gerakos and Kovrijnykh (2010). They argue that, given economic performance is stochastic, deterministic approaches to assessing accrual quality are inherently limited. Yet deterministic approaches are traditionally used in the accounting literature. Jones (1991), for example, develops a measure of accruals quality that assumes non-discretionary accruals are a deterministic and linear function of the changes in sales and the level of property, plant and equipment. Accruals unexplained by the model are assumed to represent discretionary (and, thus, lower quality) accruals. Similarly, Dechow and Dichev (2002) specify a deterministic intertemporal decomposition of cash flows and then measure accruals quality as the estimation error in a regression of changes in working capital on past, current, and future cash flows from operations.

As summarized in Dechow et al. (2010), the deterministic benchmarking measures popular in empirical accounting research have fundamental limitations. Deterministic benchmarking treats reporting bias as noise, thus entangling reporting bias with the stochastic variation that arises from volatility in the firm's operating environment. Of course, grouping volatility and bias into a common regression error term may lead to the excessive identification of manipulated or low quality earnings.

Gerakos and Kovrijnykh (2010) pursue an alternative approach to deterministic benchmarking. They use two basic assumptions to specify a parsimonious stochastic model of earnings that separately accounts for both economic shocks and reporting bias. The assumptions are that (1) true earnings are persistent and (2) the goal of reporting bias is to mask the true impact of economic shocks. Persistence for the purpose of their model means that a performance shock does not totally dissipate within a single period (i.e., is not completely transitory). The goal of reporting bias is argued to result in strategic manipulation that is generally in the opposite direction to the performance shock. With these assumptions, the resulting model implies that strategic manipulation leads to a negative second lag autocorrelation in the residuals from a regression of current on lagged earnings. This approach does have the limitation that it can capture only those types of systematic earnings manipulations that can be identified as a statistical regularity in a time series. Thus, the approach does not identify infrequent manipulation events such as large write downs of assets that effectively move future expenses into the current period (i.e., "big baths"). On balance, Gerakos and Kovrijnykh (2010) offers a promising alternative to deterministic benchmarking.

Another novel approach to assessing accounting quality is developed by Hribar et al. (2010), who argue that unexplained audit fees (rather than unexplained earnings or unexplained accruals) may offer advantages in measuring accounting quality. They find the portion of audit fees left unexplained by a model of audit fee determinants is positively correlated with other measures of accounting quality. Nevertheless, Hribar et al.'s audit-fee-based measure has incremental power beyond established quality metrics for predicting fraud, restatements, and SEC comment letters and is less associated with innate firm characteristics than the established quality measures are. While Hribar et al.'s (2010) measure appears novel and attempts to capture earnings quality in a manner less confounded by innate characteristics than extant measures are, its ultimate usefulness to accounting researchers remains an open question. One might, for example, argue that Hribar et al.'s measure captures impaired auditor independence or auditor effort rather than earnings quality (despite the authors' efforts to control for these auditor characteristics).

A final point worth considering with respect to measures of earnings quality is that commonly used metrics are unlikely to capture the same underlying construct. Rather, they are likely capturing different dimensions of quality. If so, researchers should be extremely cautious about using multiple quality metrics and claiming that a consistent directional effect on all of them for some disclosure or governance variable indicates a robust finding. A nice illustration of this concern is provided by a modeling paper by Ewert and Wagenhofer (2010). They provide a clear benchmark definition of earnings quality for the purposes of their study—the amount of reduction in the market's uncertainty about the firm's terminal value due to the earnings report. They then compare this measure with widely used metrics such as value relevance, persistence, predictability, smoothness, and accrual quality. Not surprisingly, they find that each measure captures very different effects and that several of the common metrics even vary non-monotonically with the benchmark earnings quality measure.

With regard to Core's (2001) call for more use of natural language processing techniques, I agree that computational linguistics offers a promising way to capture aspects of disclosure not readily measured by other means. On the other hand, using such approaches to study disclosure decisions is not without important challenges and is not the only (nor necessarily the best) way to obtain large sample evidence about broad aspects of firms' disclosure choices.

In terms of challenges, it is not obvious what aspects of a text such as a 10-K, analyst report, or management comment are best to extract in order to assess something of interest about the impact of disclosure choices at the relatively subtle level of the words that are used. Moreover, if the disclosure is mandated and some of the content (though not the exact wording) is mandated, the attempt to infer managerial motives from word choices may be even more difficult.

Li (2008) introduces natural language processing approaches to the accounting literature by investigating the relatively simple issues of how annual report readability is associated with firm performance and earnings persistence. He measures readability using the Fog index from the computational linguistics literature as well as the more basic metric of the length of the document. The main findings are that annual reports of firms with lower or more transitory earnings are harder to read. The paper provides important initial evidence that computational linguistics can be used to assess how basic aspects of disclosure such as readability may be associated with the manager's motives to obfuscate her performance.

Li's (2008) paper also points to some of the challenges and limitations of using natural language processing techniques to examine disclosure motives and consequences. As discussed by Bloomfield (2008), while Li's evidence is consistent with the obfuscation hypothesis (that managers make bad news costly by writing annual reports that are unnecessarily long and that use excessively big words and long sentences), it may also be consistent with competing hypotheses. Li's test is based on the notion that managers can succeed in altering the market's reaction to bad news by making such news more costly to analyze via obfuscation. This reasoning depends on the market under-reacting to the obscured information, yet Li finds mixed evidence that this is the case. Even to the extent that some of Li's evidence is supportive of under-reaction being associated with obfuscation, Bloomfield notes that there is still the common problem that correlation does not prove causality (e.g., longer reports and market under-reaction may have no causal relation to each other and could instead both be caused by certain types of poor performance).

Alternative explanations for Li's findings include some of the usual suspects from the discretionary disclosure literature, raising the issue of whether disclosure measures based on computational linguistics merely allow somewhat novel aspects of disclosure to be studied subject to all of the same problems faced by more traditional disclosure metrics. For example, the Fog index and report length are measures of total disclosure that do not separate discretionary and non-discretionary (i.e., mandated) components. Thus the results may merely reflect greater difficulty in explaining bad news or transitory income components or simply more mandated requirements with regard to the discussion of such news. More subtly, a discretionary managerial motive other than obfuscation may be at work. Possibilities raised by Bloomfield (2008) include attribution (good outcomes are attributed to skill, but bad outcomes to bad luck which entails additional length to explain) and litigation avoidance (managers may write longer and more complicated reports to shield themselves from litigation when news is bad).

What if the literature proceeds to the point where the main issue becomes distinguishing among various discretionary managerial motives related to language choices? The natural language processing literature in accounting may then need to more directly confront some of the issues in the underlying literatures in linguistics, psychology and natural language processing. Using the example above of Li's findings, how is attribution measured in the foundation literatures and how (if at all) is obfuscation measured? Can basic metrics like the Fog index and length be used to distinguish among relatively subtle competing theories from linguistics and psychology for managers' choices of words? Moreover, studies in the underlying literatures have shown that people put more weight on easy to process information in making judgments (e.g., Shah and Oppenheimer, 2007). Thus, a higher Fog index may make the information less important for readers' judgments, regardless of whether managers were motivated by obfuscation or attribution when increasing a disclosure's Fog score. Similarly, a number of recent papers examine various aspects of the "tone" of management or analyst reports (e.g., Rogers and Zechman, 2010; DeFranco et al., 2010). Are these papers defining and measuring tone in the same way? Is there a standard way to define and measure this construct that separates it from other constructs that may be similar or related?

Li (2008) has been followed by a burgeoning literature in accounting using computational linguistics to explore a variety of disclosure-related issues. Li (2010) examines the information content of forward-looking statements in the Management Discussion and Analysis (MD&A) section of 10-K and 10-Q filings using a Naïve Bayesian machine learning algorithm. He finds that more positive forward-looking statements are associated with both firm fundamentals and disclosure characteristics (such as lower accruals and a lower MD&A Fog index). Li (2010) also assesses various ways of

measuring the tone of the forward-looking statements and finds that the Naïve Bayesian approach is the only one for which positive tone predicts positive future performance. Positive tone does not predict positive future performance for the three commonly used dictionaries (Diction, General Inquirer, and the Linguistic Inquiry and Word Count), raising the question of how well these general dictionaries work for analyzing the specialized language in corporate filings.

Rogers and Zechman (2010) examine the relation between disclosure tone and shareholder litigation and find that optimistic tone is targeted by plaintiffs in lawsuits. Rogers et al. are careful to use both general-purpose and context-specific text dictionaries to quantify tone, thus demonstrating that their findings of a positive association between optimistic language and litigation risk are not sensitive to the particular text dictionary used to assess tone.

DeFranco et al. (2010) follow Li (2010) in using a Naïve Bayesian computational linguistics procedure. They apply it, however, to a much different research issue—an examination of the tone of debt analysts' discussions about events that may cause debt-equity conflicts of interest. They find that the tone of bond analysts' event discussions explains the disagreement between debt and equity analysts' investment recommendations for the debt and equity securities of the sample firms. Thus, the tone of these discussions appears to be of value to debt investors.

Larcker and Zakolyukina (2010) use data from subsequent financial restatements to classify the question and answer (Q&A) section of quarterly earnings conference calls as "truthful" or "deceptive." One strength of their paper is that the word categories used in developing their classification models are based on psychological and linguistic research that demonstrates these word groupings are related to deception. They go on to show in out-of-sample tests that their models predict subsequent financial restatements significantly better than a model using discretionary accruals plus traditional control variables.

Finally, Brown and Tucker (2010) develop an MD&A modification score based on an algorithm commonly used by Internet search engines to determine similarities between documents. This measure has the advantage of measuring changes in MD&A disclosures in an automated and objective fashion, thus allowing for large sample study of MD&A changes. Changes in, rather than levels of, MD&A disclosure has the desirable feature of potentially removing the large amount of MD&A disclosure that is repeated each year. The raw score of the difference between consecutive year's MD&A documents also has several disadvantages. One of these is that the raw difference tends to have a mechanical correlation with document length, but this correlation can effectively be purged. The more fundamental shortcoming of this measure is that it provides a summary measure of how different the wording is between documents, but cannot assess the exact nature of the wording differences (e.g., differentiating between changes in descriptions of the same item due to, say, changes in tone versus changes because old items have been dropped or new ones added with, say, a consistent tone).

4.4. Other novel approaches to measuring aspects of disclosure

Despite some concerns and challenges, I agree with Core (2001) that natural language processing approaches offer promise for extending the boundaries of empirical disclosure research. It is also, however, very much worth considering other novel approaches to measuring specific aspects of disclosure and even overall disclosure levels. Recent examples include Miller's (2002) comprehensive measure of firm-initiated, public written disclosures; Dyck and Zingales (2002), Miller (2006), Bushee et al. (2010), and Soltes (2010) who each examine the role of the business press as a mechanism for transmitting firm news; Rogers (2008); and Berger et al. (2009) who develop market-based measures that aim to indirectly capture disclosure quality; and Kalay (2010) who uses a market-based measure in a somewhat different manner—to measure the proportion of uninformed versus informed traders in a firm's securities (and to then investigate how the impact of different disclosure mechanisms varies with the proportions of informed versus uninformed traders).

Miller (2002) collects all available *public* disclosures from the Dow Jones News Retrieval Service for a sample of 80 small market capitalization firms (under $1 billion) that have eight contiguous quarters in which earnings before extraordinary items exceeds that of the same quarter a year earlier. Each management communication is treated as a single disclosure bundle that can include multiple information items such as an earnings announcement and a dividend increase announcement (perhaps within the same press release). For each disclosure bundle, Miller codes each information item included in the bundle.

Miller's disclosure measurement precedes the introduction of computational linguistics approaches to the accounting literature and thus does not focus on capturing "soft" dimensions of the disclosures such as their tone. Instead, similar to some earlier work, Miller attempts to go beyond machine-readable financial data such as those on Compustat by measuring disclosure more comprehensively using hand collected data in combination with a researcher-specific coding scheme. This approach has the advantage of capturing broader aspects of disclosure choice than those available from Compustat. Nevertheless, only public, written disclosures are captured, thus omitting things like verbal guidance from management to analysts through meetings or conference calls (which are likely less important for the small capitalization firms in Miller's sample). Moreover, these types of disclosure metrics have tended to be limited to small samples because of the intensive data collection required.

Miller's measure is comprehensive in capturing all firm-initiated, public written disclosures that are available on the Dow Jones News Retrieval Service, but the use of such a measure raises two related issues. When, why and to what extent does the business press screen or filter firm-initiated disclosures and when, why and to what extent does the press initiate news about firms on its own? Zingales (2000) speculates that readers rely on business press reporting to form opinions when they consider the information provided by the press to be reliable. Dyck and Zingales (2002) argue that the media

play a role in reducing contracting costs because the media aggregates and credibly communicates information to the public. Dyck and Zingales (2002) and Miller (2006) contend that the press also plays a corporate governance role and that, partially as a result, there is a demand from media consumers for investigative reporting. Miller (2006) shows that investigative reporting of accounting fraud has information content, but that press coverage of firms and frauds is skewed towards cases of broader interest to readers and situations that have lower costs to identify and investigate.

Bushee et al. (2010) investigate the business press's role as an information intermediary that packages and disseminates existing information on the one hand, while also creating new information through investigative journalism. They find that the business press's dissemination role around earnings announcements reduces information asymmetry around these short windows. Soltes (2010) focuses exclusively on the dissemination role of the press and comes to conclusions similar to those in Bushee et al., but over annual windows. The focus on annual windows by Soltes requires massive data collection of all firm-initiated announcements over each year in the sample, but has the advantage of facilitating a stronger attempt than is possible in Bushee et al. (2010) to identify the causal impact of the press on information asymmetry and to quantify the impact specifically attributable to the press's role as a transmission mechanism.

Rogers (2008) examines the relation between trading by insiders and their choice of disclosure "quality." He introduces a measure of disclosure quality using changes in market liquidity and applies this new disclosure quality metric to three types of management disclosure: forecasts, conference calls, and press releases. Rogers approach essentially reverses the usual structure of empirical disclosure studies, where some traditional measure of disclosure quality is often shown to be associated with market liquidity measures. Given the empirical evidence that higher quality disclosure improves market liquidity, Roger's approach assumes that changes in market liquidity provide a reasonable proxy for disclosure quality.

An obvious advantage of this approach is the ability to infer disclosure quality for large samples of firms using machine-readable data. The potential concern with this approach (and other approaches that attempt to infer disclosure quality or quantity from a market measure) is that changes in market liquidity may arise for reasons other than changes in disclosure quality. Rogers (2008) and other papers with related approaches thus have two potential avenues for providing assurance that the indirect measure used does indeed capture variation in disclosure quality. First, control variables can be used to attempt to hold constant factors other than disclosure quality that affect the disclosure quality metric. Second, validation tests can be conducted to examine whether the disclosure quality metric varies with disclosure choices in ways consistent with established theory or evidence. Rogers (2008) takes both steps and provides reasonable validation of his liquidity-based measure of disclosure quality.

Berger et al. (2009) present preliminary work on a somewhat related idea, which uses the ratio of firm-specific return variation to firm-specific cash-flow variation as a comprehensive measure of a firm's information quality. The reasonableness of measuring information quality as the ratio of the idiosyncratic volatility of returns to the idiosyncratic volatility of reported cash flows depends on one's view about the meaning of idiosyncratic stock returns. The finance literature has traditionally regarded firm-specific returns as "noise" rather than as information (Roll, 1988; West, 1988). A competing view has emerged in the past decade based on research that attempts to link firm-specific returns to economic efficiency. Morck et al. (2000) argue that less comovement (or synchronicity) in stock returns across firms is consistent with more firm-specific information being impounded in price.

Berger et al.'s contribution is to decompose firm-specific stock returns into firm-specific cash flow news and information quality. In other words, high firm-specific return variation can arise from either high firm-specific cash flow variation or high quality information. It is thus necessary to control for firm-specific cash flow variation (i.e., the underlying fundamentals of the firm) in order to use firm-specific returns to measure information quality. Thus, one challenge with Berger et al.'s (2009) approach is the difficulty in controlling for firm-specific cash flow variation. Similar to Rogers (2008), Berger et al. (2009) do not merely use control variables to attempt to hold constant effects other than information quality. They also perform validation tests. An additional issue that arises with their measure, however, is that it is a measure of the total information quality embedded in the firm's stock price. Thus, anything that makes price more informative should affect the Berger et al. (2009) measure, regardless of whether it is a disclosure or simply informed trading by those with private information. The breadth of the measure may have advantages for some research questions, but its very breadth makes validation of the measure more difficult.

Kalay (2010) examines variation in suboptimal option exercise activity as a proxy for variation in the sophistication of the firm's investors and finds that sophisticated investors concentrate their trading activity in firms with increased levels of information asymmetry and decreased share liquidity. In addition, he examines three different disclosure mechanisms to investigate how each impacts sophisticated versus unsophisticated investors. The results indicate that sophisticated investors concentrate their trading activity in firms that issue earnings guidance on a regular basis, whereas less sophisticated investors are more prevalent in firms with increased levels of newswire dissemination and superior investor relation activities.

The final areas of empirical voluntary disclosure research reviewed by BCLW are management forecasts and conference calls, with their review covering four issues: the decision to provide a management forecast, the choice of characteristics of the management forecast, the decision to bundle management forecasts with earnings announcements, and the use of conference calls. I offer two thoughts related to these research questions. First, the increased bundling of management forecasts with earnings announcements has multiple implications. BCLW note that Rogers and Van Buskirk (2008) explain the need to disentangle earnings news from forecast news and provide a method for doing so. I would add that, from a policy perspective, the frequent bundling of these two disclosures suggests that it may be difficult to effectively regulate

one of these activities independently of the other. Finally, conference calls receive relatively little attention in BCLW's review, but this is a research area that seems particularly well suited to their call for more natural language processing work.

5. Mandatory disclosure models

BCLW begin their discussion of disclosure regulation models by reviewing the traditional economic rationales for disclosure regulation: financial externalities, real externalities, agency costs, and economies of scale. These rationales result in regulation being difficult to justify, as is true with respect to regulation in a host of areas according to standard economic theory. While healthy skepticism of the potential benefits from regulation of disclosure (and other activities) is understandable, casual empiricism suggests that heavy regulation of disclosure (and other activities) is common in the most developed economies. The question thus arises as to whether something is missing from the existing models given that they often tend to indicate there is no net benefit from regulation.

BCLW then review models related to two decisions faced by accounting regulators—how much conservatism accounting rules should display and the extent to which fair values versus historical costs should be used. BCLW note that more conservative accounting standards can be characterized by a larger difference in informativeness between upward and downward revisions, but that this is a simplified view because it applies only to a single asset. With multiple assets, measures are aggregates and may therefore have different information properties. The potential for aggregation to create different information properties highlights the importance of deriving the information properties of *aggregate* reports based on conservative accounting that is applied at the *individual* asset level. Similarly, requiring higher verifiability before reporting gains need not reduce informativeness given that individual assets are aggregated in financial statement reports. Higher verifiability would reduce informativeness if all gains and losses were reported separately, but once gains and losses are aggregated it could actually be more informative overall to report (less informative) gains that are grouped with (more informative) losses.

I wholeheartedly agree with BCLW's view of the importance of increasing the emphasis in future modeling work toward the information properties of the aggregate measures that financial statements report rather than merely the information properties of different measurement systems on individual assets (and liabilities) prior to their aggregation in the financial statements. On the other hand, I also think that one can offer almost the opposite advice to empiricists. Empirical research on both mandated and voluntary disclosure overwhelmingly examines disclosures that are the result of an aggregation process (e.g., actual or forecasted net income or earnings per share, line items on financial statements, segment footnote information, and most issues discussed by management in settings like conference calls) without making any attempt to discover how discretion was used in going from the original, disaggregated data to get to the aggregated data reported publicly.

Observing data at a finer level than that reported publicly is, of course, difficult. Segment reporting research has been able to use retroactive restatement of segment reports to the *management approach* from the original reporting under the *industry approach* to infer something about the extent, causes and consequences of discretionary aggregation of segment data under the old segment reporting rule (Berger and Hann, 2003, 2007). More recently, confidential U.S. Census Bureau data have been used to examine how the private disaggregated data observable by firm managers get aggregated into reported segment footnotes in U.S. manufacturing industries (Bens et al., 2011). Gaining a better understanding of the role of discretion in aggregation for disclosures other than segment reports would represent a large step forward in understanding disclosure decisions, particularly those related to the published financial statements. While difficult, it should be possible either by exploiting mandated changes in financial reporting or by gaining access to private, internal, disaggregated data from sources such as the Census or private data vendors.

BCLW conclude their review of models of disclosure regulation with a discussion of models that examine how pure historical cost accounting versus pure fair value accounting compare to each other when the reported accounting numbers affect real decisions such as production choices. A result that emerges from such models is that pure fair value accounting has the potential to lead to suboptimal actions. Allen and Carletti (2008) and Plantin et al. (2008) both model situations where decreasing prices combine with pure fair value accounting to force asset fire sales which then drive down prices even further. By allowing accounting numbers to affect asset sale decisions, models such as these produce some thought-provoking results.

Two other theory papers that provide insights for the choice of mandatory reporting regime when accounting affects real decisions are Kanodia et al. (2005) and Gigler et al. (2009). Kanodia et al. examine the effect of imprecision in measuring the firm's investment level when there are two sources of information asymmetry. Insiders have private information about profitability (that affects the distribution of cash flows arising from investment) and the firm's investment level can only be measured and disclosed with noise. In this setting, an optimal level of imprecision exists (with respect to both the reporting of profitability and the reporting of investment) that sustains the first-best investment level. Given information asymmetry about both profitability and investment, removing all imprecision about one without addressing the measurement error of the other would alter the market's cash flow expectations in a way that would cause the firm to invest sub-optimally. Moreover, Kanodia et al. (2005) demonstrate that as the amount of asymmetric information about the data underlying insiders' actions increases so should the imprecision in measuring and reporting these actions.

Gigler et al. (2009) explore the impact of more frequent mandatory reporting when there are multiple market imperfections. Information asymmetry exists between managers and investors about project profitability. Whether the firm has chosen to invest in the short-term or long-term project also cannot be credibly disclosed. The long-term project has a bigger net present value, but the short-term project produces stochastically higher cash flows in the early periods and stochastically lower cash flows in the later periods. Gigler et al. show that more frequent disclosure in this setting results in more efficient market prices (i.e., prices better reflect the firm's underlying cash flows). Nevertheless, more frequent disclosure leads the manager to engage in short-term projects, which may not maximize shareholders' welfare.

As discussed by Sapra (2010), two broad insights emerge from papers such as Allen and Carletti (2008), Plantin et al. (2008), Kanodia et al. (2005) and Gigler et al. (2009). First, in a world of multiple sources of information asymmetry, regulation that restricts disclosures related to one source of asymmetric information may reduce rather than improve welfare. The reason is that restrictions on disclosures related to one activity may have unintended consequences with regard to the (possibly unregulated) disclosures related to other sources of asymmetric information. For example, in Kanodia et al. (2005) the two sources of information asymmetry are profitability and investment and removing all imprecision from disclosures about one without doing so for the other leads to less efficient firm investment decisions.

Second, the decision to measure and disclose an activity may affect the activity. In other words, the firm's actions and, hence, the firm's distribution of cash flows may be altered by the acts of measuring and disclosing them. For example, in Gigler et al. (2009) more frequent disclosure alters the cash flow distribution by leading the manager to select short-term projects.

The two broad insights developed in the models of how mandated disclosure interacts with investment decisions thus lead to the theme that mandatory disclosures which result in higher transparency need not improve social welfare and may instead reduce it. The notion that more transparency in disclosure can reduce welfare when disclosures affect production decisions is of obvious importance to standard setters and researchers. In parallel with the earlier discussion of discretionary disclosure decisions, the papers summarized in this section indicate that incorporating production effects of disclosure choices raises doubts about whether the costs are less than the benefits of disclosure rules that aim to achieve goals such as increased transparency.

6. Empirical mandatory disclosure research

Although models that allow for investment effects of accounting numbers hold considerable promise, taking the predictions from these stylized models to the data presents difficulties. As always, much of what is held constant in the models is not constant in the empirical domain and is often difficult to control for or to estimate accurately. Moreover, the endogenous relation between accounting rules and investment decisions that is at the heart of what makes these models interesting results in endogeneity issues that can be difficult to satisfactorily address empirically.

To take but one example, Bhat et al. (2010) examine how changes in commercial bank mortgage-backed-securities (MBS) holdings relate to liquidity-driven changes in MBS prices and mark-to-market accounting during the 2007 financial crisis. Bhat et al. define feedback as the tendency of banks to sell in the face of liquidity-driven price declines. They interpret their evidence as consistent with such a feedback effect from mark-to-market accounting rules. A major difficulty with structuring and interpreting these tests is, however, that liquidity in the MBS market is likely endogenously related to the behavior of commercial banks. If the liquidity changes in the MBS market during the period tested are not simply exogenous shocks but are endogenously determined by the commercial banks' actions, disentangling cause from consequence and identifying the role of any potential feedback effect becomes considerably more complicated.

Rather than focusing on empirical studies such as Bhat et al. that examine either the fair value accounting or conservatism issues emphasized in their review of mandatory disclosure models, BCLW focus their empirical review of mandated disclosure papers on those that examine a major change in disclosure regulation. The two changes reviewed in detail are the introduction of Reg FD and the Sarbanes-Oxley legislation (SOX). For the Reg FD studies, BCLW conclude that the overall findings of these papers are mixed with regard to the pre-Reg FD to post-Reg FD changes in: pre-earnings-announcement disclosure, the informativeness of analysts' earnings forecasts and stock recommendations, the extent of information asymmetry, and cost of capital effects. With regard to SOX, BCLW's main point is that the event study evidence on the wealth effect of SOX enactment produces conflicting inferences.

While it is useful in a broad review of research related to the financial reporting environment to devote attention to disclosure regulation and mandated disclosures, BCLW are not able to focus in depth on these research areas. A more focused review of these areas is provided by Leuz and Wysocki (2008). By concentrating on disclosure regulation only, that review is able to raise a more interesting and complete set of unanswered questions and is also able to provide more specific guidance on how one might move forward in addressing these open issues.

The approach in Leuz and Wysocki (2008) is to review the theory work and empirical evidence about firm-level and market-wide consequences of disclosure regulation using an approach that integrates work across the disciplines of accounting, economics, finance and law. The survey finds little conclusive evidence on market-wide impacts of disclosure, but considerably more developed evidence on firm-level consequences. The paper concludes by identifying seven unanswered questions about disclosure regulation and providing suggestions for how future research can productively proceed in these areas. The unanswered questions are in broad areas, such as obtaining a better understanding of why disclosure regulation is pervasive in advanced economies, doing more investigation of the real and macroeconomic

outcomes of regulation (capitalizing on recent regulatory changes such as mandatory IFRS adoption), and learning more about the interactions among and competition between various regulators around the world.

The ambitious questions posed at the end of Leuz and Wysocki's review of disclosure regulation are certainly appealing and the authors provide some examples of the areas of empirical research that could begin to address these important issues. I think it is also important, however, to consider some of the challenges researchers confront in trying to answer such questions and in trying to provide research findings that are of real and lasting value in arriving at policy decisions about disclosure regulation.

One simple concern is with regard to the robustness of many of the empirical findings from studies of disclosure regulation and voluntary disclosure choices. For example, my discussion of the difficulties with studies of the impact of proprietary information costs on discretionary disclosure illustrates that even fairly widely accepted results (like disclosure is reduced by proprietary costs) do not seem to hold up to close scrutiny.

Another issue is whether the data are simply too scarce with regard to the highest level policy decisions about disclosure regulation to be of much use in attempting to study how such decisions might be improved. In other words, while there are lots of mandated changes in individual accounting standards, there are very few sweeping changes in broad regulation along the lines of the U.S. securities acts of 1933 and 1934 or the mandated adoption of IFRS. Empirical research can and does attempt to very cleverly exploit these major regime changes and the variation observable within any given major regime change in terms of different affected groups, differences in timing of the change across groups, and so on. Nevertheless, it seems reasonable to wonder if even the best empirical research about these infrequent past major regime changes can really inform debate about how best to proceed when the next major reform opportunity arises.

Finally, a more pragmatic concern is whether it is realistic for disclosure regulations to be written in a manner that either anticipates evolving complexities in business practices and transactions or provides sufficient flexibility to incorporate these evolving complexities without the need for changes in the regulations. If not, disclosure regulations will likely tend to become increasingly reactive rather than proactive and research aimed at providing stable principles for improved disclosure regulation might prove futile.

7. Conclusion

Disclosure research has made significant progress in the past decade and a number of exciting recent developments provide optimism that even greater progress will be made in the decade ahead. My discussion of BCLW's review has focused on a few key issues and recent developments that I think merit particular attention.

First, assessing the relative contribution to the information reflected in stock prices of various accounting disclosures is just one example of an accounting research question in which a low frequency item of interest (in this case, quarterly stock returns) is related to higher frequency variables (in this case, the accounting disclosure events). Thinking about the most powerful ways to examine associations between low and high frequency data promises to allow much more flexible exploration of such issues as the relation between earnings and returns, the properties of analysts' forecasts, and the relation between disclosure and the cost of capital.

The bulk of my discussion dealt with voluntary disclosure. With regard to theory work in this area, an important issue is whether further modeling of pure exchange economies is likely to provide significant additional insights. If not, incorporating real effects on production choices into the models will become increasingly important. While this has already been done with some success in the area of mandated disclosure, it is admittedly more challenging when studying accounting choices made endogenously by managers.

Empirical work on voluntary disclosure has some major problems, but these problems also create opportunities and some very intriguing new approaches are reinvigorating work in this area. My discussion provides considerable detail about the difficulties with studies that aim to assess whether proprietary information costs reduce discretionary disclosure. The problems I focus on in the discussion are conflicting evidence, measurement error in commonly used proxies of product market competition (the overwhelmingly common metric used to capture proprietary information costs), failure to control for benefits of discretionary disclosure likely to be highly correlated with the proprietary costs, and difficulties disentangling the proprietary cost hypothesis from an alternative agency or entrenchment cost hypothesis. The last of these concerns provides an important reason why there are considerable difficulties in using publicly available data to study the question of how much information is being withheld from a firm's disclosures. Studies that make use of confidential, internal firm data (such as those available for manufacturing industries from the Census Bureau) thus are likely to become increasingly necessary to make progress in understanding the motives for non-disclosure.

An additional issue worth mentioning is that most empirical research about the impact of proprietary information costs on disclosure seems very much unaware of the variety of predictions on this issue offered in the theory literature (as reviewed very nicely by BCLW). Instead of the common empirical prediction that proprietary costs prevent full disclosure, alternative predictions emerge from various models as follows. Full disclosure can emerge as an equilibrium with proprietary costs (because non-disclosure leads to an inference that news is bad, which is costly). Proprietary costs can also result in disclosure of both sufficiently favorable and sufficiently unfavorable news, but not intermediate news (Sujis, 2005).

Another challenging field of voluntary disclosure research involves studies of the relation between disclosure and the cost of capital. One difficulty is providing a convincing theoretical justification for the prediction that cost of capital should

be linked to disclosure choices. Another concern is the empirical challenge of measuring the cost of capital with reasonable precision. A third significant difficulty is measuring disclosure quality. For those wishing to focus on earnings quality as their disclosure metric, I think two issues are of paramount concern. First, it is difficult to use multiple earnings quality measures and claim that all of them should have the same relation with an outcome variable of interest (such as cost of capital). Earnings quality measures such as value relevance, persistence, predictability, smoothness, and accrual quality do not all capture the same underlying economic construct. Second, deterministic approaches to assessing earnings quality or accrual quality are inherently limited. Thus, the approach taken by Gerakos and Kovrijnykh (2010) in specifying a parsimonious stochastic model of earnings that separately accounts for economic shocks versus reporting bias offers a promising new approach.

For broader measures of disclosure quality than just earnings quality, other approaches are needed. Natural language processing techniques represent one innovative way to capture both broad aspects of disclosure and aspects of disclosure not readily measured by other means. Li's (2008) introduction of these techniques into the accounting literature has been followed recently by a small boom in papers using a related set of approaches. As this novel literature reaches the point of trying to distinguish among various managerial motives related to language choices, underlying issues from the literatures in linguistics, psychology and natural language processing will likely need to be addressed. For example, how should tone be measured, and how can it be separated from related or similar constructs?

Computational linguistics need not be the only approach to capturing non-earnings aspects of disclosure quality. Examinations of the role of the business press as a means for disseminating firm news show promise as do approaches that develop market-based measures that aim to indirectly capture disclosure quality.

In the field of mandatory disclosure and disclosure regulation, models have tackled difficult and controversial issues, such as how much conservatism accounting rules should display and the extent to which fair values versus historical costs should be used. An important point made by BCLW in this regard is that information properties of a given accounting system that obtain at the individual asset level may very well not be the same information properties that the given accounting system will display at the level of aggregate accounting reports. Thus, theorists need to work toward modeling the information properties of the aggregate measures that financial statements report.

Perhaps ironically, I think that virtually the opposite advice is in order for empiricists. The archival empirical data we observe are almost always the output of a considerable amount of aggregation. To really understand the decisions that lead to this aggregated output, we need to learn more about how managers use their discretion in going from the disaggregated data they observe internally to arrive at the highly aggregated items that get reported externally.

My final comments with regard to disclosure regulation build on suggestions in a review by Leuz and Wysocki (2008). While I commend them for raising ambitious and thoughtful questions at the end of their very detailed, cross-disciplinary review, I mention three challenges researchers confront in trying to perform research aimed at informing disclosure regulation. First, as a research community, we have relatively few broadly robust empirical findings about the impact of disclosure regulation or about voluntary disclosure choices. Second, the most sweeping changes in broad disclosure regulation are both rare and quite specific to a particular set of economic and societal circumstances. The question that thus arises is whether studies of past changes such as SOX or mandated IFRS adoption can be generalized to the next major policy revision. Finally, I question what I think is an implicit assumption in some research aimed at informing disclosure policy regulation—namely, that the regulatory framework can be proactive rather than reactive even in the face of increasingly complex business practices.

In closing, I offer a quick thought related to the portion of BCLW's review dealing with information intermediaries. I did not cover this part of their review in my discussion because I wanted to focus in some depth on firms' (and thus managers') disclosure choices and the regulation of firm disclosures. The one minor thought I do offer is that the focus on sell-side equity analysts may leave some readers unaware of recent research that examines other information intermediaries. BCLW are aware of this and mention in their conclusion that research should continue to build on the recent work of Johnston et al. (2009) and DeFranco et al. (2010) by further considering the role of debt markets in the firm's information environment. Another area worthy of attention is research on the role of the business press, which is summarized briefly in my discussion. The past few years have also seen research related to the roles of credit rating agencies (Beaver et al., 2006; Cheng and Neamtiu, 2009), sell-side debt analysts (DeFranco et al., 2010; Gurun et al., 2009; Johnston et al., 2009), and buy-side analysts (Groysberg et al., 2008).

References

Allen, F., Carletti, E., 2008. Mark-to-market accounting and liquidity pricing. Journal of Accounting and Economics 45 (2–3), 358–378.

Ali, A., Klasa, S., Yeung, E., 2009. The limitations of industry concentration measures constructed with Compustat data: implications for finance research. Review of Financial Studies 22 (10), 3839–3871.

Ball, R., 2010. Does anticipated information impose a cost on risk-averse investors? Working Paper, University of Chicago.

Ball, R., Easton, P., 2010. Dissecting earnings timeliness. Working Paper, Notre Dame University and University of Chicago.

Ball, R., Shivakumar, L., 2008. How much new information is there in earnings? Journal of Accounting Research 46 (5), 975–1016.

Bamber, L., Cheon, Y., 1998. Discretionary management earnings forecast disclosures: antecedents and outcomes associated with forecast venue and forecast specificity choices. Journal of Accounting Research 36 (2), 167–190.

Basu, S., 1997. The conservatism principle and the asymmetric timeliness of earnings. Journal of Accounting and Economics 24 (1), 3–37.

Beaver, W., Shakespeare, C., Soliman, M., 2006. Differential properties in the ratings of certified versus non-certified bond-rating agencies. Journal of Accounting and Economics 42 (3), 303–334.

Bens, D., Berger, P., Monahan, S., 2011. Discretionary disclosure in financial reporting: an examination comparing internal firm data to externally reported segment data. The Accounting Review 86 (2), 1–33.
Berger, P., Chen, H., Li, F., 2009. Firm specific information and the cost of equity capital. Working Paper, University of British Columbia, University of Chicago, and University of Michigan.
Berger, P., Hann, R., 2003. The impact of SFAS No. 131 on information and monitoring. Journal of Accounting Research 41 (2), 163–223.
Berger, P., Hann, R., 2007. Segment profitability and the proprietary and agency costs of disclosure. The Accounting Review 82 (4), 869–906.
Beyer, A., Cohen, D., Lys, T., Walther, B., 2010. The financial reporting environment: review of the recent literature. Journal of Accounting and Economics 50 (2–3), 296–343.
Beyer, A., Guttman, I., 2010. Voluntary disclosure, disclosure bias and real effects. Working Paper, Stanford University.
Bhat, G., Frankel, R., Martin, X., 2010. Panacea, pandora's box, or placebo: feedback in bank holdings of mortgage-backed securities and fair value accounting. Working Paper, Washington University.
Bloomfield, R., 2008. Discussion of "Annual report readability, current earnings, and earnings persistence". Journal of Accounting and Economics 45 (2–3), 248–252.
Botosan, C., Harris, M., 2000. Motivations for a change in disclosure frequency and its consequences: an examination of voluntary quarterly segment disclosures. Journal of Accounting Research 38 (2), 329–353.
Botosan, C., Plumlee, M., 2005. Assessing alternative proxies for the expected risk premium. The Accounting Review 80 (1), 25–53.
Botosan, C., Stanford, M., 2005. Managers' motives to withhold segment disclosures and the effect of SFAS No. 131 on analysts' information environment. The Accounting Review 80 (3), 751–771.
Bradshaw, M., Brown, L., 2006. Do sell-side analysts exhibit differential target price forecasting ability? Working Paper, Boston College and Georgia State University.
Brown, S., Tucker, J., 2010. Large-sample evidence of firms' year-over-year MD&A modifications. Working Paper, University of Florida.
Bushee, B., Core, J., Guay, W., Hamm, S., 2010. The role of the business press as an information intermediary. Journal of Accounting Research 48 (1), 1–19.
Cheng, M., Neamtiu, M., 2009. An empirical analysis of changes in credit rating properties: timeliness, accuracy and volatility. Journal of Accounting and Economics 47 (1–2), 108–130.
Core, J., 2001. A review of the empirical disclosure literature: discussion. Journal of Accounting and Economics 31 (1–3), 441–456.
Dechow, P., Dichev, I., 2002. The quality of accruals and earnings: the role of accrual estimation errors. The Accounting Review 77 (1), 35–59.
Dechow, P., Ge, W., Schrand, C., 2010. Understanding earnings quality: a review of the proxies, their determinants and their consequences. Journal of Accounting and Economics 50 (2–3), 344–401.
DeFranco, G., Vasvari, F., Vyas, D., Wittenberg-Moerman, R., 2010. Debt analysts' views of debt-equity conflicts of interest. Working Paper, London Business School, University of Chicago, University of Minnesota, and University of Toronto.
Dyck, A., Zingales, L., 2002. The corporate governance role of the media. In: The Right to Tell—The Role of Mass Media in Economic Development. The World Bank Institute, Washington, DC, pp. 107–137.
Easton, P., Monahan, S., 2005. An evaluation of accounting-based measures of expected returns. The Accounting Review 80 (2), 501–538.
Elton, E., 1999. Presidential address: expected return, realized return, and asset pricing tests. The Journal of Finance 54 (4), 1199–1220.
Ewert, R., Wagenhofer, A., 2010. Earnings quality metrics and what they measure. Working Paper, University of Graz.
Fama, E., French, K., 2002. The equity premium. The Journal of Finance 57 (2), 637–659.
Gerakos, J., Kovrijnykh, A., 2010. Reporting bias and economic shocks. Working Paper, University of Chicago.
Gigler, F., Kanodia, C., Sapra, H., Venugopalan, R., 2009. An equilibrium analysis of the costs and benefits of more frequent financial reporting. Working Paper, University of Chicago, University of Illinois, Urbana-Champaign, and University of Minnesota.
Gode, D., Mohanram, P., 2003. Inferring the cost of capital using the Ohlson-Juettner model. Review of Accounting Studies 8 (4), 399–431.
Gox, R., Wagenhofer, A., 2009. Optimal impairment rules. Journal of Accounting and Economics 48 (1), 2–16.
Groysberg, B., Healy, P., Chapman, C., 2008. Do buy-side analysts out-perform the sell-side? Financial Analyst's Journal 64, 25–39.
Guay, W., Kothari, S., Shu, S., 2005. Properties of implied cost of capital using analysts' forecasts. Working Paper, Boston College, MIT, and University of Pennsylvania.
Gurun, U., Johnston, R., Markov, S., 2009. Sell-side debt analysts and market efficiency. Working Paper, University of Texas, Dallas and Ohio State University.
Ghysels, E., Santa-Clara, P., Valkanov, R., 2006. Predicting volatility: getting the most out of return data sampled at different frequencies. Journal of Econometrics 131 (1–2), 59–95.
Ghysels, E., Sinko, A., Valkanov, R., 2007. MIDAS regressions: further results and new directions. Econometric Reviews 26 (1), 53–90.
Hail, L., Leuz, C., 2006. International differences in cost of capital: do legal institutions and securities regulation matter? Journal of Accounting Research 44 (3), 485–531.
Harris, M., 1998. The association between competition and managers' business segment reporting decisions. Journal of Accounting Research 36 (1), 111–128.
Hribar, P., Kravel, T., Wilson, R., 2010. A new measure of accounting quality. Working Paper, University of Iowa and University of Texas, Dallas.
Johnston, R., Markov, S., Ramnath, S., 2009. Sell-side debt analysts. Journal of Accounting and Economics 47 (1–2), 91–107.
Jones, J., 1991. Earnings management during import relief investigations. Journal of Accounting Research 29 (2), 193–228.
Kalay, A., 2010. Investor sophistication, disclosure and the information environment of the firm. Working Paper, University of Chicago.
Kanodia, C., Singh, R., Spero, A., 2005. Imprecision in accounting measurement: can it be value enhancing? Journal of Accounting Research 43 (3), 487–519.
Karuna, C., 2010. Industry product market competition and corporate voluntary disclosure: evidence from forward-looking line items at the industrial segment level. Working Paper, University of Houston.
Lambert, R., Leuz, C., Verrecchia, R., 2007. Accounting information, disclosure and the cost of capital. Journal of Accounting Research 45 (2), 385–420.
Larcker, D., Zakolyukina, A., 2010. Detecting deceptive discussions in conference calls. Working Paper, Stanford University.
Leuz, C., Wysocki, P., 2008. Economic consequences of financial reporting and disclosure regulation: a review and suggestions for future research. Working Paper, MIT and University of Chicago.
Li, F., 2008. Annual report readability, current earnings, and earnings persistence. Journal of Accounting and Economics 45 (2–3), 221–247.
Li, F., 2010. The information content of forward-looking statements in corporate filings—a naïve Bayesian machine learning approach. Journal of Accounting Research 48 (5), 1049–1102.
Miller, G., 2002. Earnings performance and discretionary disclosure. Journal of Accounting Research 40 (1), 173–204.
Miller, G., 2006. The press as a watchdog for accounting fraud. Journal of Accounting Research 44 (5), 1001–1033.
Morck, R., Yeung, B., Yu, W., 2000. The information content of stock markets: why do emerging markets have synchronous stock price movement? Journal of Financial Economics 58 (1–2), 215–260.
Myers, S., Majluf, N., 1984. Corporate financing and investment decisions when firms have information that investors do not have. Journal of Financial Economics 13 (2), 187–221.
Naranjo, P., 2010. The appropriateness of realized returns and implied cost of capital as proxies for expected returns. Working Paper, MIT.
Plantin, G., Sapra, H., Shin, H., 2008. Marking-to-market: Panacea or pandora's box? Journal of Accounting Research 46 (2), 435–460.
Rogers, J., 2008. Disclosure quality and management trading incentives. Journal of Accounting Research 46 (5), 1265–1296.
Rogers, J., Stocken, P., 2005. Credibility of management forecasts. The Accounting Review 80 (4), 1233–1260.
Rogers, J., Van Buskirk, A., 2008. Bundled forecasts and selective disclosure of good news. Working Paper, Ohio State University and University of Chicago.

Rogers, J., Zechman, S., 2010. Disclosure tone and shareholder litigation. Working Paper, Ohio State University and University of Chicago.
Roll, R., 1988. R-squared. The Journal of Finance 43 (3), 541–566.
Sapra, H., 2010. The economic trade-offs in the fair value debate. Journal of Law, Economics and Policy 6 (2), 193–218.
Shah, A., Oppenheimer, D., 2007. Easy does it: the role of fluency in cue weighting. Judgment and Decision Making Journal 2 (6), 371–379.
Soltes, E., 2010. Disseminating firm disclosures. Working Paper, Harvard University.
Sujis, J., 2005. Voluntary disclosure of bad news. Journal of Business Finance and Accounting 32, 1423–1435.
Tang, V., 2009. Proprietary costs and privately held firms' financing choice between public offerings and private placements. Working Paper, Georgetown University.
Verrecchia, R., 1983. Discretionary disclosure. Journal of Accounting and Economics 5 (3), 179–194.
Verrecchia, R., Weber, J., 2006. Redacted disclosure. Journal of Accounting Research 44 (4), 791–814.
West, K., 1988. Dividend innovations and stock price volatility. Econometrica 56 (1), 37–61.
Wysocki, P., 2009. Assessing earnings and accruals quality: U.S. and international evidence. Working Paper, University of Miami.
Zingales, L., 2000. In search of new foundations. The Journal of Finance 55 (4), 1623–1653.