Review of Accounting Studies
https://doi.org/10.1007/s11142-020-09544-x

Do going concern opinions provide incremental information to predict corporate defaults?

Elizabeth Gutierrez[1] · Jake Krupa[2] · Miguel Minutti-Meza[2] · Maria Vulcheva[3]

Published online: 24 June 2020

Abstract

Investors, regulators, and academics question the usefulness of going concern opinions (GCOs). We assess whether GCOs provide incremental information, relative to other predictors of corporate default. Our measure of incremental information is the additional predictive power that GCOs give to a default model. Using data from 1996 to 2015, initially we find no difference in predictive power between GCOs alone and a default model that includes financial ratios. However, there is an imperfect overlap between GCOs and other predictors. We show that GCOs increase the predictive power of several models that include ratios, market variables, probability of default estimates, and credit ratings. Using a model that includes ratios and market variables, GCOs increase the number of predicted defaults by 4.4%, without increasing Type II errors. Our findings suggest that GCOs summarize a complex set of conditions not captured by other predictors of default.

Keywords Going concern opinions · Predictive accuracy · Type I and II errors · Audit quality

1 Introduction

In this study, we assess whether going concern opinions (GCOs) provide incremental information, relative to other predictors of corporate default. Our empirical measure of incremental information is the additional predictive power that GCOs give to a statistical model of default. The standard auditor's opinion gives a pass/fail statement regarding a company's compliance with Generally Accepted Accounting Principles

✉ Miguel Minutti-Meza
mminutti@bus.miami.edu

1 School of Economy and Business, Universidad de Chile, Santiago, Chile

2 University of Miami, Coral Gables, FL, USA

3 Florida International University, Miami, FL, USA

(GAAP), but it does not provide much company-specific information. The GCO is a notable exception, in that it warns investors when a company is in financial distress.[1]

Diverse stakeholders, including debt and equity investors, auditing regulators, accounting standard setters, and academics view auditors' GCOs and default prediction as important issues. For instance, GCOs receive considerable attention from the business press. A keyword search of terms related to GCOs in ProQuest retrieved 1526 press articles published in the period from 2000 to 2018, including 177 articles during the peak of the financial crisis (2008 and 2009).[2] The public scrutiny of the auditor's responsibility to issue a GCO increases after failures of high-profile businesses and during periods of economic downturn. In the aftermath of the dot-com bubble and during deliberations for the Sarbanes-Oxley Act of 2002, Weiss Ratings Inc. submitted to the United States Senate a report showing that auditors only issued GCOs for 42% of the 228 publicly traded companies that filed for bankruptcy protection in the 18 months starting in January 2001 (Akers et al. 2003). Similarly, following the 2008 financial crisis, a report by Duff & Phelps Corp. noted that auditors only issued GCOs for 63% of the public companies with annual revenues above $25 million that filed for bankruptcy in 2008 (Chasan 2009, 2012).

Auditing regulators and standard setters in the United States and other jurisdictions have strived to guide auditors and management in making going concern assessments. The Public Company Accounting Oversight Board (PCAOB) has issued Staff Audit Practice Alerts No. 3, 9, and 13 (PCAOB 2008; 2011; 2014), and a sub advisory group document on going-concern considerations and recommendations (PCAOB 2012). The Financial Accounting Standards Board (FASB) issued the Accounting Standards Update (ASU) 2014–15, requiring managers to assess going-concern uncertainty (FASB 2014).[3] Recently, after several corporate failures in the United Kingdom, the Financial Reporting Council revised the audit standard on GCOs (FRC 2019). Under the revised standard, auditors will be required to more actively challenge management's going concern assessments, test the supporting evidence, and evaluate the risk of management bias. From the perspective of a wide set of market participants, GCOs therefore play a central role as an independent and objective signal of financial distress.

Using event studies and default prediction models, the literature has documented mixed findings regarding the incremental information content of GCOs. A stream of research finds that the market reacts negatively to the announcement of GCOs (e.g., Menon and Williams 2010). However, a recent study by Myers et al. (2018) shows that the market reaction to GCOs is largely attributable to contemporaneous information released by companies around GCOs and specifically to concurrent earnings

[1] Auditing regulators worldwide have recently tried to increase the information content of the auditor's report by requiring the discussion of key or critical audit matters (Gutierrez et al. 2018). However, GCOs focus on the client's financial condition.

[2] The machine-based search criteria required the phrase "going concern*" in combination with the stem term "audit*" (e.g., audited, auditor, etc.) in the title or body of the article. For example, the press followed the General Motors case and Deloitte's decision to issue a GCO for the company's 2008 fiscal year-end (Johnson 2009).

[3] An entity must provide certain disclosures if "conditions or events raise substantial doubt about (the) entity's ability to continue as a going concern." The ASU applies to all entities and is effective for annual periods ending after December 15, 2016 (FASB 2014). In the Basis for Conclusions, ASU 2014–15 notes that in 2010 the FASB going-concern project considered proposing early-warning disclosures about going-concern uncertainties.

announcements. After isolating the effect of earnings announcements, GCOs seem to convey little incremental information to investors on their release date.

A different approach, which examines the information content of GCOs, focuses on whether they predict a company's default (e.g., the company cannot continue as a going concern, due to bankruptcy or market delisting followed by bankruptcy). Seminal research in this area compares the predictive power of GCOs to that of statistical models of probability of default (PD models). Early studies claim that GCOs are inferior to PD models that use financial information as inputs (Altman and McGough 1974; Altman 1982; Levitan and Knoblett 1985; Koh 1991). However, a landmark study by Hopwood et al. (1994) examines bankruptcy data for distressed firms from 1975 to 1985 and demonstrates that GCOs are just as good as PD models that use financial ratios and client size as inputs.[4]

The findings of Hopwood et al. (1994) are a starting point for our analysis. However, we go beyond their study in three ways. First, the past 20 years provide new data on GCOs and allow us to examine whether recent economic and regulatory conditions have changed GCOs. Second, the literature on default prediction has proposed increasingly powerful models and new techniques to assess a model's predictive power (e.g., Shumway 2001; Pepe et al. 2009; Duan et al. 2012). Third, we examine an issue overlooked by the auditing literature: whether GCOs increase the predictive power of a default model in case there is an imperfect overlap between the auditor's opinion and other predictors of default.

There are two reasons to believe that GCOs will provide incremental information, relative to other predictors of corporate default. First, the auditor has access to private information about the client. Second, the auditor may consider qualitative criteria beyond financial ratios or market variables that may imperfectly capture forward-looking information. However, inherent constraints to the auditor's approach in issuing GCOs may advantage a PD model in predicting default. The auditor's responsibility is limited to assessing relevant conditions and events prior to the date of the report and does not include forecasting (AS 2415 PCAOB 2017, previously Statement on Auditing Standards AU 341 and SAS No. 59). Typically, auditors rely on a checklist approach and individual judgment. The auditor may also be subject to time pressure and incentive issues that can affect going concern judgments.

We examine a large sample of distressed companies from 1996 through 2015.[5] Our analysis shows that bankruptcy and other forms of default are rare. (In the remainder of

[4] The Hopwood et al. (1994) study examined 134 bankrupt firms and 160 nonbankrupt firms. The authors used a PD model based on six key financial ratios and firm sizes. The model was estimated using data from 1974 to 1981 and fitted to out-of-sample data from 1982 to1985. The two main takeaways are that inferences in previous studies were attributable to using (a) nonrepresentative samples, with 50% bankrupt and 50% nonbankrupt companies, and (b) mixed samples with both distressed and nondistressed companies. Hopwood et al.'s (1994) study is a widely used methodological reference in GCO research. However, the economic environment and the regulatory landscape for auditors have changed since its publication (e.g., the passage of the Private Securities Litigation Reform Act of 1995 and the Sarbanes-Oxley Act of 2002, along with the creation of the Public Company Accounting Oversight Board in 2002).

[5] We obtain data on corporate defaults from multiple sources, including the Center for Research in Security Prices (CRSP), Compustat, Bankruptcy.com, UCLA-LoPucki Bankruptcy Research, and the Credit Research Initiative (CRI). The Risk Management Institute at the National University of Singapore is a nonprofit organization that owns the proprietary CRI database. This database includes macroeconomic, financial, and default-related information. We complement these data with financial statement variables, stock prices, and auditor information from various sources, including a text-based search to identify GCOs before auditor opinions were available in Audit Analytics.

the paper, we refer to these events collectively as default.) Only 1.8% of distressed companies default in the next year. Regarding the relation between GCOs and default, approximately 50% of companies that default in a given year have received a prior GCO, whereas 90% of companies with a GCO do not default in the next year.

When we compare default prediction based on a PD model that uses financial ratios to default prediction of GCOs, we find that in some years the model outperforms GCOs in the true positive rate (i.e., correctly predicted defaults using a classification threshold). However, in most years, GCOs outperform the model in the true negative rate (i.e., correctly predicted nondefaults). Using all years in our sample, GCOs and PD models have similar accuracy in predicting defaults. Nevertheless, there are noticeable differences in Type I (false positive) and Type II (false negative) error rates. This analysis extends the Hopwood et al. (1994) results to recent years and confirms that GCOs *independently* convey information that results in predictive power that is similar to a basic PD model. However, we also demonstrate that GCOs and other predictors of distress do not convey exactly the same information. Hence we examine an issue overlooked elsewhere: whether the imperfect overlap between GCOs and PD models can be exploited to generate a superior joint model. This imperfect overlap may be attributable to auditors' access to private information and exercise of professional judgment or to other factors that are difficult to observe.

In our main analyses, we assess the incremental predictive power that GCOs can contribute to several models of corporate default. The first model includes financial ratios and client size. The second adds stock market information. The third uses client information and company-level PD estimates from an external source, the CRI database of the Risk Management Institute at the National University of Singapore. The CRI company-level PD estimates are based on a stochastic model that combines levels and changes in financial ratios, market variables, and macroeconomic indicators (Duan et al. 2012). Finally, the fourth model uses client information and credit ratings. Our use of multiple benchmarks mitigates concerns that our findings rely entirely on a given statistical model. Moreover, the model with credit ratings also relies on rating agencies' methodologies, which include not only statistical analyses but also credit analysts' judgments (i.e., another third-party expert). Our two criteria to assess the incremental predictive power of GCOs are the increase in (i) the in-sample area under the receiver operating characteristic curve (AUROC), and (ii) the out-of-sample predictions computed using cross validation.[6]

Using a comprehensive model that includes financial ratios, client size, and market variables, the AUROCs are 0.898 and 0.883 with and without a GCO variable, respectively. The incremental AUROC of 0.015 is statistically significant ($p < 0.01$). The additional AUROC translates into an out-of-sample estimate of a 4.4% increase in the number of defaults correctly identified (23 firm-year observations) and a 0.09% reduction in the number of defaults incorrectly identified (24 firm-year observations). Hence including GCOs in the default model decreases the model's Type I errors without increasing its Type II errors. We reach similar inferences using all other

[6] The AUROC criterion has two useful aspects. (i) It summarizes a model's true positive rates (i.e., correctly predicted defaults) and false positive rates (i.e., incorrectly predicted defaults) in a simple indicator that ranges between 0.5 and 1.0. (ii) Two models' AUROCs can be compared using well-defined statistical tests, in sample and out of sample.

benchmark models in our analyses. Our combined findings indicate that GCOs provide incremental information, relative to other predictors of corporate default.

In additional analyses, we investigate whether the incremental information in GCOs varies across 12 partitions that capture differences in (a) client characteristics (size, information environment, and levels of financial distress), (b) GCO properties (first-time GCO and distress conditions mentioned by the auditor in the opinion's text), (c) auditor and engagement characteristics (Big N, first-year engagement, probability of opinion shopping, competition in the local audit market, growth in the local audit market, and client importance), and (d) economic cycles (2006–2008, 2009–2011, and 2012–2014).

We document that GCOs provide incremental information for relatively small companies, companies with poor information environments, and companies with both high and low levels of financial distress. We find that GCOs issued for the first time and GCOs citing both common and uncommon distress conditions provide incremental information. Next, we show that GCOs provide incremental information when they are issued for recurring engagements (nonfirst year), when local audit market competition, market growth, and client importance are low. The results based on this set of partitions suggest that time pressure and auditors' incentives affect going concern judgments. However, we find that GCOs provide incremental information, regardless of whether the auditor is Big N or non-Big N and whether there is a high or low likelihood of opinion shopping. Finally, we show that GCOs provide incremental information in the period leading up to the financial crisis (2006 to 2008).

Our findings suggest that GCOs are a summary signal of a complex set of conditions not captured entirely by other publicly known predictors of default. We do not view GCOs either as entirely new information or as a simple aggregation of public information. Our study's results are consistent with both possibilities. Based on our analyses, a combination of auditor judgment and statistical modeling may enhance the accuracy of going concern assessments, an interesting possibility that may be the subject of future research. Lastly, our evidence contributes to the bankruptcy prediction literature, demonstrating that GCOs can increase the predictive power of PD models.

2 Background

2.1 Standards on GCOs

Under U.S. audit standards, the auditor must evaluate the going concern assumption (AS 2415 "Consideration of an Entity's Ability to Continue as a Going Concern" PCAOB 2017). The auditor has a responsibility to evaluate whether there is substantial doubt about the entity's ability to continue as a going concern for a reasonable period of time, not to exceed one year beyond the date of the financial statements being audited (referred to as a reasonable period of time) (AS 2415.02 PCAOB 2017). However, there is no precise definition of "substantial doubt." The GCO standard includes four broad categories of events that may warrant such substantial doubt about a client's continuation as a going concern: negative trends, other indications of possible financial difficulties, internal matters, and external matters (AS 2415.06

PCAOB 2017). Moreover, the standard does not mention how the auditor should collectively interpret and assess these events, and it is silent about the use of statistical models. The auditor must therefore rely on professional judgment to conclude that a client's probability of continuing is sufficiently low to issue a GCO.

2.2 Type I and type II error rates in GCOs

A large body of research has studied the determinants of GCOs, documenting that client, auditor, and macroeconomic characteristics affect the incidence of GCOs. (See Carson et al. 2013 for a thorough literature review.) Some historical developments have impacted the rate of GCOs, such as the issuance of SAS No. 59 ("Auditor's consideration of an entity's ability to continue as a going concern"), the Private Securities Litigation Reform Act in 1995, and the Sarbanes-Oxley Act in 2002 (SOX). In recent times, Carson et al. (2013) shows that the incidence of GCOs has steadily increased from 9.8% in 2000 to 17% in 2010.

A number of studies focus on whether an auditor is accurate in predicting a client's default (e.g., Geiger et al. 2005; Geiger and Rama 2006; Feldmann and Read 2010; Blay et al. 2011; Myers et al. 2014). These studies benchmark GCOs against future bankruptcies. A Type I misclassification arises if the auditor issues a GCO and the client does not fail in the following year. A Type II misclassification arises when the auditor does not issue a GCO and the client does fail in the following year. In general, approximately half of the companies going bankrupt in the United States do not receive a prior GCO. In contrast, over two-thirds of companies with a GCO do not go bankrupt in the following year.[7]

Some stakeholders interpret the relatively large Type I and Type II error rates as evidence that GCOs have low information content. The review by Carson et al. (2013, 366) notes that "the issue of interest for regulators, creditors, lawyers, and other financial statement users is why auditors have failed to provide warning of impending bankruptcy for companies going bankrupt." An alternative explanation is that bankruptcies are simply too difficult to predict by using available information. This issue motivates us to use a variety of PD models, relying on existing information, as a benchmark for the incremental information value of GCOs.

GCOs can be incrementally informative under two mutually non-exclusive scenarios. First, the auditor's opinion aggregates public information that is not fully captured by the determinants of a default model (e.g., financial ratios), and, second, the auditor's opinion can indirectly convey private information from the auditor's discussions with management.

[7] The cited GCO accuracy studies use different sample periods and sources for GCOs and bankruptcy events. Geiger et al. (2005) and Geiger and Rama (2006) obtain GCOs from the SEC's EDGAR database and bankruptcy events from the *Bankruptcy Almanac*. Feldmann and Read (2010) compile a list of companies that filed for bankruptcy from BankrutcyData.com and then locate audit reports for those companies in the SEC's EDGAR database. Myers et al. (2014) use the Audit Analytics database to obtain relevant audit opinion data and the *Bankruptcy Almanac* to obtain bankruptcy events. Finally, Blay et al. (2016) use the Audit Analytics database to collect audit opinions and BankrutcyData.com to identify bankruptcy events.

2.3 Comparing GCOs to PD models

Early studies claim that GCOs are inferior to PD models that use the client's financial information as input (Altman and McGough 1974; Altman 1982; Levitan and Knoblett 1985; Koh 1991). These studies report that PD models predict bankruptcy correctly in approximately 85% of cases, while GCOs are correct only in approximately 45% of cases. However, subsequent studies by Hopwood et al. (1994) and Mutchler et al. (1997) find an insignificant difference between GCOs and PD models.[8] The collective evidence in prior "auditor versus PD model" research is limited in several respects. First, this research uses somewhat dated datasets ending in 1994 at the latest. Second, it focuses on a hand-collected subsample of bankrupt companies along with a similar (small) subsample of nonbankrupt companies. Third, the PD models include as inputs only financial ratios and client size.

Presently, audit firms do not use statistical models systematically as a diagnostic tool in going concern assessments. Based on our conversations with practitioners and standard setters, the most typical approach is a checklist that includes indicators of financial distress suggested by the GCO standard. This approach does not always facilitate the aggregation and weighing of multiple signals of distress. Even if the auditor uses the same inputs as a PD model, the PD model can systematically aggregate large amounts of historical data to generate a specific prediction and a confidence interval (i.e., a numeric probability that can be used to rank a client within a population of distressed companies).

2.4 Evolution of PD models

A vast body of literature in economics, finance, accounting, and other disciplines focuses on understanding and modeling the likelihood of a company's default or credit risk. Research has resulted in a number of approaches that can be classified as structural and reduced-form models. Structural models rely on an explicit relation between default risk and capital structure. These models, pioneered by Merton (1974), employ option-pricing theory in corporate debt valuation. Despite their theoretical appeal, a key disadvantage of structural models is that they are both analytically complex and computationally intensive (Jarrow 2009). Moreover, it is possible to include a distance-to-default determinant variable from a structural model like that of Merton (1974) in a reduced form model (e.g., Duan et al. 2012).

Reduced-form models view defaults as random, exogenous events. These models have evolved over time, and several studies document the incremental effect of including new determinants (e.g., accounting, market, and economy-wide data) or of using advances in methodology (e.g., logit, probit, hazard rate, and Poisson models). In the 1960s, models relied on discriminant analysis and generated credit scores, which allowed for ordinal rankings of companies to indicate their relative chance of default (Beaver 1966, 1968; Altman 1968). In the early 1980s, Ohlson (1980) and Zmijewski (1984) used models with logit and probit regressions, employing accounting ratios and

[8] A related study by Hopwood et al. (1989) focuses on various forms of modified opinions and bankruptcy. They find that, among modified opinions, GCOs have statistical significance in a PD model that also includes financial ratios.

firm size in year t as determinants of default in year $t+1$. Overall, the early models depended on accounting-based inputs. In fact, Hopwood et al. (1989, 1994) use a model based only on accounting ratios similar to these early models.

Subsequent models by Shumway (2001), Vassalou and Xing (2004), Chava and Jarrow (2004), Hillegeist et al. (2004), Bharath and Shumway (2008), and Campbell et al. (2008) have extended the prediction of default to multiple periods by introducing duration analysis and considering market variables (e.g., cumulative past returns, return volatility, etc.). These enhancements have resulted in incremental predictive power. The study of management's going concern disclosures by Mayew et al. (2015) is an example of recent work in accounting using this type of model.

More recently, a model by Duan et al. (2012) implements a stochastic forward intensity approach that (a) incorporates the effect of both defaults and other types of firm exits that may indicate distress, such as mergers and acquisitions,[9] (b) includes information from multiple periods in the form of state and trend variables, and (c) includes firm-level market characteristics and economy-wide risk factors. Duan et al. (2012) demonstrate that their model is incremental to the most recent approaches in predicting defaults. The Risk Management Institute at National University of Singapore has adopted the Duan et al. (2012) methodology to generate a proprietary database of PD estimates for many companies in multiple countries.[10]

We advance the literature by examining the incremental predictive power of GCOs, relative to various PD models. The first PD model is comparable to the approach used by Hopwood et al. (1994). The second PD model uses a combination of financial ratios, client size, and market variables from Shumway (2001) and is similar to the reduced-form models developed in the 2000s. The third PD model includes financial ratios and independent PD estimates calculated by the RMI at the National University of Singapore, using state-of-the-art methodology. Finally, the fourth model includes credit ratings in a reduced-form model. Using multiple benchmarks mitigates concerns that our findings rely on a given statistical model. In the next section, we introduce our research questions and methodology.

3 Methodology

3.1 Do GCOs provide incremental information, relative to financial ratios and other predictors of default?

Our main research question is: Do GCOs provide incremental information relative to other predictors of corporate default? In answering this question, our starting point is the comparison of predictive power between GCOs *independently* and a default model that uses financial ratios and basic determinants of GCOs and default (Hopwood et al. 1994; DeFond et al. 2002; Mayew et al. 2015). Next we examine whether including GCOs increases the model's

[9] This approach addresses a potential problem with previous default prediction methodologies. In prior methodologies, firms that disappear for reasons other than defaults are left out of the model, which can result in censoring biases (Duan et al. 2012).

[10] National University of Singapore, RMI, CRI database. Available at http://rmicri.org (accessed 27 January 2017).

predictive accuracy. Hence we estimate the following three logistic regression models.

$$P\left(DEFAULT_{i,t+1}\right) = f\left(GCO_{i,t}, Fixed\ Effects_{i,t,\varepsilon i,t}\right); \tag{1}$$

$$P\left(DEFAULT_{i,t+1}\right) = f\left(Financial\ Ratios_{i,t}, Fixed\ Effects_{i,t,\varepsilon i,t}\right); \tag{2}$$

$$P\left(DEFAULT_{i,t+1}\right) = f\left(GCO_{i,t}, Financial\ Ratios_{i,t}, Fixed\ Effects_{i,t,\varepsilon i,t}\right); \tag{3}$$

where $DEFAULT_{i,t+1}$ is an indicator variable equal to one if company i defaults in year $t+1$ and 0 otherwise and $GCO_{i,t}$ is an indicator variable equal to one if the auditor issues a going concern opinion for firm i in year t and 0 otherwise. *Financial Ratios* are key financial ratios, capturing a company's financial condition, and other control variables, including $ROA_{i,t}$, defined as net income divided by total assets; $LEVERAGE_{i,t}$, defined as total liabilities divided by total assets; $WCAP_{i,t}$, defined as current assets minus current liabilities (working capital) divided by total assets; $CURRENT_{i,t}$, defined as current assets divided by current liabilities; $CASH_{i,t}$, defined as cash and cash equivalent holdings divided by total assets; $CFO_{i,t}$, defined as cash flow from operating activities divided by total assets; $SIZE_{i,t}$, defined as the natural logarithm of total assets; $NEGEQUITY_{i,t}$, an indicator variable equal to one if total liabilities exceed total assets and 0 otherwise; and $BIGN_{i,t}$, an indicator variable equal to one if the company has a Big N auditor and 0 otherwise. $Fixed\ Effects_{i,t}$ are the Fama-French 12 industry classification and year fixed effects.[11] We cluster standard errors by company.

Our main criterion to compare the models' predictive power is the area under the receiver operating characteristic curve (AUROC). The receiver operating characteristic curve (ROC) is a plot of the true-positive rate versus the false-positive rate (i.e., the Type I Error) for different cutoff thresholds. Each point on the ROC plot represents a true-positive and false-positive rate pair corresponding to a particular decision threshold. To generate the ROC plot, we compute estimated probabilities that each observation i belongs to the default =1 group after fitting each logistic model. Next, after choosing a specific probability threshold (for example a 0.6 threshold), it is possible to classify all observations with an estimated probability greater and lower than the threshold as predicted default and nondefault, respectively. Nevertheless, given that the model is not perfect, applying a threshold results in some classification errors with respect to actual outcomes, and it is possible to determine the true-positive and false-positive rates at each possible threshold (i.e., create a two-by-two contingency table). After repeating these steps over many thresholds, it is possible to create a smooth ROC curve by plotting the true-positive and false-positive rates.[12]

[11] The Hopwood et al. (1994) model included ROA, LEVERAGE, CURRENT, CASH, two other ratios dividing current assets by total sales and total assets, and the natural logarithm of total sales as a proxy for size.

[12] There is no specific cutoff probability threshold prescribed by the auditing standards. An auditor is required to evaluate the going concern uncertainty after concluding that there is substantial doubt about the entity's ability to continue as a going concern for a reasonable period.

A model with perfect discrimination has an ROC plot that passes through the upper left corner (100% true-positive rate and 0% false-positive rate). The closer the ROC plot is to the upper left corner, the higher the overall accuracy of the model (e.g., Pepe et al. 2009). Therefore there is a relation between the overall rates of correct classification and the area under the ROC. The AUROC is a common summary statistic to compare models of discrete outcomes, also known as classifiers (Fawcett 2006; Pepe et al. 2009). The AUROC has been used in the accounting literature to evaluate the predictive ability of deceptive language in restatement cases (Larcker and Zakolyukina 2012) and the predictive ability of industry versus other determinants of litigation (Kim and Skinner 2012). It was also advanced as an alternative to pseudo-R^2 in models of auditor choice (Minutti-Meza 2013).

The AUROCs' possible values are in the interval (0.5, 1), where a value of 0.5 indicates no predictive ability and a value of 1 indicates perfect predictive ability. The AUROC can be interpreted as the probability that a randomly chosen default observation is rated or ranked as more likely to default than a randomly chosen nondefault observation. If defaulting firms receive a GCO before they defaulted, then auditors have discriminated well between healthy and distressed firms and the value of the AUROC is high. We assess the statistical significance of the difference in AUROCs between our models using a nonparametric Wald test (based on bootstrap with 1000 replications), where the null hypothesis is that both AUROC values are equal (Janes et al. 2009).[13]

Besides using the in-sample AUROCs, we also examine the out-of-sample predictive accuracy of our models. We employ K-fold cross-validation, a methodology from data mining (Witten and Frank 2005) that has been implemented in accounting studies (Kim and Skinner 2012; Larcker and Zakolyukina 2012). Cross-validation helps assess how accurately a predictive model will perform in practice or will generalize to an independent dataset, mitigating over-fitting problems. To perform K-fold cross-validation, the observations are first divided into subsets called folds, using random sampling without replacement from our whole sample. The choice of the number of folds is arbitrary; however, 10 folds have been widely used in practice (Witten and Frank 2005). As such, we divide our sample into 10 folds, with each fold containing roughly 3152 observations from our full sample. Next, we fit each of our models on nine out of the 10 folds, excluding the kth fold. Then we use the fitted models and predict the kth fold that was left out. This step of fitting a model on nine folds and estimating the kth fold is repeated nine additional times so that each fold has been left out and estimated once. We then compute the AUROC for each fold of out-of-sample predictions. This methodology creates an out-of-sample AUROC prediction for each of the 10 folds in our sample. Following Kim and Skinner (2012), we report the mean out-of-sample AUROC to further assess the predictive ability of the models in our main analysis.

[13] We use the Stata command rocreg to compare AUROCs. This command calculates standard errors based on bootstrap and does not assume that the ROCs are independent. This command is a modification of roccomp (Cleves 2002) that estimates standard errors following DeLong et al. (1988). As noted by Kim and Skinner (2012), although AUROC has advantages, a disadvantage is that it naturally increases as predictors are added to the model, in a manner analogous to unadjusted R-squares. There is no broadly accepted way of adjusting for this problem. However, our out-of-sample analyses, focusing on overall correct classification, are not affected by this problem. Finally, our cross-sectional analyses (Table 6) demonstrate that adding GCOs does not mechanically increase AUROC in all cases.

3.2 Do GCOs provide incremental information relative to financial ratios, market variables, and other predictors of default?

As explained in Section 2.4, the literature on PD models has evolved to include market variables to increase reduced-form models' predictive accuracy (Shumway 2001; Mayew et al. 2015). As such, we add market variables to our benchmark model (Eq. 2) and assess the incremental predictive power of the model after including GCOs. We estimate the following two logistic regression models, with and without $GCO_{i,t}$.

$$P\left(DEFAULT_{i,t+1}\right) = f\left(Market\ Variables_{i,t}, Financial\ Ratios_{i,t}, Fixed\ Effects_{i,t,\varepsilon i,t}\right) \quad (4)$$

$$P\left(DEFAULT_{i,t+1}\right) = f\left(GCO_{i,t}, Market\ Variables_{i,t}, Financial\ Ratios_{i,t}, Fixed\ Effects_{i,t,\varepsilon i,t}\right) \quad (5)$$

where all the variables are as defined above, with the addition of three market variables from Shumway (2001) and Mayew et al. (2015): $REL\_MKTCP_{i,t}$, defined as the logarithm of the firm's market capitalization at time t divided by the index's total market capitalization at time t; $EX\_RET_{i,t}$, defined as the cumulative returns of firm i minus the cumulative returns of the market over the 12 months leading up to the filing date for year t; and $SIGMA_{i,t}$, defined as standard deviation of the residuals from a model of monthly firm returns on market returns for the 12 months leading up to the filing date for year t.

3.3 Do GCOs provide incremental information relative to a company-level PD estimate from a recent and comprehensive default model and other predictors of default?

As explained in Section 2.4, the Duan et al. (2012) approach is a recent development in default prediction. We obtain company-level PD estimates from the Risk Management Institute at the National University of Singapore. These PD estimates were computed using the Duan et al. (2012) approach, which includes market factors, economy-wide factors, financial ratios, and other company characteristics. We include the PD estimates to our benchmark model (Eq. 2) and assess the incremental predictive power of the model after including GCOs. We estimate the following two logistic regression models, with and without $GCO_{i,t}$.

$$P\left(DEFAULT_{i,t+1}\right) = f\left(CONT\_PD_{i,t}, Financial\ Ratios_{i,t}, Fixed\ Effects_{i,t,\varepsilon i,t}\right) \quad (6)$$

$$P\left(DEFAULT_{i,t+1}\right) = f\left(GCO_{i,t}, CONT\_PD_{i,t}, Financial\ Ratios_{i,t}, Fixed\ Effects_{i,t,\varepsilon i,t}\right) \quad (7)$$

where all the variables are as defined above, with the addition of the $CONT\_PD_{i,t}$, defined as the estimated PD for the subsequent 12 months for firm i, calculated at the closest month-end before the signature date of the audit opinion for fiscal year t.

3.4 Do GCOs provide incremental information relative to credit ratings and other predictors of default?

Finally, we also compare the incremental information of GCOs, relative to credit ratings, another widely used public signal of upcoming default. This analysis does not rely on the explicit modeling of PDs. Instead, we assess the relative predictive power of the professional judgment of two third-party experts, auditors, and credit rating agencies and control for financial ratios and other client characteristics. We include credit ratings to our benchmark model (Eq. 2) and assess the incremental predictive power of the model after including GCOs. We estimate the following two logistic regression models, with and without $GCO_{i,t}$.

$$P\left(DEFAULT_{i,t+1}\right) = f\left(Credit\ Ratings_{i,t}, Financial\ Ratios_{i,t}, Fixed\ Effects_{i,t,\varepsilon i,t}\right) \quad (8)$$

$$P\left(DEFAULT_{i,t+1}\right) = f\left(GCO_{i,t}, Credit\ Ratings_{i,t}, Financial\ Ratios_{i,t}, Fixed\ Effects_{i,t,\varepsilon i,t}\right) \quad (9)$$

where information in $Credit\ Ratings_{i,t,}$ is captured by one variable indicating the level of Standard & Poor's (S&P) credit ratings and one of three alternative variables representing the change in ratings: $RATING_{i,t,}$ a categorical variable representing the credit rating of the firm at fiscal year-end t, ranging from 1 (AAA) to 10 (D); $DOWN_{i,t,}$ an indicator variable equal to one if the credit rating for firm i has been downgraded during the 12 months of fiscal year t; $DOWN\_LVL_{i,t}$, a count variable indicating the number of levels in credit ratings that firm i has been downgraded during the 12 months of fiscal year t; and $INVST\_CH_{i,t}$, an indicator variable equal to one if the credit rating for firm i has been downgraded from investment grade to non-investment grade during the 12 months of fiscal year t.

4 Data description

4.1 Sample selection

Our sample selection begins with all public companies included in Compustat North America in the years 1996 through 2015. We start in 1996 because this year is the first with machine-readable files in the SEC EDGAR database. We require a valid link to both CRSP and Audit Analytics (using CIK numbers). We keep only nonfinancial companies (i.e., delete companies with SIC codes between 6000 and 6999) and observations with nonmissing financial and market variables required to estimate our models. We then limit our analysis to distressed companies, with negative net income or negative cash flow from operations (following DeFond et al. 2002).[14] Table 1 shows

[14] As a robustness test, we confirm that our results hold when we include both distressed and nondistressed firms in a larger sample, finding that the combined model (Equation 5) has a statistically greater AUROC than any other model.

a detailed summary of the number of observations at each step of the data selection and merging process. Our final dataset includes 31,527 firm-year observations.

4.2 Auditor opinions

We use text-mining techniques to identify GCOs prior to the release of Audit Analytics, which begins tracking audit opinions in 2000. Following the technique of Butler et al. (2004), we first find audit opinions from 1996 through 2015 from 10-K filings on the SEC EDGAR website. Next, our algorithm searches for the key words “going concern,” “bankruptcy,” or “material business uncertainty” (Butler et al. 2004). Audit opinions that mention these key words receive a 1 for the GCO variable and a 0 otherwise. To assess the validity of our algorithm, we compare the results of the text-mining program to the Audit Analytics’ GCO variable. For the years of overlap between our search and Audit Analytics—from 2000 through 2015—there is a 0.933 correlation between our classification and the Audit Analytics’ classification of GCOs. Thus, in our final sample, the $GCO_{i,t}$ variable is comprised of the results from our text mining algorithm for the years prior to 2000 and then Audit Analytics for the years after 2000.[15]

4.3 Corporate default events

We focus on default events, including filing for bankruptcy, because these events are likely to challenge the going concern assumption. The auditing standards do not specifically refer to bankruptcy as the only event that constitutes an exception to the company’s ability to continue as a going concern (AS 2415.06 PCAOB 2017). We begin identifying default events, including bankruptcies, using the CRI database. The database includes a list of default events for more than 4000 firms since 1990, originating from Bloomberg, Compustat, CRSP, Moody’s, TEJ, exchange websites, and news sources. CRI classifies default events into hard and soft defaults. Hard defaults are those from which the company cannot recover. These defaults are administrations, arrangements, Chapter 11 proceedings, and liquidations, among others. Soft defaults are those from which the company could reemerge and that usually relate to debt defaults, for example, coupon or principal payment defaults, debt restructurings, loan payment defaults, etc. We include hard defaults and soft defaults that are followed by a hard default within 365 days in our analyses. Moreover, CRI identifies other events, such as delistings from stock exchanges, changes in industry, and mergers and

[15] We examined the non-overlapping GCO observations between our machine-based and the Audit Analytics (AA) classification. In total, there are 137 firm-years with a misclassification (out of 2181 GCOs available in AA). First, 37 were GCOs in our machine-based classification and not in the AA classification. Second, 100 were GCOs in the AA classification and not in our machine-based classification. Although these misclassifications constitute a small number of opinions, the biggest discrepancy resides in the second case, our machine-based algorithm failing to identify some GCOs. Out of the 100 misclassifications in the second case, we read 20 opinions and determined that the formatting of the text (e.g., page or line breaks in the middle of a sentence, unreadable text codes, etc.) in the 10-K filing prevented our algorithm from properly identifying the GCO. However, this issue would ultimately cause a reduction in GCOs’ predictive power by inducing noise in this independent variable. Also, this is a problem only for the period between 1996 and 1999, when we rely on our classification. Finally, our inferences are robust to limiting our sample to only the years covered by AA, starting with the year 2000.

Table 1 Sample selection

	Total Obs.	$GCO_{i,t}$	$DEFAULT_{i,t+1}$
Intersection of Compustat, CRSP, & Audit Analytics from 1996 to 2015	94,466	3796	783
- Firms defined as financial	(13,025)	(240)	(86)
- Firms with missing data to compute control variables	(6459)	(548)	(78)
- Firms not defined as distressed	(43,455)	(137)	(44)
Main Sample	31,527	2871	575
- Firm with missing CRI data	(6970)	(764)	(158)
CRI Analysis Sample	24,583	2127	424

This table shows the sample selection for the observations in the main analysis using financial and market variables. Then the sample selection is shown for our analyses using CRI's probability of default measure. Distressed firms are defined as companies with either negative net income or negative operating cash flows, following DeFond et al. (2002). Financial firms are defined as those firms having SICs between 6000 and 6999

acquisitions ("nondefault exits") for the firms in the database. From these other events, we include delistings with subsequent bankruptcies, liquidations, and reorganizations as default indicators. Next, we turn to more common data sources used in prior studies, including the UCLA-LoPucki Bankruptcy Research Database, Bankruptcy.com, CRSP, and Compustat to supplement our default data. As such, we include both bankruptcies and other default events (e.g., delistings followed by bankruptcy) in our analysis. Our final sample includes 575 subsequent default events.[16]

We address two timing issues in our default definition. First, if a company has multiple default events, we use the date of the first event as the beginning of the default period. Second, we exclude firm-year observations where a default event occurs between the fiscal year end and the filing date, because the auditors would observe the default before issuing their GCO in these cases.[17]

4.4 Probability of default (PD) estimated by CRI

We obtain the PD estimates from the CRI database generated by the Risk Management Institute at National University of Singapore, which imposes an additional data restriction on our sample for those analyses. For these analyses, we merge the data using Compustat Global, relying on SEDOL and ISIN as identifiers. Our final sample for the CRI PD analyses includes 24,583 firm-year observations and spans years 1996 through 2015. CRI computes PD estimates for public companies over different time horizons (one month to five years) based on a forward-intensity probability model developed by Duan et al. (2012). We use one-year PD's at the closest month-end to the audit opinion date to make the estimates comparable to the auditor's horizon in assessing going concern uncertainty. The forward intensity model is implemented by maximizing a

[16] Our results are robust to the exclusion of CRI default data and to using only bankruptcies identified by the common data sources used in the literature.

[17] Our results are robust to using the 12-month CRI PD calculated at fiscal year-end, which aligns the end of the PD horizon with the end of the auditors' GCO assessment horizon, and the six-month CRI PD calculated at the month end closest to the filing date.

pseudo-likelihood function. It includes a combination of ten firm-specific and two economy-level variables. The firm-specific variables used in the model are measures of volatility-adjusted leverage, liquidity, profitability, relative size, market misvaluation/ future growth opportunities, and idiosyncratic volatility. The first four firm characteristics enter the model as both level and trend variables. The economy-level variables are stock index returns and interest rates. The prediction accuracy of the CRI PD forecasts has been tested in-sample using accuracy ratio and ROC (CRI 2013). The range of possible accuracy ratio values is zero to one, where zero is a completely random system and one is a perfect rating system. The model used by CRI achieves an accuracy ratio of greater than 0.7 and an AUROC of greater than 0.8 at the one-year horizon for most of the countries covered by CRI including the United States. Finally, a potential benefit of a PD model over a GCO (in isolation) is that the former provides a continuous likelihood of default.

5 Results

5.1 Descriptive statistics

Table 2, Panel A, provides the full-sample descriptive statistics for the variables used in our main analyses. The average return on assets and cash flow from operations ($ROA_{i,t}$ and $CFO_{i,t}$) are negative (at −0.270 and − 0.127, respectively), because our sample is comprised of only distressed firm-years. The mean leverage is 49.2% ($LEVERAGE_{i,t}$), and 5.9% of the firm-years have negative equity ($NEGEQUITY_{i,t}$). Approximately 75% of the observations have a Big Four auditor ($BIGN_{i,t}$). We also show descriptive statistics for continuous PD estimates from the CRI database ($CONT\_PD_{i,t}$), which are available for a limited sample of 24,583 firm-year observations.

Panel B presents descriptive statistics separately for those firm-year observations with and without a GCO. Observations with $GCO_{i,t}$ equal to one have higher calculated PDs—with a mean of 5.0%—compared to observations with $GCO_{i,t}$ equal to zero, which have a mean PD of 1.2%. Approximately 10.0% of the firms that receive a GCO in our sample default a year after the auditor report ($DEFAULT_{i,t+1}$), while defaults in the non-GCO group average just 1.0%. The $GCO_{i,t}$ equal to one group also has comparatively (a) lower return on assets, cash flow from operations, working capital and current ratios, and levels of cash holdings; (b) smaller size and lower frequency of Big Four auditors; and (3) higher leverage and higher frequency of negative equity. These differences in means are all statistically significant at the 1% level. As expected, our descriptive statistics indicate that firms with GCOs exhibit generally worse financial conditions than firms without GCOs.

Panel C presents descriptive statistics separately for $DEFAULT_{i,t+1} = 0$ and $DEFAULT_{i,t+1} = 1$ observations. Among defaulting firm-years, 49.7% have $GCO_{i,t} = 1$, while among nondefaulting firm-years only 8.4% receive a going concern opinion. Observations that default also have significantly higher PDs, lower return on assets, smaller cash flows, higher leverage, and lower current and working capital ratios. More of the defaulting firm-years have negative equity. Interestingly, companies going into default are significantly larger, even in the year leading up to the default.

Table 2 Descriptive statistics

Panel A: Full sample

Variable	*Mean*	*S.D.*	*25 Perc.*	*Median*	*75 Perc.*
$DEFAULT_{i,t+1}$	0.018	0.134	0.000	0.000	0.000
$GCO_{i,t}$	0.091	0.288	0.000	0.000	0.000
$REL\_MKTCP_{i,t}$	−11.885	1.733	−13.107	−11.924	−10.735
$EX\_RET_{i,t}$	−0.035	0.818	−0.503	−0.233	0.137
$SIGMA_{i,t}$	0.046	0.030	0.025	0.038	0.058
$CONT\_PD_{i,t}$†	0.015	0.033	0.001	0.004	0.013
$ROA_{i,t}$	−0.270	0.434	−0.338	−0.112	−0.026
$LEVERAGE_{i,t}$	0.492	0.324	0.236	0.452	0.683
$WCAP_{i,t}$	0.303	0.309	0.079	0.285	0.530
$CURRENT_{,t}$	3.752	4.612	1.310	2.203	4.124
$CASH_{i,t}$	0.296	0.289	0.049	0.188	0.494
$CFO_{i,t}$	−0.127	0.298	−0.180	−0.031	0.041
$SIZE_{,i,t}$	4.874	1.854	3.540	4.664	5.997
$NEGEQUITY_{i,t}$	0.059	0.236	0.000	0.000	0.000
$BIGN_{i,t}$	0.751	0.432	1.000	1.000	1.000
N. Obs. =	31,527				

Panel B: Partition by going concern

	$GCO_{i,t} = 0$			$GCO_{i,t} = 1$			*Difference in Means*
Variable	*Mean*	*S.D.*	*Median*	*Mean*	*S.D.*	*Median*	
$\boldsymbol{DEFAULT_{i,t+1}}$	**0.010**	**0.100**	**0.000**	**0.100**	**0.300**	**0.000**	**0.090*****
$REL\_MKTCP_{i,t}$	−11.742	1.706	−11.772	−13.307	1.312	−13.351	−1.564***
$EX\_RET_{i,t}$	−0.003	0.823	−0.204	−0.356	0.690	−0.520	−0.353***
$SIGMA_{i,t}$	0.044	0.028	0.037	0.067	0.038	0.058	0.023***
$CONT\_PD_{i,t}$†	0.012	0.026	0.003	0.050	0.064	0.023	0.039***
$ROA_{i,t}$	−0.218	0.349	−0.096	−0.785	0.745	−0.560	−0.567***
$LEVERAGE_{i,t}$	0.472	0.303	0.437	0.702	0.440	0.643	0.230***
$WCAP_{i,t}$	0.330	0.287	0.306	0.031	0.381	0.021	−0.299***
$CURRENT_{,t}$	3.929	4.693	2.321	1.989	3.224	1.063	−1.941***
$CASH_{i,t}$	0.302	0.290	0.197	0.240	0.270	0.120	−0.062***
$CFO_{i,t}$	−0.096	0.248	−0.024	−0.432	0.511	−0.247	−0.336***
$SIZE_{,i,t}$	5.001	1.823	4.787	3.608	1.675	3.344	−1.392***
$NEGEQUITY_{i,t}$	0.045	0.208	0.000	0.201	0.401	0.000	0.156***
$BIGN_{i,t}$	0.766	0.423	1.000	0.599	0.490	1.000	−0.167***
N. Obs. =	28,656			2781			

Panel C: Partition by default

	$DEFAULT_{i,t+1} = 0$			$DEFAULT_{i,t+1} = 1$			*Difference in Means*
Variable	*Mean*	*S.D.*	*Median*	*Mean*	*S.D.*	*Median*	
$\boldsymbol{GCO_{i,t}}$	**0.084**	**0.277**	**0.000**	**0.497**	**0.500**	**0.000**	**0.414*****
$REL\_MKTCP_{i,t}$	−11.868	1.734	−11.902	−12.796	1.385	−12.865	−0.928***
$EX\_RET_{i,t}$	−0.026	0.820	−0.225	−0.553	0.505	−0.642	−0.527***
$SIGMA_{i,t}$	0.046	0.030	0.038	0.063	0.035	0.056	0.017***
$CONT\_PD_{i,t}$†	0.014	0.030	0.004	0.086	0.078	0.054	0.072***
$ROA_{i,t}$	−0.266	0.429	−0.110	−0.483	0.606	−0.239	−0.217***
$LEVERAGE_{i,t}$	0.486	0.320	0.446	0.826	0.379	0.801	0.340***
$WCAP_{i,t}$	0.308	0.306	0.290	0.020	0.348	0.032	−0.288***
$CURRENT_{,t}$	3.786	4.622	2.229	1.971	3.657	1.132	−1.815***
$CASH_{i,t}$	0.299	0.289	0.193	0.150	0.221	0.058	−0.149***

Table 2 (continued)

$CFO_{i,t}$	−0.125	0.297	−0.031	−0.188	0.373	−0.047	−0.063***
$SIZE_{,i,t}$	4.866	1.854	4.651	5.294	1.815	5.154	0.428***
$NEGEQUITY_{i,t}$	0.056	0.230	0.000	0.249	0.433	0.000	0.193***
$BIGN_{i,t}$	0.751	0.432	1.000	0.758	0.429	1.000	0.007
N. Obs. =	30,952			575			

+Denotes a variable constructed using CRI data which has a sample size of 24,583. Refer to Table 1

This table includes descriptive statistics for the variables in our analysis. Panel A uses all observations in our sample. Panels B and C partition the sample by going concern and by default, respectively, and report the difference in means. Throughout the table, ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10 levels, respectively, and the symbol + denotes a variable constructed using CRI data, which has a sample size of 24,583 firm-years with 2127 firm-years having a GCO and 424 firm-years having default events, as shown in Table 1. Refer to Appendix 1 for variable definitions

Figure 1 illustrates the frequencies of the GCOs by year and the subsequent defaults in our full sample, spanning the period from 1996 through 2015. Our full sample consists of 31,527 distressed firm-years, and there are 2871 observations with GCOs, or 9.1% of the sample. The annual frequency of GCOs ranges from 6.5 to 12.9%. The highest frequency of GCOs occurs from 2000 to 2002 and from 2007 to 2008, coinciding with economic downturns.[18] In contrast, the frequency of default is comparatively low. There are 575 subsequent default observations or 1.8% of the sample.[19] The incidence of defaults is highest during economic downturns in 2000–2001 and 2007–2008.

5.2 True-positive and true-negative rates using a threshold classification

Figure 2 presents the results of a descriptive analysis comparing the true positive and true negative rates when classifying companies that default (1/0) using either GCOs (1/0) or the estimated PDs from a logit model that uses financial ratios (Eq. 2). The model is fitted on three-year rolling windows, and the resulting estimates are used to predict the PD for the subsequent year. If the resulting PD is in the top decile (Panel A and C) or top two deciles (Panel B and D) of all PDs for that year, it is designated as having a predicted default equal to one.

The solid lines in Panels A and B present the true-positive rate of GCOs year over year, calculated as the percentage of default events in year $t+1$ which received a GCO in year t, while the solid lines in Panels C and D show the true-negative rates of GCOs, calculated as the percentage of nondefault events in year $t+1$ which did not receive a GCO in year t. The dotted lines in each panel columns are the resulting true-positive and true-negative rates from the model. Panel E displays the overall true-positive and true-negative rates over the full sample period and compares the statistical difference in these ratios.

[18] In general, our sample consists of relatively large firms with data in the Compustat, CRSP, and Audit Analytics databases, which are less likely to receive GCOs than the full population of U.S. listed firms (see Carson et al. 2013).

[19] Similarly, Blay et al. (2016) report that 1.79% of their financially distressed sample firms go bankrupt.

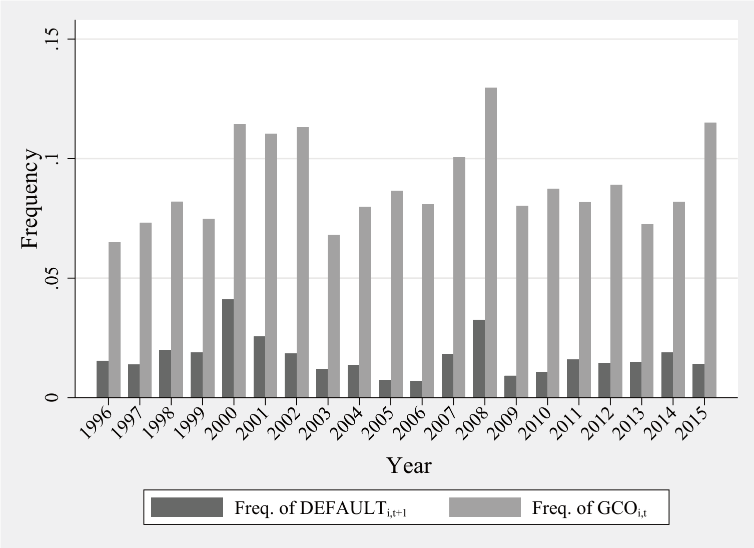

Fig. 1 Frequency of GCOs and subsequent defaults by year

Figure 2 illustrates the tradeoff between Type I and Type II errors. In terms of true positives (i.e., correctly predicted defaults), there are a handful of years where the PD model with financial ratios outperforms the GCO, especially when the larger cutoff is used (i.e., the top two deciles). However, using the top two deciles, the GCO outperforms the PD model with financial ratios in every year when considering true negatives (i.e., correctly predicted nondefaults). While the PD model with financial ratios is more accurate in terms of identifying true positives, the many false positives reduce the rate of correctly identified nondefaults. Without knowing the economic cost of false positives and the optimal classification threshold, it is unclear whether this is an acceptable trade-off. As such, we turn to the AUROC analyses to assess the predictive ability of each of our models, since the AUROC aggregates accuracy considering both true-positive and true-negative rates.

5.3 GCOs incremental information relative to financial ratios, market variables, and other predictors of default

Table 3, Panel A, shows the results for the logistic models described in Eqs. (1–5), including the coefficient estimates, the in-sample AUROCs, and the mean out-of-sample AUROCs calculated using k-fold cross validation. In Column (1), the $GCO_{i,t}$ variable alone has a positive coefficient of 2.413 ($p < 0.01$). This coefficient translates to an odds ratio of 11.173 (i.e., for a company receiving a GCO, the odds of a default event are 11.173 times greater than for a company not receiving a GCO). The AUROC for the GCO standalone model is 0.819. In Column (2), six of the eight predictors are associated with subsequent defaults ($p < 0.05$ or lower). The AUROC for the model is 0.828. In Column (3), the $GCO_{i,t}$ variable also has a positive coefficient of 2.332 ($p < 0.01$). This coefficient translates to an odds ratio of 10.298 (i.e., companies receiving a GCO are 10.298 times more likely to experience default than those not receiving a GCO, after controlling for financial ratios). The AUROC for the model is 0.862, using market

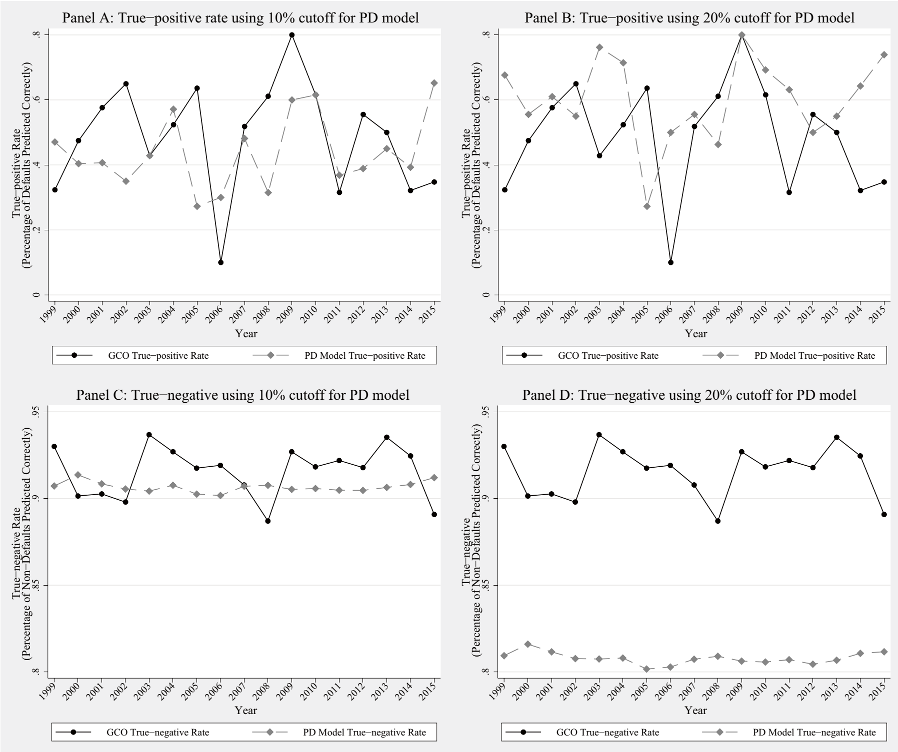

Fig. 2 True-positive and True-negative rates using a threshold classification. These panels compare the true-positive rate (i.e., correctly predicted defaults) and true-negative rate (i.e., correctly predicted nondefaults) of predicting default in year $t+1$ calculated from two measures: (1) receiving a GCO in year t and (2) being in the top 10 or top 20% of the predicted PD in year t derived from a model using financial ratios. The model is estimated using logistic regressions with three-year rolling windows (years t-3 to t-1) specified as:

$$P(DEFAULT_{i,t+1}) = f(ROA_{i,t},\ LEVERAGE_{i,t},\ WCAP_{i,t},\ CURRENT_{i,t},\ CASH_{i,t},\ CFO_{i,t},\ SIZE_{i,t},\ NEGEQUITY_{i,t},\ BIGN_{i,t},\ Industry\ Fixed\ Effects_{i,t},\ e_{it}).$$

The resulting coefficients were then used to predict a probability of default in year $t+1$ based upon the financial ratios for year t. The true-positive rate and true-negative rate of the resulting probability of default were determined based on whether the predicted probability fell within the top 10 or top 20% among all companies for year t. Please refer to Appendix 1 for variable definitions.

variables improves the model's predictive power, consistent with the findings of Shumway (2001). In Column (4), the $REL\_MKTCP_{i,t}$ and $EX\_RET_{i,t}$ variables, along with five other predictors, are associated with subsequent defaults ($p < 0.01$). The AUROC for the model is 0.883. Finally, in Column (5), the $GCO_{i,t}$ variable has a positive coefficient of 1.701 ($p < 0.01$). This coefficient translates to an increase in the odds ratio of 5.479 for companies with a GCO. In all Columns (1–5), the in-sample and out-of-sample AUROCs are above 0.8, indicating acceptable predictive classification (Hosmer and Lemeshow 2000, p. 162).

Turning to the models' relative predictive power, first we compare GCOs *alone* (Column 1 AUROC = 0.819) to the PD model that includes financial ratios, client size, and auditor type as predictors of default (Column 2 AUROC = 0.828). The difference

Table 3 GCOs' incremental information relative to other predictors of corporate default

Panel A: Logistic regression results

	Dependent Variable = $DEFAULT_{i,t+1}$				
	GCO	*Fin Ratios (FR)*	*GCO & FR*	*Market & FR*	*Full Model*
Variables	(1)	(2)	(3)	(4)	(5)
$GCO_{i,t}$	**2.413*****		**2.332*****		**1.701*****
	(26.59)		**(18.03)**		**(12.22)**
$REL\_MKTCP_{i,t}$				−0.444***	−0.387***
				(−12.09)	(−9.63)
$EX\_RET_{i,t}$				−2.027***	−1.666***
				(−7.14)	(−6.25)
$SIGMA_{i,t}$				0.401	−1.508
				(0.25)	(−0.88)
$ROA_{i,t}$		−0.411***	−0.204	−0.177	−0.070
		(−3.21)	(−1.56)	(−1.32)	(−0.52)
$LEVERAGE_{i,t}$		1.603***	1.708***	1.477***	1.582***
		(6.66)	(7.52)	(6.23)	(6.77)
$WCAP_{i,t}$		−1.395***	0.072	−1.037***	0.029
		(−5.45)	(0.31)	(−4.26)	(0.12)
$CURRENT_{i,t}$		0.064***	0.051***	0.070***	0.066***
		(3.84)	(2.98)	(4.00)	(3.71)
$CASH_{i,t}$		−0.500	−0.745**	−0.070	−0.367
		(−1.52)	(−2.31)	(−0.21)	(−1.10)
$CFO_{i,t}$		−0.538**	−0.350*	−0.946***	−0.758***
		(−2.49)	(−1.65)	(−4.26)	(−3.46)
$SIZE_{i,t}$		0.148***	0.258***	0.488***	0.513***
		(5.63)	(9.18)	(11.88)	(11.35)
$NEGEQUITY_{i,t}$		−0.220	−0.322	−0.210	−0.291
		(−0.96)	(−1.47)	(−1.02)	(−1.40)
$BIGN_{i,t}$		−0.041	−0.054	0.003	−0.028
		(−0.33)	(−0.43)	(0.02)	(−0.22)
Constant	−4.363***	−5.621***	−6.746***	−13.628***	−13.251***
	(−14.71)	(−16.55)	(−18.71)	(−20.77)	(−18.30)
Industry & Year FEs	Yes	Yes	Yes	Yes	Yes
N. Obs.	31,527	31,527	31,527	31,527	31,527
Pseudo-R^2	0.167	0.151	0.217	0.242	0.276
In-sample AUROC	**0.819**	**0.828**	**0.862**	**0.883**	**0.898**
Out-of-sample AUROC	0.811	0.819	0.854	0.876	0.891
In-Sample		(2)–(1)	(3)–(2)		(5)–(4)
Incremental AUROC		**0.009**	**0.034*****		**0.015*****

Table 3 (continued)

Panel B: Out-of-sample economic significance

Market & FR Model (Without $GCO_{i,t}$) Panel A, Column 4					*Full Model (With $GCO_{i,t}$) Panel A, Column 5*				
		$DEFAULT_{i,t+1}$					$DEFAULT_{i,t+1}$		
		=1	*=0*				*=1*	*=0*	
Model $Prediction_{i,t}$	*=1*	1.12%	7.81%	8.93%	*Model $Prediction_{i,t}$*	*=1*	1.20%	7.73%	8.93%
	=0	0.71%	90.36%	91.07%		*=0*	0.63%	90.44%	91.07%
		1.83%	98.17%	100.00%			1.83%	98.17%	100.00%
True-positive rate (1.12 / 1.83) = 61.20%					True-positive rate (1.20 / 1.83) = 65.57%				
True-negative rate (90.36 / 98.17) = 92.04%					True-negative rate (90.44 / 98.17) = 92.13%				

Table 3, Panel A includes results from the following logistic regressions.

$P(DEFAULT_{i,t+1}) = f(GCO_{i,t}, \textit{Fixed Effects}_{i,t}, e_{it})$; (Column 1).

$P(DEFAULT_{i,t+1}) = f(\textit{Financial Ratios}_{i,t}, \textit{Fixed Effects}_{i,t}, e_{it})$; (Column 2).

$P(DEFAULT_{i,t+1}) = f(GCO_{i,t}, \textit{Financial Ratios}_{i,t}, \textit{Fixed Effects}_{i,t}, e_{it})$; (Column 3).

$P(DEFAULT_{i,t+1}) = f(\textit{Market Variables}_{i,t}, \textit{Financial Ratios}_{i,t}, \textit{Fixed Effects}_{i,t}, e_{it})$; (Column 4).

$P(DEFAULT_{i,t+1}) = f(GCO_{i,t}, \textit{Market Variables}_{i,t}, \textit{Financial Ratios}_{i,t}, \textit{Fixed Effects}_{i,t}, e_{it})$. (Column 5).

Robust z-statistics are shown in the parentheses. Standard errors are clustered by company. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10 levels, respectively. Variable descriptions are included in Appendix 1.

The in-sample AUROC difference row includes the difference in the in-sample AUROC (the area under ROC) for the identified logistic models, and ***, **, and * indicate statistical significance of the AUROC differences resulting from the nonparametric Wald test at 0.01, 0.05, and 0.10 levels, respectively

Panel B displays the contingency table of predicted defaults using the model specified in columns 4 and 5 of Table 3 and subsequent defaults in time $t+1$

True-positive rate measures the proportion of actual defaults that are correctly identified as such. True-negative rate measures the proportion of actual nondefaults that are correctly identified as such. Predicted probabilities are generated using a k-fold cross validation technique, which results in out of sample predicted probabilities for each observation in the sample, following Witten and Frank (2005). To dichotomize the predicted probabilities of default from the fitted models, those observations with a predicted probability of default from the model in the top 9% of predicted probabilities in their fold are assigned a predicted value of 1. Observations outside the top 9% of predicted probabilities are assigned a predicted value of 0. The 9% cutoff is chosen to mirror the auditors' average GCO issuance rate over our sample period of 9.1%

in the two models' AUROCs is 0.009 and not statistically significant ($p > 0.10$). This comparison suggests that, *independently*, GCOs convey information that results in similar predictive power, compared to a basic PD model. This finding extends the Hopwood et al. (1994) results to recent years.

Next, we examine the incremental predictive power after adding the $GCO_{i,t}$ variable to the basic PD model (Eqs. 2 and 3). We find that the AUROC increases by 0.034, from 0.828 to 0.862 ($p < 0.01$). This finding indicates that GCOs have incremental information, relative to financial ratios and other indicators of default. Figure 3 illustrates the ROC plots for the GCO standalone model, the PD model with financial ratios, and the joint PD model with GCOs and financial ratios (Eqs. 1–3). The figure shows that the joint PD model is closer to the upper left region of the graph, indicating a larger AUROC and greater predictive power than the other two models.

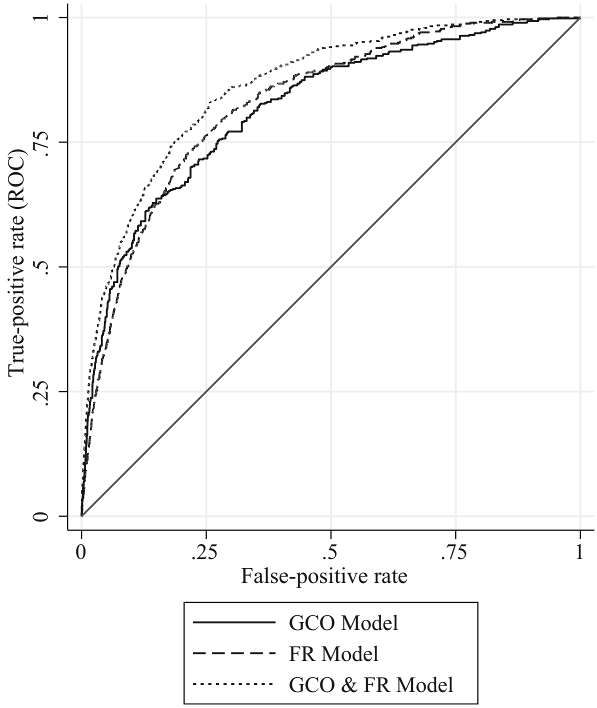

Fig. 3 ROC curves for eqs. (1), (2), and (3). This figure shows the ROC curves for all three equations: (1) GCO model, (2) financial ratio (FR) model, and (3) GCO and financial ratio model shown in Table 3. The receiver operating characteristic curve (ROC) is a plot of the true-positive rate (i.e., correctly predicted defaults) versus the false positive rate (i.e., incorrectly predicted defaults or Type I error) for different cut-offs. Each point on the ROC plot represents a true-positive rate and false-positive rate pair corresponding to a particular decision threshold. In this analysis, a model with perfect predictive power will produce curves near the upper left corner, while a random guess will be on the diagonal line. The AUROC is the area under the depicted curves

Finally, we examine the incremental predictive power after adding the $GCO_{i,t}$ variable to the model with financial ratios, client size, and market variables (Eqs. 4 and 5). We find that the AUROC increases by 0.015, from 0.883 to 0.898 ($p < 0.01$). This finding indicates that GCOs have incremental information relative to financial ratios, market variables, and other indicators of default.[20]

As a robustness test (untabulated), we estimate the models in Table 3 without industry and year fixed effects (FE). Adding GCOs still increases the AUROC of the models without FE (Columns 3 and 5). Specifically, without FE, the AUROCs of the models in Columns 4 and 5 are 0.873 and 0.885, respectively. These results suggest that the FE have a relatively small impact on the AUROCs.

[20] Cook and Ramadas (2019) argue that precision-recall curves may be preferable to ROC when there are a small percentage of cases of interest (e.g., low incidence of default). We also examined whether adding the $GCO_{i,t}$ variable to the model with financial ratios, client size, and market variables increases the area under the precision-recall curve (Equations 4 and 5). We find that including the $GCO_{i,t}$ variable increases the area under the precision recall curve from 0.174 to 0.217, and the model with the $GCO_{i,t}$ variable has a higher precision along all the recall scale (zero to one).

5.4 Out-of-sample performance of the model that includes financial ratios, client size, and market variables

While the increase of the in-sample AUROC is statistically significant in Column (5), we turn to out-of-sample performance to translate this increase into real terms. In other words, we quantify how many additional correct classifications of defaults (and nondefaults) result from the inclusion of the $GCO_{i,t}$ variable in the PD model. We employ our k-fold cross validation procedure to generate out-of-sample predictions for each observation. After each of the 10 folds is predicted from the other nine folds, we dichotomize the resulting predicted PD by selecting a cutoff threshold. We select a 9% cutoff to mirror the auditors' average GCO issuance rate over our sample period, which is 9.1%.[21] As such, if the estimated PD falls within the top 9% of predicted probabilities for that fold, we assign it a value of one to indicate a predicted default and a zero otherwise.[22]

Table 3, Panel B, shows the out-of-sample performance of the model that includes financial ratios, client size, and market variables, with and without the $GCO_{i,t}$ variable (Eqs. 4 and 5). The panel shows a contingency table for each alternative. Including a $GCO_{i,t}$ variable translates into an out-of-sample estimate of 4.4% increment (65.57–61.20% in each contingency table's true-positive rate) in the number of defaults correctly identified (23 firm-year observations). Also, including a $GCO_{i,t}$ variable results in a 0.09% reduction (92.04–92.13% in each contingency table's true-negative rate) in the number of defaults incorrectly identified (24 firm-year observations). Hence including GCOs in the default model decreases the model's Type I errors without increasing Type II errors. These findings confirm that GCOs provide incremental information, relative to other predictors of corporate default.

5.5 GCOs incremental information relative to a company-level PD estimate

Table 4 shows the results for the models in Eqs. (6 and 7). In Column (1), the $CONT\_PD_{i,t}$ variable has a positive coefficient of 14.976 ($p < 0.01$), indicating that the CRI company-level PD estimate is associated with subsequent default, controlling for financial ratios and other determinants in our basic PD model. We find that the AUROC increases by 0.013, from 0.881 to 0.894 ($p < 0.01$). This finding confirms that GCOs have incremental information, relative to other indicators of default.

In untabulated analyses, we also examine the out-of-sample performance of the model that includes the CRI company-level PD estimate, with and without the $GCO_{i,t}$ variable (Eqs. 6 and 7). Including a $GCO_{i,t}$ variable translates into an out-of-sample estimate of 5% increment in the number of defaults correctly identified (20 cases out of 398 defaults). Also, including a $GCO_{i,t}$ variable results in a 0.1% reduction in the number of defaults incorrectly identified (20 cases out of 24,162 nondefaults). Hence

[21] We thank Tyler Shumway (discussant) for suggesting this analysis and a percentage cutoff that mirrors the issuance rate of GCOs.

[22] Our inferences remain consistent when altering the cutoff threshold to the top 10, 20, and 30% of predicted probabilities. The pattern of increasing true-positive rates with the inclusion of GCOs remains similar at each cutoff, with the caveat that there are mechanically diminishing true-negative rates from the loosening of the cutoff threshold as illustrated in Figure 2.

Table 4 GCO incremental information relative to CRI probability of default

	Dependent Variable = $DEFAULT_{i,t+1}$	
	CRI Model	*Combined Model*
Variables	(1)	(2)
$\boldsymbol{GCO_{i,t}}$		**1.758*** **
		(10.16)
$\boldsymbol{CONT\_PD_{i,t}}$	**14.976*** **	**12.240*** **
	(19.35)	**(14.42)**
$ROA_{i,t}$	0.176	0.218
	(1.07)	(1.34)
$LEVERAGE_{i,t}$	1.415***	1.520***
	(4.67)	(5.11)
$WCAP_{i,t}$	−1.069***	−0.002
	(−3.48)	(−0.01)
$CURRENT_{i,t}$	0.083***	0.075***
	(4.54)	(4.03)
$CASH_{i,t}$	−0.011	−0.280
	(−0.03)	(−0.76)
$CFO_{i,t}$	−1.000***	−0.791***
	(−3.78)	(−3.03)
$SIZE_{i,t}$	0.087**	0.184***
	(2.37)	(4.79)
$NEGEQUITY_{i,t}$	−0.397	−0.456*
	(−1.54)	(−1.74)
$BIGN_{i,t}$	−0.078	−0.113
	(−0.53)	(−0.75)
Constant	−8.224***	−9.140***
	(−7.75)	(−8.54)
Year & Industry FEs	Yes	Yes
N. Obs.	24,583	24,583
Pseudo-R^2	0.236	0.268
In-sample AUROC	**0.881**	**0.894**
Out-of-sample AUROC	0.867	0.882
In-sample		(2)–(1)
Incremental AUROC		**0.013*** **

This table includes results from the following logistic regressions.

$P(DEFAULT_{i,t+1}) = f(CONT\_PD_{i,t}, \text{ Financial Ratios}_{i,t}, \text{ Fixed Effects}_{i,t}, e_{it})$; (Column 1).

$P(DEFAULT_{i,t+1}) = f(GCO_{i,t}, CONT\_PD_{i,t}, \text{ Financial Ratios}_{i,t}, \text{ Fixed Effects}_{i,t}, e_{it})$ (Column 2).

Robust z-statistics are shown in the parentheses. Standard errors are clustered by company. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10 levels, respectively. Variable descriptions are included in Appendix 1.

The in-sample AUROC difference row includes the difference in the in-sample AUROC (the area under ROC) for the identified logistic models, and ***, **, and * indicate statistical significance of the AUROC differences resulting from the nonparametric Wald test at 0.01, 0.05, and 0.10 levels, respectively.

including GCOs in this default model also decreases the model's Type I errors without increasing Type II errors.

5.6 GCOs incremental information relative to credit ratings

We generate a separate sample for this analysis, beginning with our main sample spanning the years 1996 through 2015. We then remove firms with no S&P credit rating in the Compustat North America database to compute our credit rating change variables defined above. Our final credit rating sample is comprised of 4171 observations. Table 5, Panel A, presents a summary of our sample selection process. Panel B provides descriptive statistics for the variables used in these analyses. The incidence of subsequent default in this sample is 4.6%, and the incidence of GCOs is 4.2%. Panel C shows the GCO frequency and subsequent default frequency by level of credit rating. The highest percentages of GCOs and defaults are at the B and CCC rating levels: approximately 2.0 and 1.2% of GCOs and 2.6 and 1.1% of defaults, respectively.

Table 5, Panel D, shows the results for models in Eqs. (8 and 9). Columns (1), (3), and (5) show three default models where information in credit ratings is captured by one variable indicating the level of ratings ($RATING_{i,t,}$) and by one of three alternative variables indicating changes in ratings ($DOWN_{i,t}$, $DOWN\_LVL_{i,t}$, and $INVST\_CH_{i,t}$). From the variables indicating changes in credit ratings, only $DOWN_{i,t}$ in Columns (1) and (2) has a statistically significant coefficient.

Next, the models with $GCO_{i,t}$ in Columns (2), (4), and (6) show a positive and statistically significant coefficient for this variable ($p < 0.01$). Given that $GCO_{i,t}$ and $DOWN_{i,t}$ are indicator variables, we can make a direct comparison of the magnitude of their coefficients. In Column (2), as an indicator of subsequent default, a GCO is approximately 3.5 times stronger than a credit rating downgrade. However, the results in Columns (4–6) suggest that GCOs and other variables subsume much of the information in credit rating downgrades. These results complement the findings of Funcke (2014), indicating that auditors use credit ratings in going concern assessments.

Focusing on the models' relative predictive power, adding a $GCO_{i,t}$ variable results in incremental AUROCs of 0.020, 0.019, and 0.019 in Columns (2), (4), and (6), respectively. The increases in AUROC are statistically significant ($p < 0.10$ or lower). These results confirm that GCOs have incremental information, relative to credit ratings. However, credit ratings have certain limitations in predicting default (Duan and Van Laere 2012). Large debt issuers, including Enron and Lehman Brothers, defaulted without a warning by major credit rating agencies. These agencies also failed to foresee the Latin American debt crisis and the European sovereign debt crisis.

In untabulated analyses, we also examine the out-of-sample performance of the models in Columns (1 and 2), which show the strongest coefficients for the ratings variables. Including a $GCO_{i,t}$ variable translates into an out-of-sample estimate of 5% increase in the number of defaults correctly identified (19 cases out of 190 defaults). Also, including a $GCO_{i,t}$ variable results in a 0.5% reduction in the number of defaults incorrectly identified (19 cases out of 3981 non-defaults). Hence including GCOs in this default model also decreases the model's Type I errors without increasing Type II errors.

6 Additional cross-sectional analyses

We investigate whether the incremental information in GCOs varies across 12 partitions that capture differences in the characteristics of the client, the auditor, and the

Table 5 GCO incremental information relative to credit ratings

Panel A: Sample Selection

	Total Obs.	*$GCO_{i,t}$*	*$DEFAULT_{i,t+1}$*
Main Sample from Table 1	31,527	2871	575
- Firms missing credit rating data	(27,356)	(2497)	(385)
Credit Rating Sample	4171	174	190

Panel B: Descriptive Statistics

Variable	*Mean*	*S.D.*	*25 Perc.*	*Median*	*75 Perc.*
$DEFAULT_{i,t+1}$	0.046	0.209	0.000	0.000	0.000
$GCO_{i,t}$	0.042	0.200	0.000	0.000	0.000
$DOWN_{i,t}$	0.160	0.367	0.000	0.000	0.000
$DOWN\_LVL_{i,t}$	0.189	0.497	0.000	0.000	0.000
$INVST\_CH_{i,t}$	0.040	0.195	0.000	0.000	0.000
$ROA_{i,t}$	−0.098	0.200	−0.109	−0.041	−0.013
$LEVERAGE_{i,t}$	0.751	0.252	0.591	0.722	0.879
$WCAP_{i,t}$	0.115	0.176	0.011	0.096	0.207
$CURRENT_{i,t}$	1.859	1.745	1.067	1.530	2.162
$CASH_{i,t}$	0.095	0.118	0.019	0.054	0.126
$CFO_{i,t}$	0.039	0.082	0.000	0.043	0.082
$SIZE_{i,t}$	7.628	1.327	6.611	7.542	8.591
$NEGEQUITY_{i,t}$	0.128	0.334	0.000	0.000	0.000
$BIGN_{i,t}$	0.943	0.232	1.000	1.000	1.000
N. Obs. = 4171					

Panel C: GCOs and defaults by credit ratings

$RATING_{i,t}$	*$GCO_{i,t}$ N. Obs.*	*$GCO_{i,t}$ Freq.*	*$DEFAULT_{i,t+1}$ N. Obs.*	*$DEFAULT_{i,t+1}$ Freq.*	*Total N. Obs.*
AAA	0	0.000	0	0.000	0
AA	0	0.000	0	0.000	17
A	1	0.000	1	0.000	159
BBB	0	0.000	3	0.001	568
BB	14	0.003	23	0.006	1192
B	82	0.020	109	0.026	1940
CCC	51	0.012	47	0.011	247
CC	7	0.002	3	0.001	13
C	0	0.000	0	0.000	0
D	19	0.005	4	0.001	35
Total	174	0.042	190	0.046	4171

Panel D: Logistic regression results

	Dependent Variable = $DEFAULT_{i,t+1}$					
	DOWN Models		*DOWN_LVL Models*		*INVST_CH Models*	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
$GCO_{i,t}$		**2.422*** **		**2.455*** **		**2.455*** **
		(8.08)		**(8.23)**		**(8.30)**
$DOWN_{i,t}$	**0.645*** **	**0.598*** **				
	(3.10)	**(2.66)**				
$DOWN\_LVL_{i,t}$			**−0.017**	**−0.146**		
			(−0.12)	**(−0.86)**		
$INVEST\_CH_{i,t}$					**−0.390**	**−0.664**
					(−0.85)	**(−1.28)**

Table 5 (continued)

$RATING_{i,t}$	0.381***	0.190*	0.487***	0.343***	0.476***	0.272***
	(4.30)	(1.88)	(4.45)	(2.90)	(5.46)	(2.68)
$ROA_{i,t}$	−0.999***	−0.887***	−1.116***	−1.052***	−1.133***	−1.049***
	(−3.68)	(−3.08)	(−4.11)	(−3.66)	(−4.07)	(−3.58)
$LEVERAGE_{i,t}$	0.203	0.521	0.254	0.570	0.268	0.590
	(0.40)	(1.01)	(0.50)	(1.12)	(0.53)	(1.15)
$WCAP_{i,t}$	−3.079***	−0.759	−3.213***	−0.858	−3.215***	−0.850
	(−4.76)	(−1.05)	(−4.97)	(−1.21)	(−4.95)	(−1.19)
$CURRENT_{i,t}$	−0.009	−0.073	0.001	−0.065	0.001	−0.067
	(−0.08)	(−0.48)	(0.01)	(−0.43)	(0.01)	(−0.44)
$CASH_{i,t}$	−1.148	−2.118	−1.537	−2.591**	−1.541	−2.477*
	(−0.90)	(−1.64)	(−1.21)	(−2.02)	(−1.22)	(−1.94)
$CFO_{i,t}$	−4.747***	−3.661**	−4.934***	−3.795**	−4.936***	−3.815**
	(−3.07)	(−2.14)	(−3.19)	(−2.22)	(−3.19)	(−2.25)
$SIZE_{i,t}$	−0.116	−0.084	−0.068	−0.010	−0.061	−0.022
	(−1.53)	(−1.05)	(−0.84)	(−0.12)	(−0.82)	(−0.28)
$NEGEQUITY_{i,t}$	0.492	0.350	0.438	0.298	0.434	0.298
	(1.34)	(0.91)	(1.21)	(0.79)	(1.20)	(0.79)
$BIGN_{i,t}$	−0.123	−0.321	−0.141	−0.343	−0.148	−0.356
	(−0.41)	(−1.09)	(−0.48)	(−1.19)	(−0.51)	(−1.24)
Constant	−4.290***	−3.586***	−5.157***	−4.851***	−5.138***	−4.390***
	(−4.10)	(−3.26)	(−4.35)	(−3.96)	(−5.02)	(−4.02)
Industry & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N. Obs.	4171	4171	4171	4171	4171	4171
Pseudo-R^2	0.245	0.302	0.239	0.297	0.240	0.298
In-sample AUROC	**0.863**	**0.883**	**0.863**	**0.882**	**0.864**	**0.882**
In-Sample		(2)–(1)		(4)–(3)		(6)–(5)
Incremental AUROC		**0.020****		**0.019***		0.019*

Panels A, B, and C display the sample formation and characteristics of firm-year observations that have credit ratings. Panel D includes results from the following logistic regressions.

$P(DEFAULT_{i,t+1}) = f(Credit\ Ratings_{i,t}, Client\ Characteristics_{i,t}, Fixed\ Effects_{i,t}, e_{it})$; (Column 1, 3, 5).

$P(DEFAULT_{i,t+1}) = f(GCO_{i,t}, Credit\ Ratings_{i,t}, Client\ Characteristics_{i,t}, Fixed\ Effects_{i,t}, e_{it})$. (Column 2, 4, 6).

Robust z-statistics are shown in the parentheses. Standard errors are clustered by company. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10 levels, respectively. Variable descriptions are included in Appendix 1.

The in-sample AUROC difference row includes the difference in the in-sample AUROC (the area under ROC) for the identified logistic models, and ***, **, and * indicate statistical significance of the AUROC differences resulting from the nonparametric Wald test at 0.01, 0.05, and 0.10 levels, respectively.

economic environment. For each partition, we estimate two separate default models that include financial ratios, client size, and market variables, with and without a $GCO_{i,t}$ variable (Eqs. 4 and 5). Although research has examined some auditor characteristics that influence the incremental information of GCOs (e.g., Big N versus non-Big N auditor), our analyses complement the literature by using a large and up-to-date sample of GCOs and by relying on a consistent framework to assess incremental information. We aim to determine whether including the $GCO_{i,t}$ variable results in incremental AUROC within each partition.[23]

6.1 Client characteristics

We examine three partitions based on variation in client characteristics: (1) size, where large companies are in the top quartile of $SIZE_{i,t}$ and small companies are in the lowest quartile of $SIZE_{i,t}$; (2) information environment, where we identify companies with and without financial analyst following; and (3) level of distress, where companies in low distress are in the top quartile of $CFO_{i,t}$ and companies in high distress are in the lowest quartile of $CFO_{i,t}$.

We expect that GCOs will be especially informative for smaller companies and companies with poor information environments (i.e., low analyst following). We do not have a clear expectation for companies in different levels of distress. GCOs may not provide much incremental information in cases where a client is clearly in financial distress. However, unobservable factors known by the auditor—such as client restructuring plans—may make GCOs especially informative in cases where a client is in high distress. Consistent with these expectations, in Table 6, Column (3), we document that including the $GCO_{i,t}$ variable results in incremental AUROCs for relatively small companies (0.013, $p < 0.10$), companies with no analyst following (0.025, $p < 0.01$), and companies with both high (0.016, $p < 0.10$) and low (0.014, $p < 0.10$) levels of financial distress.

6.2 Properties of the GCO

We examine two partitions based on variation in the properties of the GCO: (1) first-time GCO, where we identify companies with and without a prior GCO, and (2) common versus uncommon distress conditions mentioned by the auditor in the opinion's text. Audit Analytics provides a classification of the auditor's reasons to issue GCOs, based on the text of the audit opinion. Using this classification, available starting in the year 2000, 80% of the GCOs are issued for reasons linked to "net/operating loss (including recurring losses)" and "negative cash flow from operations." We label these opinions as "typical GCOs." The remaining 20% of GCOs are issued for reasons linked to a variety of other topics, the most frequent of which are "working

[23] As explained in our methodology section, there is a well-defined statistical test that determines the change in predictive power when a new variable is added to a model. Unfortunately, we are unaware of a similar test that compares the marginal effect of adding a new variable to two models estimated in non-overlapping samples. This test implies "double differencing." A limitation of our partition analyses is that they only provide inferences regarding whether GCOs add incremental information to both or only one partition. We find that GCOs add incremental predictive power in eight of our 12 partitions. In four partitions (i.e., level of distress, common versus uncommon distress conditions mentioned by the auditor, Big N, and high likelihood of opinion shopping), we conclude that the results do not support the notion that partitioning the sample along those criteria affects the incremental information of GCOs. Finally, we note that another limitation of these analyses is that they rely on subsamples of the data, where the low number of GCOs and defaults may reduce statistical power.

capital/current ratio deficit," "need for additional financing," and "accumulated/retained earnings deficit." We label these cases as "nontypical GCOs."[24]

Using the two partitions above, we aim to isolate cases in which GCOs could provide relatively more information. We expect that GCOs will be incrementally informative when they are issued for the first time and for nontypical distress conditions. As expected, In Table 6, Column (3) we show that GCOs issued for the first-time result in an incremental AUROC (0.013, $p < 0.01$). However, GCOs citing both common (0.007, $p < 0.05$) and uncommon (0.013, $p < 0.01$) distress conditions provide incremental information.

6.3 Auditor and engagement characteristics and economic cycles

We examine six auditor and engagement characteristics: (1) Big N, where we identify companies with and without Big N auditors and expect Big N auditors to issue GCOS with comparatively higher incremental information value (Myers et al. 2014); (2) first-year engagement, where we identify companies that switch and do not switch auditors in the current year and expect non first-time audits to be subject to comparatively lower time pressure and to have incrementally informative GCOs (Cassell et al. 2017); (3) probability of opinion shopping, where we use a similar approach to that of Newton et al. (2016) and Lennox (2000) to isolate opinion shopping around high/low likelihood of receiving a GCO[25] and expect cases with low likelihood to have incrementally informative GCOs; (4) competition in the local audit market, where we use a similar approach to that of Newton et al. (2016) to partition our sample into the top and bottom quartile of local competition (i.e., distance between one audit firm and the next closest firm in terms of MSA-level audit fees) and expect cases with low market competition to have incrementally informative GCOs; (5) local-office growth, where we use a similar approach to that of Bills et al. (2015) to partition our sample into the top and bottom quartile of auditor growth (i.e., percentage change in audit fees at the auditor-year-MSA level) and expect cases with low growth pressure to have incrementally informative GCOs; and (6) client importance for a local office, where we use a similar approach to that of Li (2009) to partition our sample into the top and bottom quartile of client importance (i.e., ratio of client-specific fees to MSA-level

[24] It is difficult to empirically separate expected and unexpected GCOs. The determinants of GCO and default have a high degree of overlap (e.g., negative ROA and stock returns). By controlling for these joint determinants in a model that includes GCOs, the incremental AUROC arguably captures the orthogonal information component of the auditor's opinion. This approach is similar to estimating a first-step model, predicting GCOs, and next including the residual (i.e., unexpected GCOs) in a second-step model of probability of default. Unfortunately, without strong exogenous variables to predict GCOs in the first step, the two-step procedure can actually result in worse inferences, due to the noise in two-step estimators.

[25] First, we model the likelihood of receiving a GCO in year t, where the independent variables are financial ratios, size, GCO in year t-1, and an indicator variable equal to one if the auditor is subsequently dismissed (DISMISS). Next, we estimate two probabilities, P1 and P0, based on the fitted model parameters for each firm-year observation, when DISMISS =1 and DISMISS = 0, respectively. Finally, we identify cases where the client seeks a favorable opinion: (1) the likelihood of receiving a GCO is higher when staying with the current auditor (P0 > P1), but the client switched auditors; and (2) the likelihood of receiving a GCO is lower when staying with the current auditor (P1 > P0), and the client did not switch.

Table 6 Comparison of GCO incremental information in various cross-sections

Cross-section	*Expectation of GCO's incremental information*	(1) *Market & FR Model (Without $GCO_{i,t}$)* *AUROC*	(2) *Full Model (With $GCO_{i,t}$)*	(3) *Cols. (2)–(1) Difference*	*No. Obs.*
Client characteristics					
Large companies [a]	Low	0.925	0.926	0.001	7889
Small companies [a]	High	0.853	0.866	0.013*	7889
Analyst following	Low	0.940	0.942	0.002	14,133
No Analyst following	High	0.853	0.878	0.025***	17,394
High distress [b]	?	0.884	0.900	0.016***	7889
Low distress [b]	?	0.918	0.932	0.014**	7889
Properties of GCO					
GCO in prior year	Low	0.878	0.885	0.007	1059
No GCO in prior year	High	0.889	0.902	0.013***	30,468
Typical GCO topic	?	0.892	0.899	0.007**	25,269
Non-typical GCO topic	?	0.877	0.890	0.013***	23,794
Auditor and engagement characteristics					
Big N auditor	High	0.901	0.912	0.011***	23,700
Non-Big N auditor	Low	0.867	0.881	0.014**	7827
First year engagement	?	0.912	0.915	0.003	1909
Non-first year engagement	?	0.899	0.915	0.016***	17,081
High opinion shopping	Low	0.874	0.886	0.012***	15,241
Low opinion shopping	High	0.897	0.916	0.019***	16,063
High audit market competition	Low	0.885	0.894	0.009	4790
Low audit market competition	High	0.903	0.916	0.013*	4358
High audit market growth	Low	0.913	0.919	0.006	4193
Low audit market growth	High	0.905	0.921	0.016**	3986
High client importance	Low	0.909	0.925	0.016	4469
Low client importance	High	0.910	0.922	0.012**	4517
Economic cycles					
Fiscal year ended 2006–2008	?	0.866	0.882	0.016*	4581
Fiscal year ended 2009–2011	?	0.922	0.925	0.003	4027
Fiscal year ended 2012–2014	?	0.890	0.898	0.008	4055

This table includes the AUROC comparison for the identified cross-sections. The AUROC's were calculated from performing the logistic regressions from Table 3, Columns 4 and 5, on the identified subset of the sample. Column 3 reports the difference in AUROCs between the models in the identified columns. In these columns, ***, **, and * indicate statistical significance of the AUROC differences resulting from the nonparametric Wald test at 0.01, 0.05, and 0.10 levels, respectively

[a] Size is calculated as the log of total assets of a firm. Firms that are in the lower quartile of size in our sample—that is, below \$34.5 million in assets—are classified as small. Firms that are in the upper quartile of size in our sample—that is, above \$402.9 million in total assets—are classified as large

[b] Level of distress is calculated by ranking operating cash flows as a percentage of total assets ($CFO_{i,t}$) from low to high. Firms that are in the lower quartile of this ranking—that is, firms with operating cash flows lower than −18.0% of total assets—are classified as severely distressed. Firms that are in the upper quartile of this ranking—that is, firms with operating cash flows greater than 4.1% of total asset—are classified as slightly distressed. Results are robust to using net income as a percentage of total assets to rank level of distress

auditor fees for each year) and expect cases with low client importance to have incrementally informative GCOs, given that, in these cases, independence is relatively less likely to be compromised.[26]

Using the six partitions above, we investigate whether variation in auditors' time pressure and incentives affects going concern judgments. Consistent with our expectations, we show in Table 6, Column (3), that GCOs provide incremental information when they are issued for recurring engagements (non-first year) (0.016, $p < 0.01$), and also when local audit market competition (0.013, $p < 0.10$), local-office growth (0.016, $p < 0.05$), and client importance are low (0.012, $p < 0.05$). Collectively, it is arguably less likely that auditor resources and independence will be compromised in these cases. However, we find that GCOs provide incremental information, regardless of whether the auditor is a Big N (0.011, $p < 0.01$) or non-Big N (0.014, $p < 0.05$) firm, and whether there is a high (0.012, $p < 0.01$) or low (0.019, $p < 0.01$) likelihood of opinion shopping.

Finally, we examine variation in economic cycles. We focus on three periods with respect to the financial crisis: leading up to the crisis (2006–2008), during it (2009–2011), and afterward (2012–2014). In Table 6, Column (3), we find that GCOs provide incremental information in the period leading up to the crisis (0.016, $p < 0.10$).

7 Other robustness and sensitivity tests

7.1 Timing of the opinion and subsequent default

An interesting issue is whether the auditor's work extends to subsequent events and whether GCOs are incrementally useful in predicting defaults that happen shortly after year-end. Our main analyses already consider this possibility by removing defaults that occur between year-end and the date of the auditor's opinion. The opinion's date is typically close to the annual report date. In untabulated analyses, we examine a partition of defaults that occur in the first half and the second half of the year after a GCO. As expected, we find that defaults in the second half of the year are comparatively more difficult to predict, the concordance between GCOs and defaults goes down from 61.4 to 34.2%. Nevertheless, we find that GCOs add incremental information to both half-year partitions using the model with market variables in (Eqs. 4 and 5). Overall, GCOs are informative within the full 12 months prescribed by the auditing standards.

Another question of interest is whether GCOs help to predict defaults that happen more than 12 months after the opinion. We include GCOs in year $t-1$ to Eqs. (4 and 5). The GCOs in year t-1 do not result in a statistically significant increase in the AUROC (above the improvement resulting from including GCOs in year t). This finding suggests that GCOs have information that helps predict default over a one-

[26] Other auditor incentives, beyond independence constraints, may also affect the results for the partitions based on client importance. For instance, the auditor may exert high effort for highly important clients, and this effort can result in incrementally informative opinions.

year horizon and that GCOs in any given year likely subsume information from the prior years' opinions.

7.2 Other predictors of default

In the model with market variables (Eqs. 4 and 5), we also use a more comprehensive list of predictor variables, including three PD estimates (1) the CRI company-level PD, (2) the Ohlson O-Score, and (3) the Z-score. We find that the AUROC increases by 0.008, from 0.901 to 0.909 ($p < 0.05$). This inference is robust to the exclusion of the predictor variables one at a time.[27]

The CRI company-level PD estimate in Table 4 (Eqs. 6 and 7) is based on a framework that uses levels and trends of the predictor variables. However, we also estimate Eqs. (4) and (5) using each variable's levels in year t and changes between years $t-1$ and t (for those variables that do not already inherently represent a change). We find that the coefficients of some of the variables' changes are statistically significant (i.e., changes in working capital, cash flow from operations, and size). We find that the AUROC increases by 0.011, from 0.890 to 0.901 ($p < 0.01$).

Finally, we calculate continuous PD estimates based on the fitted values of Eq. (4) for year t, after estimating the model's coefficients using three-year rolling windows, from year $t-1$ to year $t-3$. We add GCOs in year t to the continuous PDs to determine whether GCOs are incrementally informative. This approach considers only two summary signals of default, our own PD estimates, based on prior information, and the GCOs. We find that the AUROC increases by 0.016, from 0.858 to 0.874 ($p < 0.01$) after adding the $GCO_{i,t}$ variable to the continuous PD model.

7.3 Hazard rate model

Logistic regressions are similar to discrete-time hazard models (Shumway 2001; Mayew et al. 2015). However, logistic regressions do not perfectly mirror discrete-time hazard models, specifically in the magnitude of the estimated coefficients. The hazard function estimates the likelihood that a company will survive to a certain point in time (i.e., year t) based on its survival to an earlier time. This analysis is equivalent to using the predictor variables in multiple years (i.e., lags) to predict a default at any given point in time. This analysis mitigates concerns that the long-term trend of the predictor variables is an important omitted variable. We also use a hazard rate model to estimate Eqs. (4) and (5). Using the transformed survival probabilities generated from the

[27] Our main models (Equations 4 and 5) include overlapping determinants of the O-Score and the Altman Z-score, such as company size, working capital, current ratio, and ROA. Nevertheless, as a robustness test, we also examine whether GCOs provide incremental predictive ability to both the O-Score and the Altman Z-score, similar to our analyses based on the CRI company-level PD estimate. Including GCOs increases the AUROC of a model with the O-Score, Z-Score, and the non-overlapping control variables in the model with financial ratios and market variables. Existing evidence demonstrates that the CRI company-level PD estimate is a better predictor of performance than any other common prediction scores (Duan et al. 2012). Finally, a limitation of using all determinants is that they impose additional data constraints.

hazard rate models, we find that the AUROC increases by 0.017, from 0.805 to 0.822 ($p < 0.01$) after adding the $GCO_{i,t}$ variable. These results confirm that GCOs are incrementally informative, relative to the levels and trends of other predictors of default.

7.4 Including nondistressed observations in the sample

We also conduct our main analyses without conditioning on financial distress. (See the discussion of this restriction in our sample selection.) Specifically, using Eqs. (4) and (5), we find that the AUROC increases by 0.011, from 0.890 to 0.901 ($p < 0.01$) after including the $GCO_{i,t}$ variable. These results confirm that GCOs are incrementally informative, regardless of conditioning on financial distress. However, for our main analyses, we follow the common practice by limiting our sample to distressed firm-year observations (e.g., DeFond et al. 2002).

8 Conclusion

The auditor has the responsibility to assess the client's going concern uncertainty. Although auditing standards suggest broad factors and guidance to issue a GCO, investors, regulators, and researchers continue to raise concerns regarding the auditor's role in warning investors about business failures. We investigate whether GCOs provide incremental information in predicting a client's default. Our measure of incremental information is the additional predictive power that GCOs give to statistical default models.

Our findings have four limitations. First, our results must be interpreted relative to the models in our analyses. However, by using multiple benchmarks, we mitigate concerns that our findings rely entirely on a given statistical model. Second, our results do not speak to the cost-benefit trade-off of Type I and Type II errors, especially given the inherent difficulty in quantifying the cost of a false positive. Yet the inclusion of a GCO variable in our most comprehensive PD model results in increases to both true-positive and true-negative rates. Third, we cannot speak to the direct cost of implementing a statistical model in the audit process. We envision a joint approach where auditors may combine a client-specific PD estimate with existing methodologies, but we leave quantifying the cost of such an implementation to future research. Fourth, it is fundamentally difficult to separate the GCOs' informational role from its potential "self-fulfilling prophecy" effect. Vanstraelen (2003) and Louwers et al. (1999) document that distressed companies with a GCO are more likely to file for bankruptcy in the following year, relative to later years. Similarly, a recent study by Gerakos et al. (2017) finds support for both a prediction and a small inducement effect of GCOs.

Initially, we find no difference in predictive power between GCOs alone and a default model that includes financial ratios. However, we also demonstrate that GCOs and other predictors of distress do not convey the exact same information. There is an imperfect overlap between GCOs and other predictors.

We show that GCOs increase the predictive power of several models that include ratios, market variables, probability of default estimates, and credit ratings. Specifically, GCOs increase the number of predicted defaults by 4.4%, without increasing Type II errors, using a model that includes ratios and market variables. Our collective findings suggest that GCOs summarize a complex set of conditions not captured by other predictors of default.

Acknowledgements The authors are grateful for helpful comments from Peter Easton (editor), two anonymous reviewers, Pietro Bianchi, Khrystyna Bochkay, Lauren Cunningham, David Hay (discussant), Mark Maffett, Linda Myers, Dhananjay Nanda, Sundaresh Ramnath, Tyler Shumway (discussant), Eric Weisbrod, Michael Willenborg, Peter Wysocki, PCAOB staff, and seminar participants at the AAA Auditing Section Midyear Meeting 2016, University of Chicago, Florida Accounting Symposium 2016, University of Miami, University of Minnesota Empirical Accounting Research Conference 2016, Nanyang Technical University, National University of Singapore, PCAOB/JAR Conference on Auditing and Capital Markets 2017, and University of Toronto. The authors also thank Yamin Hao and Taylor Wiesen for their excellent research assistance. The authors thank the Risk Management Institute at the National University of Singapore for providing the CRI default data.

Data availability Data is obtainable from the sources described in the text and is available upon request.

Appendix 1

Variable Definitions

Variable		*Definition*	*Source*
$DEFAULT_{i,t+1}$	=	indicator variable equal to one if company *i* has a default event from the filing date of the annual report for fiscal year *t* to the subsequent fiscal year end in *t* + *1* and 0 otherwise.	CRI, CRSP, Compustat, Bankruptcy.com, UCLA-LoPucki Bankruptcy Research Database
$GCO_{i,t}$	=	indicator variable equal to one if a company *i* has a GCO at fiscal year-end *t* and 0 otherwise.	Audit Analytics and text mined SEC 10-K Filings
$CONT\_PD_{i,t}$	=	Estimated PD for the subsequent 12 months for firm *i* calculated at the closest month end before the signature date of the audit opinion for fiscal year *t*.	CRI database
$REL\_MKTCP_{i,t}$	=	Log (Firm Market Capitalization [PRC*SHROUT] / Index Market Capitalization [TOTVAL]).	CRSP
$EX\_RET_{i,t}$	=	Cumulative Firm Returns – Cumulative Market Returns over the 12 months leading up to the filing date for year *t*.	CRSP
$SIGMA_{i,t}$	=	Standard deviation of the residuals from a monthly return model of firm returns on market returns for the 12 months leading up to the filing date for year *t*.	CRSP
$ROA_{i,t}$	=	Net Income [NI] / Total Assets [AT].	Compustat North America
$LEVERAGE_{i,t}$	=	Total Liabilities [LT] / Total Assets [AT].	Compustat North America

Variable		Definition	Source
$WCAP_{i,t}$	=	(Current Assets [ACT] – Current Liabilities [LCT]) / Total Assets [AT].	Compustat North America
$CURRENT_{i,t}$	=	Current Assets [ACT] / Current Liabilities [LCT].	Compustat North America
$CASH_{i,t}$	=	Cash and Cash Equivalents[CHE] / Total Assets [AT].	Compustat North America
$CFO_{i,t}$	=	Cash Flow from Operating Activities [OANCF] / Total Assets [AT].	Compustat North America
$SIZE_{i,t}$	=	Log(Total Assets[AT]).	Compustat North America
$NEGEQUITY_{i,t}$	=	Indicator variable equal to one if Total Liabilities [LT] exceed Total Assets [AT] and 0 otherwise.	Compustat North America
$BIGN_{i,t}$	=	Indicator variable equal to one if the company has a Big N auditor and 0 otherwise.	Audit Analytics and Compustat (pre-2000)
$RATING_{i,t}$	=	Categorical variable representing the credit rating [CR_LVL] of the firm at fiscal year-end t, ranging from 1 (AAA) to 10 (D).	Compustat North America
$DOWN_{i,t}$	=	Indicator variable equal to one if the credit rating [CR_LVL] for firm *i* has been downgraded during the 12 months of fiscal year *t*.	Compustat North America
$DOWN\_LVL_{i,t}$	=	Count variable indicating the number of levels in credit ratings [CR_LVL] that firm *i* has been downgraded during the 12 months of fiscal year *t*.	Compustat North America
$INVST\_CH_{i,t}$	=	Indicator variable equal to one if the credit rating [CR_LVL] for firm *i* has been downgraded from investment grade to non-investment grade during the 12 months of fiscal year *t*.	Compustat North America

Relevant Compustat and CRSP variable names in brackets

References

Akers, M., Maher, M., & Giacomino, D. (2003). Going-concern opinions: Broadening the expectations gap. *The CPA Journal, 73*(October), 38–42.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23*(4), 589–609.

Altman, E. I. (1982). Accounting implications of failure prediction models. *Journal of Accounting, Auditing and Finance, 6*(1), 4–19.

Altman, E. I., & McGough, T. P. (1974). Evaluation of a company as a going concern. *Journal of Accountancy, 138*(6), 50–57.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research, 4*, 71–111.

Beaver, W. H. (1968). Market prices, financial ratios, and the prediction of failure. *Journal of Accounting Research, 6*(2), 179–192.

Bharath, S. T., & Shumway, T. (2008). Forecasting default with the Merton distance to default model. *Review of Financial Studies, 21*(3), 1339–1369.

Bills, K. L., Swanquist, Q. T., & Whited, R. L. (2015). Growing pains: Audit quality and office growth. *Contemporary Accounting Research, 33*(1), 288–313.

Blay, A., Geiger, M., & North, D. (2011). The auditor's going-concern opinion as a communication of risk. *Auditing: A Journal of Practice and Theory, 30*(2), 77–102.

Blay, A., Moon, J., & Paterson, J. (2016). There's no place like home: The influence of home-state going-concern reporting rates on going-concern opinion propensity and accuracy. *Auditing: A Journal of Practice & Theory, 35*(2), 23–51.

Butler, M., Leone, A. J., & Willenborg, M. (2004). An empirical analysis of auditor reporting and its association with abnormal accruals. *Journal of Accounting & Economics, 37*(2), 139–165.

Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. *Journal of Finance, 63*(6), 2899–2939.

Carson, E., Fargher, N., Geiger, M., Lennox, C., Raghunandan, K., & Willekens, M. (2013). Audit reporting for going-concern uncertainty: A research synthesis. *Auditing: A Journal of Theory and Practice*, 32 (supplement 1), 353–384.

Cassell, C. A., Hansen, J. C., Myers, L. A., & Seidel, T. A. (2017). Does the timing of auditor changes affect audit quality? Evidence from the initial year of the audit engagement. Journal of Accounting, Auditing & Finance, 1–27.

Chasan, E. (2009). Going-concern warnings increasingly less accurate. Reuters. https://www.reuters.com/article/goingconcern-study/going-concern-warnings-increasingly-less-accurate-idUSN2338896220090923. 28 Oct. 2019.

Chasan, E. (2012). Going concern opinions on life support with investors. ***The Wall Street Journal CFO Report***. https://blogs.wsj.com/cfo/2012/09/12/going-concern-opinions-on-life-support-with-investors/. Accessed 28 Oct. 2019.

Chava, S., & Jarrow, R. A. (2004). Bankruptcy prediction with industry effects. *Review of Finance, 8*(4), 537–569.

Cleves, M. A. (2002). From the help desk: Comparing areas under receiver operating characteristic curves from two or more probit or logit models. *The Stata Journal, 2*(3), 301–313.

Cook, J., & Ramadas V. (2019). When to consult precision-recall curves. Working paper. Public Company Accounting Oversight Board.

Credit Research Initiative (2013). NUS-RMI Credit research initiative technical report. Version 2013: 2013 Update 2b. Global Credit Review, 3 (1), 77–167.

DeFond, M., Raghunandan, K., & Subramanyam, K. (2002). Do non-audit service fees impair auditor independence? Evidence from going concern audit opinions. *Journal of Accounting Research, 40*(4), 1247–1274.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics, 44*(1), 837–845.

Duan, J. C., Sun, J., & Wang, T. (2012). Multiperiod corporate default prediction: A forward intensity approach. *Journal of Econometrics, 170*(1), 191–209.

Duan, J. C., & Van Laere, E. (2012). A public good approach to credit ratings –from concept to reality. *Journal of Banking & Finance, 36*(12), 3239–3247.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(1), 861–874.

Feldmann, D., & Read, W. (2010). Auditor conservatism after Enron. *Auditing: A Journal of Practice & Theory, 29*(1), 267–278.

Financial Accounting Standards Board (2014). Disclosures about an entity's ability to continue as a going concern. Accounting standards update (ASU) no. 2014–15. Norwalk, CT. https://www.fasb.org/jsp/FASB/Document_C/DocumentPage?cid=1176164329772&acceptedDisclaimer=true. Accessed 28 Oct. 2019.

Financial Reporting Council (2019). International standard on auditing (UK) 570 (revised September 2019). London, England: The Financial Reporting Council Limited. https://www.frc.org.uk/getattachment/13b19e6c-4d2c-425e-84f9-da8b6c1a19c9/ISA-UK-570-revised-September-2019-Full-Covers.pdf. Accessed 28 Oct. 2019.

Funcke, N. (2014). Credit rating changes and auditor reporting accuracy. Working paper. Erasmus University Rotterdam.

Geiger, M. A., Raghunandan, K., & Rama, D. V. (2005). Recent changes in the association between bankruptcies and prior audit opinion. *Auditing: A Journal of Practice & Theory, 24*(1), 21–35.

Geiger, M. A., & Rama, D. V. (2006). Audit firm size and going-concern reporting accuracy. *Accounting Horizons, 20*(1), 1–17.

Gerakos, J., Hahn, R., Kovrijnykh, A., & Zhou, F. (2017). Prediction versus inducement and the informational efficiency of going concern opinions. Working paper. University of Chicago and Arizona State University.

Gutierrez, E., Minutti-Meza, M., Tatum, K. W., & Vulcheva, M. (2018). Consequences of adopting an expanded auditor's report in the United Kingdom. *Review of Accounting Studies, 23*(4), 1543–1587.

Hillegeist, S. A., Keating, E. K., Cram, D. P., & Lundstedt, K. G. (2004). Assessing the probability of bankruptcy. *Review of Accounting Studies, 9*(1), 5–34.

Hopwood, W., McKeown, J. C., & Mutchler, J. F. (1989). A test of the incremental explanatory power of opinions qualified for consistency and uncertainty. *The Accounting Review, 64*(1), 28–48.

Hopwood, W., McKeown, J. C., & Mutchler, J. F. (1994). A reexamination of auditors versus model accuracy within the context of the going-concern opinion decision. *Contemporary Accounting Research, 10*(2), 409–431.

Hosmer, D., & Lemeshow, S. (2000). *Applied logistic regression* (Second ed.). Hoboken: John Wiley & Sons.

Janes, H., Longton, G., & Pepe, M. S. (2009). Accommodating covariates in ROC analysis. *The Stata Journal, 9*(1), 17–39.

Jarrow, R. A. (2009). Credit risk models. *Annual Review of Financial Economics, 1*(1), 37–68.

Johnson, S. (2009). The growing concern over "going concern." **CFO.com**. https://www.cfo.com/accounting-tax/2009/03/the-growing-concern-over-going-concern/. Accessed 28 Oct. 2019.

Kim, I., & Skinner, D. (2012). Measuring securities litigation risk. *Journal of Accounting and Economics, 53*(1–2), 290–310.

Koh, H. C. (1991). Model predictions and auditor assessments of going concern status. *Accounting and Business Research, 21*(84), 331–338.

Larcker, D., & Zakolyukina, A. (2012). Detecting deceptive discussions in conference calls. *Journal of Accounting Research, 50*(2), 495–540.

Lennox, C. (2000). Do companies successfully engage in opinion-shopping? Evidence from the UK. *Journal of Accounting and Economics, 29*(3), 321–337.

Levitan, A. S., & Knoblett, J. A. (1985). Indicators of exceptions to the going concern assumption. *Auditing: A Journal of Practice and Theory, 5*(1), 26–39.

Li, C. (2009). Does client importance affect auditor independence at the office level? Empirical evidence from going-concern opinions. *Contemporary Accounting Research, 26*(1), 201–230.

Louwers, T. J., Messina, F. M., & Richard, M. D. (1999). The auditor's going-concern disclosure as a self-fulfilling prophecy: A discrete-time survival analysis. *Decision Sciences, 30*(3), 805–824.

Mayew, W. J., Sethuraman, M., & Venkatachalam, M. (2015). MD&a disclosure and the firm's ability to continue as a going concern. *The Accounting Review, 90*(4), 1621–1651.

Menon, K., & Williams, D. (2010). Investor reaction to going concern audit reports. *The Accounting Review, 85*(6), 2075–2105.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance, 29*, 449–470.

Minutti-Meza, M. (2013). Does auditor industry specialization improve audit quality? *Journal of Accounting Research, 51*(4), 779–817.

Mutchler, J. F., Hopwood, W., & McKeown, J. M. (1997). The influence of contrary information and mitigating factors on audit opinion decisions on bankrupt companies. *Journal of Accounting Research, 35*(2), 295–310.

Myers, L. A., Schmidt, J. J., & Wilkins, M. S. (2014). An investigation of recent changes in going concern reporting decisions among big N and non-big N auditors. *Review of Quantitative Finance and Accounting, 43*(1), 155–172.

Myers, L. A., Shipman, J. E., Swanquist, Q. T., & Whited, R. L. (2018). Measuring the market response to going concern modifications: The importance of disclosure timing. *Review of Accounting Studies*, 4 23 (4), 1512–1542.

Newton, N., Persellin, J., Wang, D., & Wilkins, M. (2016). Internal control opinion shopping and audit market competition. *The Accounting Review, 91*(2), 603–623.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*(1), 109–131.

Pepe, M. S., Longton, G., & Janes, H. (2009). Estimation and comparison of receiver operating characteristic curves. *The Stata Journal, 9*(1), 1–16.

Public Company Accounting Oversight Board (2008). Staff audit practice alert no. 3 audit considerations in the current economic environment. Washington, D.C. https://pcaobus.org/Standards/QandA/12-05-2008_APA_3.pdf.

Public Company Accounting Oversight Board (2011). Staff audit practice alert no. 9 assessing and responding to risk in the current economic environment. Washington, D.C. https://pcaobus.org/Standards/QandA/12-06-2011_SAPA_9.pdf. Accessed 28 Oct. 2019.

Public Company Accounting Oversight Board (2012). Investor sub advisory group. Going concern considerations and recommendations. Washington, D.C. https://pcaobus.org/News/Events/Documents/03282012_IAGMeeting/Going_Concern_Considerations_and_Recommendations.pdf. Accessed 28 Oct. 2019.

Public Company Accounting Oversight Board (2014). Staff audit practice alert no. 13 matters related to the auditor's consideration of a company's ability to continue as a going concern. Washington, D.C. https://pcaobus.org/Standards/QandA/09222014_SAPA_13.pdf. Accessed 28 Oct. 2019.

Public Company Accounting Oversight Board (2017). Consideration of an Entity's ability to continue as a going concern. Washington, D.C. https://pcaobus.org/Standards/Auditing/Pages/AS2415.aspx. Accessed 28 Oct. 2019.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business, 74*(1), 101–124.

Vassalou, M., & Xing, Y. (2004). Default risk in equity returns. *Journal of Finance, 59*(2), 831–868.

Vanstraelen, A. (2003). Going-concern opinions, auditor switching, and the self-fulfilling prophecy effect examined in the regulatory context of Belgium. *Journal of Accounting, Auditing and Finance, 18*(2), 231–254.

Witten, I. H., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann series in data management systems, Second edition. Morgan Kaufmann.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research, 22*, 59–82.