Short Selling and Default Prediction: International Evidence

Mark Maffett*
University of Chicago Booth School of Business

Edward Owens
Simon School of Business, University of Rochester

Anand Srinivasan
National University of Singapore

March, 2013

Abstract
We examine how short selling affects the accuracy of default prediction models. Our results indicate that a dynamic multiperiod logit model accurately predicts 18 percentage points more actual occurrences of default in countries where short-selling is widely practiced. Moreover, our analyses indicate that short selling has virtually no impact on the proportion of inaccurately classified non-default observations. In addition, we find that the lower predictive accuracy in short sale constrained environments is attenuated by the introduction of put option trading. Taken together, these results provide evidence that short selling increases the informativeness of equity market indicators of financial distress without affecting the level of uninformative speculative trading. Finally, in countries that face significant short selling restrictions, our results indicate that greater availability of other, non-market-based, sources of information significantly improves default prediction accuracy, providing evidence of an enhanced role for corporate transparency in settings with significant capital market frictions.

We are grateful to the Risk Management Institute for providing the dataset on cross-country default used in this study. Srinivasan acknowledges research support from research grants from the Risk Management Institute at the National University of Singapore. Financial support from the University of Chicago Booth School of Business and the University of Rochester Simon School of Business is gratefully acknowledged. We are grateful for helpful comments received from Bill Beaver, Mike Minnis, Shiva Rajgopal, Wendy Wilson, and workshop participants at Emory University, the University of Rochester, and the 2012 FARS Mid-year Meeting.
* Corresponding author. Tel.: +17737029656; mark.maffett@chicagobooth.edu.

1. Introduction

The effect of short selling on the informativeness of equity prices is heavily debated by academics, regulators and politicians. From an efficiency perspective, theoretical research suggests that restrictions on short selling reduce the incentive to acquire private information, leading to overpricing (Miller 1977), excess volatility (Hong and Stein 2003) and a delay in the speed with which prices reflect private information (Diamond and Verrecchia 1987). A large body of empirical research supports these predictions (e.g., Jones and Lamont 2002; Bris et al. 2007; Chang et al. 2007; Saffi and Sigurdsson 2011; Boehmer and Wu 2012; Kaplan et al. 2013).

Yet, since the onset of the 2008 financial crisis, stock exchange regulators around the world have placed numerous restrictions on short selling and continue to contemplate others. Many such regulators have taken the view that short sales have the potential to generate "selling pressures or unusual volatility causing significant downward spirals [in equity prices] (ESMA 2012/715)." In an effort to limit these concerns, over the past 5 years, many EU member states have elected to place significant restrictions on short selling. For example, on November 1st, 2012 Spanish securities market regulators, with the backing of the European Securities and Market Authority (ESMA), announced their intention to introduce emergency measures to ban short sales given "certain adverse situations… exist[ing] at present that constitute a serious threat to the financial stability of, and confidence in, the financial market in Spain (ESMA 2012/715)." At the same time the ESMA issued a sweeping regulation on short selling and certain aspects of credit default swaps.[1]

---

[1] The regulation included, for example, efforts to: increase the transparency of short positions held by investors in certain EU securities; reduce risks associated with naked short selling; and to ensure that EU member states have clear powers to intervene in exceptional situations to reduce systematic risks and risks to financial stability and market confidence arising from short selling. See the ESMA's website for further details: (http://www.esma.europa.eu/page/Short-selling).

In support of this view, prior theoretical research on speculative trading demonstrates that the ability to sell a stock short can create the opportunity for predatory trading which may lead to overly large (relative to the underlying fundamentals) declines in share prices. For example, Goldstein and Grumbel (2008) shows that, if prices in the secondary market also play an important role in a firm's investment allocation decisions, this can create a feedback effect which leads an uninformed trader to sell (or sell short) the firm's stock in hopes of influencing the firm's investment decisions. The primary implication of the model is that this form of speculative trading results in a decrease in the informativeness of stock prices. Brunnermeier and Pedersen (2005) provides further support for this notion by demonstrating that, if the trading behavior of some (large) investors can be anticipated, this creates the incentive for speculative traders to sell (or short sell) the stock in anticipation of the investor's liquidation driving further, non-fundamentals-based, price declines. However, little empirical work examining this potential role for short selling yet exists.

In this paper, we contribute to this debate by examining the effect of short selling on the ability to accurately assess a firm's likelihood of default using publicly available information. The degree to which short selling is actively practiced is likely to be particularly relevant for assessing a firm's likelihood of default given the role short sales play in conveying negative information. Prior research demonstrates, both theoretically (Diamond and Verrecchia 1987) and empirically (Boehmer and Wu 2012), that constraints on short selling affect not only the average level of price informativeness, but also have a relatively larger effect on the incorporation of bad news into stock prices.[2] Given the central importance of accurate default prediction to numerous corporate external financing decisions, an understanding of how short selling affects default

[2] Constraints on short-selling can arise from numerous sources including: explicit county-level legal prohibitions, an inactive securities lending market, or the absence of exchange-traded stock options.

prediction is of first-order importance. Yet, surprisingly little research exists on the effects of short selling constraints on default prediction. The goal of our paper is to help fill this gap in the literature.

Ex ante, the impact of short selling on the predictive accuracy of default models is unclear. As discussed above, theory suggests that limitations on short selling may reduce the informativeness of equity prices (Diamond and Verrecchia 1987). An extensive body of prior research on credit risk modeling has developed a variety of methods for empirically estimating a firm's risk of default using capital markets-based sources of information such as: stock return, stock volatility, market capitalization and distance to default.[3] The common implication of this prior research is that the informativeness of these capital market-based measures should have a strong impact on the accuracy of default prediction. If short selling encourages increased efforts in price discovery and improves price efficiency, it should increase default model predictive accuracy.[4]

Alternatively, the theoretical models of Goldstein and Grumbel (2008) and Brunnermeier and Pedersen (2005), and the aforementioned views of many international securities regulators, demonstrate a potential role for short selling in uninformative speculative trading. Decreases in market value driven by speculative, rather than informed trading, make prices less informative with respect to a firm's fundamental value. If short sales facilitate such speculative trading, an active short selling market may have the effect of making some relatively fiscally sound firms

[3] See for example: Beaver 1966; Altman 1968; Ohlson 1980; Shumway 2001; Chava and Jarrow 2004; Hillegeist et al. 2004; Duffie et al. 2007; Barath and Shumway 2008; Duan et al. 2012. Equity market-based characteristics are also a critical input in practical methods of default prediction, such as the structural model of default developed by Merton (1974) and its industry application (the widely used KMV model).

[4] It is important to note that the theory on short selling constraints does not suggest that a given piece of negative information is *never* incorporated into price, but rather that it is incorporated with a delay. However, f news about the firm arrives on an ongoing basis throughout the year, we expect that this leads to constant differences in the average informativeness of equity prices across regimes where short selling is and is not practiced.

appear to have a high likelihood of default, thereby reducing the ability to make meaningful distinctions across firms on the basis of solvency.

An investigation of default model predictive accuracy offers a novel method of assessing the impact of short selling on the informativeness of equity prices. The accuracy of a default prediction model can be expressed both in terms of its sensitivity (i.e., how well it identifies actual incidences of default) as well as its specificity (i.e., how well it identifies actual incidences of non-default). These two fundamental aspects of predictive accuracy speak directly to the debate on the merits of short-selling. If ability to short sell increases incentives to acquire and trade based upon private information, as suggested by Diamond and Verrecchia (1987), short selling will increases the informativeness of equity prices and potentially the sensitivity of default prediction models. On the other hand, if the ability to short sell increases speculative non-information-based trading, as suggested by Goldstein and Grumbel (2008), Brunnermeier and Pedersen (2005), and the potential for downward spirals in prices as suggested by numerous securities market regulators, it will decrease the informativeness of equity prices and potentially the specificity of default prediction models.

We use a cross-country database of defaults covering 28 countries over a period of 20 years to directly investigate how short selling affects default prediction. The primary advantage of this international sample is that it provides significantly more cross-country heterogeneity in the practice of short selling than could be observed in any single country. We employ as our primary measure of this variation a country-level indicator of the practice of short selling constructed from the dataset in Bris et al. (2007). To corroborate the inferences drawn from the Bris et al. (2007) measure, we use an alternative proxy for short selling restrictions based on the incorporation of negative firm specific information into stock price.

To estimate the likelihood of default, we employ a dynamic multiperiod logit model that includes both market (relative size, prior return and return volatility) and accounting (return-on-assets and leverage) inputs following the reduced-form approach suggested in Shumway (2001) as well as several additional variables suggested by the structural approach of Duffie et al. (2007). Given that there is virtually no extant research on cross-country heterogeneity in default prediction, the existence of which is a necessary condition for our empirical analysis, we begin by descriptively documenting the cross-country performance of our default model using both a pooled global sample as well as separate estimation for each of our sample countries. Not surprisingly, we document substantial cross-country heterogeneity in the predictive accuracy of our default model as captured by both the percentage of actual defaults in the highest deciles of predicted default probability as well as the area under a receiver operating characteristic curve (hereafter, and "ROC" curve).

We next investigate the extent to which the observed cross-country heterogeneity in default prediction accuracy can be explained by variation in the practice of short selling. If the ability to short sell increases the incentive to acquire and trade based on private information, we predict lower default prediction accuracy in countries that restrict short selling. Our empirical results confirm this prediction. For those firms not facing significant restrictions, we find that our default prediction model accurately predicts 18 percentage points more actual occurrences default. The difference in the areas under ROC curves across partitions based on short selling constraints is nearly 10%. Inferences are similar if we instead use the extent to which negative firm specific information is incorporated into stock price as an alternative measure of short selling constraints.

Although the ability to sell a stock short may increase the extent to which negative private information is incorporated into price, if short selling also increases the prevalence of predatory speculative trading as suggested by the models of Brunnermeir and Pedersen (2005) and Goldstein and Grumbel (2008), it is possible that more, relatively financially healthy, firms may be incorrectly classified as highly likely to default.[5] We next investigate the possibility that, despite the improved accuracy in identifying actual incidences of default, short selling may also encourage non-informative speculative trading, leading to an increase in the proportion of non-default observations being inaccurately classified as highly likely to default. Our results provide little evidence that short-selling leads to more uninformative speculative trading as captured by the proportion of financially healthy firms that are identified as being highly likely to default.

Next, we conduct a series of analyses intended to more directly identify the effect of short selling constraints on default prediction by addressing the possibility that our results reflect an omitted variable correlated with country-level short selling constraints. First, we control for static differences at the country-level by including country fixed effects in our default models. Although we expect the largest extent of variation in short sale constraints to be driven by country-level aspects of the market infrastructure, investigating differences in predictive accuracy based on within-country variation in the ability to sell a stock short helps mitigate the concern that an unmodeled country-level variable may confound our inferences. Inferences including country-level fixed effects in our default model are consistent with short selling leading to increased default model predictive accuracy. As an alternative method of identifying within-country variation, rather than partitioning firms based on country-level differences in short selling, we instead partition our sample firms based on the firm-level negative

[5] Our analysis implicitly assumes that short selling has no direct effect on a firm's likelihood of default. To the extent this is not true in for our sample, we expect this to bias against finding a significant association.

informativeness measure. These results further suggest that heterogeneity in short selling explains a significant portion of variation in default model predictive accuracy, again indicating that unmodeled country-level sources of variation are unlikely to explain our findings.

As a further attempt to directly identify the effect of short selling on default prediction, we next examine differences in default prediction before and after the introduction of put options. Diamond and Verrecchia (1987) suggests exploiting time-series variation in the severity short sale constraints arising from the introduction of put options (arguably, a more cost effective means in establishing a short position) as a way of measuring the effect of short sale constraints. To achieve this identification, we exploit the fact that two of the countries in our sample identified as non-short sale practicing countries introduced put option trading during our sample period, while the remaining four non-short sale practicing countries never allowed put option trading. The results of this analysis indicate that, in those countries with significant short selling restrictions, the introduction of exchange-traded put options significantly increased the predictive accuracy of our default model. This finding provides additional evidence that restrictions on short selling, rather than the presence of some form of unmodeled heterogeneity, drive the documented differences in default prediction.

As a final attempt to assess the plausibility of our finding that short sale constraints explain a significant extent of heterogeneity in default model predictive accuracy, we examine the role of other, publicly available, sources of information in mitigating the negative effects of short selling constraints. Bushman et al. (2004) develop a framework that characterizes the availability of firm-specific information (hereafter “corporate transparency”) as a function of the three facets, corporate reporting, private information acquisition and information dissemination, which collectively produce, gather, validate and disseminate information about the firm.

Although, in an efficient capital market, prices are likely to be a summary statistic for all three sources of information, as discussed previously, in the presence of constraints on short-selling, there is likely less incentive to acquire private information, potentially making prices less informative. The informativeness of a firm's financial statements and the information gathered by the news media are likely to be less affected by constraints on short selling. Given that our default model also contains financial statement variables, it is possible that, in the presence of short sale constraints and correspondingly less informative equity prices, market participants rely more heavily the other, non-market based, variables in order to access default likelihood. Consistent with this prediction, using a measure of corporate transparency constructed following Bushman et al. (2004) as an aggregate measure of these non-market based sources of information, we document that for those firms facing short selling constraints, greater corporate transparency can serve to increase the accuracy of default prediction.

Overall, our results suggest that constraints on short selling significantly reduce default model predictive accuracy while having a minimal effect on the proportion of non-defaulting firms incorrectly identified as highly likely to default. These findings make several contributions to the existing literature. Foremost, our results shed light on the implications of short sale constraints. Consistent with the theoretical arguments of Miller (1977) and Diamond and Verrecchia (1987), an extensive empirical literature demonstrates that short sale constraints can inhibit the incorporation negative information into prices. Our findings highlight a significant decrease in default prediction accuracy as a novel consequence of the loss in informativeness created by short sale constraints. Moreover, some theoretical models of speculative trading (Goldstein and Grumbel 2008; Brunnermeir and Pedersen 2005) suggest that the ability to short sale may lead prices to over- or under-shoot their fundamental values, suggesting the possibility that widely practiced short selling may over-predict the incidence of default. Our results provide little support for this prediction. Given the ongoing debate in the European Union and elsewhere, over the pros and cons of allowing short selling, thse results are likely to be of particular interest to policy makers and regulators.

Second, ours is the first paper (of which we are aware) to document significant heterogeneity in the predictive accuracy of default models across countries. The vast majority of prior literature on default prediction has focused on model predictive accuracy largely in the context of US data or in other singe country settings.[6] Prior U.S.-based literature has also given some consideration to the relative importance of accounting and market-based sources of default risk information, with the general conclusion being that market based variables often render accounting variables insignificant for default prediction.[7] We document that, while market- and accounting based sources of default risk information behave as predicted in the vast majority of countries in our sample, there is significant heterogeneity in the predictive accuracy of these commonly used default models across countries, particularly with respect to the market-based predictors. Moreover, our results identify cross-country differences in the practice of short selling an important determinant of this heterogeneity.

Finally, we document that, in countries with relatively high corporate transparency, the direct inclusion of accounting information in default prediction models can help to overcome the effects of short sale constraints on predictive accuracy. Prior default prediction literature does not clearly articulate why, other than model misspecification, public financial reporting information should have any effect on default prediction incremental to market information. We shed light on this issue by identifying market frictions, such as constraints on short selling, as an explanation

[6] See for example, Jones and Lamont 2002; Bris et al. 2007; Chang et al. 2007; Saffi and Sigurdsson 2011; Boehmer and Wu 2012; Kaplan et al. 2013.
[7] See for example, Chava and Jarrow (2004), and Hillegeist et al (2004).

for the incremental explanatory power of accounting-based sources of default risk information relative to market-based sources. Moreover, our findings highlight improved default prediction accuracy in the presence of significant market frictions as an additional capital market benefit of greater corporate transparency.

The remainder of the paper proceeds as follows. In Section 2 we describe our research design; in Section 3 we provide a description of our data source and sample selection criteria; in Sections 4 and 5, we present our empirical results; in Section 6 we conclude.

2. Research Design

Addressing our primary question of interest, how short-selling constraints affect default prediction, requires that we construct empirical proxies for both the extent to which it is possible to short a given firm in a particular market as well as for assessing a firm's likelihood of default. In this section, we discuss our empirical proxies for short-selling restrictions and default likelihood.

2.1. Measures of short selling constraints

We consider two alternative measures of short selling constraints. Our first measure draws from the data in Bris et al. (2007) to construct a measure of the extent to which short-selling is practiced in a particular country. Bris et al. (2007) collect data from practitioners and regulators on the practice of short sales in 59 countries. An important aspect of the particular dataset gathered by Bris et al. (2007) is that it focuses on the prevalence of the actual practice of short selling from the perspective of investors (e.g., Morgan Stanley and Goldman Sachs), rather than the explicit regulations regarding the permissibility of short selling. This distinction is important because, as pointed out by Bris et al. (2007), although short-selling is *legally permitted* (in some form) in most countries, it is *commonly practiced* in significantly fewer. It is only with widespread practice that we expect short selling to have a meaningful impact on the informativeness of equity prices for default prediction.[8] Table 1 reports whether short selling is commonly practiced in each sample country based on the classifications in Bris et al. (2007).

Our second proxy for short selling constraints is a country-level measure of the co-movement of firm-level returns with country-specific negative market returns. The main implication of Diamond and Verrecchia (1987) is that restrictions on short selling reduce the extent to which prices reflect private firm-specific information, particularly with respect to negative news. Bris et al. (2007) hypothesize and find evidence consistent with the notion that, in the presence of short selling restrictions, less idiosyncratic risk is incorporated into prices, particularly with respect to negative news. This finding suggests that a country's downside market model r-squared (i.e., the r-squared from a regression of firm return on negative market returns) is likely to be higher in countries with significant short selling constraints.

We construct a measure of the co-movement between firm returns and negative market returns by first estimating a "downside" market model regression (e.g., Bris et al. 2007) for each firm-year in country *c* using weekly stock return:

$$r_{i,c,t} = \alpha_0 + \alpha_1 r^{-}_{m,c,t} + \varepsilon_{i,c,t}, \qquad (2)$$

where $r_{i,c,t}$ is firm *i*'s weekly return in year *t*, and $r^{-}_{m,c,t}$ is the corresponding weekly return on the market index in country *c* when it is negative (and takes the value of zero if the weekly market index return is positive). We utilize the r-squared from the estimation of Eq. (2), $R^{2-}_{i,c,t}$, as our

[8] This measure also clearly has some drawbacks, two of which are particularly relevant for our study. Foremost, any generalized (i.e., forced to take a value of 0 or 1) country-level measure of this form is necessarily requires some subjective judgment about what constitutes common practice (e.g., Bris et al. (2007) characterize Hong Kong as a country where short selling is permitted only as of 1996 even though it was allowed for a subset of stocks in 1994). Second, Bris et al. (2007) measures the practice of before the end of our sample period. Nonetheless, we believe this measure is still likely to be effective in partitioning our sample countries on the basis of the relative extent to which short-selling is practiced and that and any recent developments likely create noise and thus bias against finding significant differences across partitions formed on the basis if the Bris et al. (2007) measure.

measure of firm-year downside synchronicity, which is decreasing in the degree to which negative firm-specific information is impounded into firm *i*'s stock price (i.e., high downside synchronicity is positively correlated with short sales constraints). To construct our country-level measure, we first average $R^{2-}_{i,c,t}$ across all firms in country *c* within year *t* to obtain country-year downside synchronicity ($R^{2-}_{c,t}$) , then average the country-year measures across all years *t* to arrive at a single country-level measure of downside synchronicity (i.e., R^{2-}_{c}), which we refer to as *MktModelR2C*. When comparing default model predictive accuracies across short-sale partitions, we deem sample observations with below (above) median *MktModelR2C* to have low (high) short-sale constraints. Table 1 reports *MktModelR2C* by sample country, as well as whether a given sample country is categorized as having "Low" or "High" short-sale constraints based upon this measure.

2.2. Estimating the likelihood of default

Prior literature suggests two approaches to empirically estimating a firm's likelihood of default using publicly available information, a reduced form and a structural approach. In the reduced form modeling of defaults, it has become common practice to assess a company's likelihood of default using a multi-period logit model that includes a combination of both market- and accounting-based default risk measures [e.g., Shumway (2001); Chava and Jarrow (2004); Campbell et al. (2008); Beaver (2012)]. The structural approach, as exemplified in Duffie et al. (2007), develops a time series model of the covariates likely to affect a firm's likelihood of default and then specifies a particular parameterization of these covariates to explain actual incidences of default. The structural approach in Duffie et al. (2007) suggests the inclusion of both a variety of dynamic firm-specific and macroeconomic covariates.

We follow the reduced form multi-period logit approach of Shumway (2001), supplemented by those variables suggested by the structural approach in Duffie et al. (2007).[9] Specifically, we estimate the following logistic regression where the probability of default for firm *i* in year *t* is estimated:

$$P(DEFAULT_{i,t+1} = 1) = \frac{1}{1+e^{-z}}, \tag{1}$$

$$\begin{aligned} z = {} & \alpha_0 + \alpha_1 LERET_{i,t} + \alpha_2 LSIGMA_{i,t} + \alpha_3 LRSIZE_{i,t} + \alpha_4 ROA_{i,t} + \alpha_5 LTA_{i,t} \\ & + \alpha_6 DTD_{i,t} + \alpha_7 RFRATE1YR_{i,t}. \end{aligned}$$

DEFAULT is an indicator variable that equals one if the firm defaults in year *t*+1 and equals 0 otherwise. The firm-year predictive variables are measured at the end of year *t*, where year *t* is the most recently available data prior to *t+1*. Under this dynamic methodology, Shumway (2001) recommends using a multiperiod logit model, where each year a firm survives is included as a non-failure observation, and default observations are included as a failure observation only in the year of failure. Accordingly, *DEFAULT* = 0 includes all firm-year observations for firms that never default, as well as all firm-year observations for defaulted firms in years prior to the year immediately preceding their default.[10] We delete all firm-years of data for defaulted firms after their default year. Standard errors are clustered at the firm level to account for the lack of independence between firm-year observations.

LERET, *LSIGMA*, and *LRSIZE* are market-based predictive variables, where *LERET* is lagged twelve-month cumulative abnormal stock return (i.e., firm return less the market index

---

[9] Our objective is not to identify the "best" empirical default prediction model, but rather to explore cross-sectional variation in the performance of default prediction models in general across countries. To this end, we employ what we believe is a widely accepted empirical default prediction model in the academic literature.

[10] We follow the vast majority of literature (e.g., Shumway 2001, Chava and Jarrow 2004, Campbell et al. 2008) and assume that any censoring (i.e., firms that leave the sample for reasons other than default) is non-informative. While this may not be strictly true, for censoring to cause an issue in our study it would need to be the case that the informativeness of censoring differs across our sample partitions, which is less plausible. Moreover, Duffie et al. (2007) show that the effect of censoring is minimal on one-year ahead default prediction, which is our focus. Nonetheless, we repeat our analyses after excluding all firms that are censored, and inferences are unaffected.

return), *LSIGMA* is lagged twelve-month return volatility, and *LRSIZE* is the logarithm of a firm's market capitalization relative to the aggregate sample market capitalization. We use market data as of the end of the month following the month of financial statement data availability.[11] This allows the market time to incorporate the financial statement data. *ROA* and *LTA* are accounting-based predictive variables (e.g., Shumway 2001; Beaver et al. 2005), where *ROA* is a measure of profitability (return-on-assets), and *LTA* is a measure of leverage (total liabilities divided by total assets). *DTD*, distance-to-default, is a default risk measure generated from the theoretical underpinnings of the Black-Scholes-Merton structural model of default probabilities using both accounting and market data. *RFRATE1YR* is the interest rate in percent of the firm's country's one-year government treasury security.

We initially estimate three versions of Eq. (1). First, we estimate a specification that omits the Duffie et al. (2007) variables (i.e., *DTD* and *RFRATE1YR*) and thus closely resembles the Shumway (2001) and Beaver et al. (2005, 2012) specifications. Next, we estimate a specification that omits *LSIGMA*, *LRSIZE*, *ROA*, and *LTA*, thus closely matching the model used in Duffie et al. (2007).[12] Finally, we estimate the full Eq. (1) specification (i.e., the "combined" model), which essentially uses all of the predictors from both the Shumway (2001) approach and the Duffie et al. (2007) approach. All variables are further discussed in the Appendix.

2.3. Assessing model predictive accuracy

The theoretical literature on short selling suggests two aspects of default model predictive accuracy are likely to be of interest. First, if as suggested by the prior theoretical literature such as Diamond and Verrecchia (1987), the active practice of short selling increases privately

[11] For example, if financial statement data are available 04/17/2004, we use market data as of 05/31/2004.
[12] Note that when estimating this second specification, we separately include firm cumulative return (*LRET*) and the return on the market index (*INDEXRET*) in place of *LERET*, to more closely match the specification used in Duffie et al. (2007).

informed trading, we expect the proportion of actual defaults correctly identified by our prediction model (i.e., true positives) to be greater in countries without short sale constraints. Second, even if a higher level of informed trading increases the extent to which stock prices efficiently convey information about actual defaults, theoretical literature such as Brunnermeier and Pedersen (2005) suggests that it can also increase predatory speculative trading and cause downward pressure on prices. Such downward price pressure could potentially lead to a greater incidence of non-default observations being identified as likely default observations (i.e., a higher incidence of false positives).

To compare the predictive accuracy of our models, we adopt two related approaches suggested by the prior literature, that incorporate each of the aforementioned aspects of default model predictive accuracy. First, we measure a model's predictive accuracy as the fraction of actual sample defaults (or non-defaults) with a predicted probability of default falling in the top three predicted probability of default deciles for all sample firm-years (e.g., Beaver et al. 2005; 2012).[13] That is, after using a given model to estimate predicted default probabilities for all sample firm-years, we rank the predicted default probabilities into deciles and note the decile into which each sample observation falls. We then construct the variable $ACCUR_m$ (i.e, predictive accuracy of model m) as the cumulative percentage of sample default (non-default) observations that fall in the top three deciles when default probabilities are estimated using model m. When we examine differences both in $ACCUR_m$ across sample partitions and in *ACCUR* across different models within the same partition, we test the statistical significance of these differences using a Monte Carlo randomization methodology.

[13] Although we discuss in the text results only for the first three deciles, we present results for all ten deciles.

Second, we assess model predictive accuracy with the area under receiver operating characteristic curves (e.g., Chava and Jarrow 2004).[14] ROC curves are cumulative probability curves across the entire sample population (ordered by estimated default probability) that simultaneously consider how a model performs in terms of both sensitivity (i.e., how accurately the model classifies actual incidences of default) and specificity (i.e., how accurately the model classifies non-default observations), where the area under the curve is increasing in model accuracy. The area under a ROC curve is generally expressed relative to the unit square area, where a value of 0.5 reflects a random model with no predictive ability, and a value of 1.0 indicates perfect predictive ability. The total area under the ROC curve reflects the tradeoff between increasing sensitivity and decreasing specificity (i.e., the tradeoff between how well our prediction model identifies actual default versus suggesting a similarly high likelihood of default for non-default observations).

3. Data and sample selection

3.1. NUS Credit Research Initiative

Our firm-level default data comes from the Risk Management Institute (RMI) at National University of Singapore. In July 2009, RMI launched the non-profit Credit Research Initiative (CRI) to promote independent transparent research in the credit risk arena.[15] The foundation of the CRI is a database of over 53,000 listed firms in 46 countries across the Asian-Pacific, North American, Western European and Latin American regions. The proprietary database that underlies this output, includes extensive panel data on firm stock price, financial statement data,

[14] Moody's uses a similar tool called Cumulative Accuracy Profiles to assess model performance (Sobehart et al. 2001).
[15] For more information on this initiative, refer to http://rmicri.org/home/.

and events of defaults from 1990 to the present, categorized by default class. It is this underlying proprietary database from which we draw our sample data.[16]

The CRI research team collects default events from numerous sources, including Bloomberg, Compustat, CRSP, Moody's, exchange web sites and media outlets. Because definitions of credit default can vary across national jurisdictions and between data sources, CRI continuously attempts to normalize to a common set of default definitions. In the version of the dataset we use, default events recognized by CRI include "1) bankruptcy filing, receivership, administration, liquidation, or any other legal impasse to the timely settlement of interest and/or principal payments; 2) a missed or delayed payment of interest and/or principal, excluding delayed payments made within a grace period; 3) debt restructuring/distressed exchange." Delistings or "other exits" are not considered as defaults initially, but are reclassified as defaults if a firm experiences a default within one year of the delisting. Technical defaults (i.e., covenant violations) are not included in the definition of default. In addition to these general categories, CRI separately examines cases that require special attention to determine whether a default event has actually occurred.

3.2. Sample selection and descriptive statistics

We begin with all default observations in the CRI database, which gives us an initial default sample of 12,771 default observations. However, the CRI default dataset provides a separate observation for each instrument that is defaulted upon by a given occurrence of firm default (e.g., if a firm has two loans outstanding at the time of bankruptcy filing, the bankruptcy filing would generate two observations in the default dataset). Accordingly, we delete all such "duplicate" observations, leaving 8,258 distinct firm-default observations. Because of the structure of our empirical tests, in the case where a given firm has multiple defaults in the

[16] This data source has been used in prior published work by Duan et al. 2012.

database we retain only the first default occurrence for a given firm, which reduces our default sample to 5,562 firm-level observations. We next delete banks and utilities, leaving 5,416 defaults.

Our analyses require both accounting data and market data. Our primary source for accounting data is the CRI 'financial statements' dataset. Specifically, we require measures of return on assets (*ROA*) and leverage (*LTA*). We attempt to supplement any missing CRI financial statement data with Worldscope data, where we merge the CRI data with Worldscope based on ISIN.[17] After merging the financial statement data into the default sample, the 5,416 defaults yield 31,703 firm-year observations (i.e., 5,416 default-year observations and 26,287 non-default-year observations). The remaining 533,422 firm-year observations in the financial statements data set provide the potential pool of additional non-default-year observations (i.e., all firm-year observations for firms that never defaulted), yielding a total sample of 565,125 firm-year observations.

As discussed above, we utilize three market-based measures in our prediction models. For excess firm return (*LERET*) and return volatility (*LSIGMA*), our primary data source is the CRI 'pd' dataset, which contains data on closing monthly stock price. For the third market variable, *LRSIZE*, our primary data source is Datastream, because the use of Datastream allows us to directly obtain market capitalization in a common currency (i.e., U.S. dollars) across all sample observations, which obviates the need to engage in currency conversion on the CRI price data. Distance to default (*DTD*) and the one-year government risk free rate (*RFRATE1YR*) are likewise obtained directly from the CRI database.

[17] Supplementation with data from Worldscope/Datastream adds very few additional observations to the sample, which provides comfort that the CRI data are fairly comprehensive.

We delete all observations with missing values for any of our predictive variables, and likewise delete observations where *LSIGMA* equals zero (i.e., firms with no price change over the prior twelve months), as well as all observations from countries with no defaults in the dataset. These deletions result in a final analysis sample of 335,231 firm-year observations comprised of 2,153 default-year observations and 333,078 non-default-year observations from fiscal years 1989 through 2012. The defaults that underlie the default-year observations span the years 1991 through 2012. Finally, we Winsorize the financial statement variables at the upper and lower 2.5%, and Winsorize *LERET* and *LSIGMA* at the upper 2.5% only, because these variables have natural lower bounds.

Table 1 presents the total number of sample observations and number of sample default observations by country. Clearly, the United States is dominant in the sample, both in total number of observations (24.94%) and in default observations (43.33%).[18] Asia-Pacific nations are also well represented in the sample. Figure 1 presents our sample default frequency by year. As expected, our sample exhibits a pronounced spike in default frequency in the years surrounding 2000 and 2008, which roughly coincide with the aftermath of the Asian financial crisis, the 2001 recession and the 2008 financial crisis.

Table 2 presents aggregate sample descriptive statistics for the variables we use in our default prediction models. Although we use a distinctly different sample and time period, the overall distributional characteristics of our accounting and market-based predictor variables appears to be generally similar to those reported in prior studies (e.g., Beaver et al. 2012).

---

[18] In untabulated robustness tests, we find that our key inferences are robust to exclusion of the United States from the analysis.

4. International default model estimation

We begin our analysis by estimating default prediction models (i.e., Eq. 1) using our global pooled sample, with results presented in Table 3. Basic relations revealed by the pooled estimation reported in Panel A are generally consistent with prior literature that has estimated similar models using U.S.-only data (e.g., Shumway 2001). Specifically, focusing on column (3) for discussion, firms that have higher lagged abnormal stock returns (*LERET*), a larger relative size (*LRSIZE*), a greater return-on-assets (*ROA*), and a larger distance to default (*DTD*) are less likely to default, whereas firms with higher return volatility (*LSIGMA*) and greater leverage (*LTA*) are more likely to default.

Panel B of Table 3 presents the number of default observations falling in each decile of predicted default probability, where each decile is computed from the combined default and non-default firm-year observations, and is ranked in descending order (i.e., decile 1 has the highest predicted default probability). For example, in the combined model, 59.03% of the default-year observations fall in the highest predicted default probability decile, whereas only 0.70% (i.e., 100%-99.30%) of default firms appear in the lowest predicted default probability decile. As described earlier, our key focus in terms of model predictive accuracy (*ACCUR*) is the cumulative percentage of default firms in the highest three deciles (e.g., Beaver et al. 2005). Looking at this statistic in Panel B reveals that the combined model predictive accuracy ($ACCUR_C$) is 80.59%, which is similar in magnitude to the comparable statistic in Beaver et al. (2005). We note that the predictive accuracy of the combined model is only slightly higher than that of model 1 in column (1) (80.03%), which suggests that the addition of distance to default and the risk-free rate contributes little incremental predictive power over and above the core market-based and accounting-based variables used by Shumway (2001), Beaver et al. (2005, 2012), among other studies. Nonetheless, for the remainder of the study we work with the combined model specification.

Table 4 presents results from separate estimation of Eq. (1) for each sample country having greater than ten default observations in the sample. Generally speaking, the relations between the model predictive variables and default incidence are consistent in sign and significance across the country-level estimations. For example, four of the predictive variables never enter a country-level regression with a statistically significant sign that is opposite to the predicted sign (note that we do not have a prediction for the sign of *RFRATE1YR*). The exceptions are *LRSIZE* (statistically significant positive coefficient in 13.6% of sample countries) and LERET (statistically significant positive coefficient in 4.5% of sample countries).

In general, the mean accuracy (*ACCUR*) across countries is generally consistent with results from our pooled estimation in Table 3 (86.65%). However, we note wide variation in predictive accuracy among countries, ranging from 63% in the Philippines to 100% in Denmark, Italy, and Norway.[19]

5. Effect of short sale constraints

We now turn to the primary focus of our analysis; how short selling affects default prediction accuracy. Again, we examine two specific conjectures. First, we examine whether short-selling constraints prevent default risk-relevant information from being incorporated into market variables, thus lowering the model's predictive power with respect to identifying incidences of default (i.e., the classification of true positives). Second, we investigate the related possibility that short sale based trading may be speculative rather than informed, leading a to a

[19] We note that Denmark, Italy, and Norway have relatively few sample defaults, which could explain their high predictive accuracy.

larger proportion of non-default observations being classified in the highest default likelihood categories (i.e., the classification of false positives).

5.2. Primary empirical results

To examine the effects of short selling on default model predictive ability, we estimate Eq. (1) separately for countries with and without significant short-selling restrictions, where Panel A of Table 5 presents associated regression statistics.

Columns (1)-(3) of Panel A report results from partitioning the sample based on the country-level short selling practice measure from Bris et al. (2007). Specifically, we identify those countries where short selling is not practiced and where short selling is practiced, and estimate Eq. (1) separately for each group of countries. As reported in column (3), there are indeed significant differences in the predictive variable coefficient magnitudes across partitions. In particular, the magnitude of the coefficients on the market-based variables (e.g., *LERET*, *LSIGMA*, *DTD*) are smaller in the high short sale constraint countries, consistent with short sale constraints impeding the predictive ability of market-based variables. Interestingly, the coefficient magnitudes on the accounting variables (e.g. *ROA*, *LTA*) are larger in short sale constraint countries, which suggests that direct accounting inputs may take on increased predictive importance in the presence of short sale constraints. A look at the comparative r-squareds of the models in columns (1) versus (2) suggests that the default prediction model has a better overall fit in countries with relatively low short sale constraints. However, we leave formal assessment of the relative overall predictive abilities to the analysis in Panels B, C and D.

Columns (4)-(6) repeat the analysis using country level downside market model r-squared as a proxy for short sales constraints (*MktModelR2C*), where high (low) short sale constraint countries have below (above) median *MktModelR2C*. Similar to the results in columns (1)-(3), the magnitude of coefficients on market-based (accounting-based) predictors are smaller (larger) in countries with relatively high short sale constraints. In most cases the short-sales practice-based partitioning in columns (1)-(3) generates larger differences in magnitudes than does partitioning on *MktModelR2C*, along with a relatively larger difference in overall model fit (i.e., 0.195-0.078 vs. 0.190-0.111). These patterns are consistent with the Bris et al. (2007) practice-based measure providing a relatively sharper indicator of constraints that affect the ability of market-based variables to predict default.

In Panel B of Table 5, we move to a formal test of differences in overall model predictive ability across short sale constraint partitions. As reported in Panel B, using either measure of short sale constraints, overall model predictive accuracy is statistically significantly larger in the low constraint partition, where we assess significance of the differences using Monte Carlo randomization tests. For example, using the partitioning based on the Bris et al. (2007) measure of short sales constraints, the model classifies 86.36% (68.61%) of the sample default observations within the top three predicted default probability sample deciles.

In Panel C of Table 5, we examine differences across short-sale partitions in the rate of false positives (i.e., the proportion of incorrectly classified non-default sample observations). Our first observation is that the non-default observations are essentially uniformly distributed across predicted default probability deciles across both partitions using either measure of short-sales constraints. More importantly, these patterns result in an insignificant difference across short-sale partitions in the proportion of non-default observations classified in the top three predicted default probability deciles. This provides little evidence that short selling increases the false positive rate in default prediction.

Panel D assesses differences in model predictive abilities across partitions using areas under receiver operating characteristic curves rather than the decile classification methodology. Areas under ROC curves simultaneously incorporate the differences in sensitivity and specificity that are separately reflected in Panels B and C. Figure 2 presents the Panel D results graphically. Again, inferences are consistent, in that using either measure of short sale constraints, overall model predictive ability is significantly larger in the low constraint partitions. We once again note that the Bris et al. (2007) measure of short-selling practice appears to be a more severe friction with respect to default predictive accuracy relative to the more general measure of short sales constraints, which might also capture other frictions.

In Table 6, we present the results of several robustness tests concerning predictive accuracy differences across short sale constraint regimes. For brevity, we present predictive accuracy differences using only areas under receiver operating characteristic curves. Panel A presents a specification of Eq. (1) that includes industry fixed effects, to account for potential differences across industries in the performance of the default prediction model. As reported, the predictive accuracy difference across short sales partitions remains strongly statistically significant. Similarly, Panel B presents results from estimating Eq. (1) with the inclusion of country fixed effects to account for differences in the performance of the default prediction model across countries other than from differences in short sale constraints. Again, our inferences are unaltered, although the magnitude of the differences in predictive accuracy across constraint partitions is mitigated (ROC curve area difference of 0.0351 vs. 0.0960 in the model without country fixed effects, as in Panel C of Table 5).

Finally, Panel C presents predictive accuracy differences across short sale constraint partitions, but rather than using a country-level constraint measure, we use a firm-level constraint measure. Specifically, we partition the sample based on the median value of *MktModelR2F*, the firm-specific downside market-model r-squared from the estimation of Eq. (2) (i.e., $R^{2-}_{i,c}$, the average of $R^{2-}_{i,c,t}$ for firm *i* across all years *t*). Although the results are naturally weaker than when using the country-level partitions, there is indeed a statistically significant difference in predictive accuracy across firm-level constraint partitions (ROC curve area difference of 0.02, p-value 0.05). This test provides a very strict test of the effect of short sale constraints, and reveals that *within-country*, firm-specific short sale constraints diminish default model predictive accuracy.

5.3. Constraint relief - put option introduction in the presence of short sale restrictions

Extant literature points out that put options can be used to mimic a short position in a stock, i.e., the existence of put options mitigates the friction caused by short sales constraints (e.g., Bris et al. 2007). In this section, we exploit this dynamic to enhance the identification of the effect of short sales on model predictive accuracy. In our sample, there exist two countries which do not allow short sales, but which introduced the trading of put options during the sample period. In particular, Malaysia and South Korea introduced put options in December 2000 and January 2002, respectively (Charoenrook and Daouk 2005). If our primary findings are indeed associated with short sales constraints as we infer, we expect to see the negative effects of short sales constraints on model predictive accuracy mitigated after the introduction of put options in these two countries.

To test this dynamic, we estimate Eq. (1) for Malaysia and South Korea separately for the pre-put option period (*POSTPUT* = 0) and post-put option period (*POSTPUT* = 1), and report predictive accuracy statistics based on ROC curve areas in Table 7. As expected, the performance of the prediction models significantly improves in the *POSTPUT* = 1 period (ROC curve area difference of 0.0553, p-value 0.06). To provide some assurance that this result is not merely picking up a time trend, we repeat this analysis using four countries which do not allow short sales and never introduce put options (China, Phillippines, Taiwan, and Thailand) and again for the remainder of sample countries where put options are traded throughout our sample period. For this analysis, we choose June 2001 as the line of delineation for a pseudo-*POSTPUT* partitioning variable, and report the results of estimating the default prediction models in these countries in the pseudo-*POSTPUT* partitions in the second and third rows of Table 7. As reported, there are no significant differences in model predictive accuracy in these countries across the pre-versus post-June 2001 periods, which suggests that the differences reported in Panel A are indeed attributable to the put option relief of short sales frictions.

5.4. The effect of corporate transparency

Recall from our Table 5 analysis that whereas the magnitude of the coefficient estimates on market-based variables decreases in the presence of short sale constraints, in general the magnitude of the coefficient estimates on accounting variables increases in the presence of short sale constraints. This suggests that the direct inclusion of accounting information in default prediction can help to overcome the detrimental effects of short sale constraints. If this inference is correct, we would then expect to see that, in the presence of constraints, model predictive accuracy is higher when corporate transparency (and therefore the quality of reported accounting information) is relatively high. Stated differently, we expect that the ability of financial statement variables to overcome short sale constraints will be enhanced in settings with good corporate transparency.

To measure corporate transparency, we follow the framework used in Bushman et al. (2004), who consider corporate transparency to reflect the availability of firm-specific information to the public. Further, in this framework corporate transparency is comprised of three facets: corporate reporting (e.g., quality of financial statements and associated disclosures), acquisition and communication of private information (e.g., analyst activities), and the quality of information dissemination mechanisms (e.g., media penetration). To empirically implement this corporate transparency measure, we combine proxies from each of the three dimensions into one country-level numeric score. For the corporate reporting dimension, we utilize the Leuz (2010) country-level reporting cluster categorization. For the acquisition of information dimension, we utilize country-level analyst coverage based on I/B/E/S data. For the information dissemination dimension, we use country-level internet penetration as of 2005, where we obtain data from the World Bank. We first rank our sample countries on each dimension, then compute a weighted average rank across the three dimensions, where corporate reporting receives a weight of 0.6 and each of the other two dimensions receives a weight of 0.2.[20] When conducting empirical tests across partitions, we delineate "low" versus "high" corporate transparency based on the sample country median weighted-average rank. Table 1 reports both the corporate transparency rank and associated low versus high categorization by sample country.

To test the interactive effect of corporate transparency and short sale constraints on default prediction accuracy, we estimate the default prediction models across a two-by-two sample partition sort. That is, we estimate the models separately for low and high corporate transparency partitions within each of the *ShortSalesPracticed* subsamples. If the direct use of accounting information in default prediction indeed helps overcome short sale constraints, in the presence of short sale constraints (i.e., in the *ShortSalesPracticed* = 0 subsample) we expect to see better overall model predictive accuracy when transparency is high. Panel A of Table 8

[20] This weighting scheme is consistent with the Bushman et al. (2004) corporate transparency measure. Specifically, the Bushman et al. (2004) measure is constructed from factor analysis using four distinct measures for corporate reporting, and one measure each for the other two dimensions.

presents the results of this test. Indeed, in the presence of short sale constraints, model predictive accuracy is significantly higher when corporate transparency is high (ROC curve area difference of 0.0819, p-value < 0.01). Consistent with our intuition, Panel B reveals that when short sale constraints are relatively low, corporate transparency does not produce a significant difference in model predictive accuracy.

6. Conclusion

The financial crisis of 2008 has renewed the debate about the costs and benefits of short selling restrictions. Our paper contributes to this discussion by exploiting cross-country variation in short-sale constraints to examine the effect of these restrictions on default prediction. The context of default prediction allows for both an examination of the benefits, in the form of increased private information acquisition and improved price informativeness, and the costs, in the form of predatory speculative trading to be examined.

Moreover, while the factors that affect default prediction have been studied extensively in U.S. capital markets, little is known from prior literature about how empirical predictors of financial distress perform across countries and in particular how differences in countries' information environments affect market participants' abilities to assess a company's likelihood of default. In this paper, we investigate these issues directly using a broad sample of global defaults to examine cross-country differences in the predictive accuracy of a commonly used class of default prediction models.

Foremost, our results suggest substantial cross-country variation in the informativeness of capital markets information for default prediction. Subsequent tests indicate that short selling constraints are a statistically and economically significant determinant of this country-level heterogeneity. Our results indicate that, while the ability to sell a stock short increases the ability of market participants to better identify those firms most likely to default without imposing a significant cost in the form of overpredicting distress for non-defaulting firms. The introduction of put-options in several of our sample countries suggests a direct effect of short selling on default prediction. Finally, our results suggest that, for firms facing short-selling constraints, other non-market based sources of information can help to bolster predictive accuracy.

Although, we conduct a series of additional analyses, including a within-country firm-level analysis to address the possibility that a variable correlated with short selling constraints may affect our inferences, a causal interpretation of our results should be made cautiously.

Appendix
Variable definitions

Subscripts *i* and *t* refer to a particular firm and fiscal year, respectively. Subscript *c* refers to a country, and subscript *m* refers to a particular default prediction model (market-only, accounting-only, or combined).

$ACCUR_m$	The predictive accuracy of default prediction model *m*, measured as the cumulative percentage of default firm observations in the highest three predicted default probability deciles, where deciles are computed from combined default and non-default firm-year observations.
$DEFAULT_{i,t}$	An indicator variable that equals one if firm *i* has a default event in year *t*, and equals zero otherwise, where default events are identified from the RMI data from National University of Singapore.
$DTD_{i,t}$	Firm *i*'s distance to default; from the NUS RMI data.
$INDEXRET_{i,t}$	Twelve month cumulative return on the market index for firm *i*; calculated from the NUS RMI price data file.
$LERET_{i,t}$	Twelve month cumulative excess stock return (over the market index return) for firm *i* ending in the month following firm *i*'s financial statement availability for fiscal year *t*; calculated from the NUS RMI price data file.
$LRET_{i,t}$	Twelve month cumulative stock return for firm *i* ending in the month following firm *i*'s financial statement availability for fiscal year *t*; calculated from the NUS RMI price data file.
$LRSIZE_{i,t}$	The natural logarithm of firm *i*'s relative size, computed at the end of the month following firm *i*'s fiscal year *t* financial statement data availability; relative size is computed as firm *i*'s stock market capitalization (in U.S. dollars) divided by the aggregate sample market capitalization (in U.S. dollars), where stock market capitalization is obtained from Datastream.
$LSIGMA_{i,t}$	Standard deviation of firm *i's* monthly stock return for the twelve months ending in the month following firm *i*'s financial statement availability for fiscal year *t*; calculated from the NUS RMI price data file.
$LTA_{i,t}$	Leverage ratio for firm *i* in year *t*; calculated from the RMI data as total liabilities divided by total assets.
$MktModelR2C_c$	The country-level downside r-squared from aggregation of firm-year market model regressions. Specifically, we estimate a market-model regression using weekly returns for each firm *i* in country *c* in year *t* regressed on negative weekly returns on the market index for country *c* and obtain the r-squared from each. Then, we average the r-squareds first by year within each country, then average these country-year r-squareds by country across years.

MktModelR2F$_i$	The firm-level downside r-squared from aggregation of firm-year market model regressions. Specifically, we estimate a market-model regression using weekly returns for each firm *i* in country *c* in year *t* regressed on negative weekly returns on the market index for country *c* and obtain the r-squared from each. Then, we average the r-squareds for firm i across all sample years.
POSTPUT$_{i,t}$	For firms in Malaysia and South Korea, an indicator variable that equals one if firm *i*'s year *t* observation is after the introduction of put options in firm *i*'s country, and equals zero otherwise; put options were introduced in Malaysia on 12/01/2000, and introduced in South Korea on 01/28/2002.
PSEUDOPOSTPUT$_{i,t}$	An indicator variable applicable to all countries that either had put options traded throughout our sample period, or never had put options traded during our sample period (i.e., all sample countries other than Malaysia and South Korea); equals one if firm *i*'s year *t* observation is after 06/01/2001, and equals zero otherwise.
RFRATE1YR$_{i,t}$	The interest rate on a one-year government debt security in firm *i*'s country; from the NUS RMI data.
ROA$_{i,t}$	Return on assets for firm *i* in year *t*; calculated from the RMI data as net income divided by lagged total assets.
ShortSalesPracticed$_c$	An indicator variable that equals one if short-selling is allowed/practiced in country *c* (i.e., "high" price informativeness) and equals zero otherwise (i.e., "low" price informativeness), where the coding is done based on the analysis in Bris et al. (2007).
Transparency$_c$	Country *c*'s corporate transparency, measured as the weighted average country-rank of a corporate reporting measure (based on institutional clusters from Leuz 2010), an information collection measure (based on country-level analyst following), and an information dissemination measure (based on country-level internet penetration), with weights of 0.6, 0.2 and 0.2.

REFERENCES

Altman, E. ,1968. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. Journal of Finance, 23,4, 598–609.

Beaver, W. ,1966. Financial Ratios As Predictors of Failure. Journal of Accounting Research, 4, 71–111.

Beaver, W. H., Correia, M., McNichols, M. F. ,2012. Do Differences in Financial Reporting Attributes Impair the Predictive Ability of Financial Ratios for Bankruptcy? Review of Accounting Studies.

Beaver, W., McNichols, M., Rhie, J.-W. ,2005. Have Financial Statements Become Less Informative? Evidence from the Ability of Financial Ratios to Predict Bankruptcy. Review of Accounting Studies, 10, 93–122.

Beber, A., Pagano, M. ,2011. Short-Selling Bans Around the World. Accepted for publication in Journal of Finance.

Bharath, S. T., Shumway, T. ,2008. Forecasting Default with the Merton Distance to Default Model. Review of Financial Studies, 21,3, 1339–1369.

Boehmer, E., Wu, J. ,Julie. ,2012. Short Selling and the Price Discovery Process. Review of Financial Studies.

Bris, A., Goetzmann, W., Zhu, N. ,2007. Efficiency and the Bear: Short Sales and Markets Around the World. The Journal of Finance, 3.

Brunnermeier, M., Heje Pedersen, L. ,2006. Predatory Trading. The Journal of Finance, LX,4.

Bushman, R. M., Piotroski, J. D., Smith, A. J. ,2004. What Determines Corporate Transparency? Journal of Accounting Research, 42,2, 207–252.

Campbell, J., Hilscher, J., Szilagyi, J. ,2008. In Search of Distress Risk. The Journal of Finance, LVIII,6.

Charoenrook, A., Daouk, H. ,2005. A Study of Market-Wide Short-Selling Restrictions. Working paper.

Chava, S., Jarrow, R. ,2004. Bankruptcy Prediction with Industry Effects. Review of Finance, 8,4, 537–569.

Claessens, S., Klapper, L. ,2005. Bankruptcy Around the World: Explanations of Its Relative Use. American Law and Economics Review, 7.

Diamond, D., Verrecchia, R. ,1987. Constraints on Short-Selling and Asset Price Adjustment to Private Information. Journal of Financial Economics.

Djankov, S., Hart, O., McLiesh, C., Shleifer, A. ,2008. Debt Enforcement Around the World. Journal of Political Economy, 116,6, 1105–1149.

Duan, J.-C., Sun, Jie, Wang, T. ,2012. Multiperiod Corporate Default Prediction—A forward Intensity Approach. Journal of Econometrics.

Duffie, D., Saita, L., Wang, K. ,2007. Multi-Period Corporate Default Prediction with Stochastic Covariates. Journal of Financial Economics, 83.

Durnev, A., Morck, R., Yeung, B., Zarowin, P. ,2003. Does Greater Firm-Specific Return Variation Mean More or Less Informed Stock Pricing? Journal of Accounting Research, 41,5.

Goldstein, I., Guembel, A. ,2008. Manipulation and the Allocational Role of Prices. The Review of Economic Studies, 75,1, 133–164.

Gormley, T., Gupta, N., Jha, A. ,2011. Corporate Bankruptcy and Creditor Incentives.

Hillegeist, S., Keating, E., Cram, D., Lundstedt, K. ,2004. Assessing the Probability of Bankruptcy. Review of Accounting Studies, 9, 5–34.

Hong, H., Stein, J. ,2003. Differences of Opinion, Short-Sales Constraints, and Market Crashes. Review of Financial Studies, 16,2, 487–525.

Jones, C. M., Lamont, O. A. ,2002. Short-Sale Constraints and Stock Returns. Journal of Financial Economics, 66,2, 207–239.

Kaplan, S. N., Moskowitz, T. J., Sensoy, B. A. ,2010. The Effects of Stock Lending on Security Prices: An Experiment. National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w16335

Lang, M., Lins, K. V., Maffett, M. ,2012. Transparency, Liquidity, and Valuation: International Evidence on When Transparency Matters Most. Journal of Accounting Research, 50,3, 729–774.

Leuz, C. ,2011. Different Approaches to Corporate Reporting Regulation: How Jurisdictions Differ and Why. Accounting and Business Research.

Leuz, C., D. Nanda, Wysocki, P. ,2003. Earnings Management and Investor Protection: an International Comparison. Journal of Financial Economics, 69, 505–527.

Maffett, M. ,2012. Financial Reporting Opacity and Informed Trading by International Institutional Investors. Journal of Accounting and Economics, 54,2-3, 201–220.

Merton, R. C. ,1974. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. The Journal of Finance, 29,2, 449–470.

Miller, E. M. ,1977. Risk, Uncertainty, and Divergence of Opinion. The Journal of Finance, 32,4, 1151–1168.

Morck, R., Yeung, B., Yu, W. ,2000. The Information Content of Stock Markets: Why do Emerging Markets have Synchronous Stock Price Movements? Journal of Financial Economics, 58, 215–260.

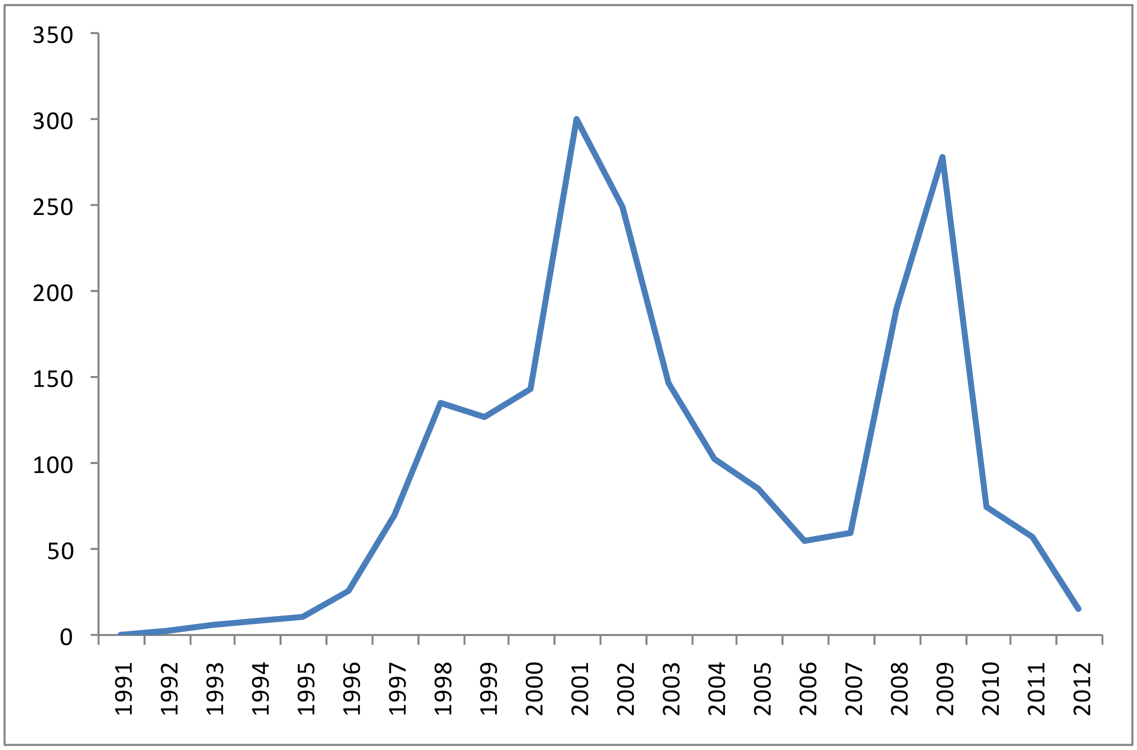
Ohlson, J. ,1980. Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research, Spring.

Saffi, P. A. C., Sigurdsson, K. ,2011. Price Efficiency and Short Selling. Review of Financial Studies, 24,3, 821.

Shumway, T. ,2001. Forecasting Bankruptcy More Accurately: A Simple Hazard Model. The Journal of Business, 74,1, 101–124. doi:10.1086/209665

Vassalou, M., Xing, Y. ,2004. Default Risk in Equity Returns. The Journal of Finance, LIX,2.

Figure 1
Sample defaults by year
Figure 1 presents calendar-year frequency of the 2,153 default observations in our sample.

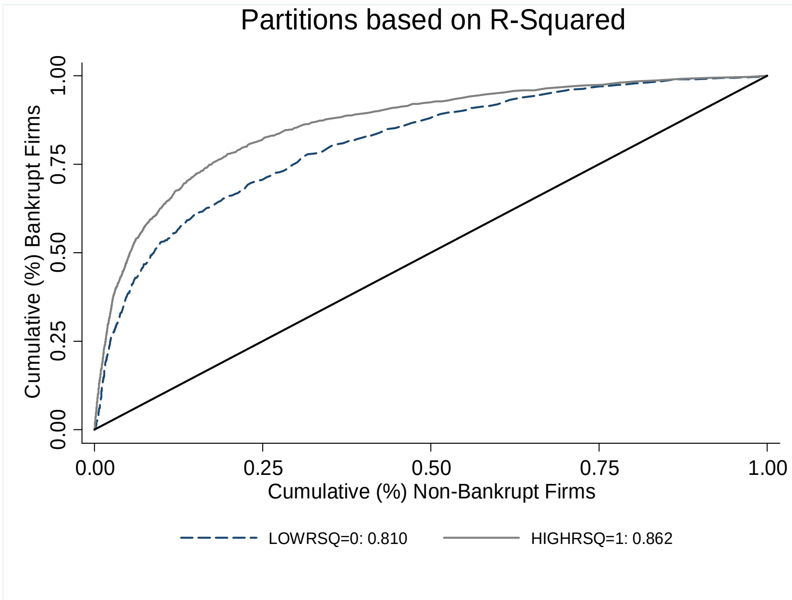


Year	Frequency	Percent	Year	Frequency	Percent
1991	1	0.05	**2002**	249	11.57
1992	3	0.14	**2003**	147	6.83
1993	7	0.33	**2004**	103	4.78
1994	9	0.42	**2005**	86	3.99
1995	11	0.51	**2006**	56	2.60
1996	27	1.25	**2007**	60	2.79
1997	71	3.30	**2008**	190	8.82
1998	136	6.32	**2009**	278	12.91
1999	127	5.90	**2010**	75	3.48
2000	143	6.64	**2011**	58	2.69
2001	300	13.93	**2012**	16	0.74
			Total	**2,153**	**100.00**

Figure 2

ROC curves corresponding to the Table 5 Panel C analysis

Figure 2 plots curves from a receiver operating characteristic analysis across sample partitions based on the existence of country-level short-selling constraints. The dark solid 45-degree line in each graph represents the appearance of a ROC curve for a random model with no predictive ability. A larger area under the curve reflects a model with greater predictive accuracy.



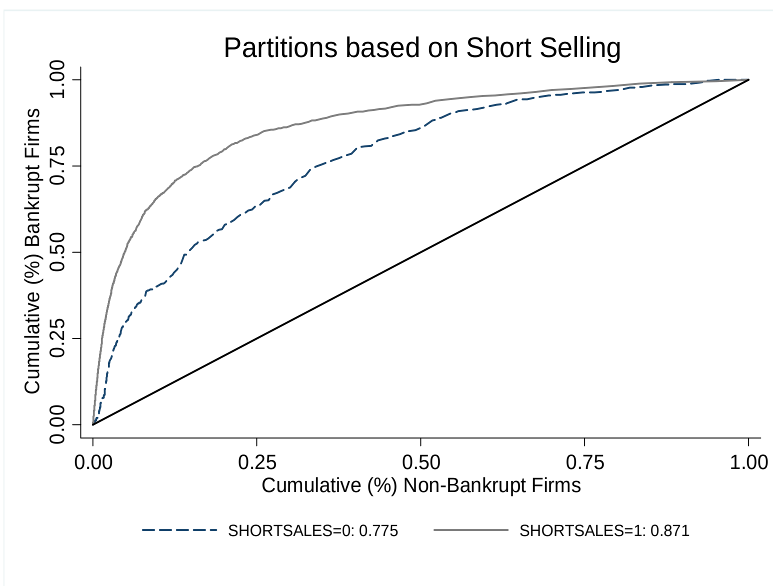


Table 1
Defaults by country

Table 1 presents the number of firm-year observations for both the total sample and the default sample, by country. Table 1 also presents information on whether short selling is practiced within the country (Bris et al. 2007), our country level market-model r-squared measure, and a country-level measure of corporate transparency.

Observations		**Total**		**Default**			
Country	**Freq**	**%**	**Freq**	**%**	***ShortSalesPracticed***	***MktModelR2C***	***Transparency***
Australia	17,145	5.11	96	4.46	Yes	0.064 (Low)	17.0 (High)
Austria	1,200	0.36	5	0.23	Yes	0.142 (High)	6.0 (Low)
Belgium	1,781	0.53	6	0.28	Yes	0.145 (High)	11.6 (Low)
Canada	10,384	3.10	66	3.07	Yes	0.066 (Low)	19.2 (High)
China	13,688	4.08	155	7.20	No	0.361 (High)	4.2 (Low)
Denmark	2,214	0.66	11	0.51	Yes	0.091 (Low)	9.4 (Low)
Finland	1,822	0.54	4	0.19	No	0.144 (High)	16.0 (High)
France	9,835	2.93	43	2.00	Yes	0.091 (Low)	10.0 (Low)
Germany	10,835	3.23	116	5.39	Yes	0.096 (Low)	11.8 (Low)
Hong Kong	12,621	3.76	38	1.76	Yes	0.150 (High)	19.4 (High)
India	18,494	5.52	26	1.21	No	0.185 (High)	14.6 (High)
Indonesia	2,965	0.88	28	1.30	No	0.122 (Low)	3.4 (Low)
Italy	3,276	0.98	14	0.65	Yes	0.214 (High)	14.4 (High)
Japan	50,551	15.08	165	7.66	Yes	0.178 (High)	14.2 (High)
Malaysia	11,922	3.56	66	3.07	No	0.239 (High)	18.0 (High)
Netherlands	2,508	0.75	18	0.84	Yes	0.136 (High)	17.4 (High)
Norway	2,256	0.67	11	0.51	Yes	0.131 (Low)	10.0 (Low)
Philippines	1,870	0.56	19	0.88	No	0.145 (High)	1.6 (Low)
Portugal	722	0.22	1	0.05	Yes	0.129 (Low)	5.0 (Low)
Singapore	7,086	2.11	26	1.21	Yes	0.202 (High)	18.2 (High)
South Korea	17,006	5.07	77	3.58	No	0.176 (High)	12.8 (High)
Spain	1,732	0.52	5	0.23	No	0.203 (High)	8.8 (Low)
Sweden	4,803	1.43	16	0.74	Yes	0.152 (High)	12.0 (Low)
Switzerland	2,862	0.85	6	0.28	Yes	0.123 (Low)	10.6 (Low)
Taiwan	12,601	3.76	29	1.35	No	0.281 (High)	12.6 (High)
Thailand	5,295	1.58	72	3.34	No	0.181 (High)	5.0 (Low)
UK	24,151	7.20	101	4.69	Yes	0.077 (Low)	21.4 (High)
US	83,606	24.94	933	43.33	Yes	0.092 (Low)	21.0 (High)
Total	**335,231**	**100.00**	**2,153**	**100.00**			

Table 2
Descriptive statistics

Table 2 presents descriptive statistics for key variables used in our study, where our sample size for all variables is 335,231 firm-year observations. *LERET*, *LRET*, and *INDEXRET* are market-based measures of cumulative annual firm excess return, firm return, and the return on the market index, respectively. *LSIGMA* and *LRSIZE* are market-based measures of return volatility and relative market capitalization, respectively. *ROA* and *LTA* are accounting-based measures of return-on-assets and leverage, respectively. *DTD* is distance to default. *RFRATE1YR* is the one year government bond rate. All variables are further defined in the Appendix.

Variable	**Mean**	**Std**	**P25**	**P50**	**P75**
LERET	0.020	0.610	-0.324	-0.077	0.205
LRET	0.099	0.662	-0.297	-0.026	0.304
INDEXRET	0.077	0.247	-0.081	0.086	0.221
LSIGMA	0.151	0.110	0.080	0.122	0.187
LRSIZE	-10.544	2.325	-12.175	-10.689	-9.072
ROA	-0.003	0.175	-0.013	0.026	0.072
LTA	0.500	0.242	0.317	0.505	0.671
DTD	3.306	2.682	1.312	2.854	4.843
RFRATE1YR	3.461	2.769	1.062	3.533	5.080

Table 3
Default prediction using the pooled global sample
Panel A of Table 3 presents results of the multiperiod logit models of Eqs. (1)-(3) using 335,231 firm-year observations. *DEFAULT* is an indicator variable that equals one if a firm-year observation is a default-year observation, and equals zero otherwise. *LERET*, *LSIGMA*, and *LRSIZE* are market-based measures of firms' cumulative annual excess return, return volatility, and relative market capitalization, respectively. *ROA* and *LTA* are accounting-based measures of return-on-assets and leverage, respectively. *DTD* is distance to default calculated via the Merton model. *RFRATE1YR* is the country-specific one-year risk-free rate. *LRET* is firm-specific cumulative annual total return, and *INDEXRET* is the corresponding country-specific cumulative annual index return. All variables are further defined in the Appendix. Robust standard errors clustered by firm are reported in parentheses. *, **, and *** indicate significance (two-sided) at the 10%, 5% and 1% levels, respectively. Panel B presents in-sample prediction accuracy results, where we tabulate the number of default observations by predicted default probability decile, where deciles are formed using both default and non-default observations.

Panel A: Logistic regression output

Model:	1	2	Combined
Dep. Var.:	*DEFAULT*	*DEFAULT*	*DEFAULT*
Column:	(1)	(2)	(3)
Intercept	-8.548***	-4.146***	-6.490***
	(0.118)	(0.055)	(0.148)
LERET	-1.191***		-0.759***
	(0.059)		(0.058)
LSIGMA	2.870***		1.790***
	(0.133)		(0.141)
LRSIZE	-0.098***		-0.035***
	(0.010)		(0.011)
ROA	-1.209***		-1.287***
	(0.079)		(0.077)
LTA	2.772***		2.091***
	(0.081)		(0.087)
DTD		-0.577***	-0.377***
		(0.024)	(0.022)
RFRATE1YR		0.001	-0.013**
		(0.006)	(0.006)
LRET		-0.742***	
		(0.089)	
INDEXRET		0.497***	
		(0.093)	
N	335,231	335,231	335,231
Pseudo-R^2	0.136	0.119	0.158

Panel B: In-sample predictive accuracy - defaults by predicted default probability decile

Model:	**1**		**2**		**Combined**	
Default Prob Decile	**N**	**Cum %**	**N**	**Cum %**	**N**	**Cum %**
1	1,220	56.67	1,138	52.86	1,271	59.03
2	330	71.99	310	67.26	303	73.11
3	173	**80.03**	175	**75.38**	161	**80.59**
4	131	86.11	143	82.03	130	86.62
5	91	90.34	112	87.23	86	90.62
6	60	93.13	85	91.18	62	93.50
7	59	95.87	71	94.47	56	96.10
8	38	97.63	59	97.21	31	97.54
9	31	99.07	31	98.65	38	99.30
10	20	100.00	29	100.00	15	100.00
Total	**2,153**		**2,153**		**2,153**	

Table 4

By-country default prediction model estimation

Table 4 presents results of the multiperiod logit default prediction model of Eq. (3) estimated by country, for each country having greater than ten defaults. The dependent variable *DEFAULT* is an indicator variable that equals one if a firm-year observation is a default-year observation, and equals zero otherwise. *ROA*, *LTA*, and *ETL* are return-on-assets, leverage, and cash flow-to-liabilities. *LERET*, *LSIGMA*, and *LRSIZE* are market-based measures of firms' cumulative annual excess return, return volatility, and relative market capitalization, respectively. *ROA* and *LTA* are accounting-based measures of return-on-assets and leverage, respectively. *DTD* is distance to default calculated via the Merton model. *RFRATE1YR* is the country-specific one-year risk-free rate. *ACCUR* is the percentage of default observations in the top three predicted default probability deciles, where deciles are formed using both default and non-default observations All variables are further defined in the Appendix. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively, based on standard errors clustered by firm. We do not report the estimated intercept or standard errors.

Variables:	***N***	***LERET***	***LSIGMA***	***LRSIZE***	***ROA***	***LTA***	***DTD***	***RFRATE1YR***	**Pseudo-R^2**	***ACCUR***
Pred. Sign		**-**	**+**	**-**	**-**	**+**	**-**	**?**		
Australia	17,145	-0.806***	1.877**	0.153***	-0.352	1.543***	-0.610***	0.458***	0.170	87.50
Canada	10,384	-0.833*	1.816*	0.049	-1.082**	2.066***	-0.510***	-0.008	0.200	84.85
China	13,688	-0.620**	-2.511	-0.201***	-7.918***	2.247***	-0.011	0.114*	0.074	67.10
Denmark	2,214	1.177**	4.873	-0.209	-5.870***	4.443**	-0.572***	0.156	0.377	100.00
France	9,835	0.097	2.108	-0.202**	-5.626**	3.713***	-0.225*	0.053	0.209	90.70
Germany	10,835	-0.367	2.088**	-0.059	-3.589***	0.052	-0.437***	0.435***	0.195	87.07
Hong Kong	12,621	-0.320	0.234	0.096	-1.880**	2.345***	-0.216	0.104*	0.100	73.68
India	18,494	0.168	0.817	0.423***	-5.382*	1.438	-0.519***	-0.126	0.133	76.92
Indonesia	2,965	-0.924**	1.479	0.400***	-1.517	0.931	-0.329*	0.018	0.125	82.14
Italy	3,276	-1.117***	1.948	0.044	-5.695	7.791***	-1.428***	-0.021	0.400	100.00
Japan	50,551	-0.441	2.270**	-0.118**	-7.205***	6.380***	-0.648***	-0.350	0.271	93.94
Malaysia	11,922	-0.295	1.172	-0.048	-6.458***	2.406***	-0.330***	-0.050	0.188	87.88
Netherlands	2,508	0.146	5.279	-0.085	-3.257	1.171	-0.533**	0.176**	0.210	88.89
Norway	2,256	-0.963	1.372	0.269	-4.257**	1.283	-1.479***	0.091	0.383	100.00
Philippines	1,870	-0.455	-0.019	0.104	-3.405**	0.351	-0.385***	-0.086*	0.092	63.16
Singapore	7,086	-1.201*	2.901	-0.001	1.434	4.359***	-0.555**	-0.095	0.183	88.46
S Korea	17,006	-0.287	0.428	-0.082	-0.528	5.101***	-0.092	0.224***	0.199	93.51
Sweden	4,803	-0.015	2.190	-0.118	-0.609	0.944	-0.590**	-0.002	0.149	87.50
Taiwan	12,601	-0.632	0.444	-0.112	-11.877***	4.255***	-0.108	0.411***	0.210	93.10
Thailand	5,295	-0.071	-0.884	0.063	-3.219**	6.260***	-0.619***	-0.016	0.304	94.44
UK	24,151	-0.827***	1.981**	-0.077*	-0.706**	1.355***	-0.207***	0.008	0.101	78.22
US	83,606	-0.742***	2.436***	-0.075***	-1.051***	2.143***	-0.489***	0.008	0.228	87.35
% Correct Sign		81.8%	86.4%	59.1%	95.5%	100.0%	100.0%			
%Sig. Incorrect Sign		4.5%	0.0%	13.6%	0.0%	0.0%	0.0%			

Table 5

Predictive accuracy and country-level short selling constraints

Panel A of Table 5 presents results of the multiperiod logit model of Eq. (3) estimated separately for short-sales partitions. Columns (1)-(3) partition based on the country-level practice of short selling. Columns (4)-(6) use country-level market-model r-squared as a proxy for short-sale constraints. *DEFAULT* is an indicator variable that equals one if a firm-year observation is a default-year observation, and equals zero otherwise. *LERET*, *LSIGMA*, and *LRSIZE* are market-based measures of firms' cumulative annual excess return, return volatility, and relative market capitalization, respectively. *ROA* and *LTA* are accounting-based measures of return-on-assets and leverage, respectively. *DTD* is distance to default calculated via the Merton model. *RFRATE1YR* is the country-specific one-year risk-free rate. All variables are further defined in the Appendix. Robust standard errors clustered by firm are reported in parentheses. *, **, and *** indicate significance (two-sided) at the 10%, 5% and 1% levels, respectively. $^{\#}$, $^{\#\#}$, and $^{\#\#\#}$ indicate differences that are significant at the 10%, 5% and 1% levels, respectively, based on Monte-Carlo randomization tests. Panel B (Panel C) presents in-sample prediction accuracy results (false positive statistics), where we tabulate the number of default observations (non-default observations) by predicted default probability decile, where deciles are formed using both default and non-default observations. Panel D presents in-sample prediction accuracy results based on the areas under receiver operating characteristic curves.

Panel A: Logistic regression output

	ShortSalesPracticed			*MktModelR2C*		
	No	Yes	Diff. (2)-(1)	High	Low	Diff. (5)-(4)
Dep. Var.:	*DEFAULT*	*DEFAULT*		*DEFAULT*	*DEFAULT*	
Column:	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	-6.005***	-6.947***		-7.596***	-6.146***	
	(0.307)	(0.174)		(0.264)	(0.177)	
LERET	-0.428***	-0.772***	-0.344$^{\#\#\#}$	-0.483***	-0.860***	-0.377$^{\#\#\#}$
	(0.092)	(0.068)		(0.085)	(0.078)	
LSIGMA	-0.104	2.324***	2.428$^{\#\#\#}$	0.599**	2.583***	1.984$^{\#\#\#}$
	(0.371)	(0.161)		(0.265)	(0.210)	
LRSIZE	0.020	-0.056***	-0.076$^{\#\#\#}$	-0.059***	-0.041***	0.018
	(0.021)	(0.012)		(0.018)	(0.013)	
ROA	-2.353***	-0.947***	1.406$^{\#\#\#}$	-1.617***	-0.835***	0.782$^{\#\#\#}$
	(0.282)	(0.087)		(0.215)	(0.090)	
LTA	2.200***	2.174***	-0.026	2.853***	1.773***	-1.080$^{\#\#\#}$
	(0.169)	(0.104)		(0.157)	(0.099)	
DTD	-0.135***	-0.484***	-0.349$^{\#\#\#}$	-0.232***	-0.441***	-0.209$^{\#\#\#}$
	(0.034)	(0.028)		(0.034)	(0.028)	
RFRATE1YR	0.007	0.057***	0.050$^{\#\#\#}$	0.025*	-0.032***	-0.057$^{\#\#\#}$
	(0.010)	(0.011)		(0.013)	(0.007)	
N	87,395	247,836		168,256	166,975	
Pseudo-R^2	0.078	0.195		0.111	0.190	

Table 5, continued

Panel B: In-sample predictive accuracy - defaults by predicted default probability decile

	ShortSalesPracticed					*MktModelR2C*				
	No		Yes			High		Low		
Default Prob Decile	N	Cum %	N	Cum %	$ACCUR_{Y\text{-}N}$	N	Cum %	N	Cum %	$ACCUR_{L\text{-}H}$
1	193	40.12	1,095	65.49		390	52.63	874	61.90	
2	83	57.38	227	79.07		97	65.72	220	77.48	
3	54	68.61	122	86.36	17.75###	69	75.03	105	84.92	9.89###
4	51	79.21	69	90.49		55	82.46	62	89.31	
5	33	86.07	39	92.82		42	88.12	44	92.42	
6	29	92.10	41	95.28		28	91.90	36	94.97	
7	16	95.43	27	96.89		30	95.95	27	96.88	
8	8	97.09	23	98.27		14	97.84	20	98.30	
9	8	98.75	19	99.40		11	99.33	15	99.36	
10	6	100.00	10	100.00		5	100.00	9	100.00	
Total	481		1,672			741		1,412		

Table 5, continued

Panel C: Non-default observations by predicted default probability decile

	ShortSalesPracticed					*MktModelR2C*				
	No		Yes			High		Low		
Default Prob Decile	N	Cum %	N	Cum %	Yes-No	N	Cum %	N	Cum %	Low-High
1	8,546	9.83	23,688	9.62		16,435	9.81	15,823	9.56	
2	8,657	19.79	24,557	19.60		16,729	19.80	16,478	19.51	
3	8,685	29.79	24,662	29.62	-0.17	16,757	29.80	16,592	29.53	-0.27
4	8,689	39.78	24,714	39.66		16,770	39.81	16,636	39.58	
5	8,707	49.80	24,745	49.71		16,784	49.83	16,654	49.64	
6	8,710	59.82	24,743	59.76		16,798	59.86	16,661	59.70	
7	8,724	69.86	24,756	69.82		16,795	69.89	16,671	69.77	
8	8,731	79.91	24,761	79.88		16,812	79.92	16,677	79.84	
9	8,732	89.95	24,765	89.94		16,815	89.96	16,683	89.92	
10	8,733	100.00	24,773	100.00		16,820	100.00	16,688	100.00	
Total	86,914		246,164			167,515		165,563		

Panel D: ROC curves

	ShortSalesPracticed				*MktModelR2C*			
	No	Yes	Diff (Y-N)	P-Value (Diff)	High	Low	Diff (L-H)	P-Value (Diff)
ROC Area	0.7747	0.8707	0.0960***	0.00	0.8098	0.8620	0.0522***	0.00
N	87,395	247,836			168,256	166,975		

Table 6
Robustness

Table 6 presents model predictive accuracy results using areas under receiver operating characteristic curves from estimation of a default prediction model of alternative specifications of Eq. (3) across short-sale partitions. In Panel A, we estimate Eq. (3) as in Columns (4)-(6) of Table 5 with the addition of industry fixed effects. In Panel B, we estimate Eq. (3) as in Columns (4)-(6) of Table 5 with the addition of country fixed effects. In Panel C, we estimate Eq. (3) as in Columns (1)-(3) of Table 5, but instead use firm-year (rather than country-level) market model r-squared to measure short-sale constraints.

Panel A: Combined Model with Industry Fixed Effects

	ShortSalesPracticed			
	No	**Yes**	**Diff (Y-N)**	**P-value (Diff)**
ROC Area	0.7772	0.8746	0.0974***	0.00
N	87,395	247,836		

Panel B: Combined Model with Country Fixed Effects

	ShortSalesPracticed			
	No	**Yes**	**Diff (Y-N)**	**P-value (Diff)**
ROC Area	0.8416	0.8767	0.0351***	0.00
N	87,395	247,836		

Panel C: Combined Model with Firm-level measure

	MktModelR2F			
	High	**Low**	**Diff (L-H)**	**P-value (Diff)**
ROC Area	0.8330	0.8533	0.0203*	0.05
N	145,749	145,740		

Table 7
Introduction of put options in the presence of short sale constraints

Table 7 presents in-sample predictive accuracy results based on three separate estimations of the multiperiod logit default prediction model of Eq. (3), where predictive accuracy is measured as the areas under receiver operating characteristic curves. The first row reports statistics for those countries with short sale constraints that introduced put options during the sample period, where the estimation is done separately for the period before (*POSTPUT* = 0) and after (*POSTPUT* = 1) those countries introduced the trading of put options. The second and third rows, respectively, repeat the analysis for countries with short sale constraints where put options are never traded, and countries that allow short sales where put options trade throughout the sample period. For both sets of countries that did not change their put option regime during the sample period, we define *PSEUDOPOSTPUT*, an indicator that delineates the sample at the same general time as *POSTPUT* (June 2001). #, ##, and ### indicate differences that are significant at the 10%, 5% and 1% levels, respectively, based on Monte-Carlo randomization tests.

	***POSTPUT* =**			
ROC Curve Area	**0**	**1**	**Diff (1-0)**	**P-value (Diff)**
ShortSalesPracticed = *No*; Puts began trading during sample pd.	0.8103	0.8656	0.0553*	0.06
(Malaysia, South Korea)				
	***PSEUDOPOSTPUT* =**			
	0	**1**	**Diff (1-0)**	**P-value (Diff)**
ShortSalesPracticed = *No*; Puts never traded during sample pd.	0.7886	0.8047	0.0161	0.53
(China, Philippines, Taiwan, Thailand)				
ShortSalesPracticed = *Yes*; Puts trade throughout sample pd.				
(refer to Table 1 for list of countries)	0.8680	0.8746	0.0066	0.48

Table 8
Corporate Transparency

Table 8 presents model predictive accuracy results using areas under receiver operating characteristic curves from estimation of the multiperiod logit default prediction model of Eq. (3) across low and high *Transparency* partitions. In Panel A, we estimate Eq. (3) separately across *Transparency* partitions in countries where short sales are not practiced. In Panel B, we estimate Eq. (3) separately across *Transparency* partitions in countries where short sales are practiced.

Panel A: *ShortSalesPracticed = No*

	Transparency			
	Low	**High**	**Diff (H-L)**	**P-value (Diff)**
ROC Area	0.7575	0.8394	0.0819***	0.00
N	25,550	61,845		

Panel B: *ShortSalesPracticed = Yes*

	Transparency			
	Low	**High**	**Diff (H-L)**	**Low**
ROC Area	0.8714	0.8721	0.0007	0.95
N	36,508	211,328		