Enterprise Risk Management and Default Risk: Evidence From the Banking Industry

Sara A. Lundqvist
Anders Vilhelmsson

Abstract
Enterprise risk management (ERM) has emerged as a framework for more holistic and integrated risk management with an emphasis on enhanced governance of the risk management system. ERM should theoretically reduce the volatility of cash flows, agency risk, and information risk—ultimately reducing a firm’s default risk. We empirically investigate the relationship between the degree of ERM implementation and default risk in a panel data set covering 78 of the world’s largest banks. We create a novel measure of the degree of ERM implementation. We find that a higher degree of ERM implementation is negatively related to the credit default swap (CDS) spread of a bank. When a rich set of control variables and fixed effects are included, a one-standard-deviation increase in the degree of ERM implementation decreases CDS spreads by 21 basis points. The degree of ERM implementation is, however, not a significant determinant of credit ratings when controls for corporate governance are included.

Introduction
Enterprise risk management1 (ERM) has emerged as a framework for more holistic and integrated risk management. An integral part of ERM is enhanced internal
control and governance of the risk management system (Committee of Sponsoring Organizations of the Treadway Commission [COSO], 2004). A call for better governance in response to corporate failings and the financial crisis (Kirkpatrick, 2009) can be attributed to ERM’s advancement in more recent years.

ERM should be able to create long-run competitive advantages and value through consistent and systematic measurement and management of firm risks and by ensuring proper information and incentives for business managers (Nocco and Stulz, 2006). ERM incorporates traditional risk management, such as risk identification and hedging, and risk governance, such as the organization, structure and monitoring of the risk management system. The objective of this study is to determine if there is a relationship between default risk and the degree of ERM implementation.

A firm’s default risk is a forward-looking measure of the firm’s own probability of default or the current and future risk facing its creditors. Credit ratings are a commonly used proxy for default risk, and many credit rating studies have focused on how quantifiable and retrospective factors, like financial ratios or macroeconomic factors, predict credit ratings. While these factors do contain a lot of information, more qualitative aspects of the firm are often ignored in default risk prediction despite the fact that they may better capture how the firm will act in the future. Corporate governance is such a qualitative aspect, but only a limited number of studies have investigated the relationship between corporate governance and credit ratings (Bhojraj and Sengupta, 2003; Ashbaugh-Skaife, Collins, and LaFond, 2006). Surprisingly, the existing studies on default risk predication have overlooked an important qualitative and forward looking factor directly related to a firm’s default risk, namely, the firm’s risk management.

Risk management theoretically decreases the volatility of cash flows, which lowers the probability of default and ultimately lowers the expected costs of financial distress (Smith and Stulz, 1985; Bartram, 2000). Through its risk management component, ERM should result in the same benefits. In addition, governance mechanisms have been found to decrease a firm’s default risk by increasing the amount of credible information available for properly evaluating the default risk of the firm and decreasing agency risk through monitoring (Bhojraj and Sengupta, 2003; Ashbaugh-Skaife, Collins, and LaFond, 2006). These same benefits should theoretically also be attained through the implementation of an ERM system; the risk governance component facilitates information and communication and provides an additional layer of monitoring but in the context of the risk management system. There can be varying degrees of ERM implementation, from superficial to comprehensive, and ERM should be the most effective when implemented to higher degrees.

There is case-specific evidence that effective ERM reduces risk and can help a firm maintain or improve its credit rating (Fraser and Simkins, 2007). The practical relationship between ERM and credit ratings/credit rating agencies is one that is mentioned frequently in the ERM literature (Beasley et al., 2008; Hoyt and Liebenberg, 2011; McShane, Nair, and Rustambekov, 2011; Pagach and Warr, 2011; Lundqvist, 2015). A direct relationship between a firm’s credit rating and ERM implementation has, to our knowledge, never been formally investigated.
This study is complementary to the empirical studies on the value effect of ERM implementation such as Hoyt and Liebenberg (2011) as well as Farrell and Gallagher (2014). While these authors study the total value effect on Tobin’s Q, we look at one specific channel that affects value, namely, the reduction of a firm’s default risk. This is a step toward a more detailed understanding of how ERM creates value.

Our sample consists of 78 of the largest banks in the world. Banks are selected given the significance of both default risk and ERM in the banking industry.

We construct a novel measure of the degree of ERM implementation by using text-based searches of annual reports for word combinations related to a number of dimensions of ERM. The degree of ERM implementation often varies over time for each bank, and there is considerable variation in degrees of ERM implementation across banks. We proxy default risk with credit default swap (CDS) spreads and credit ratings; Hilscher and Wilson (2013) find that CDS spreads reflect the two aspects of default risk, raw default probability and systematic default risk, while credit ratings reflect mainly systematic default risk.

We find a significant and negative relationship between the degree of ERM implementation and a bank’s CDS spread. A one-standard-deviation increase in degree of ERM implementation decreases the CDS spread by approximately 50 basis points (bp), 21 bp when controlling for bank characteristics and corporate governance. With degree of ERM implementation as the sole determinant of credit ratings, a one-standard-deviation increase in degree of ERM implementation increases the likelihood of having an AAA- or AA-rated bank by 12 percent. However, when controlling for corporate governance characteristics of the bank, the degree of ERM implementation is no longer a significant determinant of the credit rating. In being consistent with Hilscher and Wilson (2013), the CDS and credit rating results jointly suggest that the degree of ERM implementation lowers default risk by primarily decreasing the probability of default and to a lesser extent by reducing systematic default risk. The lack of relationship in the credit rating sample may alternatively be because we are capturing how credit rating agencies incorporate ERM into the ratings process, as primarily a part of their corporate governance assessment.

**Enterprise Risk Management**

During the credit crisis in 2007, “winning” risk management practices were those that included cooperative organizational structures and firm-wide sharing of information about risk (Jorion, 2009). As a result, comprehensive risk management frameworks now emphasize the importance of good governance of the risk management system (e.g., Institute of International Finance, 2007; Basel Committee on Banking Supervision, 2008; Financial Services Authority, 2008). The Basel Accords, recommendations on banking regulation, have followed suit; in their proposed enhancements of the Basel II framework (specifically, Pillar 2 that pertains to the supervisory review process) they provide enhanced guidance for firm-wide governance and risk management (Basel Committee on Banking Supervision, 2009). “The purpose of this supplemental Pillar 2 guidance is to address the flaws in risk management practices revealed by the crisis, which in many cases were
symptoms of more fundamental shortcomings in governance structures at financial institutions” (Bank for International Settlements, 2009, para. 4).

ERM is a framework for achieving a better governed risk management system. Risk governance mechanisms, like hiring a chief risk officer (CRO), have become the new focus in risk management (Aebi, Sabato, and Schmid, 2012). Lundqvist (2014) finds that the holistic organization of the risk management function is the main identifier of ERM. Key dimensions of this organizational process are, for example, the establishment of a risk committee, the hiring of a senior risk manager, firm-wide communication regarding risk management, and/or the creation of a risk management philosophy (Lundqvist, 2014). Holistic organization is essentially synonymous with risk governance—the direction, control, and structure of the risk management system. The results from Lundqvist (2015) suggest that risk governance is implemented based on the need for more comprehensive governance in a firm and not as a superficial means to appease stakeholders.

This speaks to the fact that there can be varying degrees of ERM implementation, from superficial to comprehensive. For example, the COSO (2004) framework is made up of eight components of ERM implementation; if all of the components are present and functioning properly, an entity of any size can have effective ERM (COSO, 2004). There is also empirical support that the effectiveness of ERM implementation is dependent on the degree of ERM implementation; where degree is described as maturity (Farrell and Gallagher, 2014), quality (Baxter et al., 2013), and fullness (Gates, 2006). Therefore, when ERM is implemented to a high degree and in a comprehensive manner, it should be the most effective at creating value for the firm.

However, there has been difficulty pinpointing the exact mechanism with which the implementation of ERM actually creates value. Generally, traditional risk management and governance-related benefits are central to the argumentation. For example, Nocco and Stulz (2006) argue that ERM can create long-run competitive advantages by creating value on both the macro level, by helping the firm maintain access to capital markets and other resources, and the micro level, by creating a “way of life” for managers and employees at all levels of the firm. Macro benefits are arguably related to the standard corporate finance theories of risk management; risk management mitigates costs by reducing the underinvestment problem, mitigating financial distress costs, coordinating investment and financing strategies, and mitigating costs resulting from agency problems associated with managerial incentives (Bartram, 2000). The “way of life” provides micro-level benefits that are unique to ERM. This “way of life” is established by enforcing direction, control, and structure on the risk management system—implementing risk governance. Therefore, the firm can benefit from ERM in terms of enhanced governance of the risk management system and from its more basic risk management purposes.

Empirical evidence is inconsistent in its support of the argument that ERM is value creating, though perhaps skewed in its favor. Though there are a limited number of empirical studies on ERM due to its more recent evolution, a main strand in the ERM literature is the investigation of the impact of ERM implementation on a firm. The majority of papers focus on how ERM creates value for an enterprise where value is
defined in terms of excess stock market returns (Beasley, Pagach, and Warr, 2008; Gordon, Loeb, and Tseng, 2009; Baxter et al., 2013), an overall measure of value proxied by Tobin’s Q (Hoyt and Liebenberg, 2011; McShane, Nair, and Rustambekov, 2011; Baxter et al., 2013), perceived value such as better decision making and profitability (Gates et al., 2012), performance measures like buy-and-hold returns and return on equity (Aebi, Sabato, and Schmid, 2012), return on assets (ROA) (Baxter et al., 2013), and cost and revenue efficiency (Grace et al., 2015).

Beasley, Pagach, and Warr (2008) and Gordon, Loeb, and Tseng (2009) find that the relationship between ERM and performance is firm specific. Hoyt and Liebenberg (2011) and Farrell and Gallagher (2014) find an ERM premium of roughly 20–25 percent in Tobin’s Q. McShane, Nair, and Rustambekov (2011) also look at the impact on Tobin’s Q, but they find that insurance firms show a positive relationship between Standard & Poor’s (S&P) ERM ratings only as the rating increases over the first three levels. Baxter et al. (2013) find that ERM, also measured by S&P ERM ratings, is positively associated with operating performance and earnings response coefficients. Gates et al. (2012) take a more qualitative approach to value and find that aspects of ERM significantly impact the perceived performance of a firm. In a similar survey of ERM practices, 1 percent of the survey respondents said they had spread ERM throughout all aspects of their operations, and that small group also claimed to have significantly higher benefits from implementation (Gates, 2006).

Aebi, Sabato, and Schmid (2012) focus on the distinct characteristic of ERM of implementing risk governance. They find that banks where the CRO directly reports to the board of directors exhibit significantly higher stock returns and returns on equity during the financial crisis of 2007/2008. Similarly, Grace et al. (2015) find that risk governance practices result in increased cost and revenue efficiency.

Pagach and Warr (2010) are overall unable to find support that ERM is value creating using a wide range of firm variables. They call for further study in the area, particularly on how ERM’s success can be measured.

The inconsistent evidence on the value of ERM may suggest that ERM is not value creating or that the costs of implementing outweigh any possible benefits. On the other hand, it could suggest that the current focus variables used to measure value are inappropriate or are simply too noisy. Like Pagach and Warr (2010), we believe the previous “value” variables to be the source of variation in the results.

By analyzing the relationship between the degree of ERM implementation and default risk, we answer Pagach and Warr’s (2010) call for further study. We isolate and analyze potential value creation through one specific channel: the reduction of a firm’s default risk. Based on the law of one price, firm value is determined by the discounted future free cash flows of the firm, where the discount factor is determined by the cost of debt and equity. The value effect of reducing default risk can be in part due to the decrease in expected financial distress costs; this has a positive effect on the cash flow of the company and increases the value of the firm when bankruptcies are costly and not temporary (nonoperational). Particularly for banks further from financial distress, the more significant value effect likely comes from the direct
increase in the expected payoff from a firm’s debt and hence the decrease in the cost of
debt.

**Default Risk and ERM**

Fraser and Simkins (2007) argue specifically that ERM can lower a firm’s cost of debt by helping a firm maintain or improve its credit rating; this results in the reduction of a firm’s overall cost of capital and an increase in its value. They give the example of Hydro One\(^2\) where analysts from S&P and Moody’s rated a new debt issue citing ERM explicitly as a factor in the rating process. Hydro One’s new debt issue took place in 2000, which is “early” in the ERM timeline. However, in 2004 senior analysts at Moody’s confirmed that ERM was a significant factor in the ratings process both in 2000 and 2004 (Aabo, Fraser, and Simkins, 2005). Aabo, Fraser, and Simkins (2005) find that ERM-related questions were incorporated into Moody’s corporate governance assessment. However, the consideration of ERM in the rating process has been made more explicit in the last 10 years.

The focus of S&P’s ERM assessment for financial firms is on five key areas: risk governance, operational risk, market risk, credit risk, and liquidity and funding. They also state that risk governance is the foundation of the evaluation structure where they focus on assessing a financial institutions risk culture, risk appetite, aggregation of risk at the enterprise level, and the quality of its risk disclosure (S&P, 2006).

In 2008, S&P announced its intent to also incorporate an ERM analysis into its corporate ratings (S&P, 2008). Two years later it clarified that the ERM assessment was simply an extension of the management assessments that had always been part of the rating process; S&P felt that its use of the term “enterprise risk management” created confusion and a misinterpretation that the announcement involved a change to the existing rating process (S&P, 2010). One of the five areas of review in the analysis of management and governance is risk management/financial management (S&P, 2012). The only mention of “enterprise risk management” takes place in this section; for corporates, S&P looks for comprehensive enterprise-wide risk management standards and tolerances as well as standards for operational performance in its review of risk management. In addition, the methodology for assessing management and governance places weight on the board retaining control as the final decision-making authority with respect to key enterprise risks (S&P, 2012). It is not however clear that the use of the term “enterprise risk management” or “enterprise risk” in relation to corporates is intended to be associated with ERM; one would suspect not given the clarification made in 2012.

For insurance firms however, there is a separate ERM assessment that covers other risks except for financial management (S&P, 2012); the connection here to ERM is much more straightforward. S&P places more of a focus on ERM for insurers than for corporates or financial institutions; for insurance firms, they have ERM ratings in addition to credit ratings. In the analysis of ERM for insurers, S&P considers five subfactors: the risk management culture, risk control, emerging risk management, emerging risk management, emerging risk management, emerging risk management, emerging risk management, emerging risk management.

\(^2\)For more on ERM at Hydro One, see Aabo, Fraser, and Simkins (2005).
risk models, and strategic risk management (S&P, 2013). This incorporation of ERM into credit ratings seems to give balanced treatment to both risk management and risk governance aspects.

For corporates, the incorporation of ERM is more related to the assessment of a firm’s management and corporate governance, while for insurance firms it is a separate and specific assessment of ERM in terms of both its risk management and risk governance components. For financial institutions it seems to be somewhere in between, with risk governance taking a central role; however, for financial institutions the incorporation of ERM has not been updated from what we can tell since 2007 (S&P, 2007).

The relationship between ERM and credit ratings/credit rating agencies is mentioned frequently in the ERM literature (Beasley, Pagach, and Warr, 2008; Hoyt and Liebenberg, 2011; McShane, Nair, and Rustambekov, 2011; Pagach and Warr, 2011; Lundqvist, 2015). The rise of interest in and implementation of ERM is often argued to be a result of the increased focus on ERM by rating agencies. S&P has determined that ERM is an important aspect in evaluating the creditworthiness of a firm. But beyond the practical, the theoretical relationship between ERM and default risk is fairly straightforward. As discussed previously, ERM can be viewed as being made up of two main components: risk management and risk governance. Therefore, the fundamental theories related to risk management and corporate governance can explain the relationship between ERM and default risk.

Theoretically, capital market imperfections create incentives for firms to implement risk management on the basis that it is value creating to do so. One way to create value is through the reduction of transaction costs of financial distress. Risk management can reduce the probability of default by decreasing the volatility of cash flows, thereby reducing expected costs of financial distress (Smith and Stulz, 1985; Bartram, 2000). Overall, empirical studies on risk management support the relationship between risk management, measured by the use of a variety of derivatives, and the probability of default, measured by leverage, interest coverage, or credit rating (Wall and Pringle, 1989; Mayers and Smith, 1990; Nance, Smith, and Smithson, 1993; Samant, 1996; Géczy, Minton, and Schrand, 1997; Gay and Nam, 1998; Brunzell, Hansson, and Liljeblom, 2011).

Despite this theoretical and empirical connection between default risk and risk management, this relationship is ignored in the studies on the determinants of default risk. There are a plethora of studies in the area of default risk, specifically the prediction of credit ratings (Blume, Lim, and MacKinlay, 1998; Kamstra, Kennedy, and Suan, 2001; Bhojraj and Sengupta, 2003; Amato and Furfine, 2004; Ashbaugh-Skaife, Collins, and LaFond, 2006; Curry, Fissel, and Hanweck, 2008; Bissoondoyal-Bheenick and Treepongkaruna, 2011; Matthies, 2013b). The determinants of credit ratings fall into three main categories: financial ratios and financial data, corporate governance mechanisms, and macroeconomic factors (Matthies, 2013a). Empirical studies have mostly looked at how the quantity of a firm’s risk, financial and

---

macroeconomic factors, affects credit ratings; this has resulted in a set of standard and robust factors that affect credit ratings, for example, leverage, liquidity, and size. These types of factors are more retrospective and do not necessarily capture the forward-looking aspects that are relevant for default risk. There have been a limited number of studies investigating corporate governance and its impact on default risk (Bhojraj and Sengupta, 2003; Ashbaugh-Skaife, Collins, and LaFond, 2006). Governance characteristics may better capture future firm activities, and governance in the context of the risk management system is an integral part of ERM.

Corporate governance can be viewed as a mechanism for reducing two risks that affect the firm’s likelihood of default: agency risk and information risk. Governance mechanisms, like monitoring, will mitigate the agency costs that occur from conflicts between managers and all external stakeholders (Ashbaugh-Skaife, Collins, and LaFond, 2006). Increased monitoring in the firm should result in better decision making by managers and increased value to all stakeholders (Bhojraj and Sengupta, 2003; Ashbaugh-Skaife, Collins, and LaFond, 2006). Weak governance will, therefore, shift the probability distribution of future cash flows downward; this increases the probability of default and should in turn decrease the credit rating (Ashbaugh-Skaife, Collins, and LaFond, 2006). Additionally, the firm’s likelihood of default is dependent on having credible information for properly evaluating the default risk as well as agency costs in a firm. Governance reduces information risk in the firm by encouraging firms to disclose information in a timely manner (Bhojraj and Sengupta, 2003).

Empirical evidence does in fact suggest that after controlling for firm-specific risk characteristics, credit ratings are negatively associated with the number of blockholders and CEO power and positively related to takeover defenses, accrual quality, earnings timeliness, board independence, board stock ownership, and board expertise (Ashbaugh-Skaife, Collins, and LaFond, 2006). Bhojraj and Sengupta (2003) also find that firms with greater institutional ownership and stronger outside control of the board enjoy lower bond yields and higher ratings on new bond issues.

Pulling from the fundamental theories related to risk management and corporate governance would suggest that ERM should negatively impact default risk through its components of risk management and risk governance. The probability of default should be reduced through the risk management itself, by reducing the volatility of cash flows, and through the proper governance of the risk management system, by reducing information risk and agency risk. Information risk and agency risk are relevant concerns in the context of a risk management system in terms of, for example, the reliability of information about firm risks and the risk taking of managers. Based on practical guidance, like COSO (2004), and previous empirical studies (Gates, 2006; Baxter et al., 2013; Farrell and Gallagher, 2014), the higher the degree of ERM implementation, the more effective it is; hence, the degree of ERM implementation should be negatively related to the default risk of a firm.

As mentioned previously, two studies address a similar relationship; McShane, Nair, and Rustambekov (2011) and Baxter et al. (2013) analyze the impact of S&P ERM ratings on firm performance and value. Additionally, Eckles, Hoyt, and Miller (2014) look for the impact of ERM on the risk taking, proxied by the stock return volatility, of
firms in the insurance industry. They find that firms implementing ERM experienced reduced firm risk as well as an increase in operating profits per unit of risk. In the ERM determinants literature, Lundqvist (2015) investigates if having publicly rated debt is related to risk governance implementation and finds that there is no relationship. However, as far as we know, there are no studies that have analyzed the direct relationship between the degree of ERM implementation and default risk. We, therefore, are able to meet the need for a new focus variable in the literature as well as address the lack of research on how ERM, and risk management in general, affects a firm’s default risk.

**Sample, Data, and Empirical Method**

We construct a novel measure of the degree of ERM implementation using text-based searches of annual reports for word combinations related to a number of ERM dimensions. We proxy default risk with CDS spreads and credit ratings. This results in two partially overlapping samples where the union sample, the set of banks that are either in the CDS sample or the credit rating sample, is made up of 78 of the world’s largest banks. We then investigate the relationship between default risk and ERM.

**Sample**

We start by using DataStream to create a list of all the banks in the world with active equity status at the time of collection and with data available for total assets in U.S. dollars (USD) at fiscal year end 2006. The year 2006 is used as the base year because generally there is access to annual reports back to 2005, which are used for the measurement of the degree of ERM implementation, and DataStream begins providing CDS data in 2007, which is one of the proxies used for default risk. This list is composed of 1,563 banks with total assets ranging from 22,000 USD to 1.9 trillion USD; the largest bank in 2006 being UBS AG. The list of banks is ordered from largest to smallest based on total assets. Beginning with the largest bank, we check for the availability of S&P credit rating data and/or CDS data, which are the proxies used for default risk, for any of the years between 2005 and 2011, which is the time span in the study. If credit ratings or CDS data are available, we then check for access to annual reports for any of the years between 2005 and 2011. We obtain the annual reports from the banks’ Web sites, and in cases where annual reports are unavailable online, we contact the bank via e-mail to try to obtain them. If CDS data are available and annual reports are available, the bank is included in the CDS sample. If credit rating data are available and annual reports are available, the bank is included in the CDS sample. If at the time of collection neither credit rating nor CDS data are available or if we cannot obtain annual reports, the bank is excluded. A bank can be in both or only one of the samples. We follow this procedure of checking for default risk data and then annual reports for each bank in order of largest to smallest until the CDS sample

---

4 There are 1,800 banks if you also count those that do not have total asset data (DataStream output #NA). In this count, there are a few cases of duplicate entries of the same bank in the database.
contains a total of 55 banks. At this point, the credit rating sample contains 72 banks; more banks have credit rating data than CDS data. The union sample is a total of 78 different banks. For 2006, the average total assets of the 78 banks are equal to approximately 790 billion USD. Thirty-nine banks are European, 1 is Russian, 17 are North American (11 U.S. and 6 Canadian), 14 are Asian, 1 is Israeli, 4 are Australian, and 2 are Brazilian. The smallest bank in the union sample is, according to the original list of banks, the 117th largest bank in terms of total assets at the fiscal year end of 2006 with total assets of 62 billion USD.

This means that 39 banks out of the 117 largest are not included in the union sample. Thirty-two of these are excluded because at the time of collection we did not have access to S&P credit ratings or data on traded CDSs during the time period. Only one was excluded because there were no annual reports available to us at the time of collection. Despite the active equity status according to DataStream at the time of collection, four banks are excluded because they have inactive dates during the time period and two appear to have been sold or acquired during the time period.

A survivor sample is necessary because annual reports for banks that merge or cease to exist during the time period were not available to us. Because we employ a survivor sample, the level of default risk is likely to be underestimated given that banks with poor default risk are likely to go bankrupt and thus drop from the sample. It may also bias the degrees of ERM implementation, but we do not expect it to have an effect on the relationship itself.

Banks are an appropriate sample given that default risk and ERM are often a focus in the banking industry. A main operation of banks is the management and control of counterparty risk, market risk, and operational risk. Regulators have stressed the importance of a firm-wide risk management system in managing a bank’s counterparty risk, market risk, and operation risk, specifically through regulation recommendation like the Basel Accords. In addition, the financial industry generally shows more extensive implementation of ERM (Beasley, Clune, and Hermanson, 2005; Pagach and Warr, 2011), so banks should provide a sample where there are ERM implementers but also varying degrees of implementation.

Measuring Default Risk

We proxy banks’ default risk using two variables: borrower credit ratings and CDS spreads. The CDS spread is the amount paid for insurance against default and is a direct market-based measure of the firm’s default risk, and credit ratings are opinions of the credit rating agency regarding a corporation’s relative default risk (S&P, 2011). Therefore, both the credit rating of the bank and the CDS spread is driven by a bank’s credit quality and level of default risk that is otherwise unobservable.

We use both measures in order to obtain a robust picture of ERM’s effect on default risk. Credit rating data are more readily available and credit ratings have long been used as an indicator of credit quality. However, credit ratings have been under scrutiny given their poor performance before and during the financial crisis. Credit ratings also have a certain stability, which means that credit rating changes may not
reflect default risk in a timely manner. This is one reason to bring in CDS prices as an additional measure of default risk; CDS prices should be a better measure of default risk because of its market-based and timely nature. One would therefore expect that credit rating changes should lag credit-spread changes (Hull, Predescu, and White, 2004).

A second reason to include both credit ratings and CDS data are that Hilscher and Wilson (2013) find that CDS spreads and credit ratings capture different aspects of an investor’s credit risk. They break investor credit risk down into two attributes: raw default probability and systematic default risk. They find that credit ratings strongly reflect variation in systematic risk but are poor predictors of corporate failure. They, however, find that estimated default probabilities are strongly related to CDS spreads and that the CDS risk premia are strongly related to their measure of systematic default risk. According to Hilscher and Wilson, CDS spreads would be a more suitable measure of investor credit risk than credit ratings since CDS spreads contain information about both attributes.

**Credit Default Swaps.** CDS mid spreads for the last day each year (results are very similar when the yearly average CDS price is used) are obtained from DataStream (DataStream code: SM). A typical bank has about 50–70 different CDS contracts traded that differ in time to maturity, currency, the definition of a credit event, and the seniority of the debt. We choose the specific contract based on picking the specifications that are most common, hence minimizing the variation in contract specification between the banks. There is never a change in the type of contract for a given bank. We select contracts with a maturity of 5 years (55 of 55 contracts), denomination in U.S. dollars (52 of 55 contracts), and senior debt (50 of 55 contracts). In cases when this specific CDS contract is not available, we select contracts according to the following rules in order of importance. If U.S. dollar denomination does not exist we use euros (3 contracts), and if senior debt does not exist we use subordinated debt (5 contracts). The definition of credit event has a regional variation, with the most common being modified restructuring in North America and modified-modified restructuring in Europe. We choose in order of preference: modified restructuring (MR) (13 of 55 contracts), modified-modified restructuring (MM) (32 of 55 contracts), and complete restructuring (CR) (10 of 55 contracts), which is sometimes called full restructuring (FR). The difference in credit event may affect the level of the CDS spread. Since the credit event is closely tied to geographic region, we capture this effect by including regional dummy variables in all regression models. Packer and Zhu (2005) investigate the pricing impact of different credit events and find the effect to be minor; for three different credit events used, the difference is largest between full restructuring and modified restructuring, but it is only 2.77 bp.

Table 1 shows descriptive statistics for the CDS prices. For the CDS sample we lose the years 2005 and 2006 due to data availability in DataStream. The maximum sample size for any given year is 54 banks. The average value of the CDS spread varies over time from 46.9 bp in 2007 to 458.4 bp in 2011. There is also a lot of cross-sectional variation between the banks with the first quartile at 86.2 bp and the third at 265.1 when looking across all years and banks.
### TABLE 1
Descriptive Statistics for the Credit Default Swap Prices, Credit Ratings, and Degree of Enterprise Risk Management Implementation Scores

<table>
<thead>
<tr>
<th></th>
<th>All Years</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Credit Default Swap (CDS) Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Observations:</td>
<td>213</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>54</td>
</tr>
<tr>
<td>CDS spreads</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>264.31</td>
<td>–</td>
<td>–</td>
<td>46.88</td>
<td>193.59</td>
<td>175.40</td>
<td>295.34</td>
<td>458.41</td>
</tr>
<tr>
<td>St. dev.</td>
<td>381.35</td>
<td>–</td>
<td>–</td>
<td>11.96</td>
<td>139.78</td>
<td>314.84</td>
<td>354.38</td>
<td>552.46</td>
</tr>
<tr>
<td>1st quartile</td>
<td>86.24</td>
<td>–</td>
<td>–</td>
<td>40.88</td>
<td>112.00</td>
<td>65.00</td>
<td>104.88</td>
<td>149.61</td>
</tr>
<tr>
<td>Median</td>
<td>139.10</td>
<td>–</td>
<td>–</td>
<td>45.00</td>
<td>140.00</td>
<td>95.31</td>
<td>159.07</td>
<td>267.26</td>
</tr>
<tr>
<td>3rd quartile</td>
<td>265.13</td>
<td>–</td>
<td>–</td>
<td>53.38</td>
<td>225.00</td>
<td>141.67</td>
<td>261.36</td>
<td>481.31</td>
</tr>
<tr>
<td>Max</td>
<td>2,646.40</td>
<td>–</td>
<td>–</td>
<td>69.00</td>
<td>785.19</td>
<td>2,135.00</td>
<td>1,519.00</td>
<td>2,646.40</td>
</tr>
<tr>
<td>Min</td>
<td>27.00</td>
<td>–</td>
<td>–</td>
<td>27.00</td>
<td>70.50</td>
<td>43.33</td>
<td>55.83</td>
<td>67.55</td>
</tr>
<tr>
<td>Degree of enterprise risk management implementation (°ERM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>49.61</td>
<td>–</td>
<td>–</td>
<td>48.10</td>
<td>47.74</td>
<td>49.96</td>
<td>51.13</td>
<td>50.20</td>
</tr>
<tr>
<td>St. dev.</td>
<td>10.01</td>
<td>–</td>
<td>–</td>
<td>9.14</td>
<td>10.30</td>
<td>9.15</td>
<td>8.83</td>
<td>11.64</td>
</tr>
<tr>
<td>1st quartile</td>
<td>45.75</td>
<td>–</td>
<td>–</td>
<td>40.00</td>
<td>44.00</td>
<td>47.00</td>
<td>49.00</td>
<td>46.00</td>
</tr>
<tr>
<td>Median</td>
<td>52.00</td>
<td>–</td>
<td>–</td>
<td>52.00</td>
<td>52.00</td>
<td>52.00</td>
<td>53.00</td>
<td>52.50</td>
</tr>
<tr>
<td>3rd quartile</td>
<td>56.00</td>
<td>–</td>
<td>–</td>
<td>54.50</td>
<td>55.00</td>
<td>56.00</td>
<td>57.00</td>
<td>58.00</td>
</tr>
<tr>
<td>Max</td>
<td>65</td>
<td>–</td>
<td>–</td>
<td>60</td>
<td>63</td>
<td>63</td>
<td>62</td>
<td>65</td>
</tr>
<tr>
<td>Min</td>
<td>12</td>
<td>–</td>
<td>–</td>
<td>28</td>
<td>19</td>
<td>21</td>
<td>19</td>
<td>12</td>
</tr>
</tbody>
</table>

|                      |           |      |      |      |      |      |      |      |
| **Credit Rating Sample** |         |      |      |      |      |      |      |      |
| # Observations:      | 442       | 46   | 53   | 62   | 69   | 70   | 71   | 71   |
| Credit ratings       |           |      |      |      |      |      |      |      |
| # AAA or AA          | 145       | 18   | 25   | 30   | 22   | 18   | 17   | 15   |
| # A                  | 234       | 24   | 24   | 29   | 41   | 39   | 39   | 38   |
| # BBB                | 51        | 3    | 3    | 3    | 6    | 13   | 11   | 12   |
| # < BBB              | 13        | 1    | 1    | 0    | 0    | 0    | 4    | 6    |
| Degree of enterprise risk management implementation (°ERM) | | | | | | | | |
| Mean                 | 47.36     | 43.26| 44.74| 46.94| 46.17| 48.73| 49.23| 50.27|
| St. dev.             | 10.96     | 11.49| 10.28| 9.68 | 11.66| 10.02| 10.71| 11.59|
| 1st quartile         | 41.00     | 36.00| 37.75| 41.00| 41.75| 44.00| 46.25| 46.00|
| Median               | 50.50     | 45.50| 47.00| 48.50| 50.00| 56.00| 52.00| 58.00|
| 3rd quartile         | 56.00     | 53.00| 53.00| 54.00| 55.00| 56.00| 57.00| 58.00|
| Max                  | 66        | 61   | 60   | 65   | 64   | 63   | 63   | 66   |
| Min                  | 12        | 15   | 24   | 24   | 17   | 21   | 18   | 12   |

*Note:* Year-end CDS prices are obtained from DataStream (DataStream code: SM). See main text for details on CDS selection. We use the Standard & Poor’s year-end historical local borrower rating obtained from DataStream (DataStream code: BSPHL). Plus and minus ratings are grouped into the same category as the major rating and all non-investment-grade banks (below BBB) are grouped into a single category. The degree of enterprise risk management implementation is constructed using text searches of annual reports. See main text for details on construction. The hypothetical maximum for the degree of enterprise risk management implementation is 83.
Credit Ratings. We use the S&P year-end historical local borrower ratings collected from DataStream (DataStream code: BSPHL). S&P uses the following ratings: AAA, AA, A, BBB, BB, B, CCC, CC, C, and D, where AAA denotes the strongest creditworthiness, and C or D denotes the weakest or that default has occurred. Ratings from AA to CCC may be modified by the addition of a plus (+) or minus (−) sign to show relative standing within the major rating categories. Firms with a rating of BBB− or higher are considered investment grade; anything below is considered speculative grade.

Ratings are grouped into four categories. We decide to treat plus and minus ratings as belonging to the same category, and we also group all non-investment-grade banks into the same category. AAA- and AA-rated banks are grouped into the same category. Table 1 shows descriptive statistics for the credit ratings. Most banks fall into the A rating, and only 13 ratings for all banks and all years are non-investment-grade. The maximum sample size for any given year is 72 banks.

Measuring the Degree of ERM Implementation

In terms of public disclosure, Liebenberg and Hoyt (2003) argue that firms typically do not disclose whether they are managing risks in an integrated manner and much of risk management disclosure is related to specific risks. This would make it difficult to identify firms that have implemented ERM or, as is our aim, to assess the degree of ERM implementation in firms from public disclosure. However, this has become increasingly easier as “disclosure about the system for monitoring and managing risk is increasingly regarded as good practice” (Organisation for Economic Co-operation and Development, 2004; Kirkpatrick, 2009). A result of this is that firms may superficially disclose typical ERM dimensions despite relatively low efforts toward comprehensive degrees of implementation. Take, for example, the hiring and disclosing of a CRO, a typical marker of ERM implementation, that could potentially be put in place rather superficially and geared toward window dressing.

To get as comprehensive of an assessment of the degree of ERM implementation from public disclosure as possible, we include a variety of dimensions of ERM implementation in our measure, from basic risks that are considered to the organization and control of the risk management system. This allows us to assess the degree of ERM implementation from the disclosure of underlying dimensions instead of looking for rare disclosures of “integration” or narrow, potentially superficial, dimensions of ERM. The aim is to create a measure of the degree of ERM implementation that captures both breadth and depth.

We created a comprehensive list of the dimensions of ERM implementation. The list is based on Desender (2011) and Lundqvist (2014), both heavily influenced by COSO (2004), and Hoyt and Liebenberg (2011); the final list consists of 83 dimensions of ERM. Both Desender (2011) and Lundqvist (2014) test their lists with professionals in the area of ERM to ensure their completeness as well as to exclude unnecessary dimensions. Our belief is that all items are relevant and necessary to be able to properly assess the degree of ERM implementation. The dimensions used in the study heavily reflect COSO’s eight components of ERM implementation; the framework states that as long as each of the components are present and functioning properly, an
entity of any size can have effective ERM (COSO, 2004). Lundqvist uses almost identical dimensions in the survey used to create the “pillars of ERM”; Lundqvist states that all four pillars should be represented when attempting to measure ERM implementation levels. Lundqvist finds the weighting of the final four pillars to be difficult given expert opinions, and the exploratory factor analysis results (similar to principal component analysis) make weighting or reducing the comprehensive set of dimensions based on loadings difficult; given this, we believe all dimensions must be included and give them all equal weight. Desender also uses an aggregate measure made up of his large number of dimensions that he similarly terms “degree of enterprise risk management implementation.”

We search the banks’ annual reports for each of the 83 dimensions of ERM implementation. As a single word seldom represents an ERM dimension well, we search for word combinations. Some dimensions may be represented by more than one set of word combinations, for example, synonym combinations. As a specific example, one of the dimensions of ERM is the statement of a risk appetite. For this, we search for “risk + appetite,” which only gives a search hit if the word “appetite” exists within ±200 characters of the word “risk.” In some cases, we also use combinations with more than two words, for example, “risk + response + plan”; in this case, we count a hit if all the words exist within ±200 characters from the first word. See the supplementary Internet material (Lundqvist and Vilhemsson, 2016) for a table with the full list of dimensions and the respective search combinations as well as the percentage of bank-year observations with a hit for each dimension for the full CDS and credit rating sample. The table also shows the percentage of bank-year observations with a hit for each dimension for two subsamples of banks: those with high degrees of implementation (90th percentile) and those with low degrees of implementation (10th percentile).

We code all search combinations that have at least one hit with a one and the others with a zero. As mentioned, some dimensions are represented by more than one set of word combinations; in these cases, a one for that dimension is an “or” function of the individual search combinations. The sum of the coded variables is the degree of ERM implementation for each bank-year. An alternative would be to use the sum of the number of hits for each search combination, but that would put more weight on search combinations that are more common, which does not necessarily mean that they are more important. As mentioned previously, the zero/one coding gives equal weight to all search combinations that form the degree of ERM implementation. The degree of ERM implementation can range in values between 0 and 83.

The average degree of ERM implementation increases over time. The union sample has an average value of 47.5 over the whole time period, meaning that on average the banks implement just under 60 percent of the ERM dimensions. There is substantial cross-sectional heterogeneity among the banks within each year, for the union sample the difference between the first and third quartile is 17 in 2005 and 12 in 2011. See Table 1 for descriptive statistics for the degree of ERM implementation for the CDS and credit rating sample separately.

For the credit rating sample, the banks with low degrees of ERM implementation (10th percentile) have an average implementation of 24.2 (29 percent); for the banks
with high degrees of implementation (90th percentile), the average degree of ERM implementation is 61.2 (83 percent). See the supplementary Internet material (Lundqvist and Vilhemsson, 2016) for more on the comparison between subsamples of banks. On average, the banks with high degrees of ERM implementation have around 2.5 times more ERM implementation than those with low degrees of implementation, and they implement 54 percent more of the ERM dimensions. This is a function of how the measure is constructed, but the construction and the hypothesis of the article imply that high-implementation firms have more effective, “better,” ERM than low-implementation firms.

There are certain dimensions that are more clearly associated with higher degrees of ERM implementation. Nine dimensions have over 80 percent more hits in the banks with high degrees of implementation (90th percentile) than the banks with low degrees of implementation (10th percentile): consideration of litigation issues, documents and record as control, review of the functioning and effectiveness of controls, risk appetite, monitoring of processes, CRO, code of conduct/ethics, audit committee responsibility, and allocated risk owners. Three key characteristics of ERM, namely, risk appetite, CRO, and allocated risk owners, show clear differentiation between high and low degrees of implementation. The other dimensions are control and governance-type dimensions that, as argued previously, should become emphasized under an ERM system, and consideration of a perhaps more atypical risk, namely, litigation risk.

Hoyt and Liebenberg (2011) similarly search public disclosure for indicators of ERM adoption. We attempted to follow their methodology in order to classify the banks as ERM implementers or non-ERM implementers, essentially creating a dummy variable that simply identifies ERM and non-ERM banks instead of assessing the degree of implementation. Hoyt and Liebenberg search public information for a number of keywords associated with ERM. We include the same keywords as part of our measure of the degree of ERM implementation, but they make up only a total of 3 points out of the possible 83. When using Hoyt and Liebenberg’s keywords exclusively to create a distinct dummy variable (“or” function of the coded 1/0 search hits), we find that in 99.3 percent of the 442 bank-years in the credit rating sample the bank would be characterized as an ERM implementer. This high percentage is largely because of the “risk committee” search term. If the dummy variable were instead constructed without the “risk committee” search term, still in 78.5 percent of the bank-years the bank would be characterized as an ERM implementer. This dummy variable then essentially classifies banks based on the existence of a CRO and in some cases the use of “enterprise risk management” or a synonym in the annual reports. Apropos the earlier argument from Liebenberg and Hoyt (2003) with regard to disclosure of risk management in an integrated manner, of the searches for “enterprise risk management” or synonyms.

---

5The keywords are: enterprise risk management, CRO, risk committee, strategic risk management, consolidated risk management, holistic risk management, and integrated risk management.

6The three dimensions are: risk committee, CRO, and “enterprise risk management” or synonyms.
management” and its synonyms, 12.5 percent of bank-year observations in the credit rating sample had at least one hit for “integrated risk management,” 10.4 percent for “enterprise risk management,” and less than 4 percent for the others.

As mentioned before, the disclosure of a CRO does seem to be an important dimension in distinguishing between high degree of ERM implementation banks and low implementation banks using our measure. However, the disclosure of a risk committee or the use of “enterprise risk management” or a synonym is not a distinguishing characteristic; the differences between the percentage of bank-year observations with a hit for the banks with a high degree of implementation (90th percentile) and low degree of implementation (10th percentile) are only 7 percent and 1 percent, respectively.

What this suggests is that in our sample, nearly all of the banks implement ERM to some degree; we see this also in our measure where no firm scores a zero, the minimum score is a 12 (14 percent), and the average score of low ERM implementation (10th percentile) is 24.2 (29 percent). In the largest banks, a measure of ERM that does not take into account comprehensiveness, degree, depth, and breadth of implementation is insufficient for capturing any meaningful variability in ERM implementation; due to this, we do not include Hoyt and Liebenberg’s measure in the analysis.

We believe a proper assessment of ERM implementation must include both breadth and depth. The benefit is that it should better measure the degree of ERM implementation without placing too much weight on typical ERM indicators that do not necessarily guarantee comprehensive implementation. Such a measure can also vary over time for each firm and has more variation between firms given that it is not a dummy variable. Our measure also has the advantage of avoiding the potential self-selection bias from survey measures used in, for example, Farrell and Gallagher (2014). By measuring each year we also get within firm variation, which is lacking in most surveys since the firm only participates once. The disadvantage is that we cannot get the private information that is available in a survey.

Our measure of the degree of ERM implementation could of course capture the disclosure of ERM, risk management in general, or even the general level of disclosure. To alleviate these concerns we add a general disclosure proxy (the total number of words in the annual report) as a control variable in the regression analysis and find that our measure of the degree of ERM implementation captures a dimension that is different from general disclosure.

Validity of the Measure. To confirm the validity of our measure of the degree of ERM implementation we conduct a difference-in-difference study comparing the change in the measured degree of ERM implementation for banks with a major change in their risk management (the treatment sample) to banks without such an event (the control sample).7 We use the hiring of a CRO as the treatment effect. We proxy the hiring year

---

7We thank an anonymous referee for suggesting to test the change in our measure after a major risk management event.
for a given bank by looking at when a given bank uses the exact phrase “chief risk officer” in the annual report for the first time. We calculated the difference in the degree of ERM implementation for the treatment sample as

\[
(\text{ERM}_t - \text{ERM}_{t-1}) - 1,
\]

with time period \( t \) being the first bank-year to have at least one hit for “chief risk officer.” We subtract 1 from the difference to control for the mechanic effect that the degree of ERM implementation is increased by one when we get a hit for “chief risk officer.” Our definition of an event means that a single bank can have at most one event resulting in the number of bank-years in the control sample being considerably larger than in the treatment sample. Banks without treatment will be all banks that either have zero occurrence of “chief risk officer” for all years or at least one occurrence already the first year in the sample. For the CDS sample this results in 7 bank-years for the treatment sample and 153 bank-years for the control sample. The difference in the treatment sample is an average increase in the degree of ERM implementation of 6.86 and in the control sample 0.96. A standard \( t \)-test for the difference of means gives a \( p \)-value of 0.03, showing that a significant risk management change has a large and statistically significant impact on our measure of the degree of ERM implementation.

For the credit rating sample, the same computations result in 19 bank-years for the treatment sample and 296 for the control sample. The difference in the treatment sample is an average increase of 3.42 and in the control sample a decrease of -0.16. A standard \( t \)-test for the difference of means gives a \( p \)-value of 0.02.

Control Variables

We control for differences among the banks in the level of overall disclosure, level of risk, profitability, bank characteristics and valuation, and corporate governance. The controls we select are those used in the bank credit rating determinants literature, for example, Curry et al. (2008) for supervisory ratings given to bank holding companies and Bissoondoyal-Bheenick and Treepongkaruna (2011) for bank credit ratings. Influenced by the findings in Bhojraj and Sengupta (2003) and Ashbaugh-Skaife, Collins, and LaFond (2006), we also include a number of measures related to corporate governance. Controls for credit ratings are fairly well established; it is however, uncommon to use CDS spreads as a cross-sectional firm proxy of default risk in the same way we do. Because a close relationship between credit ratings and CDS prices has been fairly well established (Hull, Predescu, and White, 2004; Daniels and Shin Jensen, 2005; Micu, Remolona, and Wooldridge, 2006), we employ the same set of controls for CDS spreads and credit ratings. The control variables are: number of words, total assets (TA) (measured in trillions), ROA, Tier 1 capital ratio, nonperforming loans over total assets, provision for loan losses over total assets, corporate governance score, audit committee independence, and single biggest owner.

All control variables, apart from the number of words, are collected from DataStream. Corporate governance data come from the Thomson Reuters ASSET4 ESG Database.
in DataStream. All variables are defined in the supplementary Internet material (Lundqvist and Vilhemsson, 2016), including DataStream codes, and descriptive statistics are given in Table 2.

Modeling CDS Spreads
To explain the CDS spreads we use a panel regression model of the form

$$CDS_{it} = \alpha_t + x_{it}\beta + u_{it}$$

with $\beta$ being a vector of coefficients. We find that there is significant variation in the CDS spreads over time that is not captured by bank specific variables. We model this by allowing the intercept, $\alpha_t$, to vary over time (time fixed effects). The time-varying intercept accounts for differences in the level of the CDS spread that is common to all banks and obviates the use of macroeconomic control variables. As explanatory variables, $x_{it}$, we use the same control variables described in the previous section. One potential advantage with using panel data is that including firm fixed effects can considerably mitigate the problem of unobserved heterogeneity. We have, however, a sample that is much smaller in the time dimension than in the cross-section, and we also expect there to be a lot more variation across banks than across time for a given bank. Because of this, we abstain from using firm fixed effects but instead use fixed effects for different geographical regions; see the “Endogeneity” section for further details.

Modeling Credit Ratings
As is common in the credit rating literature (see, e.g., Blume, Lim, and MacKinlay, 1998), we use an ordered probit model. We observe differences in the level of ratings between the years, just as we did with the CDS spread, and also here account for these
differences by using time fixed effects. We also use fixed effects for different geographical regions in the same manner as for the CDS spread.

Endogeneity

There are three issues of endogeneity to address in this study. The first is the issue of reverse causality, a case of simultaneity. We argue that default risk is a function of the degree of ERM implementation and that an increased degree of implementation should result in higher credit ratings and lower CDS spreads. However, a reverse argument could in fact be that the degree of ERM implementation is a function of default risk and that having high default risk, low credit ratings, and high CDS spread, should result in a need for more effective ERM. Evidence of this can be seen in the risk management determinant literature mentioned previously. The reverse causality argument predicts a positive relationship between the degree of ERM implementation and default risk, so this would bias the coefficient upward and against us finding a relationship. Finding a suitable instrument variable in order to correct for such an endogeneity problem is difficult and has its own set of disadvantages. However, we do identify a possible instrument and use two-stage least squares to alleviate this concern (see footnote 10).

Hoyt and Liebenberg (2011) use a treatment effects model since they appear to view nonrandom assignment of treated and nontreated firms (ERM implementers and nonimplementers) as their primary endogeneity problem. Unfortunately, the treatment effects model cannot solve the problem of omitted variable bias (often called unobserved heterogeneity) when the omitted variable is unobservable and can hence not be included in the equation that estimates the treatment. A common unobserved variable in corporate finance research is managerial quality. One could propose that banks with better management quality are more likely to implement higher degrees of ERM and would therefore have lower default risk, not because of the risk management but because of management quality. Since management quality is difficult to measure, it ends up missing from the model, and the degree of ERM implementation would therefore capture the impact of manager quality on default risk. We mitigate this omitted variable problem by including variables that proxy managerial quality. For example, ROA can give an indication on how efficient management is at using its assets to generate revenues. Additionally, the corporate governance variables reflect a bank’s management practices and efficiency. To further alleviate potential problems from omitted variables we take advantage of the panel data set that makes it possible to include cross-sectional fixed effects. We include fixed effects based on geographical region by including dummy variables for Australia, North America, and Asia. Europe is left out and its effect is hence captured by the intercept; Russia and South America have too few observations to be included as separate geographical regions and are therefore added to Europe.8

The final endogeneity problem is the potential measurement error in our measure of degree of ERM implementation. This results in attenuation bias that will bias

---

8Adding Russia to Asia and South America to North America or excluding Russia and South America from the data leaves the results virtually unchanged.
the coefficient toward zero. Therefore, this would bias us against finding a relationship.

**RESULTS**

**CDS Spreads and the Degree of ERM Implementation**

Descriptive statistics for the independent variables used in the CDS spread model can be found in Table 2. For detailed descriptive statistics for the degree of ERM implementation for the CDS sample, see Table 1.

The correlations, reported in Table 3, show that the degree of ERM implementation is negatively correlated with the CDS spread and that control variables, such as the ratio of nonperforming loans to total assets, show a strong positive correlation with the CDS spread. The two highest correlations among the explanatory variables are between the degree of ERM implementation and the number of words in the annual report (0.65) and between the two measures for bad loans (0.71). The high correlation between the number of words and the degree of ERM implementation is of potential concern; however, as we will see from the regression results, while both variables explain the CDS spread, only the degree of ERM implementation is significant when both variables are included in the same regression.

Table 4 shows the results for the OLS regressions where the dependent variable is the CDS spread for the end of each year (2007–2011). Year fixed effects are included for all specifications. Since the residuals have positive skewness and normality is rejected, we base our inference on bootstrapped standard errors.

In specification (1), we estimate CDS spreads as a function of solely the degree of ERM implementation. We find a significant and negative relationship. Higher degrees of ERM implementation result in lower CDS spreads, which is in line with our expectation. The magnitude of the degree of ERM implementation coefficient is −0.019, which is also economically meaningful. A one-standard-deviation increase in the degree of ERM implementation (10.01) lowers the CDS spread by 19.02 percent; since the CDS spread of the average bank is 264.31, this corresponds to a decrease of 50.3 bp.

In specification (2), we control for the number of words in the annual report in order to control for the potentially inflated number of hits due to a generally higher level of disclosure in certain banks. As the sole explanatory variable, number of words is a significant determinant of CDS spreads. However, when the degree of ERM implementation and number of words are included in the specification together, as in specification (2), number of words is no longer significant. This means that the common variation in the degree of ERM implementation and number of words is what explains CDS spreads and not the overall level of disclosure. The magnitude of the degree of ERM implementation coefficient is almost unchanged with the addition

---

9Year fixed effects (year dummies) are not presented to conserve space but can be obtained from the authors upon request.
Table 3
Variable Correlations

<table>
<thead>
<tr>
<th>Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Log CDS price</td>
<td>1.00</td>
<td>-0.15</td>
<td>-0.02</td>
<td>-0.23</td>
<td>-0.43</td>
<td>0.15</td>
<td>0.60</td>
<td>-0.25</td>
<td>-0.13</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>2. ERM</td>
<td>1.00</td>
<td>0.65</td>
<td>0.26</td>
<td>0.00</td>
<td>-0.03</td>
<td>-0.09</td>
<td>0.27</td>
<td>0.33</td>
<td>-0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. # words</td>
<td>1.00</td>
<td>0.44</td>
<td>-0.17</td>
<td>-0.04</td>
<td>0.11</td>
<td>-0.06</td>
<td>0.06</td>
<td>0.33</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Total assets (TA)</td>
<td>1.00</td>
<td>-0.16</td>
<td>0.05</td>
<td>-0.23</td>
<td>-0.13</td>
<td>-0.09</td>
<td>0.39</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Return on assets (ROA)</td>
<td>1.00</td>
<td>-0.13</td>
<td>-0.45</td>
<td>-0.38</td>
<td>0.30</td>
<td>-0.04</td>
<td>-0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Tier 1 ratio</td>
<td>1.00</td>
<td>0.11</td>
<td>0.07</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Nonperforming loans/TA</td>
<td>1.00</td>
<td>0.71</td>
<td>-0.15</td>
<td>-0.18</td>
<td>0.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Provision for loan losses/TA</td>
<td>1.00</td>
<td>-0.09</td>
<td>-0.10</td>
<td>0.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Audit committee independence</td>
<td>1.00</td>
<td>0.36</td>
<td>-0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Corporate governance score</td>
<td>1.00</td>
<td>-0.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Single biggest owner</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Variables from 2005 to 2011 year-end. Correlations computed on the average of the year by year. ERM is the degree of enterprise risk management implementation. Variable definitions can be found in the supplementary Internet material (Lundqvist and Vilhemsson, 2016).
<table>
<thead>
<tr>
<th>Explanatory Variable</th>
<th>(1) Coefficient</th>
<th>(2) Coefficient</th>
<th>(3) Coefficient</th>
<th>(4) Coefficient</th>
<th>(5) Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERM</td>
<td>-0.019***</td>
<td>-0.020***</td>
<td>-0.011***</td>
<td>-0.009***</td>
<td>-0.008**</td>
</tr>
<tr>
<td># words (in thousands)</td>
<td>0.000</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total assets (TA) (in billions)</td>
<td>-0.095***</td>
<td>-0.223***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Return on assets (ROA)</td>
<td>-0.105**</td>
<td>-0.087**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tier 1 ratio</td>
<td>-0.044**</td>
<td>-0.034</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonperforming loans/TA</td>
<td>0.008**</td>
<td>0.010**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provision for loan losses/TA</td>
<td>0.021**</td>
<td>0.018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audit committee independence</td>
<td>-0.006***</td>
<td>-0.007***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corporate governance score</td>
<td>-0.002</td>
<td>0.004*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single biggest owner</td>
<td>0.003</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Regional fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R²</td>
<td>44.9%</td>
<td>44.9%</td>
<td>69.0%</td>
<td>48.6%</td>
<td>70.8%</td>
</tr>
<tr>
<td># observations</td>
<td>213</td>
<td>213</td>
<td>197</td>
<td>172</td>
<td>162</td>
</tr>
</tbody>
</table>

Note: Coefficients from OLS panel regressions with heteroskedasticity (White) adjusted standard errors in parentheses and bootstrapped standard errors (10,000 repetitions) in square brackets. The dependent variable is the natural logarithm of the CDS spread at the end of each year (2007–2011). Nonperforming loans/TA and provision loan losses/TA are scaled by 1,000. ERM is the degree of enterprise risk management implementation. Control variable definitions can be found in the supplementary Internet material (Lundqvist and Vilhemsson, 2016). The intercept and fixed effects (year and regional dummies) are not presented to conserve space but can be obtained from the authors upon request.

Denotes significance at the 10% level, **5% level and ***1% level based on bootstrapped standard errors.
of number of words and still significant at the 1 percent level. Our measure of the
degree of ERM implementation therefore captures a dimension that is different from
general disclosure.\textsuperscript{10}

In specification (3), we add controls for bank characteristics and the risk taking of the
bank. For the degree of ERM implementation, the coefficient is smaller (−0.011) but
still significant at the 1 percent level.

We then test the corporate governance controls separately. The coefficient for the
degree of ERM implementation decreases only slightly in comparison to the previous
specification (−0.009) and is still significant. Only audit committee independence is
significant among the corporate governance variables.

In specification (5), we control for both the bank characteristics and corporate
governance and again get a significant (at the 5 percent level) and negative coefficient
for the degree of ERM implementation. The magnitude of the coefficient is
still economically important; a one-standard-deviation increase in the degree of
ERM implementation (10.01) decreases the CDS spread with 21 bp. Specification (5)
shows a relatively high coefficient of determination at more than 70 percent, and the
regional fixed effects are no longer jointly significant for this specification, indicating
that the set of control variables can explain most of the heterogeneity among the
banks.\textsuperscript{11}

To see if the results are largely driven by the inclusion of 2007, which has fewer
banks due to limited coverage in DataStream, we redo the CDS results with the
year 2007 excluded. The results are reassuringly similar when 2007 is excluded;
the significance of the degree of ERM implementation is almost completely
unchanged and the magnitude of the coefficient is slightly increased in all
specifications.

Credit Ratings and the Degree of ERM Implementation

Table 5 shows the results from the ordered probit regressions. The dependent variable
is the S&P credit rating at the end of each year (2005–2011). Marginal effects of
the coefficients are only reported for the degree of ERM implementation since this is

\textsuperscript{10}For robustness we also estimate a two-stage least squares specification using number of
words as an instrument for the degree of ERM implementation for all specifications. Number
of words is a significant determinant of the degree of ERM implementation (relevance
criteria), and we would not expect number of words to be a significant determinant of CDS
spreads (exclusion criteria). The magnitudes of the estimated degree of ERM implementation
coefficients are very similar to the OLS estimation. However, we are mindful of the problems
associated with weak instruments (Roberts and Whited, 2011) and therefore use the OLS
estimations as the main specifications.

\textsuperscript{11}Because the sample period includes the recent financial crisis, we test for an effect on credit
ratings and CDS spread of an interaction between the crisis period (2006 and 2007) and the
degree of ERM implementation for the full specifications. We do not find a significant
relationship between the interaction and CDS or credit ratings.
<table>
<thead>
<tr>
<th>Explanatory Variable</th>
<th>(1) Coefficient</th>
<th>Marginal Effects</th>
<th>(2) Coefficient</th>
<th>Marginal Effects</th>
<th>(3) Coefficient</th>
<th>Marginal Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERM</td>
<td>-0.032 (0.006)**</td>
<td>1.10%</td>
<td>-0.034 (0.007)**</td>
<td>1.07%</td>
<td>-0.020 (0.007)**</td>
<td>0.60%</td>
</tr>
<tr>
<td></td>
<td>-0.49%</td>
<td>-0.47%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.49%</td>
<td>-0.48%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.12%</td>
<td>-0.12%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># words (in thousands)</td>
<td>0.001 (0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total assets (TA) (in billions)</td>
<td></td>
<td></td>
<td>-0.328 (0.092)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Return on assets (ROA)</td>
<td></td>
<td></td>
<td>-0.090 (0.087)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tier 1 ratio</td>
<td></td>
<td></td>
<td>0.092 (0.036)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonperforming loans/TA</td>
<td></td>
<td></td>
<td>31.569 (5.344)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provision for loan losses/TA</td>
<td></td>
<td></td>
<td>61.728 (16.136)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># observations</td>
<td>442</td>
<td>442</td>
<td>388</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Coefficients from the ordered probit regressions with heteroskedasticity (White) adjusted standard errors in parentheses. The dependent variable is the S&P credit rating at the end of each year (2005–2011). ERM is the degree of enterprise risk management implementation. Control variable definitions can be found in the supplementary Internet material (Lundqvist and Vilhemsson, 2016). The intercept and fixed effects (year and regional dummies) are not presented to conserve space but can be obtained from the authors upon request. *Denotes significance at the 10% level, **5% level and ***1% level.
the variable of primary interest. Specifications are identical to those in the CDS model.

In specification (1), we use the degree or ERM implementation as the sole explanatory variable. The coefficient is significant at the 1 percent level. For interpretation we look at the marginal effects. The effect of the degree of ERM implementation for category one (AAA or AA rating) is that a one-unit increase in the degree of implementation increases the likelihood of having an AAA or AA rating with 1.10 percent, and it decreases the probability of having A, BBB, and <BBB ratings with 0.49 percent, 0.49 percent, and 0.12 percent, respectively (the changes in probability have to add to zero; any deviation stems from round off error). Therefore, a one-standard-deviation increase (10.96) in the degree of ERM implementation increases the probability of having an AAA or AA rating by roughly 12 percentage points. This result is in line with the overall expectation that higher degrees of ERM will result in higher credit ratings for banks. Controlling for the number of words or the general disclosure of the bank has very little impact on the results. When we control for bank characteristics (specification (3)), coefficient size for the degree of ERM implementation drops but maintains significance.

In specification (4) (see Table 6), we control for corporate governance characteristics of the bank. The degree of ERM implementation is no longer significant. Corporate

### Table 6
Credit Rating Sample Continued—Panel Regression

<table>
<thead>
<tr>
<th>Explanatory Variable</th>
<th>(4) Coefficient</th>
<th>Marginal Effects</th>
<th>(5) Coefficient</th>
<th>Marginal Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERM</td>
<td>-0.008 (0.007)</td>
<td>0.33%</td>
<td>-0.009 (0.011)</td>
<td>0.29%</td>
</tr>
<tr>
<td></td>
<td>-0.22%</td>
<td></td>
<td>-0.22%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.10%</td>
<td></td>
<td>-0.07%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.01%</td>
<td></td>
<td></td>
<td>0.00%</td>
</tr>
<tr>
<td># words (in thousands)</td>
<td>0.001 (0.002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total assets (TA) (in billions)</td>
<td>-0.494 (0.143)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Return on assets (ROA)</td>
<td>-0.096 (0.094)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tier 1 ratio</td>
<td>0.159 (0.046)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonperforming loans/TA</td>
<td>34.424 (8.343)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provision for loan losses/TA</td>
<td>37.843 (17.113)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audit committee independence</td>
<td>-0.010 (0.004)**</td>
<td></td>
<td>-0.016 (0.005)**</td>
<td></td>
</tr>
<tr>
<td>Corporate governance score</td>
<td>-0.013 (0.004)**</td>
<td></td>
<td>0.003 (0.006)</td>
<td></td>
</tr>
<tr>
<td>Single biggest owner</td>
<td>0.009 (0.005)*</td>
<td>0.011 (0.006)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td></td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Region fixed effects</td>
<td>Yes</td>
<td></td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td># observations</td>
<td>281</td>
<td></td>
<td>265</td>
<td></td>
</tr>
</tbody>
</table>

Note: See Table 5 notes.

---

Marginal effects from all coefficients and the estimated cutoff points can be obtained from the authors upon request.
governance variables are significant for the determination of credit ratings but not for the CDS spreads. The final specification includes all control variables; the degree of ERM implementation is not a significant determinant of credit ratings.

Why the Degree of ERM Implementation Affects CDS and Credit Ratings Differently

The results show that the degree of ERM implementation has a significant and negative relationship with CDS spreads but has no significant relationship with credit ratings. In being consistent with Hilscher and Wilson (2013), the CDS and credit rating results, taken together, suggest that higher degrees of ERM implementation lower default risk by primarily decreasing the probability of default and to a lesser extent by reducing systematic default risk. That the degree of ERM implementation is negatively related to the probability of default is in line with the theoretical argument that ERM should reduce the volatility of cash flows, agency risk, and information risk of a firm. And that ERM has no effect on systematic risk is not so surprising.

The lack of relationship in the credit rating sample may alternatively be because we are capturing an effect (or lack thereof) of the rating process. As a sole determinant of credit ratings, the degree of ERM implementation is significantly related to credit ratings, but this relationship disappears when the corporate governance variables are introduced. The CDS results and basic correlations do not suggest that the degree of ERM implementation and the corporate governance variables are capturing the same thing. However, in the credit ratings sample, the degree of ERM implementation and corporate governance seem to capture the same effect since adding corporate governance variables eliminates the effect of the degree of ERM implementation on credit ratings.

An explanation for this is that credit rating agencies view ERM as primarily a corporate governance function. As reviewed in the “Default Risk and ERM” section, how S&P incorporates ERM into the rating process varies between corporates, financial institutions, and insurers. For corporates it is quite clear that their incorporation of ERM is directly related to their assessment of management and corporate governance. For example, in their announcement in 2008 they explicitly said that their assessment of “ERM [would] add an additional dimension to [their] analysis of management and corporate governance” (S&P, 2008, p. 2). On the other hand, insurers have specific assessment criteria for ERM where there seems to be a balance in terms of how they assess risk management and risk governance as well as a clear separation between risk governance and corporate governance. Financial institutions seem to be somewhere in between; S&P clearly incorporates an assessment of risk governance in its ratings of financial institutions. However, how much S&P differentiates risk governance and corporate governance is not as clear as it is for corporates (not at all) and insurers (differentiate). Given the seeming lack of updates in S&P material in terms of how it incorporates ERM into its ratings of financial institutions, we would argue that it is possible that they are closer to corporates in the spectrum of differentiation. Therefore, the corporate governance variables and the measure of the degree of ERM implementation, in terms of the credit rating process, may measure essentially the same thing.
CONCLUSIONS

This study provides initial evidence of the effect of the degree of ERM implementation on the amount of default risk in a firm. We construct a novel measure of the degree of ERM implementation by using text-based searches of annual reports for word combinations related to a number of dimensions of ERM. To confirm the validity of our measure of the degree of ERM implementation, we conduct a difference-in-difference study comparing the change in the measured degree of ERM implementation for banks with a major change in their risk management (the treatment sample) to banks without such an event (the control sample); we use the hiring of a CRO as the treatment effect. We show that a significant risk management change has a large and statistically significant impact on our measure of the degree of ERM implementation.

We then estimate the relationship between the degree of ERM implementation and two proxies of default risk: the year-end credit rating and the year-end CDS spread. We do this for a sample of the largest banks in the world whose default risk and ERM implementation are generally closely followed.

We find evidence that higher degrees of ERM implementation are negatively related to the level of default risk, or the risk a bank’s creditors face, as measured by CDS spreads. However, we find that the degree of ERM implementation’s relationship with credit ratings is insignificant when controlling for governance characteristics.

We believe that reduction of default risk is one way to measure the success of ERM. This also suggests that it should continue to be a focus of rating agencies and banking regulation. However, default risk is only a small piece of the value creation in a firm. Therefore, while ERM may increase value through a decrease in default risk, it may have negative implications that outweigh these positive effects. This is therefore a starting point, and we suggest that the different pieces of the value puzzle become the future focus of research in ERM.

REFERENCES


Evaluating ERM CriteriaCOMM E 2007MAY_Note2014MAY.pdf


SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher’s Web site.