Risk and Ambiguity in 10-Ks: An Examination of Cash Holding and Derivatives Use*

Richard Friberg †
Stockholm School of Economics and CEPR

Thomas Seiler ‡
Stockholm School of Economics

May 26, 2017

Abstract

We explore the role of ambiguity, as opposed to risk, in explaining firms' corporate financial policies. We create text based measures of ambiguity and risk for U.S. firms between 1995 and 2013. Measured ambiguity is high in for instance high tech industries, whereas the risk measure is high for homogeneous goods. Using within-firm variation to identify effects we find that greater ambiguity is associated with greater cash holdings and more risk with a higher probability of derivatives use. The results are in line with a simple model of liquidity management with ambiguity averse agents.

Keywords: Cash holdings; Hedging; Knightian Uncertainty; Ambiguity; Risk; Textual Analysis

JEL: D81; G31; G32.

*Financial support from the Jan Wallander and Tom Hedelius Foundation and the Ann-Marget och Bengt Fabian Svartz stiftelse is gratefully acknowledged. We thank Niklas Amberg, Ruben Andrist, Marcus Asplund, Jimmy Martinez-Correa, Imran Rasul, Martin Seiler, Jorge Mario Uribe-Gil, seminar participants at Copenhagen Business School, participants at the ENTER Jamboree 2016 in Madrid, AEFIN 2016 conference in Madrid and national conference in Gothenburg for helpful comments. We thank My Hedlin for excellent research assistance.

†Corresponding author: Stockholm School of Economics, P.O. Box 6501, 11383 Stockholm, Sweden. Email: Richard.Friberg@hhs.se

‡Stockholm School of Economics, P.O. Box 6501, 11383 Stockholm, Sweden. Email: thomas.seiler@phdstudent.hhs.se

1 Introduction

Is the distinction between uncertainties with known probabilities, *risk*, and uncertainties with unknown probabilities, *ambiguity*, useful for understanding corporate liquidity policies? In this paper we use textual analysis to explore the impact of ambiguity and risk on firms' liquidity management. The notion that we can usefully distinguish risk from ambiguity(Knightian uncertainty), has deep roots stretching back to Knight (1921), but is mute in the standard subjective expected utility framework formalized by Savage (1954). The potential relevance of the distinction is nevertheless hard to dismiss given that a large body of experimental evidence, following Ellsberg (1961), indicates that many of us prefer situations with known probabilities to situations with unknown probabilities. Following the experimental literature, decision-theoretic models of choice in ambiguous situations has been a very active research area.[1] The models have found many applications in theoretical studies of corporate finance and macro-finance, as exemplified by Nishimura and Ozaki (2007) and Bidder and Dew-Becker (2016). Yet, the richness of theoretical and experimental work on ambiguity contrasts with a paucity of empirical work outside the laboratory.

It is well established that cash holdings by U.S. firms have risen sharply in the last decades and the evidence indicates that cash holding is positively related to uncertainty, when broadly defined. Holding cash to provide liquidity is versatile in the sense that it covers liquidity needs also in states that are hard to foresee or whose probability is hard to gauge. We may therefore expect that cash is suited for ambiguous situations. In the case of risk we expect there to be derivative products that can partially be used to provide liquidity. If it is costly to hold cash a firm will thus prefer to use derivatives to help provide liquidity in the case of risk. Intuitively we may therefore expect that more ambiguity raises cash holdings and more risk increase the likelihood of a firm using derivatives. We contribute to the literature by putting this intuition on a firmer ground. We first show how these intuitions can be formalized by adding a slight twist to a classic model of liquidity holdings due to Holmström and Tirole (1998) and Tirole (2006) as developed in Almeida et al. (2014). We then proceed to generate *firm-level* measures of ambiguity and risk to show that the type of uncertainty a firm faces is crucial to understand its liquidity policies. By providing empirical measures of risk and ambiguity, we help bridge the gap between the rich theoretical literature on ambiguity and the empirical analysis of liquidity management. Our results are supportive of the notion that ambiguity is a

[1] See Etner et al. (2012) for a survey of the literature.

relevant factor to understand some aspects of firm behavior and that textual analysis is a useful way to distinguish risk from ambiguity.

Our measures of risk and ambiguity are based on textual analysis of all 10-K statements for U.S. firms between 1995 and 2013. As our starting point, we take the list of "uncertainty" words compiled by Loughran and McDonald (2011) and classify terms into sub-categories. Uncertainty words associated with objective probabilities, like "variance," "volatility" or "frequently," we classify as *risk*. Those terms explicitly associated with subjective probabilities (e.g. "believe," "perhaps"), ambiguous outcomes (e.g. "ambiguous," "indeterminate") or Donald Rumsfeld's memorable term "unknown unknowns" (e.g. "sudden," "unforeseen") are classified as *ambiguity*. We use the occurrence of these words in firms' annual reports to define indices of whether the firm is rather presented as a risky or an ambiguous investment and combine them with data from Compustat. In the resulting panel dataset, we find that firms ranking higher on our ambiguity scale are more likely to be financially constrained and hold more cash. In contrast, we do not find such effects for our measure of risk. We also find firms more exposed to measurable risks to be more likely to use financial derivatives for hedging. These results hold controlling for common empirical determinants of corporate liquidity policies, as well as firm and time fixed effects.

To corroborate our results, we run a series of validation and robustness tests. First, we examine the textual data and find that the sentences containing our uncertainty terms vary considerably over the years: with an average cosine similarity of 0.2 between uncertainty sentences in different years, it becomes unlikely that we are primarily capturing boiler plate information in the annual reports. Second, we establish that our indices exhibit intuitive cross-sectional industry patterns. We find firms ranking high on our ambiguity index to be clustered in high-tech industries, like electronic or medical equipment. High risk firms are centered in competitive industries with relatively homogeneous goods, like production of paper and steel. Third, we address the question whether our ambiguity index primarily differs from the risk index in terms of duration and severity of the uncertainty captured. We estimate an ordered Probit model on S&P long-term ratings and find the marginal effects to be statistically equivalent. Simulating the rating distribution after a one standard deviation increase in risk or ambiguity, we find the estimated effects also to be economically meaningful. Fourth, we also look at the relation between our measures and financial constraints as measured by Hoberg and Maksimovic (2015). Our model predicts a clear positive relationship between constraints and ambiguity, but no such correlation between risk and constraints, which is what we also find in the data. In the Appendix, we report the results of an additional validation exercise, where we examine the relation between our indices and market risk.

On the robustness side, besides estimating a variety of specifications, we investigate potential reverse causality issues in detail. Firms that are more likely to use hedging to secure their liquidity needs might naturally use more risk related words in their annual reports. We find little evidence for this reverse causality channel to be empirically important.

The structure of our paper is as follows. The next section treats the related literature in more detail and in Section 3 we formulate our hypotheses on cash holdings under risk and ambiguity. In Section 4 we then detail the construction of our data set, the measures of risk and ambiguity and the measures of financial constraints used in our paper. We also describe our empirical strategy. In Section 5 we examine our measures of ambiguity and risk and explore their relation to financial constraints. In Section 6 we present our results on cash holdings and hedging and we conclude in Section 7.

2 Related literature

Cash holding and financial constraints. Bates et al. (2009) show that cash-to-asset ratios for U.S. industrial firms more than doubled between 1980 and 2006 and that the increase was most marked for firms in industries that had higher cash-flow volatility. The pattern of increased cash holdings is seen in many other developed countries (Iskandar-Datta and Jia (2012)) and the results in Pinkowitz et al. (2016) indicate that U.S. firms hold no more cash than comparable foreign counterparts. The literature has put forward a variety of explanatory factors for cash holdings (see Almeida et al. (2014) for a survey). In this paper we emphasize Keynes's 1936 precautionary savings motive, though more recently a variety of other explanatory factors have been suggested. Alimov (2014) show that increased product market competition increases cash holdings; Dudley and Zhang (2016) emphasize the role of social trust in firms' home markets; Fernandes and Gonenc (2016) look at the role of international and industry diversification; Liu et al. (2015) investigate ownership structure as an explanatory factor. Jiang and Lie (2016), following a theoretical argument by Jensen (1986), look at the role of entrenched managers. He and Wintoki (2016) argue the increasing sensitivity of R&D spending to cash holdings, combined with the secular increase in R&D spending, might explain higher cash holdings.

In the modeling framework that we use, the value of precautionary savings build on the logic suggested by Froot et al. (1993): firms want to manage liquidity in such a way that they are able to finance investments and ensure survival also in adverse states of the world. We thus relate to work that also examines other ways of managing liquidity such as bank credit lines (Sufi (2009)), hedging (Disatnik et al. (2014) and debt maturity (Harford et al. (2014)). They are all tools for ensuring access to funds in different states of the world, but cash holdings have a particular strength in that they provide liquidity in a relatively large set of states. For instance, the value of bank credit lines hinge on the financial health of the bank and on the bank's continued willingness to lend to the particular firm, which is not true for cash.

Ambiguity and finance. The thought experiments in Ellsberg (1961) suggested that many people are ambiguity averse - preferring to bet on the color of a ball from an urn where they know the probability (known proportions of colors) from betting on the color of a ball drawn from an urn where they don't know the probability (unknown proportions of colors). The article triggered a large number of actual experiments and, as summarized in an influential survey by Camerer and Weber (1992), the resulting evidence supported the idea that many individuals are averse to ambiguity, preferring risky choices to ambiguous choices.

The experimental evidence on ambiguity aversion has triggered substantial theoretical interest in choice by ambiguity averse agents (see e.g. Gilboa and Schmeidler (1989), Maccheroni et al. (2006)) and the decision criteria have in turn been applied to questions in financial markets. Some explore the implications of ambiguity averse firms, and Nishimura and Ozaki (2007) for instance model the case where more ambiguity leads a firm to delay investment. However, the bulk of the modeling efforts has been directed at ambiguity averse investors. Taken together the theoretical literature shows that ambiguity aversion can have large effects on model predictions in financial markets and help explain issues like the equity premium puzzle, limited stock market participation and home bias in portfolio choice (see Epstein and Schneider (2010) or Guidolin and Rinaldi (2013) for overviews).

Some experimental work in the lab support the predictions from applications of the above class of models (see e.g. Bossaerts et al. (2010) or Füllbrunn et al. (2014)) but evidence from the field is sparse. A rare example of an article that examines choice in the field is Dimmock et al. (2016) who measure ambiguity aversion with questions administered via the RAND corporation's American Life Panel of U.S. households and then relate the ambiguity aversion to actual portfolio choices, again finding qualitative predictions in line with the models of ambiguity inverse investors.

We relate closely to a sprinkling of recent articles that examine links between ambiguity aversion and cash holding. Neamtiu et al. (2014) interpret economy wide measures of variance premia and analysts' forecast dispersion as measures of macroeconomic ambiguity and find that macroeconomic ambiguity is positively associated with cash holding. Breuer et al. (2016) lay out a model in which higher ambiguity aversion by investors leads to *lower* cash holding. The intuition is that ambiguity aversion decreases the optimal level of investment in the long run and tends to make cash holding superfluous. To test this prediction empirically they turn to cross-country data and use results from surveys of students across many countries as a measure of country specific average ambiguity aversion. They interact this measure of ambiguity aversion with cash holding and control for other country and firm specific factors and interpret their results as supporting their hypotheses. We find their results intriguing but also note that unobserved heterogeneity is a concern when identification comes from cross-country differences. The model in the present paper has the opposite implication for cash holding and depending on which channel that is emphasized the theoretical predictions may thus differ. Indeed, in the model of Agliardi et al. (2016) ambiguity aversion has a nonmonotonic effect on cash holding, first decreasing and then increasing. This suggests a need for empirical work and let us turn to this now.

Measuring risk and uncertainty using text analysis. We relate to a number of articles that create firm level measures of risk building on 10-K statements. Most text based studies focus on companies' voluntary and involuntary disclosure and generate measures for its quantity, quality and content i.e. how much is reported, how accessible or readable is it and what is its general tone or sentiment. The obtained measures are then connected either to accounting data or to market returns. Combined Cole and Jones (2005), Li (2010) and Kearney and Liu (2014) provide a comprehensive overview to this literature. Prior to the availability of online corpuses, manual analysis is reviewed by Jones and Shoemaker (1994).

A key reference for text analysis of U.S. annual reports is Loughran and McDonald (2011). Their focus is on creating high quality word lists well suited for use in the analysis of financial text and their indices are also available on WRDS. They note that many word lists developed for general purposes are not well suited for financial texts and we take their word lists as our starting point.

Another important building block for our approach is the assumption that corporate risk disclosure is actually informative on firms' risk environment. A number of recent studies looked into this issue. Kravet and Muslu (2013) find risk disclosure to be statistically and economically significantly correlated with financial market indicators and analyst forecast dispersion. Campbell et al. (2014) examine the information content of mandatory risk disclosure by extending the manual classification scheme of Nelson and Pritchard (2007). They find firms with larger risk disclosure are also more risky. Moreover, if a firm is more exposed to a certain risk type, it will also disclose more information about that specific risk. Bao and Datta (2014) apply unsupervised learning algorithms to the analysis of corporate risk disclosure and find some evidence that investors react to the type of risks disclosed.

In research complementary to ours, Avramov et al. (2016) show that firm-level risk shocks (based on textual analysis of Form 10-K) are associated with effects on a number of variables such as leverage, investment and cash holdings. A crucial difference with respect to our study is that their definition of risk differs substantially from ours and focuses on negative shocks. Their dictionaries classify words like “loss” and “negative” as risk words. This is consonant with one tradition which specifies risk as outcome times probability. In contrast, we are interested in risk in the sense of a mean-preserving spread and follow Loughran and McDonald (2011) in classifying “first moment” words as negative (or positive) terms and use these measures as controls in our regression analysis.[2] Also related is Hoberg and Maksimovic (2015) who use textual analysis of 10-Ks to create a measure of product market “fluidity” which aims to capture product market threats through channels such as entry and product innovations. They establish that firms faced with a more fluid environment hold more cash and are more conservative in their management of liquidity.

3 Development of hypotheses

To fix ideas we introduce ambiguity aversion into the workhorse model of cash holding due to Almeida et al. (2014), who in turn build on Holmström and Tirole (1998) and Tirole (2006).

The setup of Almeida et al. (2014) is the following: At time 0, a firm has assets A and can invest an amount I in a business opportunity. At time 1 there is a liquidity shock ρ with probability λ. If the shock occurs, the firm needs to pay ρI to keep the investment alive. Should the liquidity needs not be met, the project will be terminated which results

[2] In a previous version of their work, Avramov et al. (2014) also include a measure of ambiguity (uncertainty in their terminology). This measure also differs substantially from our measure of ambiguity, theirs cover eight words related to uncertainty and almost exclusively covers words which capture surprises, such as “unexpected” and “unforeseen”.

in a return of 0. If the project is continued, it will produce a return RI with probability p_i. The firm can affect the probability of success of the project by either being diligent or not. If it is diligent, the probability of success is p_G, otherwise the probability of success will be lower at p_B but the firm then receives a private benefit B. The firm can borrow money at any time from outside investors, but it might be credit constrained due to the moral hazard problem. It is assumed that investors want to induce high effort in the firm and that in this case, the net present value would be positive, even after a liquidity shock.

To examine the potential effects of ambiguity aversion we make a minimal extension to the setting above and assume that there is not a unique probability p_G of a good state if the firm is diligent, but instead that each investor believes that the success probability in the high effort case belongs to a convex set of probabilities P. For concreteness we let the set P be defined as $P = [p_G - \epsilon, p_G + \epsilon]$. We can then think of a higher ϵ as reflecting higher ambiguity. For instance if $p_G = 0.8$ and $\epsilon = 0.1$ investors would believe that the probability of a successful investment would be bounded by $p = 0.7$ and $p = 0.9$.

Extending Savage's framework to include an axiom of ambiguity aversion, Gilboa and Schmeidler (1989) show that an individual under their axioms would follow a maximin expected utility rule. That is, such an individual would choose the action that yielded the highest expected utility under the most pessimistic probability p from the set P. She would thus strictly prefer action f over action g if and only if:

$$\min_{p \in P} E_p(f) > \min_{p \in P} E_p(g). \tag{1}$$

We develop the model in more detail in Appendix A but here simply note that the combination of a binding participation constraint for investors, and a binding incentive constraint for the firm, plays a key role in determining investment. The firm's investment at time 0 is I and the expected outlay at time 1 is $\lambda \rho I$. Investment is limited by the initial assets of the firm A in addition to the funds that the firm can raise from outside investors. The income that the firm can pledge to investors in the case of success is determined by the return if investment is successful (R) minus the necessary return to the firm to keep it diligent ($B/(p_G - p_B)$). With maximin expected utility investors weigh this return (if the investment is successful) by the most pessimistic probability within the probability set, $(P_G - \epsilon)$. We denote the income that the firm can credibly pledge to investors by ρ_{0AMB} and specify the key constraint in Equation (2),

$$I + \lambda\rho I \leqslant A + \underbrace{(p_G - \epsilon)(R - \frac{B}{p_G - p_B})}_{\equiv \rho_{0AMB}} I. \tag{2}$$

Assuming a binding constraint we solve for the optimal investment:

$$I^*_{AMB} = \frac{A}{1 - \rho_{0AMB} + \lambda\rho}. \tag{3}$$

It is evident that pledgeable income ρ_{0AMB} is decreasing in ambiguity ϵ and we can thus state the first result that we will explore empirically:

Hypothesis 1 Financing constraints are more severe for firms that are seen as more ambiguous. This is directly established by increasing ϵ in equation (2).

If the firm is hit by a liquidity shock in period 1 we assume that the necessary follow up investment (ρI) is greater than pledgeable income. In this class of models the firm therefore needs to keep liquidity in order to be able to invest additional funds into the project and the cash holdings of an ambiguous firm are given by:

$$C^*_{AMB} = (\rho - \rho_{0AMB}) I^*_{AMB}. \tag{4}$$

For a given investment, the higher the possible liquidity needs, the more liquidity is kept. We can directly turn to a second hypothesis that we will examine:

Hypothesis 2 Conditional on the size of the investment, firms seen as more ambiguous will hold more cash. This is directly seen from equation (4) and noting that ρ_{0AMB} decreases as ϵ increases.

Holding cash to provide liquidity is versatile in the sense that it covers liquidity needs also in states that are hard to foresee or whose probability is hard to gauge, that is, cash is suited for ambiguous situations. In the case of risk we expect there to be derivative products that can partially be used to provide liquidity. If the firm can earn a higher return on its cash holdings by not merely keeping them as a reserve it will prefer to use derivatives to help provide liquidity in adverse states. We thus predict that firms faced with more risk will use derivatives more and cash less to ensure sufficient liquidity (and if there is some fixed cost of using derivatives be *more likely* to use derivatives). In Appendix A we provide an extension to the model where we operationalize higher risk as a mean preserving spread of R to establish the final hypothesis that we examine:

Hypothesis 3 Firms faced with higher risk will use derivatives rather than holding cash to meet higher liquidity needs.

Clearly, the development of hypotheses above should not be seen as an exhaustive theoretical investigation of how ambiguity aversion can affect cash holdings. Rather, by marrying a convenient formalization of ambiguity aversion with a standard model of liquidity management we want to link ambiguity aversion to empirical predictions on cash holding in a transparent way.

Let us also comment on that we chose to introduce ambiguity aversion on the part of investors, rather than on the part of management. As we'll discuss in Section 6.1 this modeling choice makes the interpretation of empirical results more straightforward and also fits with evidence that many CEOs are overconfident (see e.g. Malmendier and Tate (2005)). In Appendix A we show that the hypotheses above, however, go through in the model if we also introduce ambiguity aversion on the part of the firm.

4 Data and empirical strategy

4.1 Data sources

We combine data from several sources. The indices on ambiguity and risk are based on Form 10-K. We download all 10-K and 10-KSB filings starting January 1 1994 and ending December 31 2014 via the SEC EDGAR database. We clean out the filings to isolate text following the recommendations by Loughran and McDonald (2011). We disregard all appendices and focus only on actual Form 10-K text, more details on how we set up the text analysis is provided in the Appendix and we describe below how we construct text based measures of uncertainty.

Using the original filing names on the EDGAR server, we match our data with the text based measures of Loughran and McDonald (2011) available on WRDS. In addition to the text data we construct, we can make use of their positive and negative word counts. Using the SEC Analysis Suite linking table via WRDS we match our text-data with both the Compustat and CRSP universe.[3] All variable definitions are discussed in the Appendix except those that are of particular importance for our work.

We follow Bates et al. (2009) and calculate a measure of cash flow variability at the industry level. For each firm and fiscal year we calculate the standard deviation of cash flow to assets for the previous 10 years (we require at least three observations per firm to calculate this firm specific volatility). We take the average of these firm specific volatilities within each industry (two digit SIC) and fiscal year and term it "industry sigma". We denote a firm as a derivatives user in a given fiscal year if either of the Compustat variables AOCIDERGL or HEDGEGL is different from zero. The first of these measures unrealized gains and losses from financial hedging which offset variations in future cash flows. This variable is available from 2001 onwards and captures commodity, foreign exchange rate and interest rate hedges designated as cash-flow hedges. This has previously been used to create a dummy variable for hedging in Chang et al. (2013). The second measure captures gain or loss on ineffective hedges and has been used by for instance MacKay (2015). To construct our beta measures, we obtain financial market data via CRSP at a daily frequency.[4]

Our final data covers the fiscal years 1995 to 2013. We drop observations with negative assets and sales. Not all variables are present for all firms and for the regressions on cash holdings and hedging in section 6 we exclude financial firms (SIC 6000-6999) and utilities (SIC 4900-4999). These regressions are run on a sample of 61,330 observations. To limit the potential for measurement error to distort findings we winsorize our data at the 1% and 99% level. Following Bates et al. (2009) we winsorize leverage to be between zero and one. All nominal variables are deflated by CPI with base 2010.

Finally, in describing the cross-sectional patterns we also use some measures from other sources that can plausibly be linked to the degree of ambiguity of a firm. We use the measure of contract-specificity of intermediate inputs at the six-digit NAICS level from Nunn (2007). We use a dummy for high technology industries which builds on work by the Bureau of Labor Statistics as described in Hecker (1999). The data on financial constraints comes from Hoberg and Maksimovic (2015).

[3] We drop links flagged to be weak (flag variable in linking table not equal 2 or 3). Where multiple links exist, we just keep the strongest. We are able to match a total of 154,748 observations.

[4] We calculate fiscal-year matched betas using holding period returns (RET). We match this with the Fama/French return series data provided via WRDS: we use the one month treasury bill return as the risk free rate and the value-weighted return on all NYSE, AMEX and NASDAQ stocks as market portfolio. We also run a three-factor model using the small minus big (SMB) and high minus low (HML) portfolio return series.

4.2 Measures of risk and ambiguity

As our starting point we use the word list that Loughran and McDonald (2011) have identified with *overall uncertain* tone. There are 298 such words. Both authors independently categorized the words on that list as corresponding to either *risk* or *ambiguity* based on the principles described below. The classifications were then compared and discussed. When we couldn't clearly classify a word in either category, we left it unclassified. For example, the word "risk" is only included as an overall uncertainty word since it it used in a broad sense in the instructions for the 10-K: under the heading "Item 1A: Risk Factors", firms are instructed to discuss issues that pertain both to risk and ambiguity. Needless to say, the classification was done prior to any analysis. The word lists that we create are reproduced in table A.4 in the Appendix.

We classify an 'overall uncertainty' word as a *risk word* if it: captures frequency (often, frequently,...), refers to statistical terms, (probability, volatility,...) or captures minor imprecision (approximate,...). We thus aim to provide a theoretically grounded text based meaure of risk which captures that risk concerns objective probabilities and moments of a probability distribution. We hypothesize that terms such as these are used to describe situations amenable to statistical analysis and which we can describe as risk, somewhat broadly defined.[5] There are 74 risk words in our word list.

We categorize an overall uncertainty word as an *ambiguity word* if it clearly expresses subjective probability (believe, perhaps,...). In the case of a single decision maker, a clearly subjective probability is one where there is no objective probability distribution and where ambiguity aversion thus can matter for decision making. As noted in the introduction, in the framework of Savage (1954) there is no difference between a subjective and an objective probability[6] but when we distinguish decisions for which there is an objective probability distribution (risk) from those where we do not have that it is clear that we want to classify expressly subjective probabilities as ambiguous. Furthermore, almost by definition, subjective probabilities can differ between different observers. The theoretical literature on ambiguity focuses on individual choice whereas corporate and investment decisions typically involve several decision makers. Crès et al. (2011) argue that several decision makers who differ in their subjective probabilities imply that corporate decision problems might have to be characterized using sets of probability distributions as

[5]See Friberg (2015) for a book length treatment

[6]More specifically this refers to the typical interpretation of Savage (1954). The original text is actually careful to note that the use of subjective probabilities is limited to what Savage calls "small worlds", relatively simple, isolated decision problems. See Binmore (2008) for a discussion.

in Gilboa and Schmeidler (1989), even though each individual might have a well defined subjective probability for each state of the world.

We also categorize words as ambiguity words if they clearly describe ambiguity or hard to interpret observations (ambiguous, indeterminate,...). This is close to the interpretation of ambiguity in the spirit of Ellsberg (1961). Finally, one sense of the notion of ambiguity is that probabilities are hard to define because it is hard to determine the possible states-of-the-world ex ante. In theoretical work this has been analyzed as unforeseen contingencies (Dekel et al. (1998)). This flavor of ambiguity is also related to what Donald Rumsfeld termed "unknown unknowns": things we don't know that we don't know. We therefore also define words capturing surprises (sudden, unforeseen,...) as ambiguity. There are 127 words on our ambiguity word list.

For each firm i in fiscal year t we define measures of risk, ambiguity and overall uncertainty as the share of risk, ambiguity or uncertainty words in the document. We examine shares rather than levels, because financial reports have become longer over time. It is thus important to scale the number of risk and ambiguity words by a measure for document length. As an alternative specification we also define sentence-based indices as the share of risk, ambiguity or uncertainty related sentences in a document. We define a sentence to relate to ambiguity or risk if it contains at least one word from either dictionary. If a sentence contains words from both dictionaries, we will count it twice. In the next section we run several tests to validate our indices.

5 Ambiguity and risk: description and validation

In this section we take a first look at our data and then turn to a variety of validation tests for our indices, as indicated in the Introduction. We show that the text-base on which our indices are built, varies considerably within firms over the years and also document intuitive cross-sectional patterns. We also examine the relation of our uncertainty indices to classical risk measures. The findings strongly suggest that our measurement approach indeed captures important aspects of firm uncertainty. As a last sanity check of our data we find, in line with the theoretical model, that there is a positive relation between ambiguity and the tightness of financial constraints.

5.1 Summary statistics and cross-secional patterns

We present descriptive statistics on some key variables in Table 1. On average 1.3 % of words are overall uncertainty words following the classification of Loughran and McDonald (2011). The shares of ambiguity and risk words are similar on average, amounting to some 0.3 % of words. Year-to-year changes in our measures of ambiguity and risk will be key to identifying effects and we see that, while the mean and median changes are close to zero, there is substantial variation in both indexes at the firm level, implying both negative and positive changes that help identify the possible effects. In robustness exercises we define ambiguity and risk at the sentence level, again the share of sentences that contain ambiguity and risk words are similar. The share of negative words is about twice as high as the share of positive words. Finally, average cash holdings amount to 19.6 % of assets and on average 21 % of firms use derivatives. Descriptive statistics on our other variables can be found in Table A.1 of the Appendix.

Table 1: Descriptive statistics for some key variables.

Variable	mean	sd	p5	p50	p95	N
Overall Uncertainty	1.313	0.333	0.769	1.313	1.862	61,330
Ambiguity	0.321	0.118	0.136	0.317	0.526	61,330
Risk	0.340	0.122	0.166	0.326	0.568	61,330
Ambiguity$_{it}$ -Ambiguity$_{it-1}$	0.009	0.057	-0.073	0.005	0.102	61,330
Risk$_{it}$- Risk$_{it-1}$	0.001	0.067	-0.098	0.000	0.107	61,330
Ambiguity (sen)	6.434	2.488	2.632	6.287	10.829	61,330
Risk (sen)	6.587	2.210	3.295	6.364	10.714	61,330
Negative	1.589	0.456	0.843	1.585	2.327	61,330
Positive	0.696	0.178	0.451	0.674	1.009	61,330
Cash ratio	0.189	0.220	0.003	0.099	0.694	61,330
Derivatives	0.221	0.415	0.000	0.000	1.000	42,287

This table reports summary statistics on the sample of observations used in the cash holding regressions reported in Table 8. All 10-K and 10-KSB filed for fiscal years 1995 up to and including 2013 available on EDGAR were downloaded and the text based measures reported in the table are based on textual analysis of these documents. The text based measures are: overall uncertainty, ambiguity, risk, positive and negative which are based on share of words (percent) and ambiguity (sen) and risk (sen) the corresponding shares based on sentence level shares. Other data is from Compustat: Cash ratio (CHE/AT) and derivatives is dummy that is 1 if unrealized gains or losses from financial hedging are reported (AOCIDERGL) and/or if ineffective hedges are reported (HEDGEGL) and 0 otherwise. Derivatives dummy only defined from 2001 onwards. Based on sample in cash holding regressions; this excludes financial firms and utilities. Corresponding summary statistics of other variables of interest are reported in Table A.1 in the Appendix.

To gain some intuitive understanding of the indices we turn to an exploration of cross-sectional patterns regarding uncertainty words. In Table 2 we present a ranking using the Fama-French 48 Industry classification. Highly uncertain industries are found in high tech industries such as those related to pharmaceuticals and electronic equipment and also industries that are exposed to political risks, like tobacco products or petroleum and natural gas. Among the industries with the highest ambiguity values we find a strong representation of high tech industries again such as pharmaceutics and electronic equipment. These are indeed common examples of markets where drastic innovations are important (see e.g. Arrow (1962)), where competition is *for* the market rather than *in* the market (Geroski (2003)) and where disruptive innovation exerts a strong influence (Christensen (1997)). These industries can then arguably be characterized by a tendency for new innovations to crowd out existing products. Ambiguity is thus expected to be relevant to describe the randomness.

High risk industries are centered around relatively homogeneous products that rely on established technologies. The Fama-French industries that rank highest on the risk index include production of products such as paper (business supplies), paperboard containers (shipping containers) and primary metal production (steel works etc). The role of drastic innovation in these industries is less pronounced and randomness is instead to an important degree likely to be linked to fluctuations in market prices.

Table 3 provides a more systematic summary of the cross-sectional patterns in the indices; we examine the pairwise correlation between the various measures of risk and uncertainty and some firm and industry characteristics. The correlation between our uncertainty measures is positive but only .09 for the correlation between ambiguity and risk. The cash ratio is positively correlated with our measure of ambiguity but negatively with our measure of risk. In contrast, derivatives use is negatively related to ambiguity but positively to risk.

The ranking of industries above suggested that high tech industries might be more prone to experience ambiguity and indeed the correlation is positive between our measure of ambiguity and a dummy for high tech industries. Finally, for a limited set of industries we have a measure of contract specificity from Nunn (2007), which shows a positive correlation with ambiguity but negative with risk. Such a pattern is to be expected if transacting in a market implies more volatile prices whereas closer cooperation with suppliers implies a role for strategic interaction, and hence the potential for greater ambiguity.

Table 2: Industry rankings in terms of our overall uncertainty, ambiguity and risk indices.

	Overall uncertainty		**Ambiguity**		**Risk**	
Rank	*Industry*	*Avg*	*Industry*	*Avg*	*Industry*	*Avg*
1	Pharmaceutical Products	1.44	Electronic Equipment	0.36	Business Supplies	0.39
2	Electronic Equipment	1.38	Medical Equipment	0.35	Shipping Containers	0.38
3	Coal	1.36	Business Services	0.34	Steel Works Etc	0.37
4	Medical Equipment	1.35	Pharmaceutical Products	0.34	Apparel	0.36
5	Tobacco Products	1.34	Computers	0.34	Coal	0.35
6	Business Services	1.34	Coal	0.33	Fabricated Products	0.35
7	Petroleum and Natural Gas	1.32	Measuring and Control Equipment	0.33	Shipbuilding, Railroad Equipment	0.35
8	Computers	1.31	Healthcare	0.33	Textiles	0.35
9	Non-Metallic and Inddustrial Metal Mining	1.30	Non-Metallic and Industrial Metal Mining	0.32	Construction Material	0.35
10	Measuring and Control Equipment	1.29	Recreation	0.32	Petroleum and Natural Gas	0.34

This table reports the average share of overall uncertainty words, ambiguity words and risk words for the 10 highest ranked Fama-French 48 industries in each category. All 10-K and 10-KSB filed for fiscal years 1995 up to and including 2013 available on EDGAR were downloaded and the text based measures reported in the table are based on textual analysis of these documents. Financial firms and utilities excluded and calculations based on average across 110,255 observations.

Table 3: Pairwise correlations between uncertainty measures and other key variables.

	Overall U.	Ambiguity	Risk	Cash Ratio	Derivatives	High Tech
Ambiguity	0.778	1.000				
Risk	0.436	0.091	1.000			
Cash Ratio	0.241	0.241	-0.090	1.000		
Derivatives	0.044	-0.057	0.173	-0.229	1.000	
High Tech	0.131	0.131	-0.063	0.327	-0.105	1.000
Contract Specificity	0.083	0.156	-0.089	0.275	-0.156	0.467

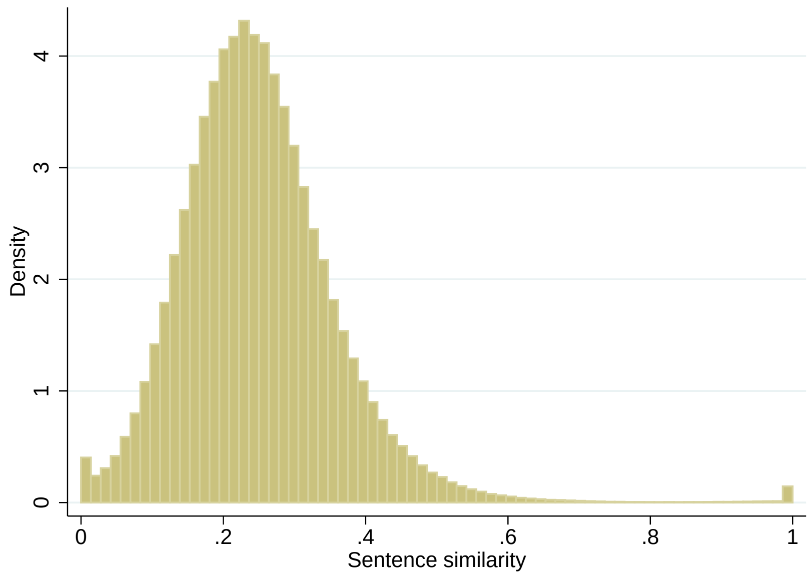
The table shows pairwise correlations between some key variables. All correlations are statistically significantly different from zero at the 1% level. Overall uncertainty, ambiguity, risk, positive and negative based on 10-K filings for fiscal years 1995 up to and including 2013. Cash ratio (CHE/AT) and derivatives dummy from Compustat (available from 2001 onwards). Financial firms and utilities excluded. The dummy for high tech industries builds on Hecker (1999)'s. Correlations with derivatives based on sample in Table 9, 45,774 observations. Contract specificity is from Nunn (2007) and builds on use of intermediate inputs at the six-digit NAICS, available for 20,494 observations.

Next we turn to a more detailed analysis of the variation in our indices. One concern with our approach is that we might just be picking up words or sentences in boiler plate text of annual reports. To tackle this concern, we excluded all exhibits from the filings before we analyzed them, so that we only use information from the actual Form 10-K document and not from any of the appendices which are more likely to be uninformative.[7]

[7]Doing so, we might miss some information – namely when a company files the sections of its annual

The remaining text, on which our indices build, changes considerably from one year to the next. Figure 1 shows the distribution of cosine similarities between identified risk and ambiguity sentences within a firm. For each identified sentence we calculate and store the highest similarity to any risk or ambiguity sentence in the 10-K for the same firm in the previous year. Only a very small percentage of sentences are identical over a year and the mean cosine similarity is around 0.2. This indicates that the text from which we identify our uncertainty words is very different each year.

Figure 1: Distribution of word-based cosine similarities among uncertainty sentences.



The figure reports a sentence level analysis of ambiguity and risk sentences using 10-K filings for fiscal years 1995 up to and including 2013. A value of 1 indicates that an identical risk and uncertainty sentence was found in the previous year for the firm in question. Based on 17,802,424 sentences.

Looking within our identified sentences, we find that about 20% of them are negated, which seems relatively high. Going through the negated sentences, however, we find that these negations do not affect the classification of sentences as risky or ambiguous. Often companies write something of the sort "We cannot exclude that ..." or "It is not inconceivable ..." thereby highlighting a risk factor, but at the same time downplaying the importance of the risk. We also look at the tone of our identified sentences and find it to be only slightly negative. This suggests that we are relatively successful at separating out first order from second order moments in terms of uncertainty.

reports in the exhibits and only references them from the main filing – but more often we drop text that is not relevant for our analysis and would constitute noise (for example bond prospectuses filed as exhibits).

5.2 Relation to credit ratings

We next consider the relationship between our uncertainty measures and credit risk. We use an ordered probit model to estimate the association between uncertain tone and credit ratings. We include interest rate coverage, operating income to sales and debt to asset ratio as controls. Because of its empirical importance, we also include the log of a firms assets as a proxy of firm size . Further, we control for the share of positive and negative words in a firm's annual report. To pick up sector-invariant time variation we include time dummies.

The results of our Probit estimations are provided in Table 4. We estimate six different models and show the standardized effect sizes for our indices. Since we have coded ratings in ascending order, a higher rating implies less credit risk. A positive coefficient estimate thus indicates an decrease in credit risk. For convenience, we do not report the coefficient estimates of the additional controls, but note that the signs are in line with expectations: a higher share of positive words, higher interest coverage and higher operational income to sales are associated with less credit risk. Equally intuitively, we find a larger share of negative words and higher debt to be positively related to credit risk.

To assess the importance of our risk and ambiguity scores for credit ratings, we run simulations of the rating distribution using our model. Table 5 shows the effect of increasing our indices by a one standard deviation shock on the prevalence of different ratings. The effect is comparable between risk and ambiguity, which is indicative that our indices do not correlate with differences in severity and duration of risk factors captured. Indeed, while a casual reading of our word lists might suggest that the ambiguity words rank higher in terms of severity, the estimated marginal effects are, if anything, larger on the risk side.

There are two main take-aways from this exercise. First, we find a negative association between our uncertainty indices and credit risk, which suggests that our indices are indeed capturing something related to the riskiness of a company. Second, a key concern with our split between risk and ambiguity could be, that our ambiguity words simply capture more severe risk factors or risk factors with a longer duration. However, in all our model specifications, we find a similar effect size of risk and ambiguity on credit quality. As an additional validation exercise, we also look at the correlations between our indices and market beta. In line with our expectations, we do find a positive correlation between the two in a variety of specifications. For brevity, we report these results in the Appendix.

Table 4: Rating estimations

VARIABLES	(1) Rating	(2) Rating	(3) Rating	(4) Rating	(5) Rating	(6) Rating
Overall U.	-0.0727***			-0.150***		
	(0.0245)			(0.0246)		
Risk		-0.0706***	-0.198***		-0.0767***	-0.180***
		(0.0170)	(0.0325)		(0.0175)	(0.0346)
Ambiguity		-0.0703***	-0.121***		-0.0998***	-0.119***
		(0.0253)	(0.0446)		(0.0252)	(0.0443)
Positive	0.177***	0.177***	0.125***	0.175***	0.179***	0.130***
	(0.0183)	(0.0185)	(0.0309)	(0.0183)	(0.0185)	(0.0311)
Negative	-0.210***	-0.206***	-0.112***	-0.198***	-0.201***	-0.116***
	(0.0222)	(0.0222)	(0.0386)	(0.0216)	(0.0218)	(0.0383)
Observations	18,689	18,689	7,540	18,689	18,689	7,540
Controls	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Inv Grade Only	NO	YES	YES	NO	NO	YES
Measures	Words	Words	Words	Sentences	Sentences	Sentences

Standardized marginal effect estimates of risk and ambiguity on long-term S&P ratings. Monthly ratings are matched by fiscal-year end. Additional controls are based on Blume et al. (1998): interest rate coverage, operating income after depreciation, operating income to sales and debt to asset ratio. Equality of the coefficient estimates of ambiguity and risk cannot be rejected at conventional significance levels.

Table 5: Rating simulations

	All ratings				*Investment grade ratings only*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Base	Overall shock	Risk shock	Ambi. shock	*Base*	Overall shock	Risk shock	Ambi. shock
		Δ/Base	Δ/Base	Δ/Base		Δ/Base	Δ/Base	Δ/Base
CCC-D	*4.2%*	9%	9%	9%				
B	*26.3%*	5%	5%	5%				
BB	*29.0%*	1%	1%	1%				
BBB	*24.4%*	-3%	-3%	-3%	*58.3%*	13%	10%	6%
A	*12.7%*	-6%	-6%	-6%	*32.7%*	-13%	-11%	-7%
AA	*2.5%*	-10%	-9%	-9%	*6.7%*	-30%	-22%	-14%
AAA	*0.9%*	-13%	-12%	-12%	*2.3%*	-41%	-31%	-20%

This table reports the results of our rating simulations. Columns (1) and (5) show the initial rating distributions for all ratings and investment grade ratings only. Columns (2), (3),(4) and (6),(7),(8) show the relative changes in the distributions after a one standard deviation increase in overall uncertainty, risk or ambiguity. Simulations are based on the ordered probit estimation in column (1)-(3) in Table 4.

Table 6: Relation between financial constraints, risk and ambiguity.

VARIABLES	(1) Fin Con	(2) Fin Con	(3) Fin Con	(4) Fin Con	(5) Fin Con
Overall U.	0.0408***				
	(0.00303)				
Ambiguity		0.134***	0.131***		
		(0.00811)	(0.00806)		
Risk		-0.0203***		-0.00602	
		(0.00668)		(0.00674)	
Ambiguity (sen)					0.00677***
					(0.000412)
Risk (sen)					-0.00107***
					(0.000382)
Observations	64,170	64,170	64,170	64,170	64,170
R-squared	0.622	0.624	0.623	0.617	0.624
FE	Firm Year	Firm Year	Firm Year	Firm Year	Firm Year

The table reports on the relation between text based measures of ambiguity and risk and financial constraints. The dependent variable is the text based measure of financial constraints from Hoberg and Maksimovic (2015): "delay investment score". Analysis based on 10-K filings starting fiscal year 1997 up to and including fiscal year 2013. Financial firms and utilities excluded. Regressions by OLS and with robust standard errors clustered by firm. *** p<0.01, ** p<0.05, * p<0.1.

5.3 Financial constraints

Our model suggests that firms which appear more ambiguous to investors will face tighter financial constraints. To explore if our data are in line with this prediction we use the text based measure of financial constraints from Hoberg and Maksimovic (2015) to the ambiguity and risk indices, while also controlling for year and firm fixed effects. The financial constraint data cover 1997 to 2013.[8] In column (1) of Table 6 we relate the measure of financial constraints to the overall uncertainty index and note that higher uncertainty for a firm is associated with the firm being more financially constrained. As

[8]Measuring financial constraints is notoriously hard. Recent evidence by Farre-Mensa and Ljungqvist (2015) show, for a broad set of commonly used measures, that supposedly constrained firms do not act as expected when hit by exogenous tax shocks. Hoberg and Maksimovic (2015) build a continuous measure of financial constraints using cosine similarities of wording in the MD&A section and on Gerard Hoberg's website the measures are updated to 2013. Based on their word lists we also created a dummy variable for financial constraints using our text data - since we did not extend this beyond the "training data" we have fewer observations for this, but qualitatively the results were the same as in the regressions reported here.

we are using firm fixed effects and year fixed effects the coefficient is identified by firm-level variation variation in overall uncertainty. In column (2) we establish a positive effect from ambiguity on financial constraints, whereas the effect of risk is negative. It may at first be surprising that risk has a negative effect but the effect is limited. The correlation between the ambiguity and the risk measure is 0.12. To the extent that multicollinearity could be driving the result that they have opposite signs we therefore include them one-by-one in columns (3) and (4), underlining that ambiguity appears to show an important link to financial constraints. In column (5) we establish that the results are robust to using the share of sentences rather than share of words.

The evidence is thus supportive of our modeling approach to introduce ambiguity on the subjective probability of success of a firm and thereby linking ambiguity and financial constraints. This result is also broadly in line with results in Hoberg and Maksimovic (2015) where they establish a positive relation between financial constraints and the extent of asymmetric information (the latter based on combination of words that capture protection with words that capture secrecy: for instance the combination "safeguard trade secrets"). We now turn to an examination of the key predictions - the relation between the measures of uncertainty on the one hand and cash holdings and derivatives use on the other.

6 Ambiguity and risk: cash holding and derivatives

6.1 Empirical strategy

We use panel-data fixed-effects regressions to investigate our hypotheses. Our regression use the following specification:

$$Y_{it} = \theta AMB_{it} + \kappa RISK_{it} + \gamma_t + \alpha_i + \beta X_{it} + \varepsilon_{it} \tag{5}$$

where Y_{it} is our outcome variable of interest for firm i in fiscal year t: It is either cash holding as a share of assets or dummy variables capturing hedging. We estimate using OLS and thus for the latter outcomes we estimate linear probability models. The coefficients θ, κ and β are parameters to be estimated. AMB_{it} is our measure for the degree of ambiguity of the firm. The variable $RISK_{it}$ refers to our textual measures of risk. The primary coefficients of interest are θ which we interpret the effect of the degree of ambiguity of the firm on the outcome variable and κ the effect of the firm's riskiness on the outcome variable.

We include time fixed effects γ_t to account for common time trends in cash holding and changes in reporting standards. For example, in 2005 the SEC changed it's requirements of risk reporting, which led to a substantial increase in the length of risk-related text in the annual reports. Our time fixed effects absorb such global changes in reporting standards.[9] We include firm fixed effects α_i to absorb differences in approaches to financial reporting across firms. For example: a firm with a very diligent approach to financial reporting might report more on risks and ambiguities than a firm which is less diligent. Including firm fixed effects, we examine the effect of changes in risk and ambiguity within a firm.

The vector X_{it} contains firm specific controls, which we select based on the existing literature. In particular X_{it} also contains the Loughran and McDonald (2011) measures of positive and negative words. We want to isolate the effects of positive and negative shocks from the effects of increases in uncertainty. Note that this is conceptually important as many definitions of risk are available and in some definitions higher risk is tantamount to negative shocks. Loosely put, we control for first moment effects to focus on second moment effects of risk and higher measured ambiguity. Finally, the statistical error term is denoted by ε_{it}.

Clearly we share many strengths and weaknesses with other research that uses textual analysis as a way of measuring variables of interest. Our use of time fixed effects and firm fixed effects are expected to take care of concerns regarding changes in tone over time or differences in tone across firms and industries. Our robustness tests using sentence level analysis and controlling for negation also take care of some other concerns that may plague an analysis which is only based on word counts. It deserves to be pointed out that there is little reason to expect a causal effect *from the word count itself* on the actions of the firm. The text should describe reality and changes in that reality may affect both the text and the actions, but there is no direct effect per se from the text on the firm's actions. In contrast, the wording in documents can impact the views of external investors, which in turn can influence the firm, for instance via a tighter financial constraint as in the model of Section 3. Such an interpretation is supported both by previous textual analysis of uncertainty measures as discussed in Section 2 and by examinations of the link between ambiguity and financing constraints in Section 5.4. We thus interpret our uncertainty indices as measures for how risky or ambiguous a firm is perceived by its investors. A stronger interpretation would be that our indices are directly related to the

[9]This is the most salient reporting change over our sample period and our results are robust to only using post-2005 data.

risk and ambiguity environment of a companies' management. In our theoretical analysis we showed that the weaker interpretation is sufficient to derive our empirical hypothesis.

Finally, one potential concern is that there may be reverse causality. Regarding ambiguity one may worry that empire-building managers might use more ambiguous language to motivate higher cash holding.[10] The use of time and firm fixed effects will partly alleviate these concerns, in essence ensuring that we rely on firm-specific variation over time in the indexes to identify effects. As a robustness check we will also run regressions where we use CEO $\times$ firm fixed effects and key results are unaffected by this, which support that such idiosyncratic CEO behavior is not driving results.

Regarding the risk index one may worry that firms need to discuss risk if they are using derivatives and that this might be driving the relation that we will estimate between hedging and risk. In section 6.4, when discussing results, we take a close look at this proposition and show that such effects appear to be of quite limited quantitative importance. Nevertheless, even if the discussion and analysis suggest that reverse causality is of limited concern one might still be worried about endogeneity of our uncertainty measures. As a final check we therefore also show that results are robust to estimations where we use instrumental variable regressions, instrumenting for our uncertainty indices. We discuss the instruments used and their validity in connection with reporting these results.

6.2 Cash holding

We now turn to an examination of the link between uncertainty and cash holding. In column (1) of Table 7 we use the measure of overall uncertainty and establish that increased uncertainty is associated with more cash holding. In column (2) we allow for separate impact of ambiguity and risk. The estimated relation between ambiguity and cash holding is positive - this is as suggested by intuition and in line with expectations from the theory that we presented. Theory also suggested that risk be dealt with by other means that by holding cash, a result that is also in line with the findings. Higher risk is associated with less cash holding. Importantly, it is within-firm variation in the uncertainty indices that identifies the coefficients.

Turning to quantitative importance of the key coefficients in column (2), the mean cash ratio in our observations included in the regressions is 18.9 % and the mean share of ambiguity words is 0.32 %. Increasing the share of ambiguity words by one standard

[10]Note that the results in the previous sections indicate that such wording is costly, increasing the cost of funds for firms. This cost should thus help limit endogeneity concerns.

deviation (0.12) would then imply that the cash ratio increased by 0.8 percentage points to 19.7 %. In the same way, increasing the mean share of risk words by one standard deviation (also 0.12) decreases cash holding by 0.3 percentage points. These magnitudes are substantial and moving a firm from one standard deviation below the mean to one standard deviation above the mean for the ambiguity index is estimated to raise cash holding by 1.6 percentage points.

We also include the shares of positive and negative words to capture the effects of "first moment" shocks. Firms that use more positive words hold more cash. This is consistent with the idea that good news imply greater investment needs in the future which would lead a liquidity constrained firm to hold more cash.[11] Negative words are also associated with higher cash holdings in the specification in column (1). Clearly, both offensive motivations in response to positive shocks (investment opportunities) and defensive motivations in response to negative shocks (building a war chest to weather tough times) can spur increased cash holding. The positive effect is the more robust across specifications however, for instance the coefficient on negative words is not statistically significant in the specification in column (2).

Our estimates might be biased due to endogeneity issues. In particular, the risk index might be driven by firms' needs to discuss derivative use. We will get back to this at the end of this section in more detail, for now we follow common practice and use lagged observations of ambiguity and risk as well as their average levels at the four-digit SIC level as instruments. Formal tests indicate that these are valid instruments: they have a significant association with the current ambiguity and risk indices at the firm level and we are not able to reject the hypotheses that the instruments are uncorrelated with the error term. The F-statistic in the first stage is high enough that Stock-Yogo test for weak instruments are rejected at the 1 percent level. The test for over-identifying restrictions is more worrisome: the Hansen J statistic is 5.92 and the p-value is 0.0518. Thus we are not able to reject the hypotheses that instruments are uncorrelated with the error term at the five percent level but the failure to reject is hardly resounding. Thus, while we take the estimates as supportive of that endogeneity is not driving results we view the estimates with caution and use uninstrumented regressions in further robustness checks.

[11] There is some resemblance here to results on the response of cash holdings to cash flows. Most work suggests that it is positive (Almeida et al. (2014)) but it can also be negative if, after controlling for the market to book ratio, a positive shock leads the firms to dis-save: draw down on cash and instead invest (see Riddick and Whited (2009)) for a model and empirical investigation of this.

Table 7: Cash holding regressions for U.S. firms 1995-2013.

VARIABLES	(1) Cash	(2) Cash	(3) Cash (IV)	(4) Cash	(5) Cash
Overall U.	0.0130**				
	(0.00540)				
Ambiguity		0.0643***	0.134***	0.0619***	0.0511***
		(0.0148)	(0.0239)	(0.0145)	(0.0140)
Risk		-0.0235**	-0.0371**	-0.0105	-0.0112
		(0.0118)	(0.0168)	(0.0117)	(0.0112)
Positive	0.0456***	0.0412***	0.0345***	0.0394***	0.0342***
	(0.00677)	(0.00677)	(0.00638)	(0.00670)	(0.00647)
Negative	0.00560**	0.00394	0.000899	0.00601**	0.00533**
	(0.00246)	(0.00250)	(0.00235)	(0.00248)	(0.00240)
Industry Sigma				0.0173	0.0242
				(0.0154)	(0.0149)
Market to Book = L,				0.00188***	0.00226***
				(0.000379)	(0.000414)
ln(Assets) = L,				-0.0161***	-0.0152***
				(0.00210)	(0.00208)
Cash Flow = L,					0.00139
					(0.00324)
Working Capital = L,					-0.00338
					(0.00290)
Capex = L,					-0.198***
					(0.0158)
Leverage = L,					-0.0883***
					(0.00658)
RnD = L,					0.00390***
					(0.000902)
Dividend = L,					0.00268
					(0.00320)
Acquisition = L,					-0.135***
					(0.00859)
Observations	61,330	61,330	59,963	61,330	61,330
R-squared	0.764	0.764	0.757	0.767	0.773
Controls	NO	NO	NO	YES	YES
FE	Firm Year	Firm Year	Firm Year	Firm Year	Firm Year

The table reports cash holding regressions using text based measures of uncertainty and positive and negative words. Text analysis based on 10-K filings fiscal years 1995 to 2013 inclusive. All other data from Compustat. Dependent variable is cash/assets (CHE/AT). The variables in columns (4) and (5) defined following Bates et al. (2009). Financial firms and utilities excluded. Regressions by OLS (GMM in column (3)) and with robust standard errors clustered by firm. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The positive effect of ambiguity on cash holding remains also in the IV regression, thus supporting the notion that reverse causality is not driving the results. The estimated coefficients are actually greater in the instrumented specification which is what we would expect to find if measurement error caused an attenuation bias. The main point that we take away from the IV-regression however is that the differential impact of ambiguity and risk is robust and that the difference is quantitatively important.

There exists a voluminous literature on cash holding, frequently using a large number of covariates (and often eschewing firm fixed effects). For comparison and robustness we add three potentially particularly relevant controls in column (4). We see that the measure of risk used by Bates et al. (2009) (the standard deviation of shocks at the two digit SIC level) does not have a statistically significant impact on cash holding when we control for firm fixed effects and our text based measures of uncertainty. We also see that increases in the market-to-book value raise cash holdings and increases in size (as measured by the natural logarithm of assets) lower them. In column (5) we include the full set of regressors from the influential article of Bates et al. (2009). The sign of coefficients are in line with expectations and we see that the coefficients on risk and ambiguity are similar to the specification with few controls in column (2).

We report some key robustness exercises in Table 8. We include the measures of positive and negative words and as before we use firm fixed effects. For ease of exposition we only report the key variables of interest. First we explore if relations change with the requirement introduced in 2005 that "risk factors" should be discussed in section 1A of the 10-K. Column (1) reports the key regression for the period before fiscal year 2005 and column (2) for the fiscal year 2005 onwards. We see that the pattern is stable but that the coefficient on ambiguity is lower in the first period.

Above we introduced text analysis in a rather simple way, examining the share of words in the full 10-K. Many other measures are possible and in particular we can base measures on the share of sentences that contain ambiguity or risk words. In column (3) we present a specification where we use this sentence based measure. It is clear that qualitative results are the same and increasing the ambiguity sentence share by one standard deviation (2.5) results in a 0.8 percentage point increase in cash holding - same as in the word based measure. Moving to the sentence level also allow us to examine negations and in column (4) we use the share of sentences with uncertainty words that are not negated as our measures of uncertainty and find results very close to those in column (2). In column (5) we examine specifications where we drop sentences found in item 8, "financial statements", and results are little affected by this. As we move to more

Table 8: Robustness of cash holding regressions, U.S. firms 1995-2013.

VARIABLES	(1) Cash	(2) Cash	(3) Cash	(4) Cash	(5) Cash	(6) Cash
Ambiguity	0.0321*	0.0587***	0.00286***	0.00273***	0.00267***	0.0607**
	(0.0184)	(0.0223)	(0.000730)	(0.000700)	(0.000835)	(0.0248)
Risk	-0.0156	-0.0349*	-0.000404	-0.000352	-0.000787	0.00577
	(0.0127)	(0.0197)	(0.000635)	(0.000618)	(0.000700)	(0.0189)
Observations	34,146	27,184	61,330	61,330	30,817	13,009
R-squared	0.812	0.837	0.773	0.773	0.806	0.832
Measure	Word	Word	Sentence	Sen w/o neg	Sen w/o Item 8	Word
Controls	YES	YES	YES	YES	YES	YES
FE	Firm Year	Firm Year	Firm Year	Firm Year	Firm Year	Firm x CEO Year
Fiscal years	1995-2004	2005-2013	1995-2013	1995-2013	1995-2013	1996-2013

The table reports cash holding regressions using text based measures of uncertainty and positive and negative words. Text analysis based on 10-K filings fiscal years 1995 to 2013 inclusive. All other data from Compustat. Dependent variable is cash/assets (CHE/AT). All other controls from Table (7) included but not reported: Positive, negative, Industry sigma and lagged values of: market to book, ln(assets), cash flow, working capital, capex, leverage, RnD, and dummy variables for acquisitions and dividend payments. Financial firms and utilities excluded. Regressions by OLS and with robust standard errors clustered by firm. *** p<0.01, ** p<0.05, * p<0.1.

narrow definitions, for instance only examining sentence shares in item 1 and 7, results tend to be unstable and frequently not statistically significant. One robustness check that deserves mention, even if it because of space constraints is not reported in Table 8, also includes the square of the ambiguity measure in the benchmark specification of Table 7, column (3). The coefficient on the squared ambiguity measure is negative (-.023) but not statistically significant at the 10 % level. The motivation for including a square term is loosely based on the model of Agliardi et al. (2016), in which, as discussed in Section 2, ambiguity aversion has a U-shaped relation to cash holding. As we examine different levels of *ambiguity*, rather than different levels of *ambiguity aversion*, this negative result should not be seen as a test of their model.

As discussed in the section that outlined our empirical strategy one might worry that an empire-building CEO might use more ambiguous words to motivate higher cash holding. To cleanse the estimates of CEO specific effects, we estimate a final specification with fixed effects at the level of CEO $\times$ firm.[12] We find a positive and significant effect of ambiguity on cash holding which is of similar size as in our benchmark regressions.

In sum, we find that a higher share of ambiguity words and sentences in the 10-Ks is

[12]Data on CEO tenure are from Execucomp.

associated with higher cash holding and that the effect is quantitatively important. No one section is key to this finding, rather it is the whole 10-K (apart from the financial statement) that matters. There is consistent evidence of a differential effect between ambiguity and risk: increases in risk are associated with no statistically significant, or in some specifications, negative effect in cash holding. Theory suggests that risk is rather managed with derivative contracts, an issue that we turn to now.

6.3 Derivatives use

We examine derivatives use with the aid of our proxy variables for whether a firm engages in hedging with derivatives. We estimate regressions using data from 2001 onwards only as the dependent variable is unavailable prior to this. In Table 9 we relate the dummy on derivatives use to our measures of uncertainty. Again we rely on the within firm variation for identification of our parameters. In column (1) we include our measure of overall uncertainty and find that it is statistically significant and positive, firms with a higher overall uncertainty index are more likely to use derivatives. The positive and negative text indices are not statistically significant in any of the regressions.

In column (2) we distinguish between ambiguity and risk and find that risk is positively related to financial hedging whereas ambiguity exhibits a negative relation. The effect of risk is quantitatively non-trivial. The mean value of our risk index is 0.32 and increasing this with one standard deviation (0.12) is thus associated with an increase in the probability of using derivatives by 2.6 percentage points, from 0.214 to 0.24. In contrast, the measure of ambiguity is negatively related to derivatives use and a one standard deviation in ambiguity lowers the probability of using derivatives by 0.9 percentage points. The effect of ambiguity is only statistically significant at the 10 % level however. In column (3) we report a specification where we, as for cash holding, use lagged values of ambiguity and risk as well as their averages at the four-digit SIC level as instruments.[13] The estimated effect of risk is somewhat increased in the instrumented specification, again a finding consistent with some measurement error in the index. It does however point to that reverse causality is not responsible for the statistically significant relation between our text based risk measure and derivatives use, a result that is also in line with the closer inspection of the relation in the next section.

[13] As in the cash holding regressions, the F-statistic is high enough that the Stock-Yogo test for weak instruments is rejected at the 1 percent level. Furthermore, we are not able to reject the hypotheses that instruments are uncorrelated with the error term at the five percent level: the Hansen J statistic is 4.62 and the p-value is 0.0992. Statistical tests thus suggest that the instruments are valid.

Table 9: The determinants of hedging. U.S. firms, linear probability estimation 2001 to 2013.

VARIABLES	(1) Deriv	(2) Deriv	(3) Deriv (IV)	(4) Deriv	(5) Deriv
Overall U.	0.0544***				
	(0.0155)				
Ambiguity		-0.0818*	-0.113	-0.0861**	-0.0745*
		(0.0423)	(0.0692)	(0.0421)	(0.0420)
Risk		0.221***	0.292***	0.204***	0.210***
		(0.0362)	(0.0543)	(0.0362)	(0.0362)
Positive	-6.04e-05	0.0196	0.0241	0.0231	0.0270
	(0.0194)	(0.0194)	(0.0183)	(0.0194)	(0.0193)
Negative	-0.00715	0.00380	0.00585	0.00226	0.00213
	(0.00740)	(0.00746)	(0.00726)	(0.00749)	(0.00747)
Industry Sigma				-0.0390	-0.0417
				(0.0351)	(0.0349)
Market to Book = L,				0.00181***	0.00116***
				(0.000318)	(0.000319)
ln(Assets) = L,				0.0353***	0.0357***
				(0.00461)	(0.00508)
Cash Flow = L,					-0.00658***
					(0.00211)
Working Capital = L,					0.00478**
					(0.00194)
Capex = L,					0.0728*
					(0.0390)
Leverage = L,					0.0796***
					(0.0144)
RnD = L,					-0.000967
					(0.000722)
Dividend = L,					0.0145
					(0.0120)
Acquisition = L,					0.132***
					(0.0318)
Observations	42,287	42,287	41,162	42,164	42,164
R-squared	0.698	0.699	0.696	0.701	0.702
Controls	NO	NO	NO	YES	YES
FE	Firm Year	Firm Year	Firm Year	Firm Year	Firm Year

The table reports regressions on derivatives use. Measures of uncertainty based on textual analysis of 10-K filings from fiscal year 1995 to 2013 inclusive. All other data from Compustat. Dependent variable is 1 if firm used derivatives and 0 otherwise, derivatives use implied by non-zero reported values of AOCIDERGL or HEDGEGL in Compustat. Financial firms and utilities excluded. Regressions by OLS (except Instrumental variable regression in col (3) which is by GMM) and with robust standard errors clustered by firm. *** p<0.01, ** p<0.05, * p<0.1.

In column (4) we include the measure of risk from Bates et al. (2009) as well as market-to-book and the natural logarithm of assets. In column (5) we include a rich set of controls, the same variables as in the corresponding regression for cash holing. Also the relation between derivatives use and other coefficients overall correspond to expectations. In particular an increase in size is likely trigger the use of derivatives (as captured by the log of assets) and there is a positive association between leverage and derivatives use. For interpretation it again deserves to be pointed out that firm fixed effects are likely to soak up much variation. The effect of risk on the probability of hedging is relatively stable across specifications.

Table 10: Robustness. The determinants of hedging. U.S. firms, linear probability estimation for 2001 to 2013.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Deriv	Deriv	Deriv	Deriv	Deriv	Deriv(Narrow)
Ambiguity	-0.0406	-0.0803	-0.00366*	-0.00359*	-8.38e-05	-0.0684
	(0.0539)	(0.0570)	(0.00213)	(0.00212)	(0.00241)	(0.0421)
Risk	0.137***	0.232***	0.0120***	0.0114***	0.0115***	0.216***
	(0.0476)	(0.0526)	(0.00210)	(0.00201)	(0.00269)	(0.0364)
Observations	15,035	27,129	42,164	42,164	22,947	42,164
R-squared	0.844	0.761	0.702	0.702	0.717	0.700
Measure	Words	Words	Sentence	Sen w/o neg	Sen w/o Item 8	Words
Controls	YES	YES	YES	YES	YES	YES
FE	Firm Year	Firm Year	Firm Year	Firm Year	Firm Year	Firm Year
Fiscal years	2001-2004	2005-2013	2001-2013	2001-2013	2001-2013	2001-2013

The table reports regressions on derivatives use. Measures of uncertainty based on textual analysis of 10-K fiscal years 1995 to 2013 inclusive. All other data from Compustat. Dependent variable is 1 if firm used derivatives and 0 otherwise, derivatives use implied by non-zero reported values of AOCIDERGL or HEDGEGL in Compustat. All other controls from Table (9) included but not reported: Positive, negative, Industry sigma and lagged values of: market to book, ln(assets), cash flow, working capital, capex, leverage, RnD, and dummy variables for acquisitions and dividend payments. Financial firms and utilities excluded. Regressions by OLS and with robust standard errors clustered by firm. *** p<0.01, ** p<0.05, * p<0.1.

We have examined several specifications and the overall pattern is stable: risk is positively related to derivatives use and ambiguity shows a weak negative relation or no statistically significant relation to derivatives use. In Table 10 we present some key robustness examinations. We first examine results before (column(1)) and after (column (2)) the change in the requirement to discuss risk factors in fiscal year 2005 and establish that qualitative patterns are stable, but that the coefficient on risk is higher in the latter period. We further establish that results are robust to sentence level analysis (column

(3)), only considering non-negated sentences (column (4)) and excluding sentences found in item 8 (column (5)). Finally, in column (6) we use a narrower definition of derivatives and again a statistically significant positive effect of the risk index on derivatives use is established.

6.4 Reverse causality?

One concern, in particular for the risk index, is the potential for reverse causality. The usage of derivative instruments may in itself trigger the need to discuss risk related concerns in the 10-K. One way to address this is via IV regressions, as we do in Tables 7 and 9. As a complement to the IV estimation we may also use changes in the set of derivatives instruments available or in hedging practices to try to gauge the extent of the problem. In their study of risk management via purchase obligations Almeida et al. (2016) identify eight industries where traded futures were introduced in 2008.[14] At the level of the individual firm this constitutes an arguably exogenous increase in the availability of derivatives. Using other firms in the same broader industries as controls we can then examine if the increase in hedging possibility is associated with a higher share of risk or ambiguity words for the firms in the industries where the set of relevant derivatives expanded.[15] In columns (1) and (2) of Table 11 we use the indices of ambiguity and risk at the firm level as dependent variable and use firm fixed effects and year effects. The key variable of interest is a dummy variable that takes the value 1 for the post 2008 period for firms in one of the industries where traded futures became available in 2008. While this is a small set of industries and just one change it is noteworthy that the effect is not statistically significant and also that the point estimate for the risk index of 0.02 (which can be related to the average risk index of 0.32) is relatively low.

Another way to examine the possible extent of reverse causality is to examine the risk index for firms that start to use derivatives in a given year, not having used derivatives in the year before. If the discussion of risk in the 10-K is largely generated by a need to discuss and motivate the use of derivatives, we would expect a marked increase in the risk index in the same fiscal year that a firm starts to use derivatives. We explore this in columns (3) and (4). As seen there is a statistically significant effect but the magnitude is small.

[14] They are all steel producing industries, NAICS: 331111, 33112, 331210, 331221, 331222, 331512, 331513, 332111.

[15] As controls we use all firms in the same three digit NAICS industries: 331 and 332.

Table 11: Reverse causality. The response of uncertainty indices to changes in the available derivative instruments or change in usage of derivatives. U.S. firms, linear probability estimation for 2001 to 2013.

VARIABLES	(1) Ambiguity	(2) Risk	(3) Ambiguity	(4) Risk
Traded futures x post	0.0251	0.0198		
	(0.0328)	(0.0391)		
Start hedge			-0.00358**	0.00537**
			(0.00167)	(0.00214)
Observations	1,263	1,263	65,101	65,101
R-squared	0.696	0.664	0.731	0.748
Controls	NO	NO	NO	NO
FE	Firm Year	Firm Year	Firm Year	Firm Year
Fiscal years	2001-2013	2001-2013	2001-2013	2001-2013

The table reports regressions on uncertainty indices on derivatives use. Measures of uncertainty based on textual analysis of 10-K fiscal years 1995 to 2013 inclusive. Regressions in columns (1) and (2) examines the effect of introducing traded futures in 8 industries in 2008 on uncertainty indices. Regressions in columns (3) and (4) report regressions where "start hedge" is a dummy variable which is 1 if the firm uses derivatives in fiscal year t that did not use derivatives in t-1. Regressions by OLS and with robust standard errors clustered by firm. *** p<0.01, ** p<0.05, * p<0.1.

Table 12: Average level of risk index, and change in the risk index, for firms who start to use derivatives in fiscal year t. U.S. firms, 2002 to 2013.

Fiscal Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Risk(t)	0.3509	0.3789	0.3805	0.387	0.361	0.352	0.3622	0.3447	0.3349	0.3512
s.d.	-0.1344	-0.1117	-0.1245	-0.1163	-0.1087	-0.0963	-0.1149	-0.1058	-0.0907	-0.1176
Risk(t-1)	0.3512	0.3416	0.3761	0.3839	0.3589	0.347	0.3516	0.3376	0.3336	0.354
s.d.	-0.1373	-0.1104	-0.1205	-0.1209	-0.1138	-0.1045	-0.1181	-0.1087	-0.0951	-0.1175
Difference	-0.0003	0.0373	0.0044	0.0031	0.002	0.005	0.0107	0.0071	0.0014	-0.0028
s.d.	-0.1022	-0.0745	-0.0739	-0.0645	-0.0609	-0.0622	-0.0598	-0.0394	-0.0484	-0.0401
T-stat	-0.0399	6.1293	0.7333	0.7304	0.4646	1.2208	2.7053	1.8923	0.291	-0.7246
P(t)	0.9682	0	0.4646	0.4659	0.6428	0.2234	0.0073	0.0611	0.7716	0.4703

The table reports average risk index for firms who start hedging in year t and the risk index in the previous year for these same firms. Standard deviations in parentheses. Final column shows t-statistics for the hypotheses that there is no difference and significance level. Measure of risk based on textual analysis of 10-K fiscal years 2001 to 2013 inclusive.

To gain further understanding of this effect we use Table 12 to examine the average level of the risk index in fiscal year t for the firms that start using derivatives in that fiscal year as well as the risk index for those same firms in the previous fiscal year. The difference is typically positive but mostly small in magnitude and only statistically significant (using a two-sided t-test) at the five percent level in 3 of these twelve years. In sum: there are indications that firms to some degree discuss risk to a greater extent when they use (have greater access to) derivatives but the effect appears to be small and the pattern is consistent with a view that the discussion of risks reflect underlying risk considerations rather than just a necessity to discuss derivatives.

7 Conclusions

We create indices of ambiguity and risk using textual information for all 10-K statements starting fiscal year 1995 and ending fiscal year 2013. A closer examination of cash holdings and derivatives yield results that are consistent with predictions from a standard model of cash holding extended to include ambiguity aversion on the part of investors. Our findings are thus consistent with the idea that agents make a distinction between ambiguity and risk and that the distinction is not only of interest to the decision theoretic literature spurred by Ellsberg (1961), but also seems to be capturing elements of decision making in the field.

Our investigation focused on liquidity management: in particular cash holding and derivatives use. More broadly capital investment as well as employment, advertising and research and development may be affected by ambiguity and risk. The model upon which we build has become a standard one for liquidity management, but as a model of investment it is likely that it fails to capture important features relating to investment (for instance time-to-build and differential impact of different sources of ambiguity). Whether investment is increasing or decreasing in risk has been the focus of much interest and some theoretical results suggest that there may be different effects from increased risk and ambiguity (Nishimura and Ozaki (2007)). However, one may expect that not only the level of investment, but also the type of investment (for instance the level of flexibility) may be affected by the extent of risk and ambiguity. We leave explorations of this for future work.

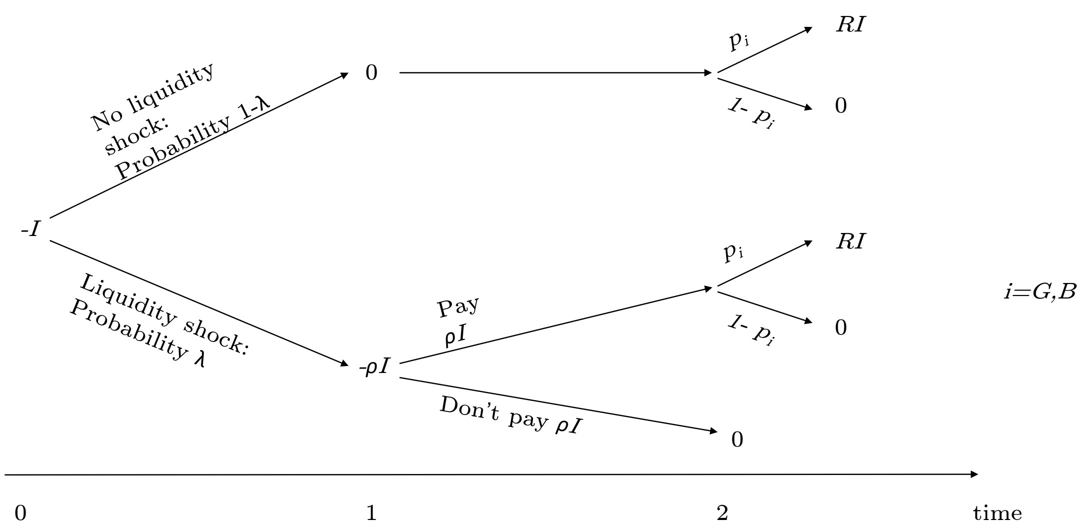
While our main aim has been to explore the relation between ambiguity and risk on the one hand, and cash holding and derivatives use on the other, we hope that the index of ambiguity that we create can be of use for future work that wants to study the links between ambiguity and decisions in other dimensions. Aggregated to industry level it might for instance be related to issues relating to the boundaries of the firm. While there has been much theoretical interest in ambiguity it has not been clear how it can be measured or operationalized in corporate finance settings - we hope to have shown the potential usefulness of a text based measure.

Appendix

A A presentation of the model in greater detail

In the main text the model is presented in a quite cursory way, partly in reflection of that the basic modeling framework will be familiar to many readers. For completeness we provide more detail on the model in this appendix. Figure A.1 summarizes the main features of the model.

Figure A.1: A framework for liquidity management.



Given our assumptions, the optimal second best contract with an outside investor characterises the amount of investment I and the distribution of the returns of the investment. To make the analysis interesting we assume that $I > A$. A key role is played by the pledgeable income: $R_f \in [0, R]$. The firm chooses effort and investment to maximize its expected return and solves the following program:

$$\begin{aligned} \max p_G R_f I - \lambda\rho I - I &\quad s.t. \\ p_G R_f &\geqslant p_B R_f + B && (IC) \\ A + (1-\lambda)(p_G - \epsilon)(R - R_f)I + \lambda((p_G - \epsilon)(R - R_f) - \rho)I &\geqslant I && (PC)\&(BC) \end{aligned}$$

The return offered to the firm needs to satisfy its incentive compatibility constraint (IC) – otherwise the firm would not choose to exert high effort. The participation constraint of the investors also serves as the budget constraint of the firm and implies that investment is bounded by the firm's own assets and the returns that it can credibly pledge to outside investors: $(R - R_f)$. Assuming (IC) is binding we have:

$$R_f = \frac{B}{p_G - p_B}$$

Using this in the participation constraint of the investors yields Equation (2) in the main text. A key feature of the model is the wedge between pledgeable income ρ_{0AMB} as defined in the main text and the expected return of the project $\rho_1 \equiv p_G R$. If the firm is hit by a liquidity shock we assume that the liquidity shock ρ is sufficiently low that it is worth to continue also in this state ($\rho < \rho_1$) but that the necessary follow up investment is greater than pledgeable income ($\rho_{0AMB} < \rho$). In period 1, if the firm has been hit with a negative liquidity shock it will thus not be able to borrow additional funds just like it is not possible to insure a house *after* it has burnt down. To be able to finance the necessary follow-up investment in this case the firm needs to have liquidity as in Equation (4).

A.1 A discussion of robustness

We introduced ambiguity in a very simple way, letting investors be ambiguity averse and letting the estimate of p_G be ambiguous. However, it is is easily established by examination of the respective incentive and participation constraints that the predictions are by and large robust to other ways of introducing ambiguity. If we assume that the firm is also ambiguity averse the results above are further strengthened since the pledgeable income decreases if the incentive constraint needs to hold for the most pessimistic probability of success. If we introduce ambiguity on both p_G and p_B the same qualitative results hold again. Similarly if we only let the firm be ambiguity averse. If we only introduced ambiguity regarding the probability of a liquidity shock, investment would be lower but there would be no direct effect on cash holding. However we believe that the probability of success of the investment conditional on effort is a quite natural candidate for a probability estimate subject to ambiguity.

There is one case where we get different results, if we only let p_B be ambiguous and assume that the firm is ambiguity averse and that this is known by investors. Then the incentive constraint will be less stringent and the firm could invest more. Only introducing ambiguity aversion on the firm level, and only on the probability of success in the case where the firm slacks off seemed like a less attractive starting point to us, not the least given substantial evidence that many CEOs are overconfident (see e.g. Malmendier and Tate (2005)). Another special case to point out is if we introduce ambiguity on p_G and investors believe that the firm is ambiguity averse when in fact it is not. Then the investors would falsely believe that the incentive constraint held while the firm instead would slack off. The implication would be that investors are consistently wrong, which, again, was less appealing as a starting point.

Let us also note that Hypothesis 2 as presented in the main text only explored cash holding for a given level of investment. As investment decreases in the extent of ambiguity one may ask if also the following holds (using ρ_0 and C^* to denote pledgeable income and cash holding in the case where there is no ambiguity aversion, $\epsilon = 0$):

$$\frac{C^*_{AMB}}{A} > \frac{C^*}{A}. \tag{6}$$

We see that the relation in equation (6) holds if

$$\frac{(\rho - \rho_{0AMB})}{1 - \rho_{0AMB} + \lambda\rho} > \frac{(\rho - \rho_0)}{1 - \rho_0 + \lambda\rho}. \tag{7}$$

Straightforward manipulation[16] establishes that the inequality in equation (7) holds if and only if

$$(\rho_0 - \rho_{0AMB})(1 - \rho + \lambda\rho) > 0. \tag{8}$$

The first term in equation (8) is positive as seen above and the second is positive by assumption if the firm is to find it optimal to continue the project in the case of a liquidity shock. We have thus established that a firm faced with greater ambiguity will hold more cash as a share of assets.

[16]Multiply the numerators to put the terms on a common denominator, collect terms and note that the denominator is positive.

A.2 Adding risk to the model

As discussed in Hypothesis 3 we consider the following extension: Assume that there is an additional term in profit (s) paid to the firm at time 2 in the case of a successful project. This term s has 0 expected value and takes the value $\underline{s}$ with probability .5 and $\overline{s}$ with probability .5. The realization of this shock is revealed at time 1, before effort is decided and before the decision is taken to make a follow up investment in the case of a liquidity shock. Since the incentive constraint is affected symmetrically both in the case of good (p_G) and bad (p_B) luck, the incentives to be diligent are unaffected, contingent on the follow up investment being made. However, in the case of a liquidity shock, a negative value of s would leave the firm too low on funds to make the required follow up investment. The firm would need to increase cash holdings in time 0 to be able to continue also in this scenario (remembering that $\underline{s} < 0$). If we assume that there is some liquidity cost q to holding cash, and that the exposure to s can be hedged with derivatives at a low enough cost, then Hypothesis 3 follows trivially.

$$C = (\rho - \underbrace{(p_G - \epsilon)(R - \frac{B}{p_G - p_B})}_{\equiv \rho_{0AMB}}) I - \underline{s}I. \tag{9}$$

B Additional data details

We use Compustat measures of cash (CHE) and assets (AT) to generate the cash ratio (CHE/AT). Size is measured as ln(AT). We in addition use the book value of equity (CEQ), share price (PRCC_F) and common shares outstanding (CSHO) to generate market-to-book value as (AT-CEQ+PRCC_F×CSHO)/AT. We use operating income (OIBDP), interest expenditure (XINT), taxes (TXT) and dividends to calculate cash flows to assets as (OIBDP-XINT-TXT-DVC)/AT. Working capital is measured by WCAP and capital expenditure by CAPX. We use long-term debt (DLTT) and debt in current liabilities (DLC) to calculate leverage as (DLTT+DLC)/AT. Research and development expenditures are captured by XRD. As in Bates et al. (2009) we set these expenditures to 0 if the variable is missing in Compustat. Acquisitions are measured by AQC/AT and in addition we create a dummy variable that takes the value 1 if the firm paid a dividend (DVC>0). In Table A.1 we present descriptive statistics on the main variables used in the cash holding regressions that are are not presented in Table 1.

Compustat is also the source for data used to relate debt ratings to our measures of

Table A.1: Descriptive statistics for variables used in cash holding and hedging regressions.

Variable	mean	sd	p5	p50	p95	N
Market to Book	2.929	5.301	0.774	1.567	8.164	61,330
ln(assets)	0.471	2.382	-3.527	0.525	4.354	61,330
Cash Flow	-0.140	0.811	-0.928	0.056	0.193	61,330
Working Capital	-0.069	0.867	-0.538	0.045	0.399	61,330
Capital Exp.	0.057	0.068	0.003	0.035	0.196	61,330
Leverage	0.249	0.255	0.000	0.191	0.820	61,330
RnD	0.405	2.089	0.000	0.000	1.126	61,330
Dividend	0.239	0.427	0.000	0.000	1.000	61,330
Acquisitions	0.024	0.065	0.000	0.000	0.164	61,330

The table presents descriptive statistics on the variables used in the cash holding regressions that are are not presented in Table 1. All variables from Compustat for fiscal years 1995-2013 inclusive. Financial firms and utilities excluded.

uncertainty. We use monthly Standard&Poor's long-term debt ratings (SPLTICRM) and match them by fiscal year and month to the firms in our data. We allow firms to have multiple securities. We code the long-term ratings as follows: we assign a value 7 to a bond if it is rated AAA, 6 for AA ratings, 5 for A ratings, 4 for B ratings, 3 for BB, 2 for BBB and 1 to everything below. Since we can't include firm fixed effects in the probit estimations, we use additional controls to absorb some of the cross-sectional heterogeneity inspired by the rating regressions in Blume et al. (1998). We calculate interest rate coverage (XINT+ OIADP)/XINT where OIADP is operating income after depreciation, operating income to sales (OIBDP/SALE) and debt to asset ratio (DLTT/AT).[17] To get the manager fixed effects in our robustness tests, we identify the CEO ID for each firm and fiscal year using the Execucomp database.

C Additional validation: relation to market beta

We next analyze the relation of our indices to betas from a market model regression and a three factor model as introduced by Fama and French (1993) and present the standardized regression coefficients in Table A.2. For our CAPM regressions we obtain financial market data via CRSP at a daily frequency. We calculate fiscal-year matched

[17] Blume et al. (1998) include average short term borrowings (BAST) in their calculation of leverage but for us this reduces sample size substantially and in the regressions reported in section 6 short term borrowings are omitted from total debt.

betas using holding period returns (RET). We match this with the Fama/French return series data provided via WRDS: we use the one month treasury bill return as the risk free rate and the value-weighted return on all NYSE, AMEX and NASDAQ stocks as market portfolio. We also run a three-factor model using the small minus big (SMB) and high minus low (HML) portfolio return series. We match betas by fiscal year to the Compsutat/CRSP universe.[18] We winsorize betas at the 1% level.

We then regress the calculated betas on our uncertainty indices. We include firm fixed effects to account for differential accounting practices across firms and time fixed effects to control for general changes in accounting standards and other common time effects. The results are shown in columns (1) and (2). In columns (3) and (4) we add the share of positive and negative terms to control for risk severity which does not change the coefficient estimates much, but reduces significance in the models using Market Beta's as outcome variable. In columns (5) to (8) we repeat this using our sentence-level indices. We see that both our indices are positively correlated with market risk, which is again supportive for our measurement approach.

We also note that albeit we can reject equality of the coefficient estimates for risk and ambiguity in most model specifications above, it is not clear, that this should be interpreted as evidence for differences in severity and duration of the risks captured. In the case of market risks, differential effects should be expected in the presence of ambiguity-averse investors. For example, Caballero and Simsek (2013) have a model where investors are more prone to shun ambiguity in turbulent times. There might be not only a flight to certainty but also a flight to risk and away from ambiguity, which would explain such an empirical pattern.

[18]We extended our daily stock data to include weekends. A full year thus corresponds to 365 days.

Table A.2: Relation to betas.

	(1) Market Beta	(2) Fama-French	(3) Market Beta	(4) Fama-French
Risk	0.0123**	0.0174***	0.00994*	0.0159***
	(0.00513)	(0.00504)	(0.00512)	(0.00504)
Ambiguity	0.0514***	0.0326***	0.0518***	0.0327***
	(0.00599)	(0.00584)	(0.00595)	(0.00583)
Positive			0.0774***	0.0559***
			(0.00985)	(0.00892)
Negative			-0.0262***	-0.0101
			(0.00861)	(0.00801)
Observations	72,472	72,472	72,472	72,472
R-squared	0.513	0.387	0.515	0.388
FE	Firm Year	Firm Year	Firm Year	Firm Year
Measures	Words	Words	Words	Words

	(5) Market Beta	(6) Fama-French	(7) Market Beta	(8) Fama-French
Risk	0.0112**	0.0163***	0.00817	0.0140***
	(0.00532)	(0.00520)	(0.00530)	(0.00521)
Ambiguity	0.0474***	0.0283***	0.0465***	0.0273***
	(0.00625)	(0.00615)	(0.00623)	(0.00614)
Positive			0.0758***	0.0549***
			(0.00987)	(0.00893)
Negative			-0.0262***	-0.0103
			(0.00862)	(0.00801)
Observations	72,472	72,472	72,472	72,472
R-squared	0.513	0.387	0.514	0.388
FE	Firm Year	Firm Year	Firm Year	Firm Year
Measures	Sentence	Sentence	Sentence	Sentence

Our uncertainty, risk and ambiguity indices are positively related to market risk. The dependent variables is calculated using 360 day rolling regressions on daily stock-market data (including weekends) and matched by fiscal-year end to the CRSP/Compustat universe. Effects are standardized. No additional controls are included.

D Details on Form 10-K Data.

U.S. public companies are required to file their Form 10-K, which essentially corresponds to the firms' annual reports, digitally with the Electronic Data Gathering Analysis and Retrieval (EDGAR) system of the SEC. The process to move files onto the EDGAR system began on April 26, 1993. On December 19, 1994, the Commission made the EDGAR rules final and applicable to all concerned parties. On January 30, 1995 phase-in commenced and as of May 6, 1996, all public domestic companies were required to make their filings on EDGAR, except for filings made in paper because of a hardship exemption. We can thus speak of a reasonably complete dataset from calendar year 1996 or fiscal year 1995 onwards. A filing contains multiple documents, one of which is the actual Form 10-K, which we will use for our analysis. The structure of Form 10-K is strictly regulated and has not changed much over the years (see Nelson and Pritchard (2007)). We provide an overview of the different sections in the 10-K in Table A.3

Using the quarterly master index files available on the SEC server, we identify all 10-K and 10-KSB filings starting January 1 1994 and ending December 31 2014. Specifically, we download all 10-K/A, 10K405, 10-K405/A, 10-KSB, 10-KSB/A, 10-KSB40 and 10-KSB40/A. For each firm and fiscal year, we keep just one document, which we select as follows. We prefer 10-K and 10-K405 over 10-KSB and 10-KSB40. We prefer any of those to the appendix filings. Only the most preferred filing is kept for each firm and fiscal year. Last, we delete all filings with less than 2000 tokens. We further drop all observations prior to fiscal year 1995, as filing via EDGAR only became mandatory in May 1996. After these limitations we have a sample of 218,516 observations.

For each filing, we select the Form 10-K document only and exclude all exhibits, PDFs, Excel files and images attached to the filing. Next, we remove all XBRL and ASCII data from the document. Parsing the document into a tree structure, allows us to identify specific HTML elements. We select and delete all tables with more than 15% numeric characters, because such tables do most likely not contain actual text. Using the Jericho Java library, we extract all remaining text from the reduced Form 10-K. If a Form 10-K document is not stored in HTML format, which is the case for older filings, we still run this procedure, but won't be able to detect and remove tables. The raw text data is split into sentences using the *Stanford Natural Language Processing Toolkit* library presented in Manning et al. (2014). We store sentences when they contain at least one word from our uncertainty word list. For each document we go through all sentences and construct word counts using the overall uncertainty, risk and ambiguity word lists.

We also calculate the highest cosine similarity of each sentence with any sentence of the firm's annual report from the previous year and identify whether the sentence contains any negation terms. As measures of document length, we count the number of sentences and words that occur in the complete Loughran and McDonald (2011) dictionary. Using this raw data we define measures of risk and ambiguity based on word counts or sentences.

Our procedure follows the recommendations by Loughran and McDonald (2011) closely, except that we base our text measures on the actual Form 10-K only, which excludes all attached exhibits. This way we loose some potentially meaningful information that was filed by reference to an appendix, but the remaining text is more likely to be relevant for our purposes. To test our procedure, we also ran it when including the exhibits. The correlation is .96 between our uncertainty word counts including exhibits and the comparable measure produced by Loughran and McDonald (2011). This provides us with a useful cross-check of our algorithm.

Table A.3: Overview of contents in 10-K document. Information obtained via: http://www.sec.gov/answers/reada10k.htm, accessed June 15, 2015.

	Item	Title	Description
Part 1	*1*	Business	Description of the company's business.
	1A	Risk Factors	List of most sigificant risks to company or its securities. Risks might apply to company or economoy as a whole. Usually it is not addressed how company manages risks.
	1B	Unresolved Staff Comments	Explanation of comments received by SEC staff on previously filed reports that have not been resolved.
	2	Properties	Company's significant properties (main plants, mines, other physical properties)
	3	Legal Proceedings	Information about significant pending lawsuits or other legal proceedings, other than ordinary litigation.
	4	empty	No information here; reserved for future regulation.
Part II	*5*	Market for Registrant's Common Equity, Related Stockholder Matters and Issuer Purchases of Equity Securities	Information on equity securities: number of holders, shares, dividends etc.
	6	Selected Financial Data	Covers selected financial information of the past five years.
	7	Management's Discussion and Analysis of Financial Condition and Results of Operation	Company's perspective on business results of past financial year; risks can be discussed i.e. assessed and how managed
	7A	Quantitative and Qualitative Disclosures about Market Risk	Exposure to market risk, such as interest rate risk, foreign currency exchange risk, commodity price risk or equity price risk. The company may discuss how it manages its market risk exposures.
	8	Financial Statements and Supplementary Data	
	9	Changes in and Disagreements with Accountants on Accounting and Financial Disclosure	Discuss any disagreements with accountants. Many investors view this disclosure as a red flag.
	9A	Controls and Procedures	Information on disclosure controls and procedures and its internal control over financial reporting.
	9B	Other information	
Part III	*10*	Directors, Executive Officers and Corporate Governance	
	11	Executive Compensation	Detailed disclosure about the company's compensation policies and programs including compensation paid to top executive officers in past year.
	12	Security Ownership of Certain Beneficial Owners and Management and Related Stockholder Matters	Information about the shares owned by the company's directors, officers and certain large shareholders, and about shares covered by equity compensation plans.
	13	Certain Relationships and Related Transactions, and Director Independence	Information about relationships and transactions between the company and its directors, officers and their family members.
	14	Principal Accountant Fees and Services	Disclose the fees they paid to their accounting firm for various types of services during the year.
Part IV	*15*	Exhibits, Financial Statement Schedules	List of the financial statements and exhibits included as part of the Form 10-K.

Table A.4: Word List Used in this Article: Consisting of a Split of Uncertainty Word List from Loughran and McDonald (2011), Continued on Next Page.

Ambiguity Words

ABEYANCE	DOUBTED	PRESUMPTIONS	UNCERTAINTY	UNPROVED
ABEYANCES	DOUBTFUL	REINTERPRET	UNCLEAR	UNPROVEN
AMBIGUITIES	DOUBTS	REINTERPRETATIO	UNCONFIRMED	UNQUANTIFIABLE
AMBIGUITY	HIDDEN	REINTERPRETATIO	UNDECIDED	UNQUANTIFIED
AMBIGUOUS	IMPRECISE	REINTERPRETED	UNDEFINED	UNRECONCILED
ANOMALIES	IMPRECISION	REINTERPRETING	UNDESIGNATED	UNSEASONABLE
ANOMALOUS	IMPRECISIONS	REINTERPRETS	UNDETECTABLE	UNSEASONABLY
ANOMALOUSLY	INCOMPLETENESS	REVISE	UNDETERMINABLE	UNSPECIFIC
ANOMALY	INDEFINITE	REVISED	UNDETERMINED	UNSPECIFIED
ARBITRARILY	INDEFINITENESS	RUMORS	UNDOCUMENTED	UNTESTED
ARBITRARINESS	INDETERMINABLE	SELDOM	UNEXPECTED	UNUSUAL
ARBITRARY	INDETERMINATE	SELDOMLY	UNEXPECTEDLY	UNUSUALLY
BELIEVE	INEXACT	SPECULATE	UNFAMILIAR	VAGARIES
BELIEVED	INEXACTNESS	SPECULATED	UNFAMILIARITY	VAGUE
BELIEVES	MAYBE	SPECULATES	UNFORECASTED	VAGUELY
BELIEVING	MIGHT	SPECULATING	UNFORSEEN	VAGUENESS
CAUTIOUS	NONASSESSABLE	SPECULATION	UNGUARANTEED	VAGUENESSES
CAUTIOUSLY	PERHAPS	SPECULATIONS	UNIDENTIFIABLE	VAGUER
CAUTIOUSNESS	PRECAUTION	SPECULATIVE	UNIDENTIFIED	VAGUEST
CONCEIVABLE	PRECAUTIONARY	SPECULATIVELY	UNKNOWN	
CONCEIVABLY	PRECAUTIONS	SPORADIC	UNKNOWNS	
CONFUSES	PRESUMABLY	SPORADICALLY	UNOBSERVABLE	
CONFUSING	PRESUME	SUDDEN	UNPLANNED	
CONFUSINGLY	PRESUMED	SUDDENLY	UNPREDICTABILITY	
CONFUSION	PRESUMES	UNCERTAIN	UNPREDICTABLE	
COULD	PRESUMING	UNCERTAINLY	UNPREDICTABLY	
DOUBT	PRESUMPTION	UNCERTAINTIES	UNPREDICTED	

Risk Words

ANTICIPATE	DEVIATING	PREDICT	RANDOMIZE	VARIANCE
ANTICIPATED	DEVIATION	PREDICTABILITY	RANDOMIZED	VARIANCES
ANTICIPATES	DEVIATIONS	PREDICTED	RANDOMIZES	VARIATION
ANTICIPATING	EXPOSURE	PREDICTING	RANDOMIZING	VARIATIONS
ANTICIPATION	EXPOSURES	PREDICTION	RANDOMLY	VARIED
ANTICIPATIONS	FLUCTUATE	PREDICTIONS	RANDOMNESS	VARIES
APPROXIMATE	FLUCTUATED	PREDICTIVE	RISKIER	VARY
APPROXIMATED	FLUCTUATES	PREDICTOR	RISKIEST	VOLATILE
APPROXIMATELY	FLUCTUATING	PREDICTORS	RISKINESS	VOLATILITIES
APPROXIMATES	FLUCTUATION	PREDICTS	RISKY	VOLATILITY
APPROXIMATING	FLUCTUATIONS	PROBABILISTIC	SOMETIME	
APPROXIMATION	IMPROBABILITY	PROBABILITIES	SOMETIMES	
APPROXIMATIONS	IMPROBABLE	PROBABILITY	VARIABILITY	
DEVIATE	LIKELIHOOD	PROBABLE	VARIABLE	
DEVIATED	OCCASIONALLY	PROBABLY	VARIABLES	
DEVIATES	ORDINARILY	RANDOM	VARIABLY	

Table A.3: *[Continued from Last Page]* Word List Used in This Article: Consisting of a Split of Uncertainty Word List from Loughran and McDonald (2011).

Unclassified Uncertainty Words				
ALMOST	CONTINGENCY	INDEFINITELY	RECALCULATE	SOMEWHERE
ALTERATION	CONTINGENT	INSTABILITIES	RECALCULATED	SUGGEST
ALTERATIONS	CONTINGENTLY	INSTABILITY	RECALCULATES	SUGGESTED
APPARENT	CONTINGENTS	INTANGIBLE	RECALCULATING	SUGGESTING
APPARENTLY	CROSSROAD	INTANGIBLES	RECALCULATION	SUGGESTS
APPEAR	CROSSROADS	MAY	RECALCULATIONS	SUSCEPTIBILITY
APPEARED	DEPEND	NEARLY	RECONSIDER	TENDING
APPEARING	DEPENDED	PENDING	RECONSIDERED	TENTATIVE
APPEARS	DEPENDENCE	POSSIBILITIES	RECONSIDERING	TENTATIVELY
ASSUME	DEPENDENCIES	POSSIBILITY	RECONSIDERS	TURBULENCE
ASSUMED	DEPENDENCY	POSSIBLE	REEXAMINATION	UNHEDGED
ASSUMES	DEPENDENT	POSSIBLY	REEXAMINE	UNSETTLED
ASSUMING	DEPENDING	PRELIMINARILY	REEXAMINING	UNWRITTEN
ASSUMPTION	DEPENDS	PRELIMINARY	RISK	VARIANT
ASSUMPTIONS	DESTABILIZING	REASSESS	RISKED	VARIANTS
CLARIFICATION	DIFFER	REASSESSED	RISKING	VARYING
CLARIFICATIONS	DIFFERED	REASSESSES	RISKS	
CONDITIONAL	DIFFERING	REASSESSING	ROUGHLY	
CONDITIONALLY	DIFFERS	REASSESSMENT	SEEMS	
CONTINGENCIES	HINGES	REASSESSMENTS	SOMEWHAT	

The table shows our split of the list of uncertainty words developed by Loughran and McDonald (2011) into risk, ambiguity or uncategorized words. Our risk indices are based on the risk words; our ambiguity indices are based on the ambiguity words and our overall uncertainty index is based on the full word list, including the unclassified words.

E Index definition details

For each document d we calculate three main indices: an overall uncertainty index based on the original Loughran and McDonald (2011) "uncertainty" word list and two sub-indices measuring risk and ambiguity. The index values vary by document d and since we keep only one 10-K filing per firm and fiscal year the indices equivalently vary by firm i and fiscal year t. Let w_j be an indicator variable that takes the value 1 for word j and 0 otherwise. We use the master word list of Loughran and McDonald (2011), containing a total of 85,131 words to the define the set of possible words and denote this set by MDL. Let n_{ijt} denote the number of occurrences of word j for firm i in fiscal year t. Given these primitives we define

$$Overall\ U._{i,t} = \left(\frac{\sum_{j\in O} n_{i,j,t}w_j}{\sum_{j\in MDL} n_{i,j,t}w_j}\right) \times 100 \text{ where } O=\{\text{Overall Uncertainty Words}\}$$
$$AMB_{i,t} = \left(\frac{\sum_{j\in AMB} n_{i,j,t}w_j}{\sum_{j\in MDL} n_{i,j,t}w_j}\right) \times 100 \text{ where } KU=\{\text{Ambiguity Words}\}$$
$$RISK_{i,t} = \left(\frac{\sum_{j\in R} n_{i,j,t}w_j}{\sum_{j\in MDL} n_{i,j,t}w_j}\right) \times 100 \text{ where } R=\{\text{Risk Words}\}$$

We thus scale our index value with the length of a document as measured by word count using the Loughran and McDonald (2011) word list. We also tested for robustness with alternative length measures as well such as the number of tokens and our results are robust to such changes.

Our sentence based measures are defined similarly; instead of using the number of words, we sum up the number of sentences per document containing a key word and scale it by the total number of sentences in the document.

References

Agliardi, E., Agliardi, R., and Spanjers, W. (2016). Corporate financing decisions under ambiguity: Pecking order and liquidity policy implications. Journal of Business Research.

Alimov, A. (2014). Product market competition and the value of corporate cash: Evidence from trade liberalization. Journal of Corporate Finance, 25:122–139.

Almeida, H., Campello, M., Cunha, I., and Weisbach, M. S. (2014). Corporate liquidity management: A conceptual framework and survey. Annual Review of Financial Economics, 6(1):135–162.

Almeida, H., Hankins, K. W., and Williams, R. (2016). Risk management with supply contracts. Working paper, SSRN 2788131.

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In The rate and direction of inventive activity: Economic and social factors, pages 609–626. Princeton University Press.

Avramov, D., Li, M., and Wang, H. (2014). Risk shocks, uncertainty shocks, and corporate policies. Working paper.

Avramov, D., Li, M., and Wang, H. (2016). Risk and corporate policies: A text-based analysis. Working paper, SSRN 2732564.

Bao, Y. and Datta, A. (2014). Simultaneously discovering and quantifying risk types from textual risk disclosures. Management Science, 60(6):1371–1391.

Bates, T. W., Kahle, K. M., and Stulz, R. M. (2009). Why do U.S. firms hold so much more cash than they used to? The Journal of Finance, 64(5):1985–2021.

Bidder, R. and Dew-Becker, I. (2016). Long-run risk is the worst-case scenario. The American Economic Review, 106(9):2494–2527.

Binmore, K. (2008). Rational Decisions. Princeton University Press.

Blume, M. E., Lim, F., and MacKinlay, A. C. (1998). The declining credit quality of U.S. corporate debt: Myth or reality? The Journal of Finance, 53(4):1389–1413.

Bossaerts, P., Ghirardato, P., Guarnaschelli, S., and Zame, W. R. (2010). Ambiguity in asset markets: Theory and experiment. Review of Financial Studies, 23(4):1325–1359.

Breuer, W., Rieger, M. O., and Soypak, C. K. (2016). Corporate cash holdings and ambiguity aversion. Review of Finance, 41(7-8):1–42.

Caballero, R. J. and Simsek, A. (2013). Fire sales in a model of complexity. The Journal of Finance, 68(6):2549–2587.

Camerer, C. and Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. Journal of risk and uncertainty, 5(4):325–370.

Campbell, J. L., Chen, H., Dhaliwal, D. S., Lu, H.-M., and Steele, L. B. (2014). The information content of mandatory risk factor disclosures in corporate filings. Review of Accounting Studies, 19(1):396–455.

Chang, F.-Y., Hsin, C.-W., and Shiah-Hou, S.-R. (2013). A re-examination of exposure to exchange rate risk: The impact of earnings management and currency derivative usage. Journal of Banking & Finance, 37(8):3243–3257.

Chen, Z. and Epstein, L. (2002). Ambiguity, risk, and asset returns in continuous time. Econometrica, 70(4):1403–1443.

Christensen, C. M. (1997). The Innovator's Dilemma: The Revolutionary Book that Will Change the Way You Do Business. Harvard Business School Press, Boston, MA.

Cole, C. J. and Jones, C. L. (2005). Management discussion and analysis: A review and implications for future research. Journal of Accounting Literature, 24:135.

Crès, H., Gilboa, I., and Vieille, N. (2011). Aggregation of multiple prior opinions. Journal of Economic Theory, 146(6):2563–2582.

Dekel, E., Lipman, B. L., and Rustichini, A. (1998). Recent developments in modeling unforeseen contingencies. European Economic Review, 42(3):523–542.

Dimmock, S. G., Kouwenberg, R., Mitchell, O. S., and Peijnenburg, K. (2016). Ambiguity aversion and household portfolio choice puzzles: Empirical evidence. Journal of Financial Economics, 119(3):559–577.

Disatnik, D., Duchin, R., and Schmidt, B. (2014). Cash flow hedging and liquidity choices. Review of Finance, 18(2):715–748.

Dow, J. and da Costa Werlang, S. R. (1992). Uncertainty aversion, risk aversion, and the optimal choice of portfolio. Econometrica: Journal of the Econometric Society, pages 197–204.

Dudley, E. and Zhang, N. (2016). Trust and corporate cash holdings. Journal of Corporate Finance, 41:363–387.

Easley, D. and O'Hara, M. (2010). Microstructure and ambiguity. The Journal of Finance, 65(5):1817–1846.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. The Quarterly Journal of Economics, 75(4):643–669.

Epstein, L. G. and Schneider, M. (2010). Ambiguity and asset markets.

Etner, J., Jeleva, M., and Tallon, J.-M. (2012). Decision theory under ambiguity. Journal of Economic Surveys, 26(2):234–270.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of financial economics, 33(1):3–56.

Farre-Mensa, J. and Ljungqvist, A. (2015). Do measures of financial constraints measure financial constraints? Review of Financial Studies.

Fernandes, N. and Gonenc, H. (2016). Multinationals and cash holdings. Journal of Corporate Finance, 39:139–154.

Friberg, R. (2015). Managing Risk and Uncertainty: A Strategic Approach. MIT Press Ltd.

Froot, K. A., Scharfstein, D. S., and Stein, J. C. (1993). Risk management: Coordinating corporate investment and financing policies. The Journal of Finance, 48(5):1629–1658.

Füllbrunn, S., Rau, H. A., and Weitzel, U. (2014). Does ambiguity aversion survive in experimental asset markets? Journal of Economic Behavior & Organization, 107:810–826.

Geroski, P. A. (2003). Competition in markets and competition for markets. Journal of Industry, Competition and Trade, 3(3):151–166.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. Journal of Mathematical Economics, 18(2):141–153.

Guidolin, M. and Rinaldi, F. (2013). Ambiguity in asset pricing and portfolio choice: A review of the literature. Theory and Decision, 74(2):183–217.

Harford, J., Klasa, S., and Maxwell, W. F. (2014). Refinancing risk and cash holdings. The Journal of Finance, 69(3):975–1012.

He, Z. and Wintoki, M. B. (2016). The cost of innovation: R&d and high cash holdings in u.s. firms. Journal of Corporate Finance, 41:280–303.

Hecker, D. (1999). High-technology employment: A broader view. Monthly Labor Review, 122:18.

Hoberg, G. and Maksimovic, V. (2015). Redefining financial constraints: A text-based analysis. Review of Financial Studies, 28(5):1312–1352.

Holmström, B. and Tirole, J. (1998). Private and public supply of liquidity. The Journal of Political Economy, 106(1):1–40.

Iskandar-Datta, M. E. and Jia, Y. (2012). Cross-country analysis of secular cash trends. Journal of Banking & Finance, 36(3):898–912.

Jensen, M. C. (1986). Agency cost of free cash flow, corporate finance, and takeovers. American Economic Review, 76(2).

Jiang, Z. and Lie, E. (2016). Cash holding adjustments and managerial entrenchment. Journal of Corporate Finance, 36:190–205.

Jones, M. J. and Shoemaker, P. A. (1994). Accounting narratives: A review of empirical studies of content and readability. Journal of Accounting Literature, 13:142.

Kearney, C. and Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. International Review of Financial Analysis, 33(0):171–185.

Keynes, J. M. (1936). The General Theory of Employment, Interest and Money. Atlantic Publishers & Distributors (P) Ltd.

Knight, F. H. (1921). Risk, uncertainty and profit. New York: Hart, Schaffner and Marx.

Kravet, T. and Muslu, V. (2013). Textual risk disclosures and investors' risk perceptions. Review of Accounting Studies, 18(4):1088–1122.

Li, F. (2010). Textual analysis of corporate disclosures: A survey of the literature. Journal of Accounting Literature, 29:143.

Liu, Q., Luo, T., and Tian, G. G. (2015). Family control and corporate cash holdings: Evidence from china. Journal of Corporate Finance, 31:220–245.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. The Journal of Finance, 66(1):35–65.

Maccheroni, F., Marinacci, M., and Rustichini, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. Econometrica, 74(6):1447–1498.

MacKay, P. (2015). Transparency of corporate risk management and performance. In Forssbaeck, J. and Oxelheim, L., editors, The Oxford Handbook of Economic and Institutional Transparency, chapter 24, pages 495–520. Oxford University Press.

Malmendier, U. and Tate, G. (2005). Ceo overconfidence and corporate investment. The journal of finance, 60(6):2661–2700.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., and McClosky, D. (2014). The Stanford CoreNLP natural language processing toolkit. In Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pages 55–60.

Neamtiu, M., Shroff, N., White, H. D., and Williams, C. D. (2014). The impact of ambiguity on managerial investment and cash holdings. Journal of Business Finance & Accounting, 41(7-8):1071–1099.

Nelson, K. K. and Pritchard, A. C. (2007). Litigation risk and voluntary disclosure: The use of meaningful cautionary language. In 2nd Annual Conference on Empirical Legal Studies Paper.

Nishimura, K. G. and Ozaki, H. (2007). Irreversible investment and Knightian uncertainty. Journal of Economic Theory, 136(1):668–694.

Nunn, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade. The Quarterly Journal of Economics, 122(2):569–600.

Pinkowitz, L., Stulz, R. M., and Williamson, R. (2016). Do us firms hold more cash than foreign firms do? Review of Financial Studies, 29(2):309–348.

Riddick, L. A. and Whited, T. M. (2009). The corporate propensity to save. The Journal of Finance, 64(4):1729–1766.

Savage, L. J. (1954). The Foundations of Statistics. Dover Publications (1972 ed), 2 revised edition.

Sufi, A. (2009). Bank lines of credit in corporate finance: An empirical analysis. Review of Financial Studies, 22(3):1057–1088.

Tirole, J. (2006). The Theory of Corporate Finance. Princeton University Press.