Internet Appendix

Narrative Ambiguity Matters

Xinbei Wei Qunzi Zhang

A Construction the *WSJ* Document-Term Matrix

We conduct data processing steps in the following order:

1. Remove all articles prior to January 1984 and after December 2024.

2. Replace all non-alphabetical characters with an empty string and set the remaining characters to lowercase.

3. Parse article text into a white-space-separated word list retaining the article's word ordering. Exclude single-letter words.

4. Exclude articles with page-citation tags corresponding to any sections other than A, B, C, or missing.

5. Exclude articles with subject tags associated with obviously non-economic content such as sports, leisure, and arts.

7. Exclude articles that contain regular data tables with little supporting text.

8. Remove common "stop" words and URL-based terms.

9. Lastly, we conduct light lemmatizing of derivative words. The following rules are applied in the order given, where 'x' is a candidate term. In each case, the stemming is only applied if the multiple terms reduce to the same stem.

(a) Replace trailing "sses" with "ss"

(b) Replace trailing "ies" with "y"

(c) Remove trailing "s"

(d) Remove trailing "ly"

(e) Remove trailing "ed." Replace remaining trailing "ed" with "e"

(f) Replace trailing "ing" with "e". For remaining trailing "ing" that follow a pair of identical consonants, remove "ing" and one consonant. Remove remaining trailing "ing"

(g) Remove words with less than 3 letters.

10. Convert an article's word list into a vector of counts for each term in the dictionary.

B Robustness Checks

B.1 Predicting International Stock Returns

The narrative ambiguity may serve as a cross-country measure of informational ambiguity, capturing the extent to which ambiguity-related concerns permeate narratives across different regions of the world. This interpretation is supported by two key observations regarding the nature of information processing in international news coverage.

First, as demonstrated in the distribution of news in *Wall Street Journal* (see Figure 1), a significant proportion of the newspaper's reports pertain to events occurring outside the United States. This indicates that the narratives contributing to the construction of ambiguity are not exclusively concentrated on the U.S. but instead reflect a broader set of geopolitical, economic, and financial uncertainties faced by a diverse set of countries. Such global reach in media coverage implies that the measured ambiguity incorporates both domestic and international information, thereby enhancing the index's potential as a global indicator of narrative ambiguity.

Second, financial markets today are increasingly interconnected, and investors across countries consume and respond to globally circulated narratives. News published by major international outlets such as the *Wall Street Journal* is not confined to domestic readers. Rather, it is widely accessed by institutional and retail investors around the world. This global dissemination of narratives implies that ambiguity embedded in such news can influence the expectations and behaviors of market participants in multiple countries.

Thus, in this section, we investigate the predictive power of our narrative ambiguity in the global perspective. The results in Table B.1 provide strong evidence of the predictive validity of narrative ambiguity measures across international equity markets. For the developed countries (see Panel A of Table B.1), narrative ambiguity exhibits a broadly positive and economically meaningful predictability in future stock market excess returns. The estimated slope coefficients are positive for all countries and statistically significant for Canada, Germany, the United Kingdom, and the United States, with France marginally significant at the 5% level. NA also generates remarkable in-sample explanatory power, with adjusted R^2 values reaching 2.545% for Canada and 1.843% for the United States. Importantly, all the out-of-sample R^2_{OOS} statistics are positive and economically significant.

Taken together, the results suggest that narrative ambiguity contains significant information for forecasting aggregate equity returns across a range of developed countries.

Panel B of Table B.1 reports results for emerging countries and documents substantial cross-country heterogeneity. Narrative ambiguity significantly predicts future excess returns in stock markets of Brazil, India, Mexico, and South Africa, with positive slope coefficients and limited out-of-sample R^2_{OOS}, while the estimated coefficient for China is statistically insignificant. Korea exhibits strong in-sample explanatory power but weak out-of-sample performance, underscoring that the predictive power of narrative ambiguity varies across emerging countries.

The results in Table B.2 provide strong evidence of the predictive validity of aligned narrative ambiguity across international equity markets. In particular, the measures derived through PLS alignment consistently outperform their raw counterparts (see Table B.1). For the G7 countries, the aligned narrative ambiguity exhibits remarkable stronger explanatory power, and the in-sample explanatory power, as measured by the average R^2 statistic, improves substantially from 1.180% for the raw NA to 2.966% for the aligned measure, NA^{PLS}. Similar results are observed for emerging markets. Overall, these findings highlight the strong predictive power of the aligned narrative ambiguity in predicting global stock returns.

B.2 Global and Local Narrative Ambiguity

Consistent with theories of globally shared risk (Gourio et al., 2013; Baker et al., 2016; Farhi and Gabaix, 2016; Lewis and Liu, 2017; Mumtaz and Theodoridis, 2017; Cesa-Bianchi et al., 2020; Caldara and Iacoviello, 2022; Chen et al., 2023), we decompose the international narrative ambiguity index into global ambiguity (NA^G) and local ambiguity (NA^L), aiming to distinguish between global-shared and country-specific sources of narrative uncertainty. This approach allows us to better capture the heterogeneous information environments faced by different markets. The decomposition enhances the interpretability of the index by isolating systematic ambiguity transmitted through global narratives from localized ambiguity rooted in domestic developments, thereby providing a more nuanced tool for examining the role of narrative-driven uncertainty in international financial markets.

Results in Table B.2 demonstrate significant differences in predictive power between global and local narrative ambiguity measures across international equity markets. Global ambiguity measure exhibits strong predictive relationships with future returns in G7 countries, with an average adjusted R^2 of 2.605%. The local ambiguity, however, generates weaker and insignificant predictive power in most developed countries' stock markets. These results indicate that, in developed markets, stock market excess returns are more sensitive to globally shared narrative ambiguity, consistent with their higher degree of financial integration. By contrast, in emerging markets, the influence of narrative ambiguity is more heterogeneous. Stock markets of Brazil, India, Korea, Mexico, and South Africa are more closely exposed to global narrative ambiguity, whereas those of China exhibit greater exposure to local narrative ambiguity.

B.3 Predicting Other Asset Returns

Although our focus is stock market, we provide additional evidence on other important assets in this section. We collect data on treasury bonds for 12 countries, future prices for 24 commodities and 4 currency forwards from the same data sources used by Huang et al. (2020) and Chen et al. (2023).[22] We calculate the monthly bond risk premium for a five-year treasury bond, following the standard procedure in Cochrane and Piazzesi (2005). For each commodity future contract, we calculate the daily excess returns on contracts with the next-nearest- or the nearest-to-delivery. Then, we obtain a time series of cumulative monthly returns by compounding these daily returns.

Owing to the limited space, we present the results for twelve Treasury bonds, four commodities, and four currencies in Table B.3. Table B.3 Panel A reports the results for Treasury bonds. As it shows, NA^{bond} exhibits significant forecasting power for all countries, and the regression adjusted R^2 ranges from 3.894% to 30.337%. Different from the implication of global narrative ambiguity on stock markets, we find that both global narrative ambiguity (NA^G) and local narrative ambiguity (NA^L) are the major forecasting contributors in bond markets. Specifically, NA^G and NA^L significantly predict bond risk premium for 9 of the 12 countries in Table B.3, and the R^2s are sizable in

[22]The 24 commodities are Aluminum, Brent oil, Cattle, Cocoa, Coffee, Copper, Corn, Cotton, Crude, Gas oil, Gold, Heating oil, Hogs, Natural gas, Nickel, Platinum, Silver, Soybeans, Soymeal, Soy oil, Sugar, Unleaded gasoline, Wheat, and Zinc.

economic magnitude.

Table B.3 Panel B reports the results for commodities. We find that $NA^{commodity}$ exhibits significant and positive predictive power for most commodity returns. Unlike the stock and bond markets, the global and local narrative ambiguity indices for commodities are defined from the perspective of the underlying asset class, rather than from a country-based perspective. The results show that the local narrative ambiguity measures perform better than global narrative ambiguity measures in forecasting the commodity returns. In particular, Brent oil, copper, and hogs have significant positive risk exposures to the local narrative ambiguity. Thus, it seems that these commodities expose to the idiosyncratic risk. While ambiguity commands a positive risk premium generally, we find negative exposure to the global NA for gold, which is usually considered as the "safe-haven" asset (Baur and Lucey, 2010). Results from Table B.3 Panel C show that, the regression slopes on NA^{currency} are significantly positive for all of the total four currency forwards. Whereas the local narrative ambiguity predicts two currencies significantly, the global narrative ambiguity predicts four out of the total four currency forwards significantly. It is worth noting that the regression coefficients on global narrative ambiguity measure are positive for three currency forwards, EUR/USD, JPY/USD, and CHF/USD, indicating positive risk premia.

B.4 Controlling Standard Stock Market Predictors

Motivated by Welch and Goyal (2008) and Rapach et al. (2019), we test whether the narrative ambiguity contains any forecasting information beyond that already conveyed by the standard stock market predictors. We first estimate a series of predictive regressions in which the narrative ambiguity index (NA) is jointly included with each of the 14 Welch and Goyal predictors individually.[23]

[23]The 14 Welch and Goyal predictors include: dividend price ratio (d/p), the difference between the log of dividends and the log of prices; dividend yield (d/y), the difference between the log of dividends and the log of lagged prices; earnings price ratio (e/p), the difference between the log of earnings and the log of prices; dividend payout ratio (d/e), the difference between the log of dividends and the log of earnings; stock variance ($svar$), the sum of squared daily returns on the S&P 500; book-to-market ratio (b/m), the ratio of book value to market value for the Dow Jones Industrial Average; net equity expansion ($ntis$), the ratio of 12-month moving sums of net issues by NYSE listed stocks divided by the total end-of-year market capitalization of NYSE stocks; treasury bills (tbl), the 3-month treasury bill: secondary market rate; long term yield (lty); term spread (tms), the difference between the long term yield on government bonds and the Treasury-bill; long term rate (ltr); default yield spread (dfy), the difference between BAA and AAA-rated corporate bond yields; default return spread (dfr), the difference between long-term corporate bond and long-term government bond returns; inflation ($infl$), the Consumer Price Index (All Urban Consumers).

As reported in Table B.4 Panel A, the coefficients on NA remain positive and highly statistically significant across all regressions, while most traditional predictors are statistically insignificant. Specifically, the estimated coefficient on NA ranges from 10.320 to 12.766 across the 14 horse-race multivariate predictive regressions, and remains statistically significant at the 1% level, while the majority of Welch and Goyal predictors become statistically insignificant once NA is included. This evidence suggests that NA captures predictive information for returns that is not subsumed by any single conventional macro-financial variable.

To further address the joint information content and potential multicollinearity among the 14 Welch and Goyal predictors, we respectively construct two aggregated measures, the equally weighted average of the 14 predictors ($Econ_{avg}$), and the first principal component (PCA) extracted from these variables ($Econ_{pca}$). When either factor is included alongside NA, the coefficients on NA remain economically large and statistically significant, indicating that the predictive power of NA is not subsumed by common macroeconomic components embedded in traditional predictors.

Then, following Rapach et al. (2019), we implement LASSO as a variable screening device and adopt the OLS post-LASSO procedure for coefficient estimation.[24] Recent studies propose OLS post-LASSO estimation to reduce the biases inherent in LASSO coefficient estimates (e.g., Efron et al., 2004; Meinshausen, 2007). The idea is to first use LASSO to reduce the dimensionality of the model; then, to alleviate shrinkage-induced bias, the coefficients for the selected predictors are re-estimated using OLS. Consistent with this framework, we use OLS post-LASSO to estimate predictive regression models for market excess returns, where the set of candidate predictors includes the lagged ambiguity index (NA) and all major predictors proposed by Welch and Goyal (2008). LASSO selection is conducted under alternative criteria, including AIC, BIC, and K-fold cross-validation. Conditional on the selected models, we re-estimate coefficients using OLS to enhance interpretability and facilitate conventional inference.

As emphasized by Rapach et al. (2019), separating variable selection from coefficient estima-

[24]Following Rapach et al. (2019), the OLS post-LASSO procedure is employed primarily for variable selection and predictive assessment rather than for formal statistical inference. While post-LASSO re-estimation via OLS mitigates the shrinkage bias inherent in LASSO coefficient estimates and improves economic interpretability, the reported t-statistics and p-values should be interpreted as descriptive rather than as conventional post-selection inference.

tion improves both economic interpretability and predictive performance relative to unrestricted OLS and pure LASSO. In line with these arguments, our post-LASSO results indicate that NA is consistently selected as an important predictor alongside a small subset of traditional variables, and the resulting models deliver substantially higher adjusted R^2 than standard benchmark regressions. Specifically, the adjusted R^2 achieves 7.028% under LASSO with AIC selection and 6.942% under K-fold cross-validation.

It is worth to note that, Leeb et al. (2015) point out the p-value for the OLS post-LASSO t-statistic is valid for making post-selection inferences for an individual coefficient. However, when testing many individual null hypotheses, conventional p-values can present a misleading picture of statistical significance. To tackle such multiple testing issue, we employ the False Discovery Rate (FDR) approach developed by Rapach et al. (2019) and Benjamini and Hochberg (2000). The results in Panel B of Table B.4 indicate that our findings are robust to this more conservative assessment. Under the LASSO (AIC) specification, 11 predictors are statistically significant based on conventional OLS post-LASSO p-values, of which 5 remain significant after controlling for the FDR at the 5% level; Importantly, NA is among these variables. Under the LASSO (BIC) specification, both selected predictors remain significant after FDR adjustment (2 out of 2), again including NA. Under LASSO with K-fold cross-validation, 13 predictors are initially significant, with 3 surviving the FDR correction, all of which include NA.

We further complement the variable selection by employing random forests, a nonlinear ensemble learning method to assess predictor importance in return predictability (Gu et al., 2020; Bryzgalova et al., 2025). Random Forests aggregate information from a large number of decision trees and are well suited to capturing nonlinearities and interaction effects among predictors without imposing parametric structure (Gu et al., 2020). Specifically, we estimate a regression random forest with the market excess return as the target variable and NA together with the full set of Welch and Goyal predictors as candidate inputs.[25] We select predictors whose permutation importance exceeds the median importance across all candidate variables. Conditional on the selected set

[25]The model is implemented using 1,000 trees, and variable importance is assessed based on out-of-bag (OOB) permutation importance, which measures the increase in prediction error when a given predictor is randomly permuted.

of predictors, we then re-estimate a linear predictive regression using OLS (post–random forest) to enhance interpretability and facilitate comparison with standard predictive regressions. Also, when Random Forests are used as an alternative, nonlinear variable screening method, 7 predictors are significant based on unadjusted p-values, and 2 remain significant after controlling for the FDR, with NA consistently retained. The empirical results show that NA is consistently selected as an important predictor, while the importance of traditional predictors varies across specifications, indicating that NA provides robust predictive information even when nonlinearities and interactions are allowed for.

Taken together, the horse-race regressions and variable-selection-methods analyses provide all consistent evidence that NA contains incremental and robust predictive information to standard stock market predictors. Importantly, this conclusion holds across linear and nonlinear models, and when model complexity is disciplined by modern machine learning techniques designed to mitigate overfitting.

B.5 Controlling the Sentiment Indices

By construction, narrative ambiguity measures the uncertain distribution of sentiment scores across news content. A natural question arises: will the stock predictability of NA remain when we control sentiment variables? Understanding this can help further elucidate the distinctive role of narrative ambiguity to conventional information embedded in sentiment.

To address this, we evaluate the robustness of NA's predictive power via sentiment-based variables. We consider two alternative sentiment indices, one proposed by Baker and Wurgler (2006) (BW) and the other being the narrative sentiment index (NS) constructed from *Wall Street Journal* articles using the BHC sentiment dictionary in combination with a bag-of-words approach, and examine their respective impacts on the equity premium. Then, we consider the volatility and skewness of the two sentiment indices.

As reported in Table B.5, narrative ambiguity continues to exhibit strong and statistically significant return predictability after accounting for these sentiment controls. The coefficient on NA remains large and highly significant at 1% significance level across regressions, even when sentiment

level, volatility, or skewness measures are included individually. Importantly, the explanatory power of sentiment variables themselves is limited. The BW sentiment level enters with a small negative coefficient, while its volatility and skewness are statistically insignificant. Similar patterns arise for the NS and its higher-order moments, all of which exhibit weak and insignificant predictive power relative to NA. Thus, after controlling for both first- and higher-order moments of sentiment, we find that the predictive power of narrative ambiguity can not be subsumed.

Taken together, these results indicate that narrative ambiguity captures a distinct dimension of market expectations related to uncertainty in economic narratives. While sentiment indices summarize the direction and intensity of market optimism or pessimism, narrative ambiguity reflects the uncertainty of distribution of prevailing narratives, which appears to contain independent and robust information for forecasting market returns.

B.6 Controlling the Geopolitical Risk, EPU, and Federal Reserve Policy

Given that geopolitical risks and Federal Reserve policy are among the most prominent drivers of uncertainty and financial market reactions (Kuttner, 2001; Baker et al., 2016; Caldara and Iacoviello, 2022; Bauer and Swanson, 2023), we consider their proxies to evaluate whether NA captures distinct informational content beyond these well-established sources of uncertainty.

In Table B.6, we explicitly account for three major classes of economic uncertainty variables emphasized in the literature. First, we control for geopolitical risk (GPR) using the index of Caldara and Iacoviello (2022), which captures uncertainty arising from geopolitical tensions, wars, and international conflicts. Geopolitical shocks have been shown to exert significant effects on asset prices and risk premia and may plausibly co-move with narrative-based measures of uncertainty. Second, we consider economic policy uncertainty (EPU) of Baker et al. (2016), a widely used measure of policy-related uncertainty that reflects fluctuations in macroeconomic and regulatory environments. EPU is particularly relevant in our setting, as narrative ambiguity in news coverage may partially reflect uncertainty surrounding policy actions or economic outlooks. Third, as a type of uncertainty measure, NA may mechanically react to Federal Reserve policy. Kuttner (2001) and Bauer and Swanson (2023) propose a measure, MPS, which captures unexpected movements in

market expectations associated with Federal Reserve policy announcements. This Fed policy–related shock is particularly suitable in our setting because it isolates announcement-driven surprises that could plausibly co-move with news narratives or investor attention. Controlling for this variable therefore directly addresses the concern that NA may simply reflect contemporaneous reactions to Federal Reserve communications or policy-related events.

Results reported in Table B.6 show that the coefficient on NA remains statistically significant and economically meaningful after controlling for GPR, EPU, and MPS. The significant p-value of NA indicates that NA is not subsumed by geopolitical events, economic policy uncertainty, or unexpected fed policy related actions. Accordingly, our results alleviate the concern that NA merely proxies for established predictors and instead support its role as an independent source of predictive information for asset returns.

B.7 Return Predictability of NA under Alternative Distributional Assumptions

We assume that the sentiment polarity is normally distributed in the construction of NA. In the appendix, we sequentially relax the distributional assumptions and execute tests assuming that the narrative polarity obey an elliptical distribution with parameters estimated from the data. Particular forms of the elliptical distribution include the normal distribution, student-t distribution, logistic distribution, exponential power distribution, and Laplace distribution (e.g., Owen and Rabinovitch, 1983). Elliptical distributions are uniquely determined by the mean and variance and can be skewed and at the same time platykurtic or leptokurtic.

First, we allow sentiment polarity to follow a Student's t distribution, which accommodates fat tails while maintaining symmetry. Under this specification, we recompute the bin probabilities and reconstruct the narrative ambiguity index (NA). As reported in Table B.7, the Student's t–based NA delivers in-sample predictive performance that is virtually indistinguishable from that of the baseline normal-based measure, with an in-sample R^2 of 1.841%, compared to 1.843% under the normal assumption. Importantly, the out-of-sample performance remains equally strong, with the R^2_{OOS} 1.594%, and the corresponding MSFE improvement is highly statistically significant. These results indicate that allowing for fat tails in sentiment polarity does not materially alter the return

predictability of narrative ambiguity.

Next, we further relax the symmetry assumption by modeling sentiment polarity using a skew-t distribution, which flexibly captures both fat tails and asymmetry. Repeating the probability calculations and ambiguity construction under the skew-t specification yields highly comparable results. As shown in Table B.7, the skew-t–based NA continues to exhibit economically meaningful and statistically significant predictive power, with an in-sample R^2 of 1.714% and an out-of-sample R^2_{OOS} of 1.513%, alongside a significant MSFE improvement.

Taken together, these sequential relaxations demonstrate that our narrative ambiguity measure is robust to increasingly flexible distributional assumptions. The stability of the narrative ambiguity and its predictive power across the normal, $t-student$, and skew-t specifications indicates that our results are not driven by restrictive parametric assumptions, but instead reflect a fundamental property of the uncertainty of distribution of sentiment-based narratives.

Table B.1: Result for Narrative Ambiguity Predicting International Stock Markets
This table reports in-sample estimation results of the following regression,

$$R_{j,t+1} = \alpha_j + \beta_j \cdot NA_t + \epsilon_{j,t+1},$$

where $R_{j,t+1}$ denotes the future value-weighted stock market excess return on country j at time $t+1$. NA_t represents the narrative ambiguity generated by news at time t. Reported are the regression slopes (β_j) and adjusted R^2 in percentage form. Also reported are the p-values. The in-sample period is March 1984 to December 2024 for developed countries and January 1995 to December 2024 for emerging markets. This table also reports the Campbell and Thompson (2008) out-of-sample R^2 statistic (R^2_{OOS}), the Clark and West (2007) MSFE statistic for the monthly out-of-sample forecasts of value-weighted stock market excess returns based on the narrative ambiguity (NA). All the predictors and regression slopes are estimated recursively using the data available at the forecast formation time t. The out-of-sample period is 1/2 of full sample period.

	β_j	p	$Adj-R^2(\%)$	$R^2_{OOS}(\%)$	MSFE
Panel A: Results for Developed Countries, 1984:03–2024:12					
Canada	14.335	0.000	2.545	1.622	4.022
France	8.761	0.045	0.621	0.867	2.134
Germany	11.605	0.008	1.225	1.638	4.063
Italy	8.553	0.088	0.391	0.687	1.687
Japan	5.225	0.215	0.111	1.026	2.530
UK	7.733	0.015	1.049	0.820	2.018
US	10.599	0.002	1.843	1.597	3.959
Panel B: Results for Emerging Countries, 1995:01–2024:12					
Brazil	17.418	0.022	1.192	2.044	3.757
China	8.916	0.209	0.163	0.781	1.417
India	14.321	0.023	1.164	0.596	1.079
Korea	33.972	0.000	5.100	-0.103	-0.185
Mexico	12.761	0.021	1.207	1.420	2.593
South Africa	12.859	0.030	1.037	1.521	2.781

Table B.2: Results for Global and Local Narrative Ambiguity Measures
This table reports estimation results of the following regression,

$$R_{j,t+1} = \alpha + \beta_j \cdot NA_t^k + \epsilon_{j,t+1},$$

where $R_{j,t+1}$ denotes the future value-weighted stock market excess return on country j at time $t+1$. $NA_{j,t}^k$ represents the global narrative ambiguity (NA_t^G) or the local narrative ambiguity ($NA_{j,t}^L$) for country j at time t. NA_t^G is measured as the first principal component of aligned narrative ambiguity ($NA_{j,t}$) and $NA_{j,t}^L$ is the residual of regressing $NA_{j,t}$ on NA_t^G for each country. Reported are the regression slopes (β_j) and adjusted R^2 in percentage form. Also reported are the p-values. The in-sample period is from March 1984 to December 2024 for developed countries and from January 1995 to December 2024 for emerging markets.

	Aligned Narrative Ambiguity			Global Ambiguity			Local Ambiguity		
	β_j	p	$Adj-R^2(\%)$	β_j	p	$Adj-R^2(\%)$	β_j	p	$Adj-R^2(\%)$
Panel A: Results for Developed Countries, 1984:03-2024:12									
Canada	0.011	0.000	4.352	0.011	0.000	4.054	0.003	0.212	0.115
France	0.009	0.001	2.180	0.009	0.001	2.044	0.002	0.400	-0.059
Germany	0.010	0.000	2.871	0.010	0.000	2.768	0.002	0.450	-0.088
Italy	0.009	0.004	1.487	0.008	0.009	1.201	0.004	0.216	0.109
Japan	0.008	0.002	1.687	0.006	0.027	0.799	0.005	0.036	0.698
UK	0.007	0.000	2.732	0.006	0.001	2.180	0.003	0.082	0.434
US	0.011	0.000	5.454	0.010	0.000	5.187	0.002	0.249	0.068
Panel B: Results for Emerging Countries, 1995:01-202412									
Brazil	0.020	0.000	3.869	0.016	0.002	2.510	0.011	0.026	1.113
China	0.013	0.005	1.954	0.008	0.099	0.481	0.011	0.022	1.194
India	0.012	0.006	1.853	0.009	0.024	1.139	0.007	0.104	0.459
Korea	0.020	0.000	3.987	0.019	0.000	3.516	0.008	0.134	0.350
Mexico	0.015	0.000	4.462	0.012	0.001	2.677	0.009	0.011	1.504
South Africa	0.012	0.002	2.444	0.012	0.003	2.157	0.004	0.287	0.038

Table B.3: Results for Other Assets
Panel A reports the results of following predictive regression,

$$r^{OA}_{j,t+1} = \alpha + \beta_j \cdot NA^{OA}_t + \epsilon_{j,t+1},$$

where $r^{OA}_{j,t+1}$ is the bond risk premium for country j at time $t+1$ and NA^{OA}_t is the bond-PLS narrative ambiguity index ($NA^{OA}_{j,t}$) for country j. It also presents the results for the global and local narrative ambiguity measures. Panel B and C show the forecasting results for commodity and currency, respectively. In each panel, reported are the slope estimate, p-statistic values, and adjusted R^2 statistics (in percentage form). All the sample periods end in December 2015.

	Aligned Narrative Ambiguity			Global Narrative Ambiguity			Local Narrative Ambiguity		
	β_j	p	$Adj-R^2(\%)$	β_j	p	$Adj-R^2(\%)$	β_j	p	$Adj-R^2(\%)$
Panel A: Results for Treasury Bond									
Canada	1.330	0.000	10.290	1.178	0.000	7.977	0.652	0.003	2.249
France	1.038	0.000	8.621	0.977	0.000	5.712	0.597	0.009	2.560
Germany	1.204	0.000	8.410	1.376	0.000	11.061	-0.275	0.196	0.183
Italy	1.318	0.000	5.667	1.396	0.000	4.814	0.536	0.130	0.573
Japan	1.018	0.000	8.059	1.008	0.000	7.898	0.457	0.013	1.408
UK	0.921	0.000	3.894	1.012	0.000	4.751	-0.030	0.899	-0.267
US	1.714	0.000	13.165	1.101	0.000	5.258	1.420	0.000	8.954
Brazil	4.937	0.000	6.951	-3.439	0.004	3.799	3.212	0.015	2.631
China	2.098	0.000	26.963	-1.945	0.000	20.116	1.085	0.002	6.612
India	2.608	0.000	28.957	1.537	0.000	11.962	2.013	0.000	17.034
Korea	2.014	0.000	30.337	1.550	0.000	20.842	1.120	0.000	8.969
SouthAfrica	2.745	0.000	14.863	1.112	0.017	2.071	2.510	0.000	12.360
Panel B: Results for Commodity									
Brent Oil	1.329	0.011	1.912	0.844	0.100	0.602	1.011	0.054	0.956
Copper	1.006	0.012	1.428	0.301	0.454	-0.118	0.970	0.016	1.309
Gold	0.741	0.002	2.423	-0.052	0.826	-0.258	0.747	0.001	2.461
Hogs	0.928	0.011	1.494	0.058	0.874	-0.264	0.953	0.009	1.591
Panel C: Results for Currency									
EUR/USD	0.472	0.004	1.923	0.459	0.005	1.805	0.111	0.502	-0.149
JPY/USD	0.530	0.001	2.446	0.415	0.013	1.393	0.352	0.035	0.936
CHF/USD	0.443	0.012	1.425	0.386	0.029	1.015	0.223	0.207	0.162
GBP/USD	0.367	0.016	1.303	-0.255	0.094	0.489	0.269	0.077	0.577

Table B.4: NA versus Welch–Goyal Predictors

Panel A reports estimation results of the following regression,

$$R_{m,t+1} = \alpha + \beta NA_t + \gamma Control_t + \varepsilon_{t+1},$$

where $R_{m,t+1}$ denotes the CRSP value-weighted excess return at time $t+1$. NA represents narrative ambiguity generated by news. $Control$ denotes the 14 financial and economic predictors proposed by Welch and Goyal (2008). Panel B and C report variable selection results using LASSO (AIC, BIC, and K-fold cross-validation) and random forests. The regressand is the CRSP value-weighted excess return, and the regressors are selected from the lagged NA and the complete set of lagged Welch-Goyal predictors. Reported coefficients are slope estimates, with p-values in parentheses. Adjusted R^2 is reported for each regression. Bold indicates significance at the 5% level according to FDR. The full sample period is March 1984 to December 2024.

	NA	dp	dy	ep	de	svar	bm	ntis	tbl	lty	tms	dfy	dfr	infl	ltr	$Econ_{avg}$	$Econ_{pca}$	$Adj-R^2$
Panel A: Multivariate Regressions																		
(1)	11.753	0.534																2.536%
	(0.001)	(0.035)																
(2)	11.715		0.541															2.583%
	(0.001)		(0.031)															
(3)	12.393			0.291														2.712%
	(0.000)			(0.021)														
(4)	10.341				0.008													2.153%
	(0.002)				(0.112)													
(5)	11.165					-0.430												1.968%
	(0.001)					(0.204)												
(6)	12.766						0.041											2.930%
	(0.000)						0.011											
(7)	10.601							-0.001										1.641%
	(0.002)							(0.996)										
(8)	10.582								-0.026									1.666%
	(0.002)								(0.729)									
(9)	10.531									-0.037								1.689%
	(0.002)									(0.628)								
(10)	10.560										-0.034							1.653%
	(0.002)										(0.810)							
(11)	10.629											0.145						1.655%
	(0.002)											(0.794)						
(12)	10.406												0.123					1.834%
	(0.002)												(0.330)					
(13)	10.320													-0.384				1.718%
	(0.002)													(0.537)				
(14)	10.596														0.065			1.825%
	(0.002)														(0.341)			
(15)	11.186															0.007		2.109%
	(0.001)															(0.128)		
(16)	10.365																0.003	2.184%
	(0.002)																(0.101)	
Panel B: Model Selection with LASSO																		
AIC	**14.494**			0.759	**0.026**	-0.470	**0.042**	0.382		**-0.441**	**-0.419**		0.089	-0.455	0.044			7.028%
	(0.000)			(0.008)	(0.002)	(0.204)	(0.274)	(0.004)		(0.000)	(0.029)		(0.552)	(0.481)	(0.587)			
BIC	**12.766**						**0.041**											2.930%
	(0.000)						(0.011)											
K-fold	**14.454**	0.092		0.695	**0.030**	-0.299	0.053	0.293		-0.450	-0.372	-1.237	0.117	-0.579	**0.049**			6.942%
	(0.000)	(0.910)		(0.017)	(0.002)	(0.454)	(0.317)	(0.052)		(0.000)	(0.056)	(0.221)	(0.444)	(0.378)	(0.545)			
Panel C: Model Selection with Random Forests																		
RF	**14.203**	-0.271	0.992	0.208			0.031		0.135	**-0.482**								4.163%
	(0.000)	(0.888)	(0.592)	(0.301)			(0.530)		(0.413)	(0.011)								

Table B.5: Control for Sentiment Indices
This table reports estimation results of the following regression,

$$R_{m,t+1} = \alpha + \beta NA_t + \gamma Sentiment_t + \varepsilon_{t+1},$$

where $R_{m,t+1}$ denotes the CRSP value-weighted excess return at time $t+1$. NA represents narrative ambiguity generated by news. *Sentiment* is the investor sentiment indices, including Baker-Wurgler Sentiment Index (BW), volatility ($VolBW$), and skewness ($SkewBW$) of Baker and Wurgler Sentiment Index, narrative sentiment index generated by news (NS), volatility ($VolNS$), and skewness ($SkewNS$) of NS. Reported coefficients are slope estimates, with p-values in parentheses. Adjusted R^2 is reported for each regression. The sample period is March 1984 to December 2023 for BW, $VolBW$, $SkewBW$ at the monthly frequency; March 1984 to December 2024 for NS, $VolNS$, $SkewNS$ at monthly frequency.

	NA	BW	$VolBW$	$SkewBW$	NS	$VolNS$	$SkewNS$	$Adj-R^2$
$R_{m,t+1}$	11.604	-0.007						2.786%
	(0.001)	(0.029)						
$R_{m,t+1}$	11.195		-0.000					1.801%
	(0.001)		(0.994)					
$R_{m,t+1}$	11.943			0.016				2.168%
	(0.001)			(0.183)				
$R_{m,t+1}$	9.831				-0.075			1.768%
	(0.005)				(0.430)			
$R_{m,t+1}$	10.681					-0.587		1.644%
	(0.002)					(0.905)		
$R_{m,t+1}$	10.562						0.005	2.016%
	(0.002)						(0.174)	

Table B.6: Control for Geopolitical Events, EPU, or Fed Policy Shocks
This table reports estimation results of the following regression,

$$R_{m,t+1} = \alpha + \beta NA_t + \gamma Policy_t + \varepsilon_{t+1},$$

where $R_{m,t+1}$ denotes the CRSP value-weighted excess return at time $t+1$. NA represents narrative ambiguity generated by news. *Policy* proxy for the economic shocks, including geopolitical risk of Caldara and Iacoviello (2022) (GPR), economic policy uncertainty of Baker et al. (2016) (EPU), and fed policy-related shocks of Kuttner (2001) and Bauer and Swanson (2023) (MPS). Reported coefficients are slope estimates, with p-values in parentheses. Adjusted R^2 is reported for each regression. The sample period is March 1984 to December 2024 for GPR, EPU at monthly frequency, and February 1988 to December 2023 for MPS at monthly frequency.

	NA	GPR	EPU	MPS	$Adj-R^2$
$R_{m,t+1}$	10.180	0.001			1.678%
	(0.004)	(0.6711)			
$R_{m,t+1}$	9.208		0.000		2.316%
	(0.007)		(0.068)		
$R_{m,t+1}$	12.192			0.004	2.531%
	(0.000)			(0.910)	

Table B.7: Stock Return Predictability by the Narrative Ambiguity under Alternative Distributional Assumptions

This table compares the predictive performance of narrative ambiguity under (1) normal distribution (NA), (2) $t-student$ distribution (NA^t), and (3) $skew-t-student$ distribution (NA^{skew-t}). The dependent variable is the one-month-ahead CRSP value-weighted stock market excess return. Reported coefficients are slope estimates, with p-value in parentheses. Adjusted R^2 is reported for each in-sample regression. This table also reports the Campbell and Thompson (2008) out-of-sample R^2 statistic (R^2_{OOS}) and the Clark and West (2007) MSFE statistic for the monthly out-of-sample forecasts of CRSP value-weighted stock market excess returns based on the narrative ambiguity (NA) and the distribution-adjusted narrative ambiguity (NA^t/NA^{skew-t}). All the predictors and regression slopes are estimated recursively using the data available at the forecast formation time t. The full sample period is March 1984 to December 2024. The out-of-sample period is 1/2 of full sample period.

	(1)	(2)	(3)
Panel A: In-sample Regression			
NA	10.599		
	(0.002)		
NA^t		10.755	
		(0.002)	
NA^{skew-t}			10.371
			(0.002)
$Adj-R^2$	1.843%	1.841%	1.714%
Panel B: Out-of-sample Performance			
R^2_{OOS}	1.597%	1.594%	1.513%
MSFE	3.959	3.954	3.748

C Additional Results

C.1 Conditional Attitudes Using VIX as an Alternative Conditioning Variable

In the previous tests, we use the historical observations to estimate the volatility of stock market. An alternative approach is to use the forward-looking volatility implied by options on the S&P 500 index, given by the VIX. Thus, we want to rule out the possibility that our narrative ambiguity NA captures the same aspects of volatility factor as the VIX.

In this section, we re-examine whether the state-dependent pricing of narrative ambiguity is sensitive to the choice of volatility proxy. We replace realized return volatility with the VIX index, estimate the following model:

$$\begin{aligned} R^M_{m,t+1} = \alpha + \beta \cdot NA_t + \gamma \cdot VIX_t + \eta \cdot (NA_t \times \vartheta_t) \\ + \sum_{i=2}^{6} \eta_i \cdot (D_{i,t} \times P_i \times NA_t \times \vartheta_t) + \varepsilon_{t+1}. \end{aligned} \tag{19}$$

We report the results in Table C.1. The findings regarding narrative ambiguity are very similar to our earlier findings in Section 3.2. In particular, the coefficients of narrative ambiguity under different market conditions are of the same magnitude and show the same pattern - increasing in the probabilities of favorable returns. That is, the narrative ambiguity premium is negative for a high expected probability of unfavorable return (loss), and the narrative ambiguity premium is positive for a high expected probability of favorable return (gain). This result indicates that the state-dependent pricing of narrative ambiguity is not merely a reflection of conventional volatility measures. Instead, narrative ambiguity captures independent information of investors' beliefs and attitudes that is distinct from volatility factors.

C.2 NA versus Disagreement Measures: Conceptual Distinction

This section provides a conceptual comparison between narrative ambiguity and commonly used disagreement measures, including the standard deviation of expected consumption growth (Buraschi and Jiltsov, 2006), mean absolute deviation of expected consumption growth (Buraschi and Jiltsov, 2007), equal weighted of disagreement (Gao et al., 2019), and aligned disagreement index (Huang

et al., 2021). While these measures are often interpreted as capturing heterogeneous belief or disagreement, they differ fundamentally in their construction. Narrative ambiguity reflects uncertainty arising from instable distribution of narratives about real-world events, whereas disagreement measures typically proxy for cross-sectional dispersion in agents' beliefs or forecasts under a shared information structure. Table C.2 summarizes the formation of these measures and their corresponding economic concepts in the model.

C.3 Return Predictability of Aligned NA Using PLS and Neural Networks

In this subsection, we extend our analysis beyond the linear PLS framework by implementing a neural network–based aggregation method. Specifically, we employ a neural network that takes the ten industry-level narrative ambiguity indices as inputs and extracts a nonlinear composite ambiguity measure, denoted as NA^{NN}. Unlike PLS, this approach allows for flexible nonlinear dependencies and interaction effects across industries, which could, in principle, be relevant during periods of heightened market stress.

We then use NA^{NN} to predict one-month-ahead market excess returns using the same empirical framework as in the baseline analysis. As reported in Table C.3, while the neural network–based measure exhibits statistically significant predictive power, its predictive performance is weaker than that of the PLS-based narrative ambiguity measure, both in-sample and out-of-sample. In particular, the PLS-based measure achieves higher in-sample R^2 and out-of-sample R^2_{OOS}, as well as larger MSFE improvements relative to the historical mean benchmark. Our finding that a nonlinear neural network aggregation does not outperform the linear PLS-based measure suggests that the pricing effects of narrative ambiguity are largely additive across industries. Industry-level narrative ambiguity appears to capture a common risk component that accumulates gradually at the market level, rather than exhibiting strong threshold effects or nonlinear interactions. Moreover, industry-level ambiguity measures are highly comoving, which limits the scope for economically meaningful nonlinear interactions. In this setting, the PLS approach provides an efficient and transparent linear aggregation that balances predictive relevance and estimation stability, explaining its superior out-of-sample performance relative to more flexible nonlinear methods.

C.4 Validation and Semantic Distinctiveness of NA

Our NA measure captures the instability and inconsistency of the probability beliefs themselves. We conduct a content validation exercise to examine whether NA captures the linguistic features of Knightian ambiguity rather than merely negative tone. We use GPT-4o to audit articles drawn from high-NA and low-NA periods, defined as the top 10 and bottom 10 months ranked by monthly NA. Each article is independently scored on two dimensions—semantic ambiguity and sentiment—based on a predefined rubric that explicitly separates vagueness and uncertainty about future outcomes from directional tone. The prompt and scoring criteria are provided in Appendix D.2.

Table C.4 shows the results of this validation exercise. First, the Spearman correlation between LLM-generated semantic ambiguity and sentiment scores is only moderate in both the high-NA group ($\rho = -0.373$) and the low-NA group ($\rho = -0.360$), indicating that while ambiguity and negative sentiment may naturally coexist in financial narratives, they still capture different information.

More importantly, articles from high-NA periods exhibit significantly greater semantic ambiguity than articles from low-NA periods. The difference is significant in both parametric and nonparametric tests, with t-statistic 2.693 and Mann–Whitney z-statistic 3.157. This result provides direct evidence that NA identifies periods of heightened semantic uncertainty and aligns with independent LLM-based assessments of vagueness and uncertainty about outcomes. We also find that high-NA periods are associated with significantly more negative sentiment than low-NA periods, with t-statistic -6.278 and Mann–Whitney z-statistic -6.679, which is consistent with the idea that informational cacophony tends to intensify in distressed states. Overall, NA captures the vagueness and Knightian uncertainty rather than merely its negative tone.

Recent NLP research has begun to quantify lexical polysemy using contextual embeddings. For example, Peters et al. (2018) introduce a new type of deep contextualized word representation that models both complex characteristics of word use (e.g., syntax and semantics), and how these uses vary across linguistic contexts (i.e., to model polysemy), and Garí Soler and Apidianaki (2021) demonstrate that BERT-derived representations encode information correlated with words' polysemy levels. These studies support using contextual embeddings to measure semantic variability in

text, which is why we implemented the BERT-based polysemy measure in addition to our narrative ambiguity index.

We implement an additional analysis based on a BERT-style contextual embedding model to quantify polysemy at the article level. Specifically, we use a pretrained Bidirectional Encoder Representations from Transformers (BERT) model proposed by Devlin et al. (2019) to obtain contextualized embeddings for each token in a *Wall Street Journal* article. For words that appear multiple times within an article, we measure polysemy as the dispersion of their contextual embeddings, following the intuition that greater variation in contextual representations reflects higher context-dependent semantic variability (e.g., Ethayarajh, 2019; Wiedemann et al., 2019). We then aggregate the word-level measures to the article level and compute a monthly average polysemy index, $Polysemy$, which we include as an additional control variable in our predictive regressions.

As shown by Table C.5, when the $Polysemy$ is added to the predictive regression, the coefficient on NA remains remarkably stable in both magnitude and statistical significance, with p-value 0.002. The results show that our narrative ambiguity remains quantitatively similar and statistically significant after controlling for contextual polysemy. That is, the predictive power of narrative ambiguity is not subsumed by lexical polysemy or semantic dispersion in word meanings.

Table C.1: Attitudes toward Narrative Ambiguity under Different Market Conditions
This table reports estimation results of the following regression:

$$R^M_{m,t+1} = \alpha + \beta \cdot NA_t + \gamma \cdot VIX_t + \eta \cdot (NA_t \times \vartheta_t) + \sum_{i=2}^{6} \eta_i \cdot (D_{i,t} \times P_i \times NA_t \times \vartheta_t) + \varepsilon_{t+1},$$

where, $R^M_{m,t+1}$ denotes the CRSP value-weighted stock market excess return at time $t+1$. NA represents narrative ambiguity generated by news. VIX_t is the option-implied expected volatility on the S&P 500 index. ϑ_t is the average absolute daily deviation of returns from the monthly average daily return of the S&P 500. The expected probability of favorable returns, P, is estimated from the monthly averages of the daily probabilities of favorable returns of the S&P 500. A return is considered favorable if it is greater than zero, where returns are assumed to be normally distributed. The range of P is divided into seven equal intervals (bins) of width 0.01, indexed by i. P_i is the midpoint of probability bin i, and $D_{i,t}$ is a dummy variable that equals one if the monthly expected probability of favorable returns of the S&P 500 in month t falls into bin i, and zero otherwise. Panel A reports the regression slope estimates and $Adj - R^2$, with p-value in parentheses. Panel B reports the estimated coefficients of ambiguity attitude, calculated as $\eta(P_i) = \hat{\eta} + \hat{\eta}_i$ using the estimated coefficients from Panel A. The sample period is January 1990 to December 2024.

	α	β	γ	η	η_2	η_3	η_4	η_5	η_6	$Adj - R^2$
Panel A: Predictive Regression										
VIX	0.000		0.000	-7496.977	15094.490	14445.824	15305.966	16308.588	15959.045	5.941%
	(0.958)		(0.418)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
Panel B: Attitudes toward NA										
P				0.47-0.48	0.48-0.49	0.49-0.50	0.50-0.51	0.51-0.52	0.52-0.53	
VIX				-7496.977	7597.513	6948.846	7808.988	8811.611	8462.067	

Table C.2: A Comparison of Narrative Ambiguity with Other Measures of Heterogeneous Belief/Disagreement
This table compares features of narrative ambiguity with those of empirical measures proxying for disagreement among heterogeneous agents. These measures include the standard deviation of expected consumption growth of Buraschi and Jiltsov (2006) (BJ), the mean absolute deviation of expected consumption growth of Buraschi et al. (2014) (BTV), the equal weighted of disagreement indices of Gao et al. (2019) (Gao), and the aligned disagreement index of Huang et al. (2021) ($Huang$), and this article. The comparison table summarizes the formation of these measures and corresponding economic concepts in the model.

	This article	Buraschi and Jiltsov (2006)	Buraschi et al. (2014)	Gao et al. (2019)	Huang et al. (2021)
The economic quantity about which the agent(s) feel uncertain / disagree	narrative polarity	expected consumption growth	expected consumption growth	(1) expected RGDP growth; (2) expected IP growth; (3) expected UNEMP; (4) expected INV growth	24 disagreement measures
Statistics underlying the measure	uncertainty of probability distribution	standard deviation	mean absolute deviation	equal weighted of standard deviation of 4 expectation variables	PLS of 24 disagreement indices
Measurement approach	narrative-based	survey-based	survey-based	survey-based	index-based
Data source	WSJ	SPF	I/B/E/S	Blue Chip	Multi-source
Variable Name	NA	BJ	BTV	Gao	$Huang$

Table C.3: Stock Return Predictability by the Aligned Narrative Ambiguity Using PLS and Neural Networks

This table compares the predictive performance of aligned narrative ambiguity constructed using a linear partial least squares (PLS) aggregation (NA^{PLS}) and a nonlinear neural network (NN) aggregation (NA^{NN}). The dependent variable is the one-month-ahead CRSP value-weighted stock market excess return. Reported coefficients are slope estimates, with p-value in parentheses. Adjusted R^2 is reported for each in-sample regression. This table also reports the Campbell and Thompson (2008) out-of-sample R^2 statistic (R^2_{OOS}) and the Clark and West (2007) MSFE statistic for the monthly out-of-sample forecasts of CRSP value-weighted stock market excess returns based on the narrative ambiguity (NA) and the aligned narrative ambiguity (NA^{PLS}/NA^{NN}). All the predictors and regression slopes are estimated recursively using the data available at the forecast formation time t. The full sample period is March 1984 to December 2024. The out-of-sample period is 1/2 of full sample period.

	(1)	(2)	(3)
Panel A: In-sample Regression			
NA	10.599		
	(0.002)		
NA^{PLS}		0.011	
		(0.000)	
NA^{NN}			1.041
			(0.000)
$Adj-R^2$	1.843%	5.454%	3.006%
Panel B: Out-of-sample Performance			
R^2_{OOS}	1.597%	2.270%	1.416%
MSFE	3.959	5.644	3.489

Table C.4: Content Validation of NA Using LLM

This table reports a content-validation exercise based on LLM-generated article-level scores for semantic ambiguity and sentiment. Panel A reports the Spearman correlation between the semantic ambiguity score and the sentiment score separately for the high-NA and low-NA groups. Panel B compares the average scores across the two groups and reports both the difference in means based on the t-test and the difference in distributions based on the Mann–Whitney rank-sum test. The high-NA group consists of articles with top 10 values of NA, and the low-NA group consists of articles with bottom 10 values of NA. *** denotes significance at the 1% level.

	High NA Group	Low NA Group	Difference / Test
Panel A: Correlation between Semantic Ambiguity and Sentiment			
Spearman correlation, ρ	-0.373	-0.360	–
Panel B: Cross-group Comparisons			
Semantic ambiguity score	3.752	3.737	$t = 2.693^{***}$
			Mann–Whitney $z = 3.157^{***}$
Sentiment score	2.423	2.470	$t = -6.278^{***}$
			Mann–Whitney $z = -6.679^{***}$
Number of articles	12879	17663	

Table C.5: Compare with Polysemy Predictor
This table compares the predictive performance of narrative ambiguity (NA) with a contextual polysemy index derived from BERT embeddings ($Polysemy$). $Polysemy$ is measured as the dispersion of contextualized token embeddings within an article and aggregated to the monthly level. The dependent variable is the one-month-ahead CRSP value-weighted stock market excess return. Reported coefficients are slope estimates, with p-value in the parentheses. Adjusted R^2 is reported for each in-sample regression. This table also reports the Campbell and Thompson (2008) out-of-sample R^2 statistic (R^2_{OOS}) and the Clark and West (2007) MSFE statistic for the monthly out-of-sample forecasts of CRSP value-weighted stock market excess returns based on the narrative ambiguity (NA) and the polysemy index ($Polysemy$). All the predictors and regression slopes are estimated recursively using the expanding data available at the forecast formation time t. The full sample period is March 1984 to December 2024. The out-of-sample period is 1/2 of full sample period.

	(1)	(2)	(3)
Panel A: In-sample Regression			
NA	10.599		10.582
	(0.002)		(0.002)
$Polysemy$		0.012	0.010
		(0.716)	(0.752)
$Adj-R^2$	1.843%	-0.178%	1.662%
Panel B: Out-of-sample Performance			
R^2_{OOS}	1.597%	0.209%	1.440%
MSFE	3.959	0.511	3.550

D Prompt

D.1 Industry Classification Standard Rules

We classified all news texts into ten categories according to the SIC industrial classification standards using the following process:

1. Instruct Chat GPT-4o model to rely on SIC code list and identify which industry category the 10,000 news items belong to.[26] The 10,000 news items are directly input into the GPT model via a carefully designed prompt. This prompt functions as a directive mechanism, channeling the model's focus towards generating industry classifications. In the operational framework of GPT models, a prompt is not merely a query or instruction posed to the model; it encompasses a broader scope, often integrating contextual information, specific input data, or illustrative examples to guide the model's response more effectively (Mann et al., 2020). The specific prompt is:

"Your task is to classify the given news text. There are 11 categories, namely ***NoDur, Durbl, Manuf, Enrgy, HiTec, Telcm, Shops, Hlth, Utils, Other, and None****. The following are the keyword descriptions designed for these ten categories:*

NoDur*: Consumer Nondurables, Food, Tobacco, Textiles, Apparel, Leather, Toys*$\{Ca1\}$;[27] ***Durbl****: Consumer Durables, Cars, TVs, Furniture, Household Appliances*$\{Ca2\}$; ***Manuf****: Manufacturing, Machinery, Trucks, Planes, Chemicals, Off Furn, Paper, Com Printing*$\{Ca3\}$; ***Enrgy****: Oil, Gas, Coal Extraction and Products*$\{Ca4\}$; ***HiTec****: Business Equipment, Computers, Software, Electronic Equipment*$\{Ca5\}$; ***Telcm****: Telephone and Television Transmission*$\{Ca6\}$; ***Shops****: Wholesale, Retail, Some Services (Laundries, Repair Shops)*$\{Ca7\}$; ***Hlth****: Healthcare, Medical Equipment, and Drugs*$\{Ca8\}$; ***Utils****: Utilities*$\{Ca9\}$; ***Other****: Mines, Constr, BldMt, Trans, Hotels, Bus Serv, Entertainment, Finance*$\{Ca10\}$; ***None****: Does not fall into the above ten categories.*

The given news text*:* $\{TEXT\}$;

Output *the category of the given text, choose one from* ***[NoDur, Durbl, Manuf, Enrgy, HiTec, Telcm, Shops, Hlth, Utils, Other, None]****, and do not make redundant explanations."*

2. Use these 10,000 classified news items and the corresponding industrial categories as input to

[26] See the full list of SIC code at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_10_ind_port.html.

[27] $Ca1$ is the keyword descriptions of this industry from SIC code list.

train the BERT model, and classify all the remaining news items into 11 categories. This procedure yields, for each article, a probability distribution over the ten industry categories, which is used as weights to construct industry-level narrative ambiguity. Here, we adopt the Fama-French ten-industry classification to align with previous literature, the None category is dropped accordingly.

D.2 Auditing Semantic Ambiguity and Sentiment

The following prompt is used to ensure the model distinguishes "Semantic Ambiguity" (uncertainty about probabilities) from "Sentiment" (directional pessimism):

You are a trained auditor of financial and economic narratives. Your task is to evaluate the given news text on two analytically distinct dimensions: ***Ambiguity and Sentiment****.*

The given news text*:* $\{TEXT\}$

Your central objective is to strictly disentangle ambiguity from sentiment.

Core rules*: - Negative or positive content does NOT by itself imply ambiguity. - Ambiguity does NOT by itself imply positive or negative sentiment. - Routine forecasting language, standard journalistic hedging, and ordinary uncertainty about future outcomes may contribute to Ambiguity when they materially affect how the text should be interpreted. - Ambiguity should reflect the overall extent to which the text is non-committal, interpretively open, or unresolved in meaning. - Sentiment should reflect the overall tone of the language, ranging from negative to positive. - Evaluate only the text itself, not the real-world situation behind it. - Do NOT use the presence or absence of specific numbers as a scoring criterion.*

Dimension 1: Ambiguity

Definition: Measure whether the text is vague, non-committal, interpretively open, or contains unresolved competing interpretations.

Scoring categories:

1 = Very Low Ambiguity: The meaning is highly explicit, direct, and settled, with almost no meaningful openness in interpretation.

2 = Low Ambiguity: The text is mostly clear, but contains limited hedging, mild openness, or small areas of unresolved meaning.

3 = Medium Ambiguity: The text contains noticeable hedging, non-committal wording, or interpretive openness, and ambiguity is a meaningful but not dominant feature.

4 = High Ambiguity: The text is clearly open-ended, non-committal, or interpretively unresolved, and this substantially affects how the text should be understood.

5 = Very High Ambiguity: The text is strongly vague, conceptually unclear, or centrally structured around unresolved competing interpretations.

Dimension 2: Sentiment

Definition: Measure the overall tone of the language, ranging from negative to positive.

Scoring categories:

1 = Very Negative Sentiment: The language is intensely negative, alarming, crisis-oriented, or overwhelmingly dominated by adverse wording.

2 = Negative Sentiment: The language is clearly negative, pessimistic, or heavily focused on adverse developments.

3 = Neutral Sentiment: The language is broadly neutral, balanced, or only mildly positive/negative overall.

4 = Positive Sentiment: The language is clearly positive, reassuring, or favorable in tone.

5 = Very Positive Sentiment: The language is strongly positive, optimistic, encouraging, or overwhelmingly dominated by favorable wording.

***Important scoring reminders**: - A text can be very negative or very positive but have low Ambiguity. - A text can be highly ambiguous but still have neutral, negative, or positive Sentiment. - Do not mechanically equate caution with high Ambiguity, but do allow repeated hedging, open-endedness, and unresolved interpretation to raise Ambiguity meaningfully. - Scores of 1 and 5 should be used for relatively clear extreme cases. - When both positive and negative language are present, score Sentiment based on the overall balance of tone rather than isolated phrases.*

***Examples**:*

Example 1 Text: The company reported weaker demand, lower margins, and declining sales. Management warned that market conditions remain difficult. Output: Ambiguity: 2; Sentiment: 1

Why: The article is clearly negative, but its meaning is straightforward. There is little ambi-

guity, while the tone is distinctly adverse.

Example 2 Text: Executives said results may improve over time, although market conditions remain uncertain and recovery will depend on broader demand trends. Output: Ambiguity: 3; Sentiment: 3

Why: The text contains ordinary cautious forecasting language and open-ended conditions. The main meaning is understandable, but ambiguity is clearly present. The tone is mixed rather than strongly positive or negative.

Example 3 Text: Some analysts said the policy would help stabilize growth, while others argued it could deepen the downturn by weakening confidence and spending. Output: Ambiguity: 5; Sentiment: 3

Why: The text presents clearly conflicting interpretations of the same development. The disagreement is central and unresolved, while the overall tone is mixed.

Example 4 Text: The company said the outlook remains uncertain and that future performance will depend on external conditions. Output: Ambiguity: 4; Sentiment: 3

Why: The text is non-committal and open-ended. Even without explicit conflict, the unresolved and conditional nature of the message substantially affects interpretation. The tone itself is mostly neutral.

Example 5 Text: Economists offered sharply different views on whether the recent slowdown reflected temporary adjustment, policy error, or the start of a broader recession, and the article did not resolve these competing explanations. Output: Ambiguity: 5; Sentiment: 2

Why: The narrative centers on unresolved competing interpretations, which strongly raises ambiguity, and the framing is also somewhat negative.

Example 6 Text: The firm faces serious losses, rising costs, and worsening market pressure, but management clearly stated that it will cut investment and reduce headcount. Output: Ambiguity: 2; Sentiment: 1

Why: The tone is strongly negative, but the message is explicit and clear. High negativity does not imply high ambiguity.

Example 7 Text: The company reported strong earnings, rising demand, and improving margins.

Management expressed confidence in the outlook for the coming quarters. Output: Ambiguity: 1; Sentiment: 5

Why: The message is direct and settled, with very little ambiguity, and the tone is strongly positive.

Now evaluate the given news text.

***Output format**:*

Ambiguity: 1/2/3/4/5

Sentiment: 1/2/3/4/5

Return only these two lines and nothing else.