Internet Appendix for: How News and Its Context Drive Risk and Returns Around the World

Charles Calomiris and Harry Mamaysky

May 21, 2018

This Supplementary Appendix contains additional appendix tables and figures referred to in the main body of the paper.

1 Distribution of never-before-seen n-grams

Figure A1 shows the distribution of unscaled conditional probabilities

$$\hat{m} = \frac{\hat{c}(w_1, w_2, w_3, w_4)}{\hat{c}(w_1, w_2, w_3)}$$

for never-before-seen n-grams. An n-gram enters the sample in any month in which it was not present in the $[t-27, t-4]$-month training window, and is then tracked for the remainder of the sample.

2 Elastic net coefficients for returns regression

Figure A2 shows rolling coefficients estimated using the elastic net decribed in the body of the paper for 1-month-ahead returns in DM and EM economies.

3 Vector autoregression analysis

Figures A3 – A6 show impulse responses from the vector autoregression described in the main body of the paper.

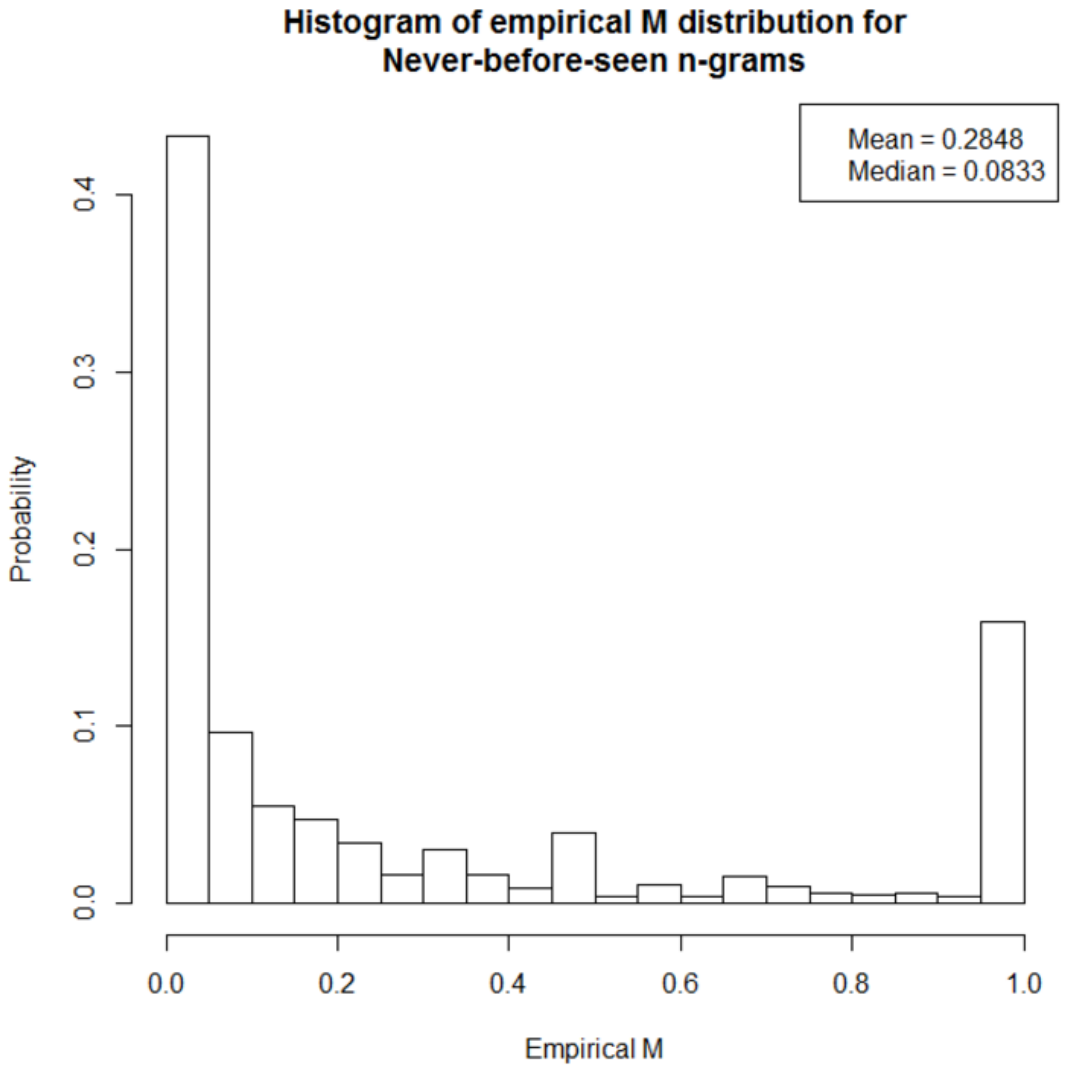

Figure A1: We observe 99.96 million never-before-seen n-grams from April 1998 through December of 2014. For a given never-before-seen n-gram in month t, we observe its empirical m in the subsample beginning in month t until the end of the full sample (December 2015). Collecting these m's across the 99.96 million never-before-seen n-grams we construct the empirical m distribution for a never-before-seen n-gram.

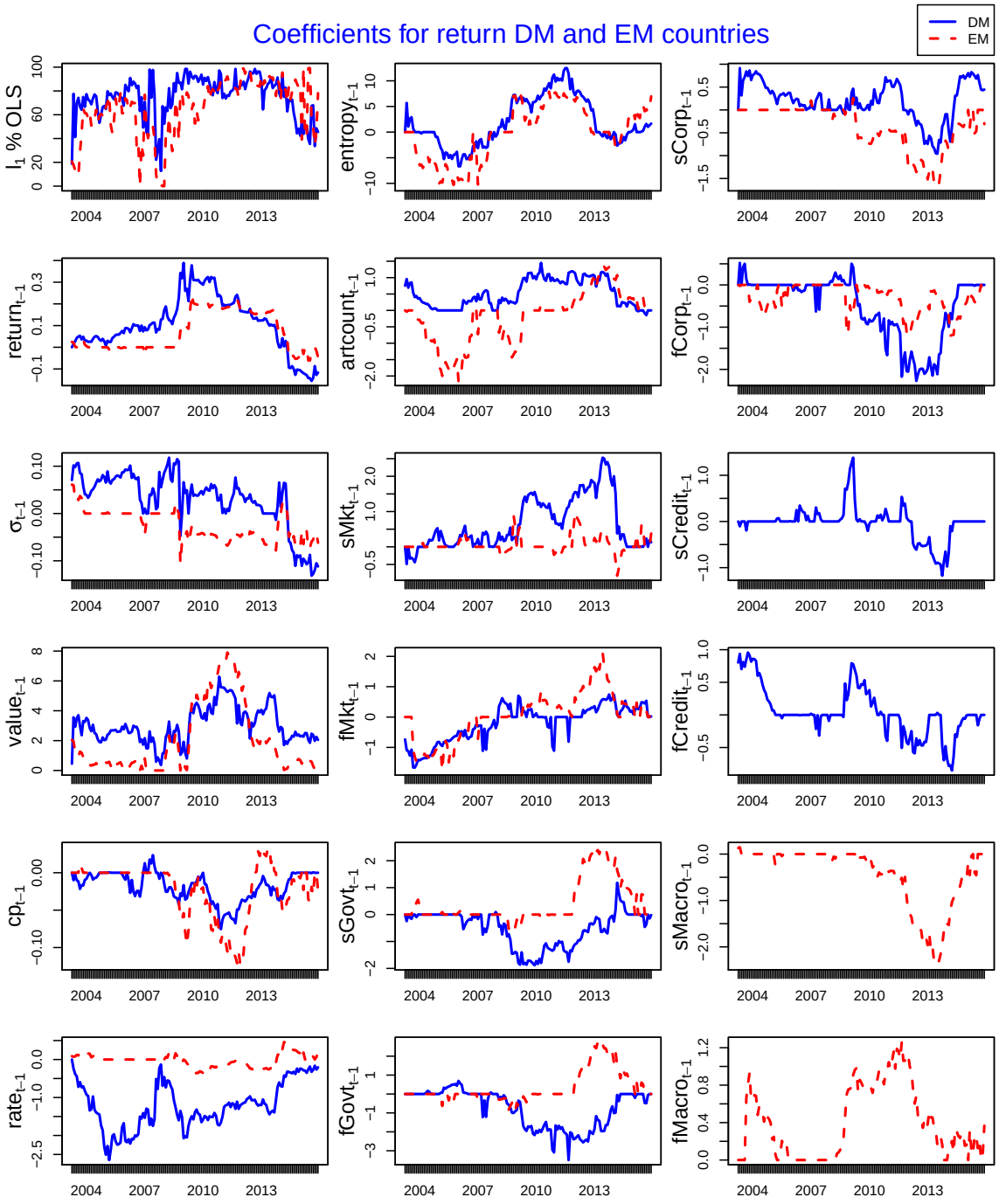

Figure A2: The charts show the time series of coefficient estimates from a rolling elastic net regression to forecast monthly returns. The chart labeled "l_1 % of OLS" gives the ratio of the elastic net coefficient l_1-norm to the OLS coefficient l_1-norm in every time period. Coefficient estimates refer to loadings on variables defined in Table 5. The elastic net regressions are run over rolling 60-month windows, with weighting parameter, λ, chosen to minimize cross-validation error. We set $\alpha = 0.75$ which represents a 0.75 weight on the lasso penalty and a 0.25 weight on the ridge regression penalty function. The out-of-sample forecasts start in March 2003 and go to December 2015.

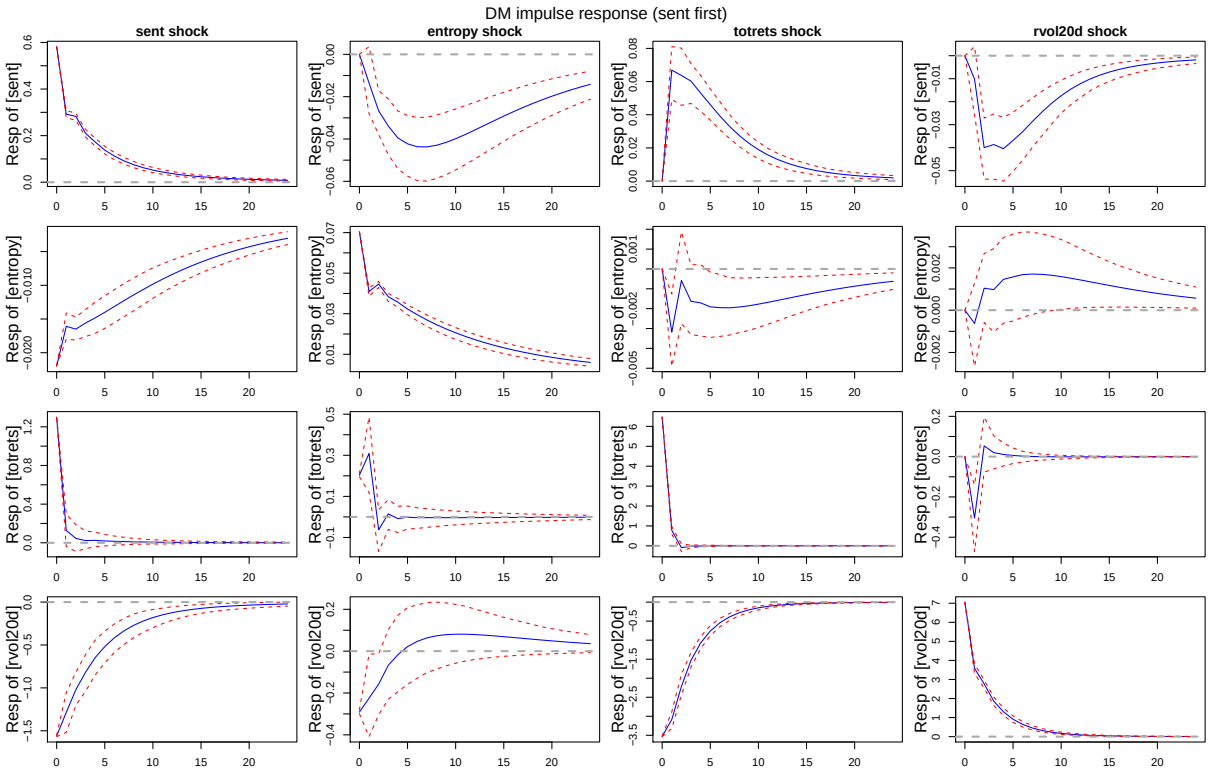

Figure A3: Shown are the impulse responses from a panel vector autoregression for the *developed markets* sample. The VAR has two lags, and includes one month returns (*totrets*), one month volatility (*rvol20d*), monthly aggregate sentiment (*sent*, which is scaled to have unit variance), and monthly entropy, as well as country fixed effects. The ordering of the VAR corresponds to the column labels in the figure. Orthogonalized shocks are computed using a Cholesky decomposition of the panel covariance matrix. Standard errors are computed by bootstraping residuals with 250 draws. Each time period represents one month.

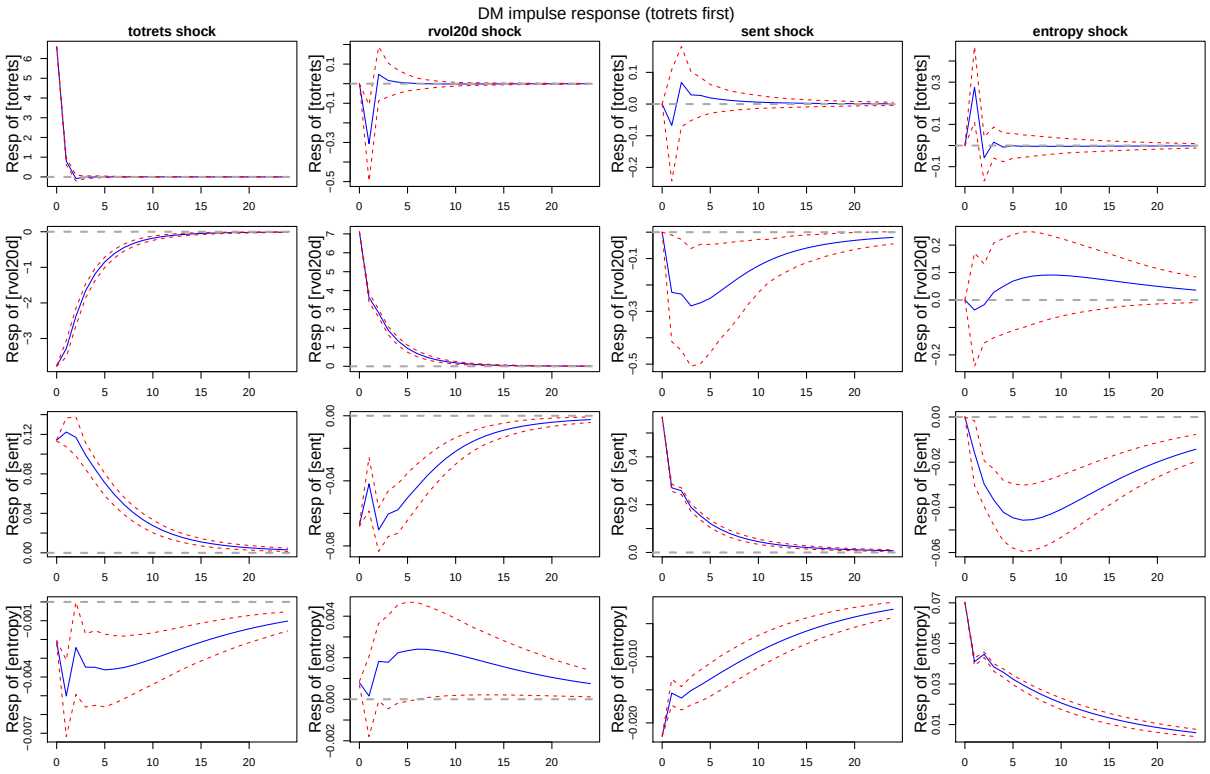

Figure A4: Shown are the impulse responses from a panel vector autoregression for the *developed markets* sample. The VAR has two lags, and includes one month returns (*totrets*), one month volatility (*rvol20d*), monthly aggregate sentiment (*sent*, which is scaled to have unit variance), and monthly entropy, as well as country fixed effects. The ordering of the VAR corresponds to the column labels in the figure. Orthogonalized shocks are computed using a Cholesky decomposition of the panel covariance matrix. Standard errors are computed by bootstraping residuals with 250 draws. Each time period represents one month.

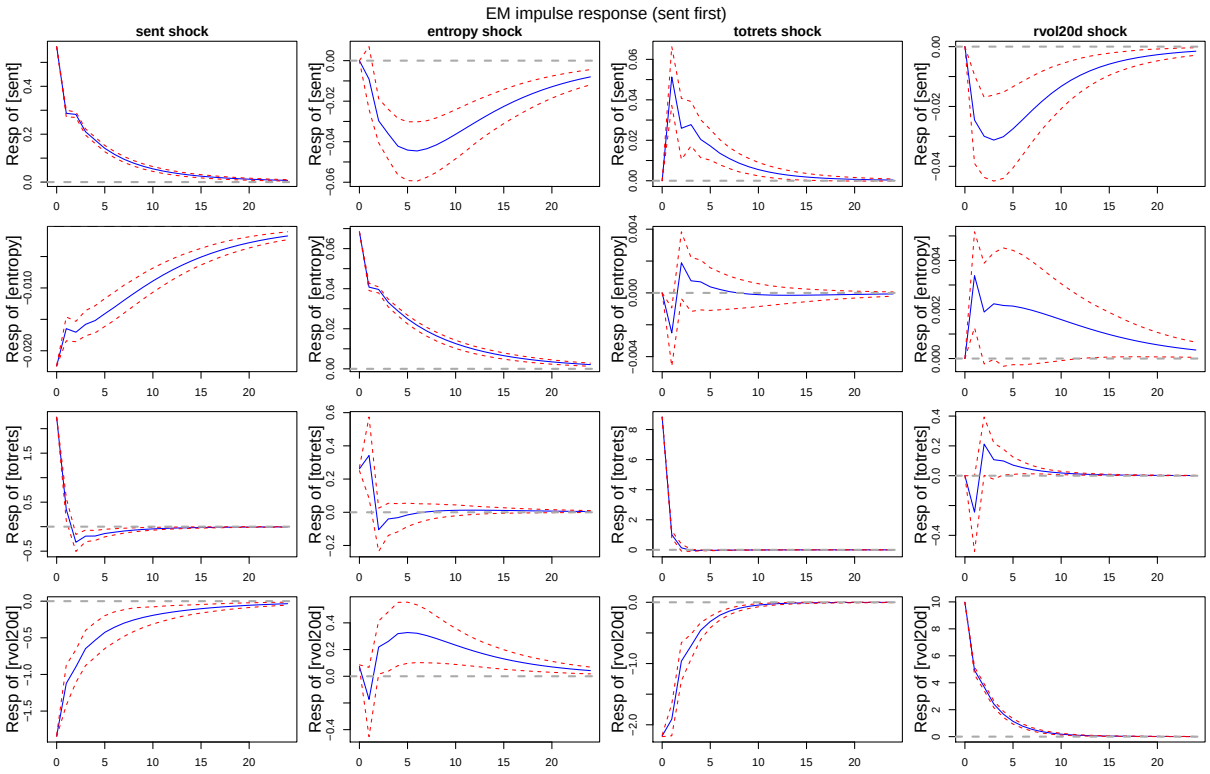

Figure A5: Shown are the impulse responses from a panel vector autoregression for the *emerging markets* sample. The VAR has two lags, and includes one month returns (*totrets*), one month volatility (*rvol20d*), monthly aggregate sentiment (*sent*, which is scaled to have unit variance), and monthly entropy, as well as country fixed effects. The ordering of the VAR corresponds to the column labels in the figure. Orthogonalized shocks are computed using a Cholesky decomposition of the panel covariance matrix. Standard errors are computed by bootstraping residuals with 250 draws. Each time period represents one month.

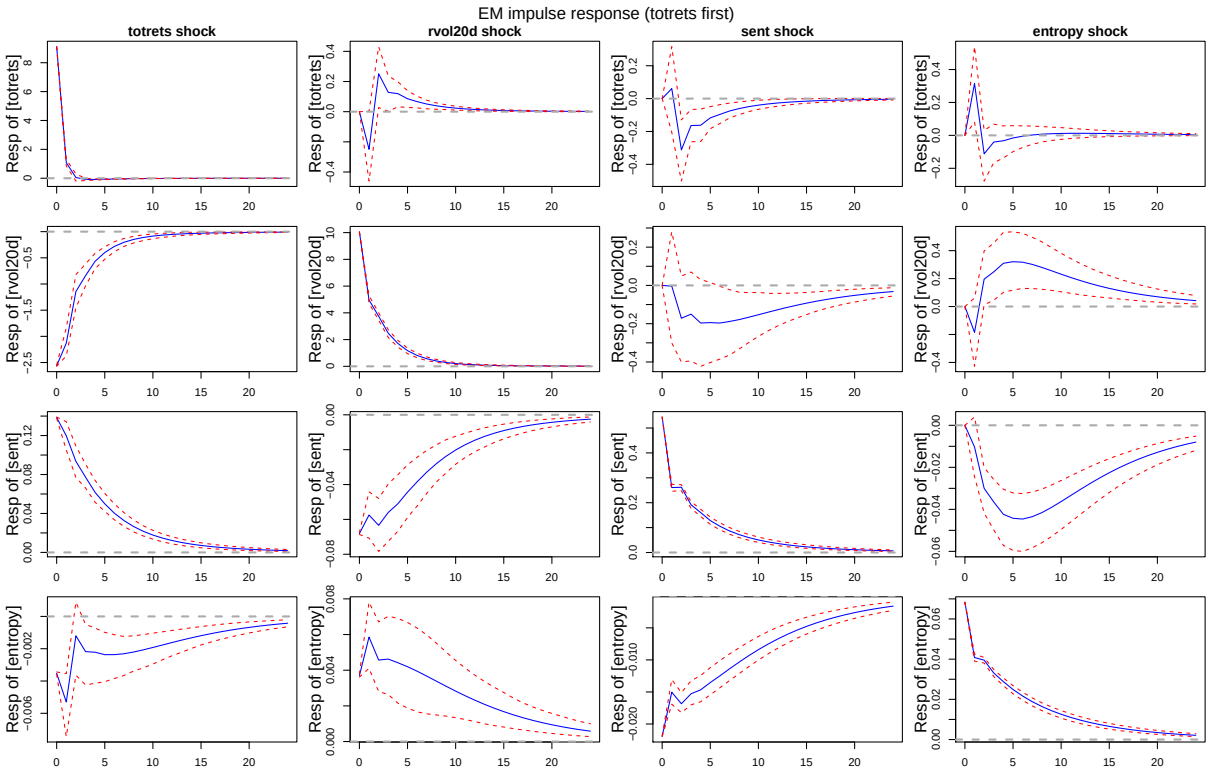

Figure A6: Shown are the impulse responses from a panel vector autoregression for the *emerging markets* sample. The VAR has two lags, and includes one month returns (*totrets*), one month volatility (*rvol20d*), monthly aggregate sentiment (*sent*, which is scaled to have unit variance), and monthly entropy, as well as country fixed effects. The ordering of the VAR corresponds to the column labels in the figure. Orthogonalized shocks are computed using a Cholesky decomposition of the panel covariance matrix. Standard errors are computed by bootstraping residuals with 250 draws. Each time period represents one month.

4 Summary of DM and EM panel regressions

Tables A1 (DM) and A2 (EM) show a one page summary of the signs of coefficient estimates in all our specifications. For DM and EM there are 12 regressions: $return$, $return^{12}$, $sigma$, and $drawdown$ for the full sample, and for the first and second subsamples.

5 Panel for 1-month returns for DM's and EM's

Tables A3 (DM) and A4 (EM) show the panel regressions for our main spefication in the paper for one-month ahead returns, $returns$.

6 Tests using the TRNA sentiment scores

Here we examine how our analysis changes when we use the sentiment indicators from the Thomson-Reuters News Analytics product, as described in the body of the paper. Table A5 gives summary statistics about correlations between TRNA and our sentiment scores across the DM and EM corpora, and across multiple topic groups. Tables A6 – A9 show our panel regressions which include either the baseline model (*Base*), our sentiment measures (*Sent*), or the TRNA sentiment measures (*TRNA*). The regressions are start in 2003 which is when the TRNA data become available. The results of our own sentiment specification (*Sent*) do not match those in the main paper because the time frame of the analyses is not the same.

Table A1: This table summarizes the coefficient estimates from all four panel specifications in the paper: *return*, $return^{12}$, *sigma*, and *drawdown*. A "+" ("-") indicates a positive (negative) coefficient estimate significant at the 10% level or better; ∅ indicates the variable is not present in a given specification. The table shows results for the developed market regressions over the entire sample, as well as the two sub-periods. Variables (all of which are defined in Table 5) include: realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of 5-year ago to current index levle (*value*), year-over-year GDP growth (*gdp*), year-over-year inflration (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and post-election dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for *Topic* in markets, government, corporate sector, commodities, credit (DM) and macro (EM).

Summary of coefficients for developed market panels

Time window	98-15 *return*	98-15 $return^{12}$	98-15 *sigma*	98-15 *drawdown*	98-07 *return*	98-07 $return^{12}$	98-07 *sigma*	98-07 *drawdown*	07-15 *return*	07-15 $return^{12}$	07-15 *sigma*	07-15 *drawdown*
$sigma_{t-1}$			+		+		+			+	+	
$sigma_{t-2}$			+				+		+	+		
$return_{t-1}$			∅	-	+		∅				∅	-
$return_{t-2}$			∅				∅				∅	
$retmi_{t-1}$	∅	∅	+	∅	∅	∅	+	∅	∅	∅	+	∅
$retmi_{t-2}$	∅	∅		∅	∅	∅		∅	∅	∅		∅
$value_{t-1}$	+	+	-	-		+	-	-	+	+	-	-
gdp_{t-1}										-		+
$gdpdeflator_{t-1}$												
cp_{t-1}	-	-		+					-	-	+	+
dcp_{t-1}					-				+	+	-	
$rate_{t-1}$	-	-	+	+	-	-	+	+	-	-	+	+
$dexch_{t-1}$						-	-					
pre												
post							+					
$entropy_{t-1}$	+			-				+	+	+		-
$artcount_{t-1}$	+											
$sMkt_{t-1}$	+			-			-		+	+		-
$fMkt_{t-1}$												
$sGovt_{t-1}$	-			+					-	-		+
$fGovt_{t-1}$				+					-			+
$sCorp_{t-1}$								-	+			
$fCorp_{t-1}$	-	-		+								
$sComms_{t-1}$												
$fComms_{t-1}$												
$sCredit_{t-1}$		+										+
$fCredit_{t-1}$					+							

Table A2: This table summarizes the coefficient estimates from all four panel specifications in the paper: *return*, $return^{12}$, *sigma*, and *drawdown*. A "+" ("-") indicates a positive (negative) coefficient estimate significant at the 10% level or better; ∅ indicates the variable is not present in a given specification. The table shows results for the emerging market regressions over the entire sample, as well as the two sub-periods. Variables (all of which are defined in Table 5) include: realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of 5-year ago to current index levle (*value*), year-over-year GDP growth (*gdp*), year-over-year inflration (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and post-election dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for *Topic* in markets, government, corporate sector, commodities, credit (DM) and macro (EM).

Summary of coefficients for emerging market panels

Time window	98-15 *return*	98-15 $return^{12}$	98-15 *sigma*	98-15 *drawdown*	98-07 *return*	98-07 $return^{12}$	98-07 *sigma*	98-07 *drawdown*	07-15 *return*	07-15 $return^{12}$	07-15 *sigma*	07-15 *drawdown*
$sigma_{t-1}$		+	+				+			+	+	
$sigma_{t-2}$		+	+				+		+	+		-
$return_{t-1}$	+		∅	-			∅			+	∅	-
$return_{t-2}$			∅				∅				∅	-
$retmi_{t-1}$	∅	∅	+	∅	∅	∅	+	∅	∅	∅	+	∅
$retmi_{t-2}$	∅	∅		∅	∅	∅		∅	∅	∅		∅
$value_{t-1}$			-	-	+	+	-	-			-	-
gdp_{t-1}	-				-	-		+	-	-		+
$gdpdeflator_{t-1}$					-				-			+
cp_{t-1}	-	-	-								-	
dcp_{t-1}			+								+	
$rate_{t-1}$		-	+	+			+		-		+	+
$dexch_{t-1}$			-	-			-			+		-
pre					-							
post					-					-		
$entropy_{t-1}$					-		+	+	+			-
$artcount_{t-1}$		-	+	+	-	-		+			+	+
$sMkt_{t-1}$							+					
$fMkt_{t-1}$					-	-	+	+		+		-
$sGovt_{t-1}$				+								
$fGovt_{t-1}$	-		+	+	-	-		+				
$sCorp_{t-1}$	-	-		+					-	-		+
$fCorp_{t-1}$	-	-	+	+		-	+	+				
$sComms_{t-1}$	+			-				-	+			
$fComms_{t-1}$									+			
$sMacro_{t-1}$												
$fMacro_{t-1}$	+	+		-					+		-	-

Table A3: Panel regressions for developed market returns. Results are shown for the base specification of the model, which excludes the text-based measures, and two text specifications, one that includes context specific sentiment (in column "Sent") and another that includes context specific sentiment interacted with entropy (in column "SentEnt"). Results are reported for the entire sample, as well as for the early and late subsamples. Observations are monthly. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include: realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of 5-year ago to current index levle (*value*), year-over-year GDP growth (*gdp*), year-over-year inflration (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and post-election dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for *Topic* in markets, government, corporate sector, commodities, credit (DM) and macro (EM). All panels include country fixed effects and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***","**", and "*" are significant at the 1%, 5% and 10% levels respectively.

Developed markets: Forecasting panel for returns

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	-0.032	-0.021	-0.020	0.114**	0.109**	0.107**	-0.115	-0.091	-0.088
$sigma_{t-2}$	0.022	0.040	0.040	-0.094*	-0.076	-0.076	0.114	0.140*	0.140*
$return_{t-1}$	0.134**	0.084	0.084	0.162**	0.121*	0.123*	0.063	0.017	0.018
$return_{t-2}$	-0.018	-0.033	-0.034	-0.048	-0.056	-0.056	0.003	-0.021	-0.023
$value_{t-1}$	1.576**	1.986***	2.004***	1.290	1.422	1.383	1.993**	2.544***	2.586***
gdp_{t-1}	-0.050	-0.124	-0.132	0.046	-0.025	-0.023	-0.140	-0.174	-0.180
$gdpdeflator_{t-1}$	0.033	-0.036	-0.038	0.115	0.063	0.062	-0.127	-0.145	-0.139
cp_{t-1}	-0.023*	-0.022**	-0.022**	0.003	0.003	0.004	-0.039**	-0.051***	-0.048***
dcp_{t-1}	0.005	0.001	0.000	-0.022**	-0.020*	-0.020*	0.040**	0.041**	0.037*
$rate_{t-1}$	-0.320**	-0.588***	-0.606***	-0.967*	-1.166**	-1.130**	-0.385	-0.664***	-0.694***
$dexch_{t-1}$	0.074	0.021	0.023	0.192	0.123	0.128	-0.007	-0.022	-0.018
pre	0.176	-0.008	-0.023	-0.048	0.017	0.013	0.437	0.066	0.032
post	0.197	0.090	0.084	-0.142	-0.151	-0.135	0.440	0.310	0.317
$entropy_{t-1}$		3.056*	3.647**		-3.631	-3.550		6.424**	7.137***
$artcount_{t-1}$		0.443*	0.416*		0.482	0.467		0.548	0.527
$sMkt_{t-1}$		0.836**	0.890**		-0.158	-0.295		1.670**	1.835***
$fMkt_{t-1}$		0.311	0.307		0.005	-0.037		0.381	0.413
$sGovt_{t-1}$		-0.973**	-0.951**		0.078	0.048		-1.592***	-1.449***
$fGovt_{t-1}$		-0.748	-0.768		0.764	0.749		-1.954***	-1.912***
$sCorp_{t-1}$		0.267	0.361		0.432	0.380		0.863*	0.903*
$fCorp_{t-1}$		-0.477*	-0.441*		0.398	0.360		-0.475	-0.413
$sComms_{t-1}$		0.236	0.192		0.326	0.355		0.069	-0.047
$fComms_{t-1}$		0.211	0.205		0.171	0.213		0.287	0.204
$sCredit_{t-1}$		0.084	0.021		0.000	0.126		-0.706	-0.813
$fCredit_{t-1}$		0.048	0.049		0.578**	0.578**		-0.358	-0.338
$R2$	0.0288	0.0463	0.0473	0.0396	0.0501	0.0499	0.0587	0.0927	0.0946
start	Apr 1998	May 1998	May 1998	Apr 1998	May 1998	May 1998	Mar 2007	Mar 2007	Mar 2007
end	Dec 2015	Dec 2015	Dec 2015	Feb 2007	Feb 2007	Feb 2007	Dec 2015	Dec 2015	Dec 2015
Nobs	4422	4406	4406	2003	1987	1987	2419	2419	2419
stderr	by time	by time	by time	by time	by time	by time	by time	by time	by time

Table A4: Panel regressions for emerging market returns. Results are shown for the base specification of the model, which excludes the text-based measures, and two text specifications, one that includes context specific sentiment (in column "Sent") and another that includes context specific sentiment interacted with entropy (in column "SentEnt"). Results are reported for the entire sample, as well as for the early and late subsamples. Observations are monthly. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include: realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of 5-year ago to current index levle (*value*), year-over-year GDP growth (*gdp*), year-over-year inflration (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and post-election dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for *Topic* in markets, government, corporate sector, commodities, credit (DM) and macro (EM). All panels include country fixed effects and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***","**", and "*" are significant at the 1%, 5% and 10% levels respectively.

Emerging markets: Forecasting panel for returns

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	-0.034	-0.025	-0.026	0.017	0.058	0.057	-0.074	-0.074	-0.076
$sigma_{t-2}$	0.039	0.046	0.045	-0.020	0.001	0.001	0.111**	0.111**	0.110**
$return_{t-1}$	0.114***	0.096**	0.099**	0.073*	0.043	0.046	0.123*	0.101	0.104
$return_{t-2}$	-0.006	-0.008	-0.007	-0.051	-0.060	-0.058	0.037	0.031	0.033
$value_{t-1}$	0.227	0.242	0.261	0.471	1.145***	1.167***	0.165	-0.106	-0.116
gdp_{t-1}	-0.091	-0.106*	-0.107*	-0.037	-0.180*	-0.178*	-0.160*	-0.132*	-0.133*
$gdpdeflator_{t-1}$	-0.036	-0.048	-0.048	-0.012	-0.117**	-0.115**	-0.117**	-0.112**	-0.115**
cp_{t-1}	-0.035**	-0.024*	-0.024*	-0.022	-0.011	-0.010	-0.021	-0.028	-0.028
dcp_{t-1}	-0.009	-0.004	-0.004	0.024	0.005	0.009	-0.029	-0.022	-0.027
$rate_{t-1}$	-0.008	-0.036	-0.034	-0.035	0.086	0.086	-0.221	-0.312**	-0.305*
$dexch_{t-1}$	0.068	0.071	0.069	0.110	0.124	0.122	0.044	0.030	0.027
pre	-0.063	-0.062	-0.066	-0.493	-0.866*	-0.891*	0.077	0.097	0.110
post	-0.557	-0.542	-0.562	-0.732	-0.978*	-1.012*	-0.474	-0.359	-0.393
$entropy_{t-1}$		2.691	2.800		-10.044**	-9.559**		7.508**	7.373**
$artcount_{t-1}$		-0.229	-0.232		-1.197*	-1.208*		0.007	0.039
$sMkt_{t-1}$		0.396	0.190		0.141	-0.217		0.186	0.039
$fMkt_{t-1}$		-0.469	-0.522		-1.562***	-1.661***		0.511	0.429
$sGovt_{t-1}$		-0.188	-0.162		-0.314	-0.365		0.005	0.001
$fGovt_{t-1}$		-0.918*	-0.872*		-1.328*	-1.304*		0.129	0.075
$sCorp_{t-1}$		-0.809*	-0.675*		0.089	0.128		-1.014**	-0.845*
$fCorp_{t-1}$		-1.031***	-0.977***		-0.766	-0.707		-0.452	-0.418
$sComms_{t-1}$		0.840***	0.813***		0.397	0.535		1.279***	1.156***
$fComms_{t-1}$		0.287	0.267		-0.079	0.033		0.833**	0.716**
$sMacro_{t-1}$		0.132	0.197		0.524	0.710		-0.704	-0.610
$fMacro_{t-1}$		0.409*	0.448*		0.247	0.343		0.565*	0.623**
$R2$	0.0144	0.0242	0.0236	-0.00486	0.0179	0.0188	0.0404	0.062	0.0597
start	Apr 1998	May 1998	May 1998	Apr 1998	May 1998	May 1998	Mar 2007	Mar 2007	Mar 2007
end	Dec 2015	Dec 2015	Dec 2015	Feb 2007	Feb 2007	Feb 2007	Dec 2015	Dec 2015	Dec 2015
Nobs	4853	4839	4839	2100	2086	2086	2753	2753	2753
stderr	by time	by time	by time	by time	by time	by time	by time	by time	by time

Table A5: We compute correlations between the CM topic-sentiment measure and the TRNA topic-sentiment measure for each country in our EM and DM data sets. This table reports the mean, minimum, maximum and interquartile range for each topic across all countries in the sample. For example, the Mean (Min) *Mkt* correlation is the average (minimum) across countries of the CM-TRNA sentiment correlation for the markets topic.

Region	Stat	Mkt	Govt	Corp	Comms	Macro	Credit
EM	Mean	0.378	0.324	0.410	0.365	0.243	
	Min	0.226	0.129	0.294	0.149	-0.079	
	Max	0.541	0.493	0.602	0.566	0.450	
	Qtr Range	0.096	0.141	0.087	0.118	0.135	
DM	Mean	0.368	0.310	0.390	0.298		0.422
	Min	0.228	0.163	0.237	-0.076		0.258
	Max	0.560	0.585	0.706	0.466		0.663
	Qtr Range	0.080	0.127	0.107	0.160		0.118

Table A6: Panel regressions for developed and emerging market samples for *returns*. Results are shown for the base specification of the model, which excludes the text-based measures, and two text specifications, one that includes context specific sentiment (in column "Sent") and another that includes context specific sentiment from the TRNA algorithm (in column "TRNA"). Results are reported starting in 2003, which is when the TRNA data become available. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include: realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of 5-year ago to current index levle (*value*), year-over-year GDP growth (*gdp*), year-over-year inflration (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and post-election dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for *Topic* in markets, government, corporate sector, commodities, credit (DM) and macro (EM). All panels include country fixed effects and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***","**", and "*" are significant at the 1%, 5% and 10% levels respectively.

Forecasting panel for DM returns

	Base	Sent	TRNA
$sigma_{t-1}$	-0.088	-0.065	-0.082
$sigma_{t-2}$	0.074	0.095	0.086
$return_{t-1}$	0.120	0.083	0.096
$return_{t-2}$	-0.005	-0.021	-0.014
$value_{t-1}$	1.792**	2.254***	2.253***
gdp_{t-1}	-0.058	-0.115	-0.107
$gdpdeflator_{t-1}$	-0.018	-0.045	-0.077
cp_{t-1}	-0.047***	-0.043***	-0.043***
dcp_{t-1}	0.054***	0.041**	0.044**
$rate_{t-1}$	-0.278	-0.561***	-0.574***
$dexch_{t-1}$	0.040	0.027	0.009
pre	0.352	0.106	0.063
post	0.276	0.149	0.027
$entropy_{t-1}$		4.953**	3.464
$artcount_{t-1}$		0.371	0.313
$sMkt_{t-1}$		1.374***	0.851
$fMkt_{t-1}$		0.547	0.307
$sGovt_{t-1}$		-1.273***	0.326
$fGovt_{t-1}$		-1.047**	-0.202
$sCorp_{t-1}$		0.299	-0.132
$fCorp_{t-1}$		-0.342	-0.221
$sComms_{t-1}$		0.140	-0.181
$fComms_{t-1}$		0.378*	0.315*
$sCredit_{t-1}$		-0.348	-0.299
$fCredit_{t-1}$		-0.214	-0.205
R2	0.0539	0.08	0.0675
start	Jan 2003	Jan 2003	Feb 2003
end	Dec 2015	Dec 2015	Dec 2015
Nobs	3469	3469	3450
stderr	by time	by time	by time

Forecasting panel for EM returns

	Base	Sent	TRNA
$sigma_{t-1}$	-0.072	-0.065	-0.066
$sigma_{t-2}$	0.077*	0.081*	0.082*
$return_{t-1}$	0.126**	0.102*	0.107*
$return_{t-2}$	0.048	0.040	0.038
$value_{t-1}$	0.247	-0.046	0.123
gdp_{t-1}	-0.089	-0.078	-0.109
$gdpdeflator_{t-1}$	-0.051	-0.059	-0.073*
cp_{t-1}	-0.054***	-0.035**	-0.032*
dcp_{t-1}	-0.012	-0.017	-0.020
$rate_{t-1}$	0.034	-0.022	-0.009
$dexch_{t-1}$	0.102	0.095	0.092
pre	0.082	0.062	-0.006
post	-0.398	-0.437	-0.456
$entropy_{t-1}$		4.207*	4.465**
$artcount_{t-1}$		-0.172	-0.160
$sMkt_{t-1}$		0.428	0.461
$fMkt_{t-1}$		0.182	-0.063
$sGovt_{t-1}$		-0.415	-0.062
$fGovt_{t-1}$		-0.253	-0.156
$sCorp_{t-1}$		-0.954**	-0.629
$fCorp_{t-1}$		-0.653**	-0.555**
$sComms_{t-1}$		0.911***	0.689**
$fComms_{t-1}$		0.478	-0.068
$sMacro_{t-1}$		-0.384	-0.150
$fMacro_{t-1}$		0.695***	0.783***
R2	0.0318	0.0509	0.0469
start	Jan 2003	Jan 2003	Feb 2003
end	Dec 2015	Dec 2015	Dec 2015
Nobs	3880	3880	3857
stderr	by time	by time	by time

Table A7: Panel regressions for developed and emerging market samples for *12 month returns*. Results are shown for the base specification of the model, which excludes the text-based measures, and two text specifications, one that includes context specific sentiment (in column "Sent") and another that includes context specific sentiment from the TRNA algorithm (in column "TRNA"). Results are reported starting in 2003, which is when the TRNA data become available. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include: realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of 5-year ago to current index levle (*value*), year-over-year GDP growth (*gdp*), year-over-year inflration (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and post-election dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for *Topic* in markets, government, corporate sector, commodities, credit (DM) and macro (EM). All panels include country fixed effects and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***","**", and "*" are significant at the 1%, 5% and 10% levels respectively.

Forecasting panel for DM 12 month returns

	Base	Sent	TRNA
$sigma_{t-1}$	0.156	0.261	0.206
$sigma_{t-2}$	0.089	0.120	0.060
$return_{t-1}$	0.261	0.080	0.241
$return_{t-2}$	-0.108	-0.254	-0.240
$value_{t-1}$	22.710***	24.168***	23.427***
gdp_{t-1}	-0.888	-1.195**	-1.106*
$gdpdeflator_{t-1}$	1.139*	0.838	0.602
cp_{t-1}	-0.622***	-0.479***	-0.506***
dcp_{t-1}	0.602***	0.417***	0.476***
$rate_{t-1}$	-2.936*	-4.591***	-4.654***
$dexch_{t-1}$	0.605	0.658	0.801
pre	1.171	-0.487	-0.641
post	0.625	-0.245	-1.222
$entropy_{t-1}$		35.451***	36.742***
$artcount_{t-1}$		-1.317	-1.231
$sMkt_{t-1}$		6.092**	7.137***
$fMkt_{t-1}$		4.088	3.085
$sGovt_{t-1}$		-6.265***	2.399**
$fGovt_{t-1}$		-2.032	2.333
$sCorp_{t-1}$		-6.247***	0.438
$fCorp_{t-1}$		-2.348	-0.555
$sComms_{t-1}$		2.207	-2.103**
$fComms_{t-1}$		3.878*	2.901*
$sCredit_{t-1}$		2.686	-7.632***
$fCredit_{t-1}$		-1.105	-0.919
R2	0.258	0.344	0.326
start	Jan 2003	Jan 2003	Feb 2003
end	Dec 2015	Dec 2015	Dec 2015
Nobs	3458	3458	3439
stderr	both	both	both

Forecasting panel for EM 12 month returns

	Base	Sent	TRNA
$sigma_{t-1}$	0.269*	0.341**	0.357**
$sigma_{t-2}$	0.432**	0.482***	0.500***
$return_{t-1}$	0.612***	0.435**	0.445**
$return_{t-2}$	0.243	0.179	0.154
$value_{t-1}$	9.129	5.694	6.990
gdp_{t-1}	-0.618	-0.488	-0.728
$gdpdeflator_{t-1}$	0.462	0.359	0.271
cp_{t-1}	-0.739***	-0.504**	-0.449**
dcp_{t-1}	0.202	0.127	0.059
$rate_{t-1}$	0.952	0.747	0.619
$dexch_{t-1}$	1.221**	1.133**	1.127**
pre	0.059	-0.128	-0.249
post	-5.344	-5.366*	-5.553*
$entropy_{t-1}$		19.046*	24.400***
$artcount_{t-1}$		-5.381***	-5.711***
$sMkt_{t-1}$		4.048	2.692
$fMkt_{t-1}$		2.535	1.181
$sGovt_{t-1}$		-6.349***	0.287
$fGovt_{t-1}$		-4.137	0.057
$sCorp_{t-1}$		-8.222**	-2.661
$fCorp_{t-1}$		-5.105**	-4.211*
$sComms_{t-1}$		1.815	1.184
$fComms_{t-1}$		1.313	0.269
$sMacro_{t-1}$		1.286	-0.542
$fMacro_{t-1}$		9.258***	9.111***
R2	0.122	0.213	0.186
start	Jan 2003	Jan 2003	Feb 2003
end	Dec 2015	Dec 2015	Dec 2015
Nobs	3880	3880	3857
stderr	both	both	both

Table A8: Panel regressions for developed and emerging market samples for *volatility*. Results are shown for the base specification of the model, which excludes the text-based measures, and two text specifications, one that includes context specific sentiment (in column "Sent") and another that includes context specific sentiment from the TRNA algorithm (in column "TRNA"). Results are reported starting in 2003, which is when the TRNA data become available. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include: realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of 5-year ago to current index levle (*value*), year-over-year GDP growth (*gdp*), year-over-year inflration (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and post-election dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for *Topic* in markets, government, corporate sector, commodities, credit (DM) and macro (EM). All panels include country fixed effects and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***","**", and "*" are significant at the 1%, 5% and 10% levels respectively.

Forecasting panel for DM volatility

	Base	Sent	TRNA
$sigma_{t-1}$	0.417***	0.390***	0.391***
$sigma_{t-2}$	0.103	0.081	0.086
$retmi_{t-1}$	0.647***	0.617***	0.640***
$retmi_{t-2}$	0.062	0.059	0.059
$value_{t-1}$	-2.049**	-2.350**	-2.391**
gdp_{t-1}	-0.204*	-0.142	-0.128
$gdpdeflator_{t-1}$	0.085	0.103	0.116
cp_{t-1}	0.048***	0.046***	0.050***
dcp_{t-1}	-0.071***	-0.062***	-0.063***
$rate_{t-1}$	0.621**	0.737**	0.766**
$dexch_{t-1}$	-0.122	-0.129	-0.123
pre	-0.004	-0.064	0.145
post	0.104	0.112	0.205
$entropy_{t-1}$		-2.725	-1.896
$artcount_{t-1}$		-0.212	-0.192
$sMkt_{t-1}$		-0.399	-0.956
$fMkt_{t-1}$		-0.351	-0.297
$sGovt_{t-1}$		0.883	-0.465*
$fGovt_{t-1}$		0.328	-0.193
$sCorp_{t-1}$		-0.235	1.249*
$fCorp_{t-1}$		-0.143	-0.325
$sComms_{t-1}$		0.022	0.217
$fComms_{t-1}$		-0.603**	-0.667***
$sCredit_{t-1}$		-0.782	-1.009*
$fCredit_{t-1}$		-0.363	-0.177
R2	0.492	0.497	0.498
start	Jan 2003	Jan 2003	Feb 2003
end	Dec 2015	Dec 2015	Dec 2015
Nobs	3469	3469	3450
stderr	by time	by time	by time

Forecasting panel for EM volatility

	Base	Sent	TRNA
$sigma_{t-1}$	0.414***	0.409***	0.405***
$sigma_{t-2}$	0.051	0.043	0.041
$retmi_{t-1}$	0.360***	0.346***	0.346***
$retmi_{t-2}$	0.088	0.090	0.076
$value_{t-1}$	-1.739***	-1.479***	-1.438***
gdp_{t-1}	-0.056	-0.058	-0.042
$gdpdeflator_{t-1}$	0.045	0.060	0.069
cp_{t-1}	-0.025	-0.049**	-0.051**
dcp_{t-1}	0.058**	0.064**	0.074***
$rate_{t-1}$	0.315***	0.371***	0.372***
$dexch_{t-1}$	-0.056	-0.047	-0.056
pre	0.542	0.592	0.626
post	-0.043	-0.059	-0.084
$entropy_{t-1}$		-3.698	-4.150
$artcount_{t-1}$		0.780**	0.720**
$sMkt_{t-1}$		0.407	0.270
$fMkt_{t-1}$		0.132	0.181
$sGovt_{t-1}$		0.544	0.874***
$fGovt_{t-1}$		0.566	0.490
$sCorp_{t-1}$		-0.052	0.070
$fCorp_{t-1}$		0.246	0.408
$sComms_{t-1}$		-0.532	-1.072***
$fComms_{t-1}$		0.235	0.558**
$sMacro_{t-1}$		0.292	-0.146
$fMacro_{t-1}$		-0.324	-0.417*
R2	0.347	0.354	0.356
start	Jan 2003	Jan 2003	Feb 2003
end	Dec 2015	Dec 2015	Dec 2015
Nobs	3873	3873	3850
stderr	by time	by time	by time

Table A9: Panel regressions for developed and emerging market samples for *drawdowns*. Results are shown for the base specification of the model, which excludes the text-based measures, and two text specifications, one that includes context specific sentiment (in column "Sent") and another that includes context specific sentiment from the TRNA algorithm (in column "TRNA"). Results are reported starting in 2003, which is when the TRNA data become available. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include: realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of 5-year ago to current index levle (*value*), year-over-year GDP growth (*gdp*), year-over-year inflration (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and post-election dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for *Topic* in markets, government, corporate sector, commodities, credit (DM) and macro (EM). All panels include country fixed effects and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***","**", and "*" are significant at the 1%, 5% and 10% levels respectively.

Forecasting panel for DM drawdowns

	Base	Sent	TRNA
$sigma_{t-1}$	0.088	-0.020	0.044
$sigma_{t-2}$	0.155	0.080	0.120
$return_{t-1}$	-0.497***	-0.339**	-0.439***
$return_{t-2}$	-0.075	0.005	-0.017
$value_{t-1}$	-13.575***	-15.192***	-14.713***
gdp_{t-1}	0.242	0.455	0.455
$gdpdeflator_{t-1}$	-0.197	-0.155	0.011
cp_{t-1}	0.296***	0.237***	0.257***
dcp_{t-1}	-0.304***	-0.204***	-0.238***
$rate_{t-1}$	2.938***	3.807***	3.796***
$dexch_{t-1}$	-0.564*	-0.541*	-0.556*
pre	0.002	0.583	0.917
$post$	-0.372	-0.332	0.215
$entropy_{t-1}$		-18.007***	-14.376**
$artcount_{t-1}$		0.226	0.279
$sMkt_{t-1}$		-6.353***	-4.777***
$fMkt_{t-1}$		-1.869	-0.633
$sGovt_{t-1}$		5.656***	-0.843
$fGovt_{t-1}$		3.990***	0.263
$sCorp_{t-1}$		2.270**	1.061
$fCorp_{t-1}$		1.747	0.725
$sComms_{t-1}$		-1.013	1.569***
$fComms_{t-1}$		-1.106	-0.621
$sCredit_{t-1}$		-0.526	1.605
$fCredit_{t-1}$		0.949	0.986
$R2$	0.31	0.391	0.346
$start$	Jan 2003	Jan 2003	Feb 2003
end	Dec 2015	Dec 2015	Dec 2015
$Nobs$	3469	3469	3450
$stderr$	both	both	both

Forecasting panel for EM drawdowns

	Base	Sent	TRNA
$sigma_{t-1}$	0.079	0.037	0.035
$sigma_{t-2}$	-0.046	-0.084	-0.091
$return_{t-1}$	-0.338***	-0.256***	-0.279***
$return_{t-2}$	-0.203**	-0.174**	-0.177**
$value_{t-1}$	-5.962**	-4.333**	-5.102**
gdp_{t-1}	0.201	0.170	0.250
$gdpdeflator_{t-1}$	0.082	0.112	0.151
cp_{t-1}	0.216**	0.105	0.091
dcp_{t-1}	0.060	0.072	0.098
$rate_{t-1}$	0.497*	0.606***	0.626**
$dexch_{t-1}$	-0.478**	-0.452**	-0.440**
pre	-0.917	-0.893	-0.738
$post$	0.576	0.217	0.334
$entropy_{t-1}$		-15.635**	-16.571**
$artcount_{t-1}$		3.255***	3.442***
$sMkt_{t-1}$		-1.652	-0.618
$fMkt_{t-1}$		-2.351*	-1.651
$sGovt_{t-1}$		3.873***	0.564
$fGovt_{t-1}$		2.239	0.011
$sCorp_{t-1}$		3.359***	0.299
$fCorp_{t-1}$		1.525*	1.360
$sComms_{t-1}$		-2.057**	-1.058**
$fComms_{t-1}$		-0.563	0.689
$sMacro_{t-1}$		-0.440	0.285
$fMacro_{t-1}$		-3.228***	-3.173***
$R2$	0.0996	0.205	0.177
$start$	Jan 2003	Jan 2003	Feb 2003
end	Dec 2015	Dec 2015	Dec 2015
$Nobs$	3880	3880	3857
$stderr$	both	both	both