How News and Its Context Drive Risk and Returns Around the World

Charles Calomiris and Harry Mamaysky
Columbia Business School

Cleveland Fed/OFR/Maryland 2017
Financial Stability and FinTech Conference

Outline of talk

- Introduction
- Text measures defined
- Empirical observations
- Empirical results
- Out-of-sample testing
- Conclusion

* Introduction

- Automated processing of natural language allows:
  - ✓Practitioners to monitor information in real time
  - ✓Researchers to systematically study how markets react to information
- Prior work has (mostly) focused on short-term responses of individual US stocks to news
  - ✓General finding of short-term underreaction to news
- Do news forecast longer horizon, country level outcomes?
  - ✓We construct news measures that seem to do so
  - ✓We think we're the first to focus on this question

How to measure country level news?

- News measures:
  - Sentiment
  - Frequency
  - Unusualness
  - News topics
- We are particularly interested in *news topics*:
  - What are country level news topics?
  - Does importance of news differ by topic?
  - Do topics differ from EM to DM?
  - How do topics evolve over time?
- In this paper, we scratch the surface…

Do news forecast country level outcomes?

- We look at four variables at the country level
  - Next month's return -- *return*
  - Next 12-month returns -- *return*^{12}
  - Next month's realized volatility -- *sigma*
  - Next 12-month maximum drawdown -- *drawdown*
- Look at stock market indexes
  - Returns measured in US$ terms – from perspective of a US-based investor
- Analyze responses separately for
  - emerging market (EM), and
  - developed market (DM) countries

Some answers, many questions…

Answers

- Short term country-level responses appear different than single name responses
- EM is different from DM
- Topics matter
- Evidence of out-of-sample predictive ability
- Evidence of regime shifts in coefficients
- Our approach compares favorably to *a priori* approaches, like Baker, Bloom and Davis' *economic policy uncertainty*

Questions

- Why do effects of sentiment depend on topic?
- Do stock index returns cause news?
- Why are EMs different from DMs?
- How much of predictability is from currency effect?
- Is the effect causal? Is it underreaction? Limited attention? Will it change?

Outline of talk

- Introduction
- Text measures defined
- Empirical observations
- Empirical results
- Out-of-sample testing
- Conclusion

* Text measures defined

Data:

- Thomson-Reuters digital news archive from 1996—2015
- 5mm EM and 12mm DM articles
- 52 countries, 28 DM and 24 EM (country list)

Text measures:

- *artcount* – number of articles per country per month
- *entropy* – "unusualness" of an article *j* (Glasserman and Mamaysky 2016)

$$H_j = -\sum_{i \in \{4-\text{grams}\}} p_i \log m_i$$

  - ❑ E.g. "Order imbalance of 10,000 shares" vs "Order imbalance of 2,000,000 shares"
- *sentiment* – the difference of positive and negative words divided by total words in article *j*:

$$s_j = \frac{POS_j - NEG_j}{a_j}$$

  - ❑ Word sentiment comes from Loughran – McDonald dictionary (examples)

Topic detection

Intuition

Find groups of words that tend to co-occur together in articles

Details

- 1242 *econ words*
  - Start w/ 237 words from index of Beim and Calomiris (2001) and find other words, bigrams and trigrams from EM corpus based on cosine similarity
  - *eg. barriers*, *currency*, *parliament*, *macroeconomist*, *IMF,* and *soybean*
- We have 2 document term matrixes:
  - EM: *5mm* articles *x 1,240* words
  - DM: *12mm* articles *x 1,242* words
- Compute cosine similarity matrixes (*1,242 x 1,242*)
  - Each word/phrase represented by *5mm* or *12mm* length vector
- Then do community detection (via modularity maximization)
- Our topics are mutually exclusive (LDA gives similar answers)

5 Topics for EM

Wordcloud

Mkt # words (mm) = 113.80

Govt # words (mm) = 76.58

Corp # words (mm) = 58.77

Comms # words (mm) = 58.90

Macro # words (mm) = 15.98

[Sample articles]

[5 topics]

[Topic similarity]

5 Topics for DM

Wordcloud

Mkt # words (mm) = 236.23

Govt # words (mm) = 111.02

Corp # words (mm) = 131.95

Comms # words (mm) = 26.86

Credit # words (mm) = 139.63

[Sample articles]

[5 topics]

[Topic similarity]

The final country level measures

- Let $f_{\tau,j}$ be the fraction of econ words in article *j* that are about topic τ
- Article sentiment is s_j
- Topic sentiment is

$$s_{\tau,j} \equiv f_{\tau,j} \times s_j$$

- Aggregate the article level measures into daily measures (weighted by number of overall words)

For a given country, we have 12 *daily* text measures:

- *entropy*
- *article count*
- *sMkt / fMkt*
- *sGovt / fGovt*
- *sCorp / fCorp*
- *sComms / fComms*
- *sMacro / fMacro* (EM)
- *sCredit / fCredit* (DM)

EM or DM specific

Outline of talk

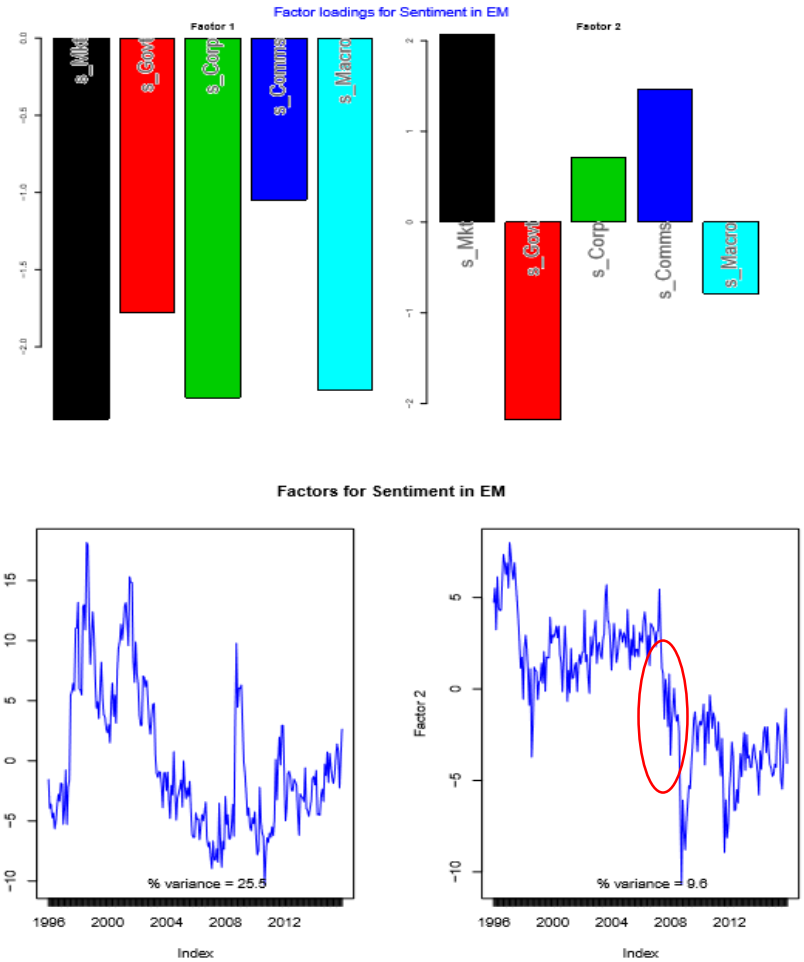
- Introduction
- Text measures defined
- Empirical observations
- Empirical results
- Out-of-sample testing
- Conclusion

* Empirical observations

EM Sentiment

- For 140 EM sentiment series (28 countries x 5 topics) we look at first 2 principal components
- PC2 – relative sentiment of Markets to Government
- Some evidence of a regime shift in PC2 around the financial crisis

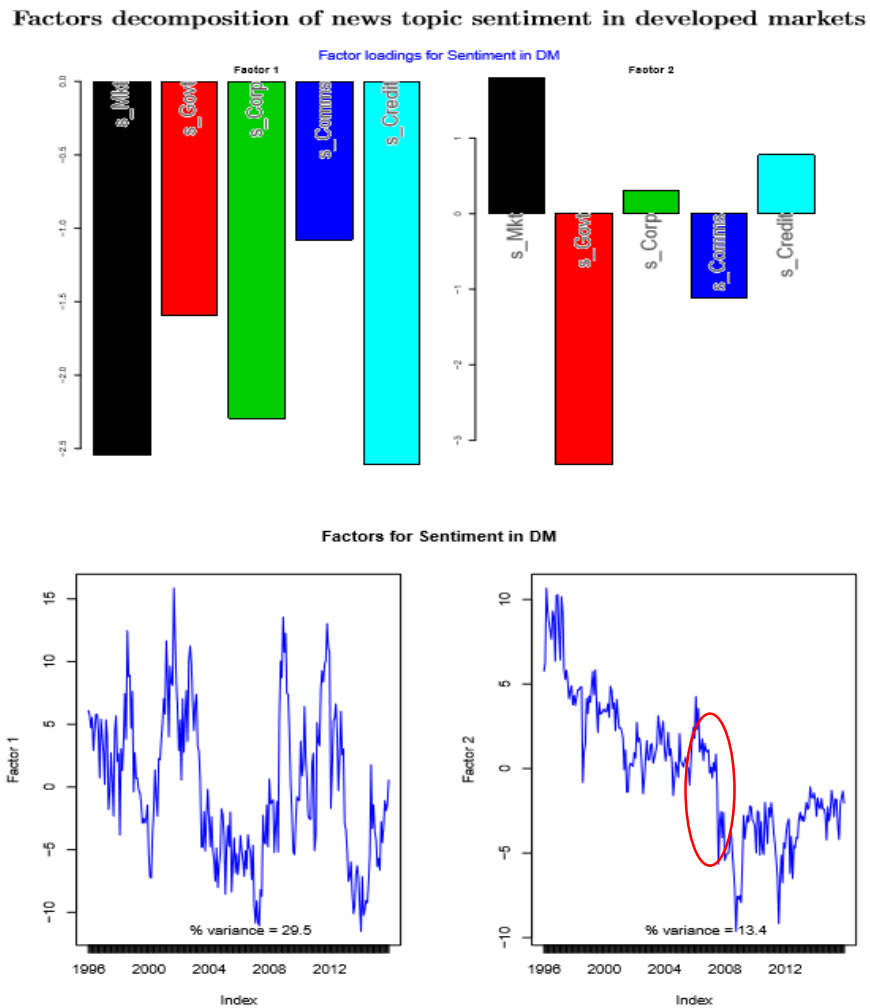
Factors decomposition of news topic sentiment in emerging markets



More principal components

DM Sentiment

- For 120 DM sentiment series (24 countries x 5 topics) we look at first 2 PCs
- PC2 – relative sentiment of Markets to Government (again!)
- Some evidence of a regime shift in PC2 a little before the financial crisis

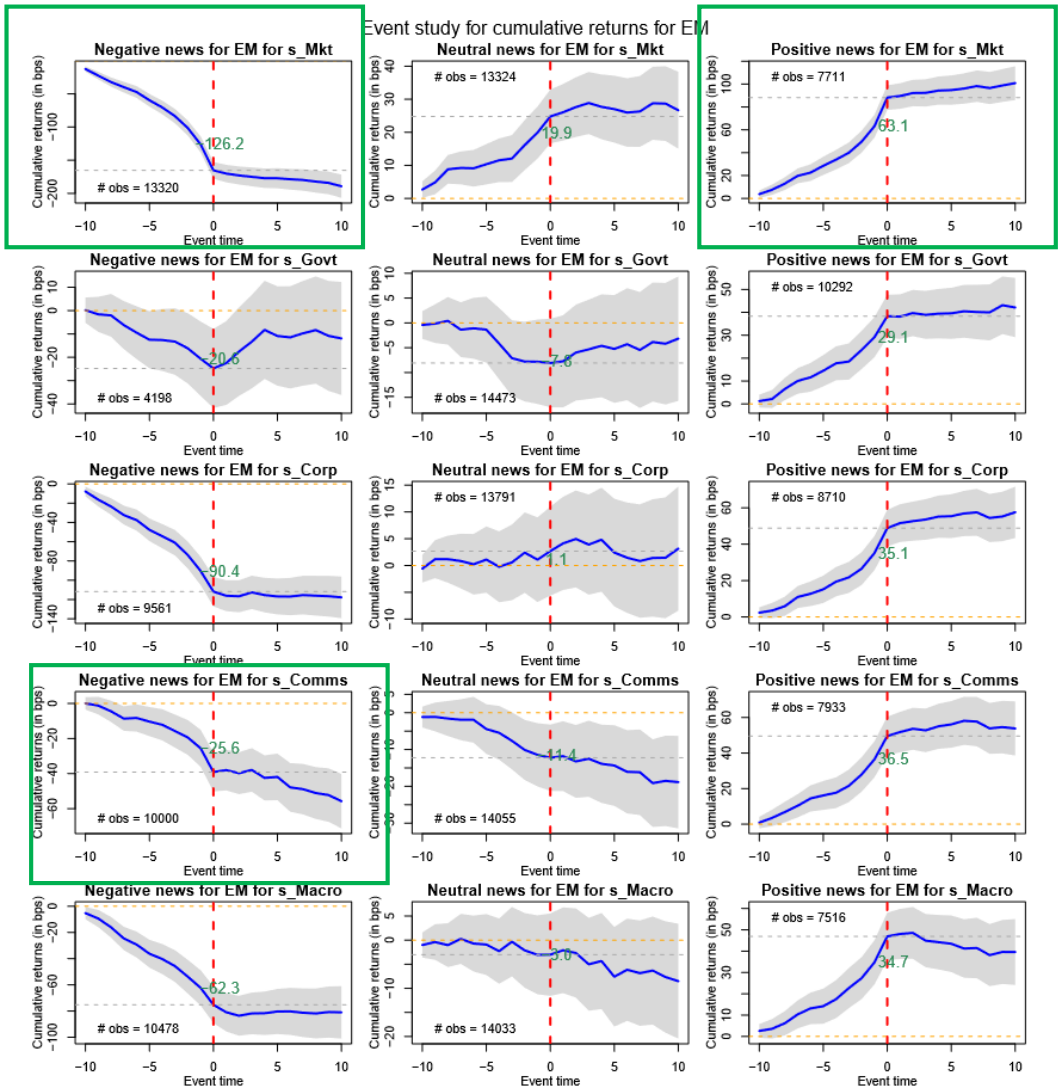


Event studies – EM

- Cumulative abnormal returns in 10 day windows around news events
- Bottom, middle and top decile of news sentiment
- Lots of caveats

Observations

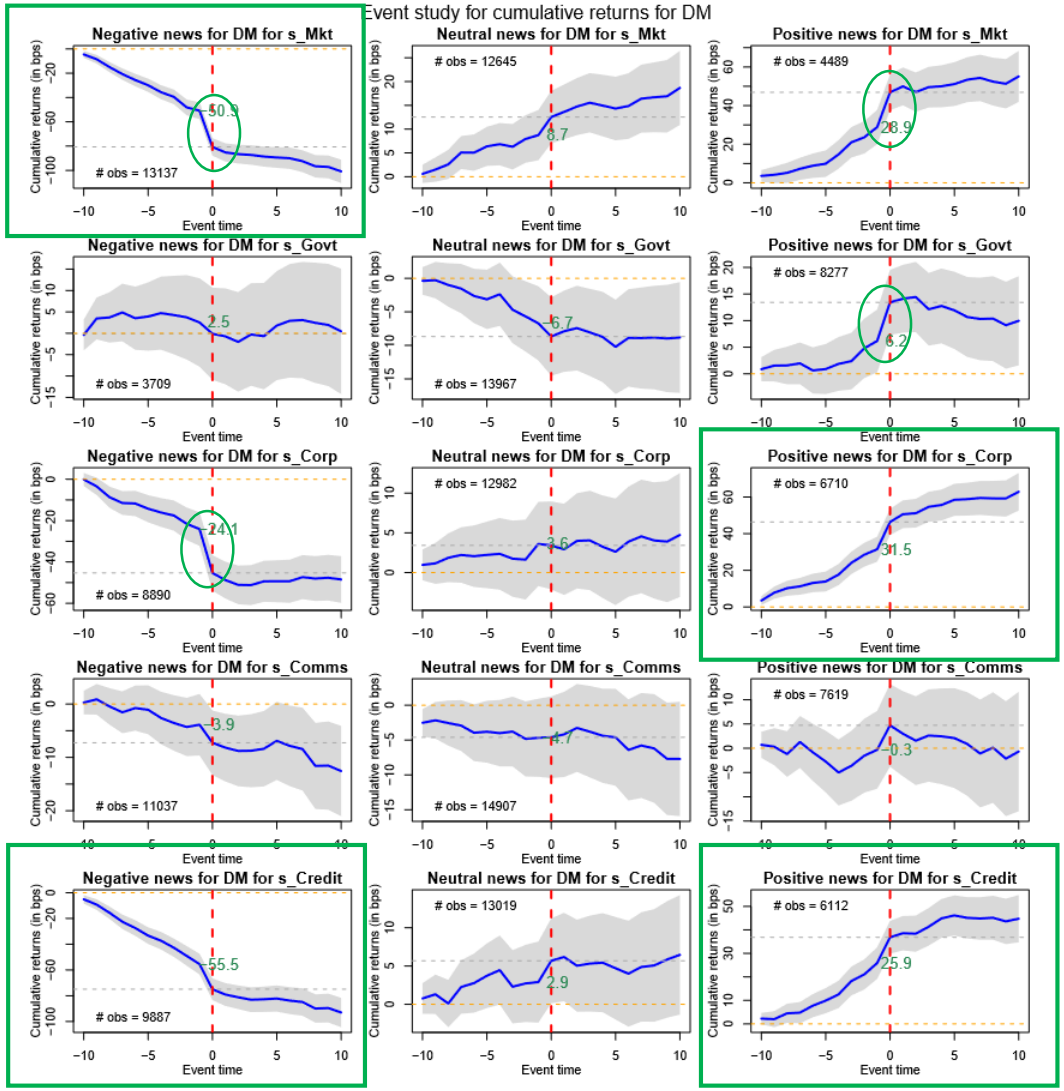
- Stock index prices drift into news (news not exogenous)
- No drift around neutral news
- Topics with post event drift:
  - Mkt (both)
  - Comms (negative)
- *Different from single name results, where there is little evidence of drift post negative news (e.g. Tetlock et al., Henderschott et al.)*
- The mechanism must differ



Event studies – DM

Observations

- News is more of a *surprise* in DM's
  - Bigger event-day jumps
- Some topics show post event drift:
  - Mkt (negative, both?)
  - Corp (positive)
  - Credit (both)



Outline of talk

- Introduction
- Text measures defined
- Empirical observations
- Empirical results
- Out-of-sample testing
- Conclusion

* Empirical results

- For EM and DM samples, run panel regressions – with country fixed effects – to forecast *t+1* observations of:
  - ✓ *return*
  - ✓ *return*^{12}
  - ✓ *sigma*
  - ✓ *drawdown*
- Control for many variables that have been shown to have forecasting power for future returns
  - The no-text measure regression is our *Baseline* model
- All text measures (except *entropy*) are normalized to unit variance
- Run regressions in the 1st and 2nd half of the sample, as well as for the overall sample

EM drawdown residuals DM drawdown residuals

Summary of results

- News matters for EM and DM!
- But results differ across EM and DM
  - Baseline R^2 lower for EM
  - % increase in R^2 from text measures larger for EM
- Sign of effects (i.e. good news or bad) is consistent across *return*, *return*^{12}, *sigma*, and *drawdown*
- Context matters:
  - positive sentiment in *Govt*, *Corp* –> bad news
  - positive sentiment in *Mkt* –> good news
- Incremental explanatory power largest for *return*^{12} and *drawdown*; explanatory power lower for *return* and *sigma*
- Evidence of state dependence, especially for *entropy*
  - Goes from a "bad" pre-crisis to a "good" post-crisis

Results detail

Outline of talk

- Introduction
- Text measures defined
- Empirical observations
- Empirical results
- Out-of-sample testing
- Conclusion

* Out-of-sample testing

- Do we have too many explanatory variables?
- What about regime shifts?
- Check out-of-sample forecasting performance
- Run rolling 5-year lasso regressions in *t-24,…,t-1* for forecasting month *t* outcomes
- Evidence of out-of-sample forecasting ability for $return^{12}$ and *drawdown*, for both EM and DM
- Effects are economically large for investors

Outline of talk

- Introduction
- Text measures defined
- Empirical observations
- Empirical results
- Out-of-sample testing
- Conclusion

* Conclusion

Useful information in text for medium-term country-level outcomes!

- Different dimensions of text matter
  - In particular, context
- Effects differ across EM and DM
- Effects differ over time
- Evidence of out-of-sample forecasting ability
- Next:
  - Currency effects?
  - Trading strategies?

Appendix

Emerging markets data summary

	mean	sd	5%	95%	AR(1)	N
return	1.044	9.316	-13.497	15.299	0.137	5549
$return^{12}$	16.796	42.392	-41.367	89.574		5549
$return^{24}$	35.742	73.016	-44.931	170.380		5297
sigma	21.482	14.282	7.085	47.108	0.538	5476
$drawdown^{12}$	17.398	17.373	-0.000	55.721		5583
$drawdown^{24}$	23.012	20.725	-0.000	64.639		5328
retmi	2.848	5.359	0.000	13.497	0.177	5549
retpl	3.891	5.993	0.000	15.299	0.122	5549
value	0.953	0.925	0.187	2.431	0.949	4917
gdp	4.157	4.836	-3.000	9.950	0.933	5964
gdpdeflator	7.133	8.984	-1.100	23.560	0.932	5964
cp	55.657	44.713	10.688	145.941	0.988	5964
dcp	1.153	7.918	-9.113	12.065	0.933	5964
rate	7.451	7.384	0.580	19.985	0.977	5964
dexch	0.365	3.695	-4.178	5.546	0.097	5955
pre	0.162	0.369	0.000	1.000	0.810	5964
post	0.157	0.363	0.000	1.000	0.802	5964
entropy	2.429	0.171	2.111	2.649	0.790	5964
artcount	26.022	30.146	3.200	73.349	0.763	5964
sMkt	-0.005	0.002	-0.009	-0.002	0.693	5964
fMkt	0.334	0.080	0.195	0.460	0.637	5964
sGovt	-0.007	0.004	-0.016	-0.002	0.565	5964
fGovt	0.290	0.096	0.163	0.480	0.596	5964
sCorp	-0.002	0.001	-0.004	-0.001	0.635	5964
fCorp	0.169	0.034	0.124	0.229	0.587	5964
sComms	-0.002	0.001	-0.004	-0.001	0.482	5964
fComms	0.159	0.064	0.082	0.289	0.696	5964
sMacro	-0.001	0.000	-0.002	-0.000	0.636	5964
fMacro	0.048	0.028	0.025	0.105	0.620	5964

Developed markets data summary

	mean	sd	5%	95%	AR(1)	N
return	0.646	6.671	-10.939	10.783	0.112	4803
$return^{12}$	8.983	27.341	-38.063	51.641		4792
$return^{24}$	18.783	42.977	-42.921	93.108		4578
sigma	19.196	11.427	7.470	41.669	0.619	4815
$drawdown^{12}$	15.305	15.465	-0.000	50.593		4817
$drawdown^{24}$	20.951	19.636	-0.000	61.803		4616
retmi	2.188	4.096	0.000	10.939	0.205	4803
retpl	2.835	3.914	0.000	10.783	0.050	4803
value	0.936	0.624	0.317	1.836	0.982	4432
gdp	2.083	3.450	-3.490	6.680	0.933	4898
gdpdeflator	1.917	2.762	-1.473	5.650	0.925	4898
cp	113.275	41.235	55.687	188.754	0.990	4898
dcp	2.503	11.916	-9.540	13.744	0.926	4898
rate	3.947	2.025	1.030	6.221	0.986	4899
dexch	0.024	7.823	-4.663	4.855	0.016	4686
pre	0.126	0.333	0.000	1.000	0.813	4899
post	0.122	0.327	0.000	1.000	0.806	4899
entropy	2.455	0.170	2.125	2.666	0.841	4899
artcount	106.700	214.462	12.787	316.793	0.547	4899
sMkt	-0.004	0.002	-0.007	-0.002	0.745	4899
fMkt	0.350	0.055	0.271	0.450	0.715	4899
sGovt	-0.004	0.002	-0.008	-0.001	0.619	4899
fGovt	0.236	0.055	0.150	0.328	0.613	4899
sCorp	-0.002	0.001	-0.003	-0.001	0.636	4899
fCorp	0.220	0.050	0.159	0.314	0.750	4899
sComms	-0.000	0.000	-0.001	-0.000	0.464	4899
fComms	0.027	0.013	0.012	0.052	0.554	4899
sCredit	-0.002	0.001	-0.004	-0.001	0.682	4899
fCredit	0.167	0.023	0.137	0.208	0.576	4899

Country list

List of EM and DM countries

	EM Countries				DM Countries		
	Country	BBG Index	TR Code		Country	BBG Index	TR Code
1	Argentina	BURCAP	AR	1	Australia	AS52	AU
2	Brazil	IBOV	BR	2	Austria	ATX	AT
3	Chile	IGPA	CL	3	Belgium	BELPRC	BE
4	China (PRC)	SHCOMP	CN	4	Canada	SPTSX	CA
5	Colombia	COLCAP	CO	5	Denmark	KAX	DK
6	Czech Republic	PX	CZ	6	DM	MXWO	–
7	EM	MXEF	EMRG	7	Finland	HEX	FI
8	Estonia	TALSE	EE	8	France	CAC	FR
9	Ghana	GGSECI	GH	9	Germany	DAX	DE
10	Hong Kong	HSI	HK	10	Greece	ASE	GR
11	Hungary	BUX	HU	11	Iceland	ICEXI	IS
12	India	SENSEX	IN	12	Ireland	ISEQ	IE
13	Indonesia	JCI	ID	13	Italy	ITLMS	IT
14	Israel	TA-25	IL	14	Japan	NKY	JP
15	Kenya	NSEASI	KE	15	Luxembourg	LUXXX	LU
16	Malaysia	FBMKLCI	MY	16	Netherlands	AEX	NL
17	Mexico	INMEX	MX	17	New Zealand	NZSE	NZ
18	Nigeria	NGSEINDX	NG	18	Norway	OSEBX	NO
19	Peru	SPBL25PT	PE	19	Portugal	BVLX	PT
20	Philippines	PCOMP	PH	20	Singapore	STI	SG
21	Poland	WIG20	PL	21	Spain	IBEX	ES
22	Russia	INDEXCF	RU	22	Sweden	OMX	SE
23	Slovakia	SKSM	SK	23	Switzerland	SPI	CH
24	Slovenia	SBITOP	SQ	24	United Kingdom	UKX	GB
25	South Africa	JALSH	ZA	25	United States	SPX	US
26	South Korea	KOSPI	KR				
27	Thailand	SET50	TH				
28	Turkey	XU100	TR				
29	Ukraine	PFTS	UA				

Text measures defined

Entropy or "unusualness"

- The entropy of an article is

$$H_j = -\sum_{i \in \{4-\text{grams in } j\}} p_i \log m_i$$

- p_i is 4-gram *i*'s fraction of all 4-grams in article *j*
- m_i is the historical conditional probability of observing the 4th word in *i* conditional on the first 3 words
  - For month *t*, we estimate this over time window *t-27,…,t-4*
- H_j is a measure of how different the language constructs in article *j* are relative to a historical sample

Text measures

Network modularity

- *Network modularity* measures the amount of connectivity in the network above what would be expected by chance
- Define a network by cosine similarity

$$similarity\ of\ i\ and\ j \equiv \frac{D_j' D_i}{\|D_i\|\|D_j\|}$$

  where D_i is the ith column of the document term matrix
- Consider a partition *C* of this network with *k* communities
- Let *e* be a $k \times k$ symmetric matrix whose $(i,j)^{\text{th}}$ element is equal to the fraction of all edge weights in the network that connect members of communities *i* and *j*
- The modularity of the network is

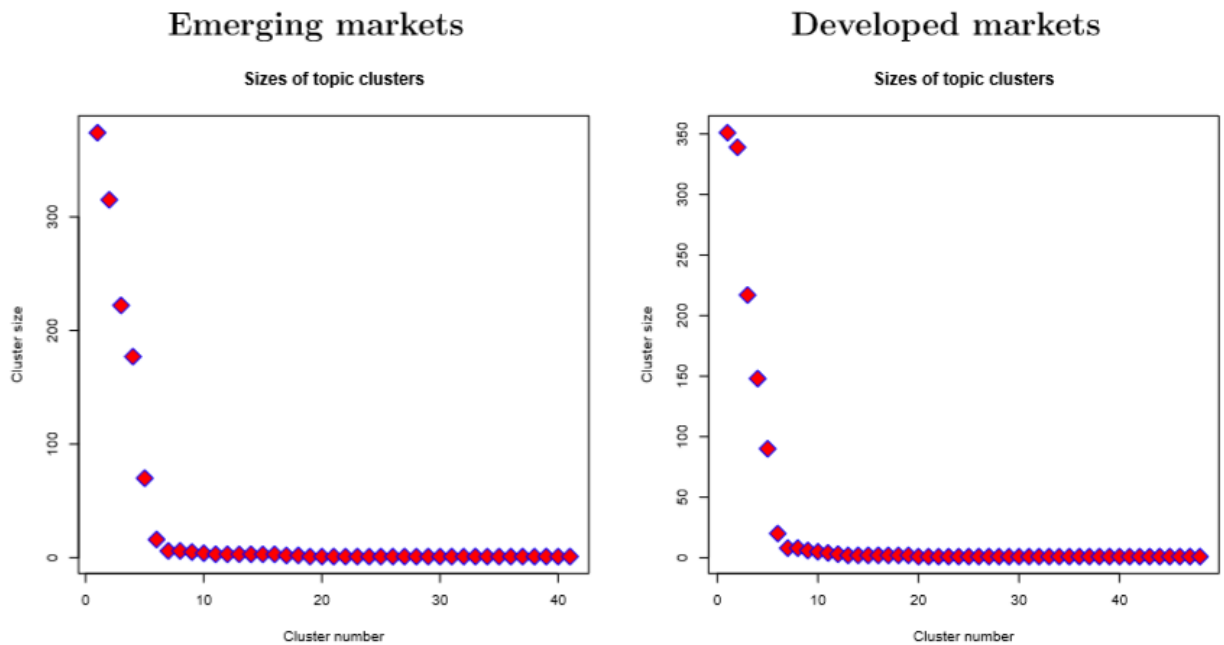
$$Q = \text{Trace } e - |e^2| = \sum_i \left(\underbrace{e_{ii}}_{\substack{\text{fraction of} \\ \text{edges entirely} \\ \text{in group } i}} - \underbrace{a_i^2}_{\substack{a_i = \text{fraction} \\ \text{of all edges} \\ \text{that involve } i}} \right)$$

  where $|\cdot|$ indicates the sum of matrix elements
- Community detection using LDA yields similar results, though is computationally more demanding

Topics

We find 5 topics

- The Louvain algorithm returns ~40 word clusters with the following numbers of words
- Place words from small clusters into big clusters



5 Topics for EM
5 Topics for DM

- How do topics evolve over time? (topic stability)

Topic stability

- Given two partitions *C* and *C'* over a set of *N* words, let $X_{i,j}$ be the *number of common words* between cluster *i* from *C* and cluster *j* from *C'*
- Let *m(i)* be an injective mapping from *C* into *C'*, then the overlap between the two clusterings is

$$\max_{m(\cdot)} \frac{1}{N} \sum_{i} X_{i,m(i)}$$

- This measures the fraction of the *N* words that are placed in the *same* paired categories in *C* and *C'*
  - If this measure is 100%, it means a perfect match between the two clusterings

Time	96-99	00-03	04-07	08-11	12-15
EM	72%	77%	74%	78%	67%
DM	70%	79%	80%	84%	78%

5 topics

Topic similarity across EM and DM

Similarity of developed and emerging market clusters

	Mkt (EM)	Govt (EM)	Corp (EM)	Comms (EM)	Macro (EM)
Mkt (DM)	0.59	0.01	0.01	0.13	0.05
Govt (DM)	0.01	0.82	0.01	0.01	0.04
Corp (DM)	0.10	0.02	0.23	0.07	0.04
Comms (DM)	0.01	0.01	0.03	0.33	0.21
Credit (DM)	0.04	0.04	0.46	0.03	0.06

Table 2: Comparison of overlap between developed and emerging market clusters obtained via the Louvain network algorithm. For two clusters, A and B, the corresponding entry in the table reports $\#(A \cap B)/\#(A \cup B)$.

5 Topics for EM
5 Topics for DM

Examples of sentiment words

- NEGATIVE: closed, fears, disrupting, losses, accused
- UNCERTAIN: almost, approximately, assumed, contingent, believed
- POSITIVE: leading, gained, strengthened, boosted, prosperity

Text measures defined

Baker, Bloom and Davis measure

To investigate the role of policy uncertainty, we first develop an index of economic policy uncertainty (EPU) for the United States and examine its evolution since 1985.2 Our index reflects the frequency of articles in 10 leading US newspapers that contain the following triple: "economic" or "economy"; "uncertain" or "uncertainty"; and one or more of "congress", "deficit", "Federal Reserve", "legislation", "regulation" or "White House".

-- Baker, Bloom and Davis (2015)

Many questions

Sample EM articles

Sample articles in each topic for emerging markets

Topic	Date	Sent	Headline
Mkt	1997-11-06	-0.22	Elbit Ltd<ELBT3.TA><ELBTF.O>Q3 loss $0.11 per share
Mkt	1996-02-16	-0.22	Uganda shilling weakens against dollar
Mkt	1999-09-06	0.12	Hungarian shares open higher on Dow gains
Mkt	2015-03-05	0.12	BUZZ-USD/THB eked out small gains
Govt	2011-03-16	-0.23	US objects to 'excessive force' in Bahrain
Govt	1997-09-18	-0.22	Tehran mayor rejects resignations of 12 mayors
Govt	2000-06-04	0.10	Clinton says Putin can build strong, free Russia
Govt	2008-04-03	0.11	Mugabe's party expects runoff, says he will win
Corp	2011-01-19	-0.25	BRIEF-Moody's downgrades Tunisia's to Baa3, outlook negative
Corp	2011-01-31	-0.25	BRIEF-Moody's downgrades Egypt to Ba2, negative outlook
Corp	2013-05-02	0.14	CORRECTED-TABLE-Philippines' sovereign credit rating history
Corp	2013-03-27	0.16	TABLE-Philippines' sovereign credit rating history
Comms	2008-09-12	-0.13	BP says Baku-Supsa oil pipeline remains shut
Comms	1996-05-09	-0.12	Russia's Novorossiisk oil port still shut by fog
Comms	2006-12-27	0.08	Great Offshore buys anchor-handling tug vessel
Comms	1997-06-26	0.08	Tunisia tender for 150,000 T U.S. wheat detailed
Macro	1996-03-07	-0.12	Hungary 1995 C/A deficit falls to $2.48 billion
Macro	2003-04-30	-0.11	Turkish Jan-Feb c/a deficit jumps to $1.178 bln
Macro	2006-03-10	0.00	Sao Paulo volta a registrar inflacao no comeco de marco
Macro	2012-09-11	0.01	CORRECTED-Lithuania current account surplus rises in June

5 Topics for EM
5 Topics for DM

Sample DM articles

Sample articles in each topic for developed markets

Topic	Date	Sent	Headline
Mkt	2012-05-21	-0.20	BRIEF-FINRA Panel awards John Galinsky $3.5 mln in compensatory damages for breach of contract against Advanced Equities
Mkt	2003-03-25	-0.20	Euro rises above $1.07 against dollar on war
Mkt	1996-01-18	0.12	UK's Clarke confident about inflation, growth
Mkt	2010-11-02	0.12	BRIEF-Metro CEO cautiously optimistic for good christmas
Govt	2009-01-08	-0.30	BRIEF-UK Serious Fraud Office to probe Madoff's UK operations
Govt	2005-09-09	-0.25	Soccer-Former secretary's claim against English FA dismissed
Govt	2014-04-29	0.13	BUZZ-GBP-4/5 on UKIP to win a seat in 2015 UK elections
Govt	2013-09-20	0.13	BUZZ-GBP-5/4 UKIP win most votes in European election
Corp	2014-07-21	-0.15	BRIEF-Valeant Pharmaceuticals contacts Quebec and U.S. regulators about Allergan's false and misleading statements
Corp	2015-12-16	-0.15	BRIEF-NQ Mobile announces termination of proposed divestment of Beijing Tianya
Corp	1996-05-26	0.13	Rangatira has 9.77 pct stake in Advantage <ADV.NZ>
Corp	2015-08-11	0.14	BRIEF-Tom Tailor to improve earnings in 2016 - CEO
Comms	2002-04-17	-0.07	Australasia port conditions - Lloyds
Comms	2012-06-13	-0.07	Cooperatives cut German 2012 wheat crop forecast
Comms	2006-10-10	0.13	TAKE A LOOK- Weekly US state crop progress reports
Comms	2006-10-16	0.13	TAKE A LOOK- Weekly US state crop progress reports
Credit	1998-11-16	-0.29	TABLE - NeoPharm Inc <NEO.A> Q3 net loss
Credit	1998-07-10	-0.27	TABLE - NDC Automation Inc <AGVS.OB> Q2 loss
Credit	2012-02-21	0.22	BRIEF-Moody's revises euramax's outlook to stable from positive
Credit	2011-04-21	0.23	BRIEF-Moody's revises Pulte's outlook to stable from positive

5 Topics for EM
5 Topics for DM

Event study econometric issues

- Our return model is estimated over entire sample
  - Evidence of alpha/news correlations in rolling windows
- News event is not independent of stock prices
- Event windows overlap
- Serial correlation within a country
- Cross-sectional correlation and asynchronicity
- Cross serial correlations

Event studies EM

Control variables

Data definitions summary

Variable	Definition
return	Total monthly stock returns (in %) including capital gains and dividend yield
$return^{12}$	Cumulative stock returns from the start of month t to the end of month $t+11$
sigma	Rolling 20-day realized volatility reported in annualized terms
drawdown	Next 12-month maximum drawdown in percent
retmi	Negative portion of returns
retpl	Positive portion of returns
value	Stock price level of five years ago divided by current stock price level
gdp	Rate of growth of real GDP
gdpdeflator	Rate of change of the GDP deflator
cp	Private sector credit to GDP ratio
dcp	First differenced of credit to GDP ratio
rate	Interest rate
dexch	Change in currency exchange rate to US Dollar, truncated at $\pm 50\%$
pre	Dummy variable set to 1 if month t is 6 or fewer months prior to an election
post	Dummy variable set to 1 if month t is 6 of fewer months after an election
entropy	Measurement of unusualness
artcount	Number of articles writing about a country given one-month period
$s[Topic]$	Sentiment in a given month due to *Topic*
$f[Topic]$	Frequency of articles in a given month due to *Topic*

Empirical results

Summary of panels

	Emerging markets				Developed markets			
	return	*return*^{12}	*sigma*	*drawdown*	*return*	*return*^{12}	*sigma*	*drawdown*
$sigma_{t-1}$		+	+				+	
$sigma_{t-2}$		+	+				+	
$return_{t-1}$	+		Ø	-			Ø	-
$return_{t-2}$			Ø				Ø	
$retmi_{t-1}$	Ø	Ø	+	Ø	Ø	Ø	+	Ø
$retmi_{t-2}$	Ø	Ø		Ø	Ø	Ø		Ø
$value_{t-1}$			-	-	+	+	-	-
gdp_{t-1}	-							
$gdpdeflator_{t-1}$								
cp_{t-1}	-	-	-		-	-		+
dcp_{t-1}			+					
$rate_{t-1}$		-	+	+	-	-	+	+
$dexch_{t-1}$			-	-			+	
pre								
$post$								
$entropy_{t-1}$					+			-
$artcount_{t-1}$		-	+	+	+			
$sMkt_{t-1}$					+			-
$fMkt_{t-1}$								
$sGovt_{t-1}$				+	-			+
$fGovt_{t-1}$	-		+	+				+
$sCorp_{t-1}$	-	-		+				
$fCorp_{t-1}$	-	-	+	+	-	-		+
$sComms_{t-1}$	+			-				
$fComms_{t-1}$								
$sMacro_{t-1}$								
$fMacro_{t-1}$	+	+		-				
$sCredit_{t-1}$						+		
$fCredit_{t-1}$								

- Ø means variable is excluded from regression
- Standard errors:
  - For *return* and *sigma* regressions we cluster by time
  - For *return*^{12} and *drawdown* cluster along both dimensions (Thompson 2011)

[Summary of results]
[EM panel summary]
[DM panel summary]

DM panel summary

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							+					
$entropy_{t-1}$	+			-				+	+	+		-
$artcount_{t-1}$	+											
$sMkt_{t-1}$	+			-			-		+	+		-
$fMkt_{t-1}$												
$sGovt_{t-1}$	-			+					-	-		+
$fGovt_{t-1}$				+					-			+
$sCorp_{t-1}$								-	+			
$fCorp_{t-1}$	-	-		+								
$sComms_{t-1}$												
$fComms_{t-1}$												
$sCredit_{t-1}$		+										+
$fCredit_{t-1}$					+							

EM panel summary

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					-					-		
$entropy_{t-1}$					-		+	+	+			-
$artcount_{t-1}$		-	+	+	-	-		+			+	+
$sMkt_{t-1}$							+					
$fMkt_{t-1}$					-	-	+	+		+		-
$sGovt_{t-1}$				+								
$fGovt_{t-1}$	-		+	+	-	-		+				
$sCorp_{t-1}$	-	-		+					-	-		+
$fCorp_{t-1}$	-	-	+	+		-	+	+				
$sComms_{t-1}$	+			-				-	+			
$fComms_{t-1}$									+			
$sMacro_{t-1}$												
$fMacro_{t-1}$	+	+		-					+		-	-

DM 12 month returns

Developed markets: Forecasting panel for next 12 month returns

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.118	0.176	0.183	0.157	0.116	0.112	0.293	0.334*	0.322*
$sigma_{t-2}$	0.053	0.095	0.097	-0.105	-0.048	-0.042	0.327*	0.377**	0.362*
$return_{t-1}$	0.166	0.033	0.029	0.211	0.108	0.106	-0.059	-0.183	-0.180
$return_{t-2}$	0.040	-0.056	-0.058	0.117	0.066	0.063	-0.072	-0.137	-0.136
$value_{t-1}$	21.315***	22.579***	22.680***	15.086***	13.254**	13.297**	25.646***	26.213***	26.299***
gdp_{t-1}	-0.484	-0.873	-0.919	0.216	0.299	0.255	-2.103***	-1.976***	-1.941***
$gdpdeflator_{t-1}$	0.883	0.490	0.468	0.566	0.517	0.477	0.032	0.104	0.112
cp_{t-1}	-0.278***	-0.223**	-0.227**	0.047	-0.010	-0.003	-0.304**	-0.297**	-0.302**
dcp_{t-1}	0.075	0.026	0.022	-0.180	-0.161	-0.161	0.268**	0.252**	0.257**
$rate_{t-1}$	-3.763***	-5.373***	-5.465***	-14.479***	-13.992***	-13.884***	-5.080***	-5.467***	-5.540***
$dexch_{t-1}$	0.038	0.024	0.026	-0.372**	-0.365**	-0.371**	0.060	0.054	0.054
pre	-0.956	-2.533	-2.585	-3.074	-3.013	-3.005	2.804	1.957	2.050
$post$	-0.475	-1.429	-1.477	-2.583	-2.930	-2.865	2.077	1.798	1.877
$entropy_{t-1}$		20.714	23.072		-23.902	-22.602		21.872*	21.303*
$artcount_{t-1}$		-0.586	-0.631		0.189	0.117		0.100	0.114
$sMkt_{t-1}$		2.765	3.017		-1.569	-2.315		5.235*	5.711**
$fMkt_{t-1}$		0.770	0.873		-4.118	-4.222		1.302	1.285
$sGovt_{t-1}$		-3.367	-3.554*		-0.203	-0.560		-3.896*	-2.792
$fGovt_{t-1}$		-3.768	-3.914		-3.370	-3.601		-6.289	-6.019
$sCorp_{t-1}$		-2.691	-1.775		3.085	3.777		-2.720	-2.916*
$fCorp_{t-1}$		-5.882*	-5.551*		-6.097	-5.965		-0.080	0.097
$sComms_{t-1}$		1.126	0.923		0.154	0.378		0.566	0.447
$fComms_{t-1}$		0.849	0.835		-1.491	-1.326		1.495	1.477
$sCredit_{t-1}$		3.703*	3.175		-1.083	-0.986		-0.852	-1.860
$fCredit_{t-1}$		0.123	0.116		0.741	0.716		-1.104	-1.312
$R2$	0.163	0.214	0.213	0.266	0.298	0.299	0.452	0.471	0.472
$start$	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
$Nobs$	4411	4395	4395	2003	1987	1987	2408	2408	2408
$stderr$	both	both	both	both	both	both	both	both	both

Summary of panels

EM 12 month returns

Emerging markets: Forecasting panel for next 12 month returns

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.189	0.257**	0.256**	-0.045	0.255	0.254	0.486***	0.422***	0.417***
$sigma_{t-2}$	0.320**	0.373***	0.369***	-0.042	0.154	0.150	0.707***	0.662***	0.655***
$return_{t-1}$	0.298	0.199	0.206	0.091	-0.197	-0.184	0.440*	0.450**	0.464**
$return_{t-2}$	0.037	0.021	0.028	-0.099	-0.210	-0.200	0.111	0.131	0.142
$value_{t-1}$	4.061	4.075	4.181	3.657	8.939**	9.021**	9.371	7.417	7.521
gdp_{t-1}	-0.972	-0.884	-0.880	-0.849	-1.918*	-1.889*	-1.614*	-1.194*	-1.175
$gdpdeflator_{t-1}$	0.503*	0.387	0.389	1.078*	0.160	0.182	-0.267	-0.058	-0.061
cp_{t-1}	-0.487***	-0.311***	-0.307***	-0.320	-0.224	-0.212	-0.030	-0.125	-0.128
dcp_{t-1}	0.234	0.230	0.215	0.481	0.348	0.360	0.082	0.226	0.183
$rate_{t-1}$	-0.732*	-0.763*	-0.764*	-1.464*	-0.300	-0.317	-0.954	-1.146	-0.991
$dexch_{t-1}$	0.550	0.576	0.575	0.039	0.221	0.197	1.001**	0.811*	0.820*
pre	-0.277	-0.214	-0.203	-2.691	-5.186	-5.340	-1.580	-1.264	-1.258
$post$	-4.406	-4.149	-4.216	1.401	-0.147	-0.346	-7.895***	-7.018***	-7.167***
$entropy_{t-1}$		1.907	0.999		-48.869	-45.438		6.525	3.590
$artcount_{t-1}$		-5.785***	-5.788***		-10.478***	-10.520***		-1.198	-1.009
$sMkt_{t-1}$		3.076	3.117		3.776	2.705		1.090	1.290
$fMkt_{t-1}$		-3.584	-3.493		-11.018**	-11.430**		8.828***	8.925***
$sGovt_{t-1}$		-1.585	-0.717		-0.671	-0.884		-0.046	0.064
$fGovt_{t-1}$		-6.820	-6.200		-10.377**	-10.517**		3.901	3.733
$sCorp_{t-1}$		-7.044**	-6.705**		-1.812	-0.842		-8.206**	-7.954**
$fCorp_{t-1}$		-7.703***	-7.661***		-7.280*	-6.887*		2.577	2.483
$sComms_{t-1}$		1.260	1.139		3.433	3.826		0.037	-0.661
$fComms_{t-1}$		1.802	1.732		5.054	5.240		3.193	2.766
$sMacro_{t-1}$		2.778	1.765		2.515	1.515		-1.895	-1.625
$fMacro_{t-1}$		5.865**	5.503**		3.544	3.078		2.863	3.042
$R2$	0.0697	0.127	0.125	0.0213	0.13	0.128	0.212	0.264	0.263
$start$	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
$stderr$	both	both	both	both	both	both	both	both	both

DM *sigma*

Developed markets: Forecasting panel for volatility

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.423***	0.398***	0.397***	0.387***	0.348***	0.348***	0.403***	0.375***	0.374***
$sigma_{t-2}$	0.136**	0.120**	0.120**	0.164***	0.148***	0.148***	0.077	0.046	0.046
$retmi_{t-1}$	0.563***	0.526***	0.525***	0.512***	0.454***	0.459***	0.580***	0.555***	0.553***
$retmi_{t-2}$	-0.009	-0.026	-0.027	-0.075	-0.087	-0.084	0.077	0.101	0.098
$value_{t-1}$	-2.163***	-2.555***	-2.591***	-1.454	-2.440**	-2.425**	-2.435**	-2.497**	-2.532**
gdp_{t-1}	-0.183*	-0.135	-0.127	-0.100	0.013	0.016	-0.180	-0.148	-0.147
$gdpdeflator_{t-1}$	-0.019	0.019	0.022	0.046	0.116	0.117	0.181	0.159	0.156
cp_{t-1}	0.011	0.008	0.008	-0.026	-0.018	-0.017	0.030	0.033*	0.032*
dcp_{t-1}	-0.018	-0.012	-0.011	0.007	0.009	0.009	-0.048**	-0.048*	-0.045*
$rate_{t-1}$	0.796***	0.872***	0.889***	1.840***	1.832***	1.846***	0.815***	0.854**	0.870**
$dexch_{t-1}$	0.053*	0.053*	0.052*	0.001	0.010	0.008	0.057*	0.056*	0.056*
pre	0.052	0.010	0.021	-0.092	-0.050	-0.047	0.026	-0.160	-0.128
$post$	0.088	0.077	0.100	0.651*	0.791**	0.797**	-0.467	-0.519	-0.501
$entropy_{t-1}$		-2.107	-2.636		-0.040	-0.850		-1.787	-2.465
$artcount_{t-1}$		0.009	0.026		0.003	-0.002		-0.609	-0.581
$sMkt_{t-1}$		-0.818	-0.859		-1.525*	-1.428*		-0.484	-0.625
$fMkt_{t-1}$		-0.338	-0.331		-0.338	-0.299		-0.019	-0.064
$sGovt_{t-1}$		0.645	0.594		0.033	0.017		1.188	1.048
$fGovt_{t-1}$		0.325	0.297		0.060	-0.007		0.736	0.688
$sCorp_{t-1}$		0.093	0.082		0.070	0.052		-0.671	-0.615
$fCorp_{t-1}$		0.057	0.023		-0.431	-0.494		-0.411	-0.443
$sComms_{t-1}$		0.086	0.133		-0.120	-0.080		0.303	0.326
$fComms_{t-1}$		-0.224	-0.200		-0.175	-0.182		-0.292	-0.267
$sCredit_{t-1}$		-0.553	-0.593		0.066	-0.032		-0.380	-0.349
$fCredit_{t-1}$		-0.102	-0.122		0.313	0.289		-0.422	-0.428
$R2$	0.471	0.477	0.477	0.447	0.458	0.457	0.454	0.459	0.46
$start$	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

EM *sigma*

Emerging markets: Forecasting panel for volatility

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.356***	0.352***	0.352***	0.280***	0.264***	0.264***	0.413***	0.402***	0.402***
$sigma_{t-2}$	0.116***	0.114***	0.113***	0.126***	0.117***	0.116***	0.044	0.030	0.029
$retmi_{t-1}$	0.529***	0.526***	0.524***	0.641***	0.650***	0.648***	0.340***	0.330***	0.329***
$retmi_{t-2}$	-0.024	-0.026	-0.026	-0.132*	-0.125	-0.126	0.098	0.113	0.113
$value_{t-1}$	-0.731**	-0.722***	-0.739***	-0.274	-0.541*	-0.551*	-1.650***	-1.450***	-1.467***
gdp_{t-1}	-0.019	-0.017	-0.015	0.010	0.078	0.079	-0.032	-0.044	-0.040
$gdpdeflator_{t-1}$	0.044	0.057	0.057	-0.005	0.064	0.066	0.088	0.102	0.102
cp_{t-1}	-0.035**	-0.047***	-0.047***	0.004	0.010	0.011	-0.076**	-0.083***	-0.084***
dcp_{t-1}	0.063***	0.062***	0.061***	0.050	0.054	0.054	0.076**	0.085**	0.084**
$rate_{t-1}$	0.367***	0.385***	0.384***	0.447***	0.399***	0.395***	0.625***	0.714***	0.706***
$dexch_{t-1}$	-0.186**	-0.178**	-0.178**	-0.242**	-0.213*	-0.215*	-0.032	-0.031	-0.032
pre	0.211	0.281	0.281	-0.091	0.342	0.345	0.718	0.762	0.753
$post$	-0.012	-0.037	-0.029	0.540	0.709	0.720	-0.135	-0.253	-0.238
$entropy_{t-1}$		0.288	0.260		10.974***	10.982***		-6.050	-6.123
$artcount_{t-1}$		0.589**	0.578*		0.261	0.265		1.108**	1.081**
$sMkt_{t-1}$		0.443	0.432		0.989*	0.943		0.215	0.185
$fMkt_{t-1}$		0.786	0.771*		1.131*	1.099*		0.059	0.021
$sGovt_{t-1}$		0.323	0.395		-0.230	-0.219		0.486	0.666
$fGovt_{t-1}$		1.065*	1.107**		0.696	0.684		0.515	0.617
$sCorp_{t-1}$		-0.068	-0.144		-0.115	-0.163		-0.166	-0.335
$fCorp_{t-1}$		0.683*	0.664*		1.445***	1.429***		-0.010	-0.057
$sComms_{t-1}$		-0.344	-0.305		-0.416	-0.304		-0.290	-0.228
$fComms_{t-1}$		0.358	0.381		0.266	0.324		0.514	0.543*
$sMacro_{t-1}$		-0.088	-0.101		-0.114	-0.168		0.124	0.120
$fMacro_{t-1}$		0.031	0.021		0.571	0.534		-0.654**	-0.677**
$R2$	0.374	0.376	0.376	0.318	0.328	0.328	0.384	0.391	0.392
$start$	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
$Nobs$	4844	4830	4830	2092	2078	2078	2752	2752	2752
$stderr$	by time	by time	by time	by time	by time	by time	by time	by time	by time

DM *drawdown*

Developed markets: Forecasting panel for drawdowns

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.114	0.028	0.024	0.074	0.034	0.037	0.016	-0.080	-0.074
$sigma_{t-2}$	0.157*	0.096	0.098	0.154**	0.098	0.096	0.064	-0.030	-0.018
$return_{t-1}$	-0.283***	-0.166*	-0.167*	-0.168*	-0.089	-0.092	-0.255**	-0.124	-0.126
$return_{t-2}$	-0.089	-0.019	-0.018	-0.026	-0.005	-0.004	-0.087	-0.035	-0.031
$value_{t-1}$	-12.351***	-13.909***	-13.980***	-5.980***	-6.590***	-6.568***	-15.914***	-16.823***	-16.983***
gdp_{t-1}	0.177	0.424	0.457*	-0.085	0.048	0.080	0.814**	0.808**	0.794**
$gdpdeflator_{t-1}$	-0.290	-0.084	-0.072	-0.355	-0.242	-0.218	0.377	0.165	0.158
cp_{t-1}	0.187***	0.156***	0.158***	-0.007	0.034	0.028	0.151**	0.147**	0.147**
dcp_{t-1}	-0.070	-0.038	-0.034	0.074	0.068	0.068	-0.143	-0.115	-0.113
$rate_{t-1}$	3.116***	3.940***	4.021***	7.668***	7.461***	7.392***	4.278***	4.514***	4.622***
$dexch_{t-1}$	-0.039	-0.025	-0.027	0.051	0.057	0.056	-0.036	-0.024	-0.025
pre	1.397	1.962	1.998	2.012	2.065	2.072	0.087	0.283	0.273
$post$	0.196	0.401	0.443	0.448	0.813	0.787	-0.444	-0.913	-0.954
$entropy_{t-1}$		-11.132*	-13.343**		11.432*	9.577		-11.476*	-12.288*
$artcount_{t-1}$		0.078	0.145		-0.064	-0.010		-0.912	-0.892
$sMkt_{t-1}$		-4.235***	-4.434***		-0.516	0.060		-6.819***	-7.337***
$fMkt_{t-1}$		-0.859	-0.882		0.690	0.817		-1.357	-1.413
$sGovt_{t-1}$		4.638***	4.366***		0.514	0.680		6.001***	4.928***
$fGovt_{t-1}$		3.942***	3.899***		0.968	1.056		5.583***	5.268***
$sCorp_{t-1}$		1.137	0.563		-1.540*	-2.022**		0.350	0.501
$fCorp_{t-1}$		2.621**	2.324**		1.457	1.341		-0.038	-0.286
$sComms_{t-1}$		-0.790	-0.512		-0.631	-0.655		0.261	0.393
$fComms_{t-1}$		-0.077	0.000		0.420	0.329		0.365	0.443
$sCredit_{t-1}$		-1.603	-1.156		-0.240	-0.384		1.669*	2.399***
$fCredit_{t-1}$		0.367	0.364		-0.389	-0.384		0.874	0.953
$R2$	0.26	0.32	0.32	0.403	0.45	0.453	0.385	0.444	0.444
$start$	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
$stderr$	both	both	both	both	both	both	both	both	both

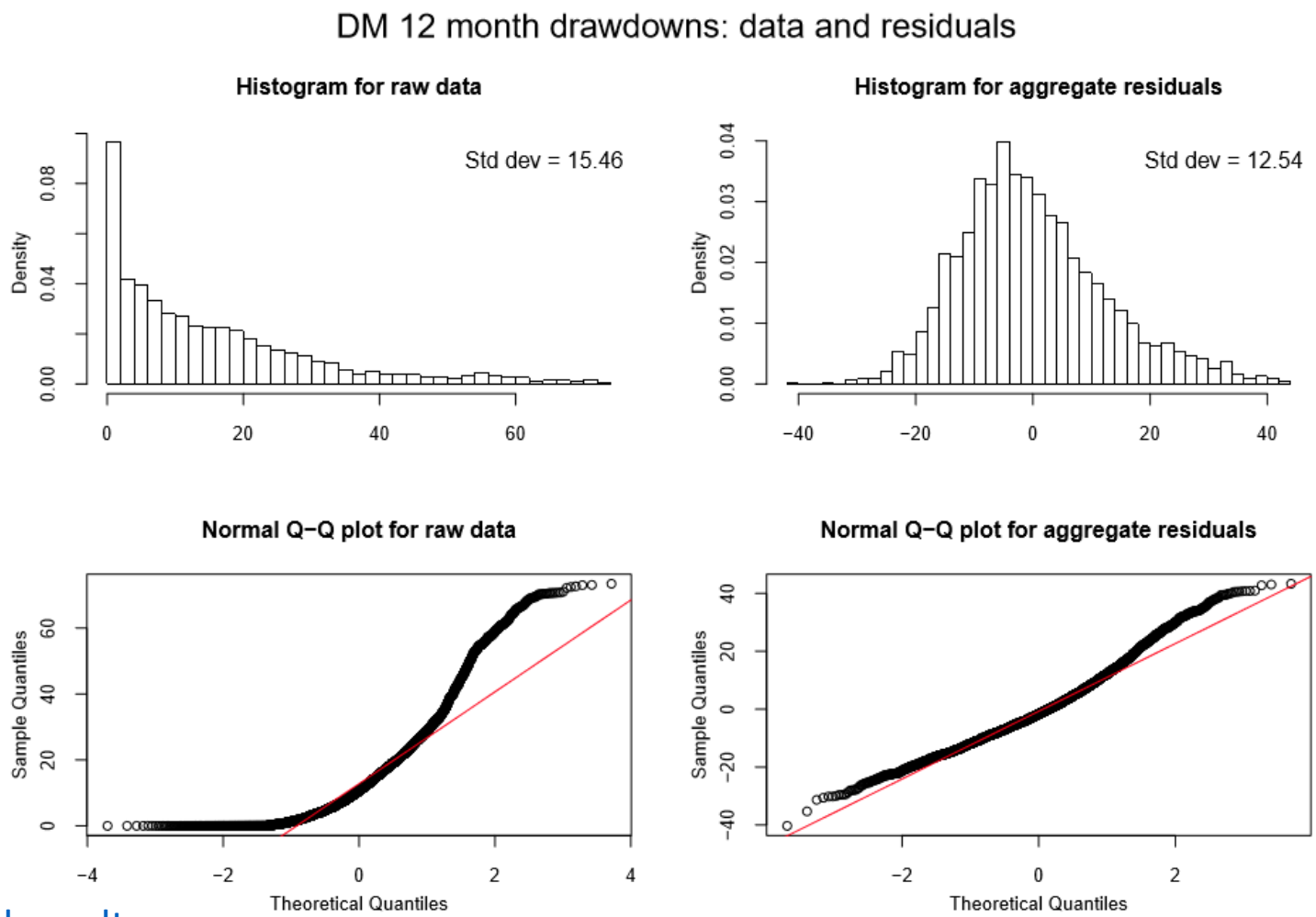
Summary of panels

EM *drawdown*

Emerging markets: Forecasting panel for drawdowns

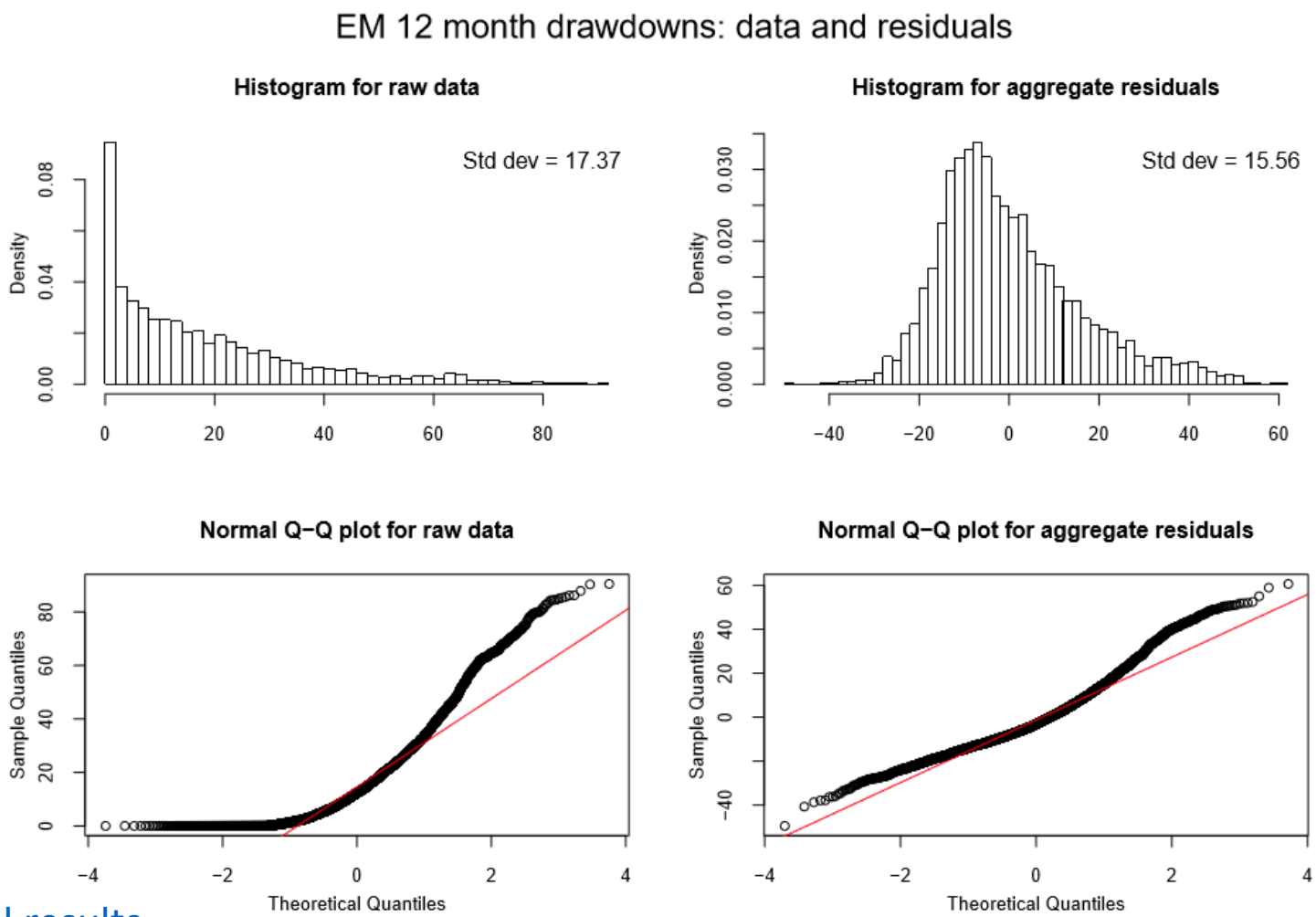
	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.092	0.066	0.069	0.107*	0.007	0.009	-0.022	-0.027	-0.024
$sigma_{t-2}$	0.023	-0.011	-0.008	0.099	0.014	0.014	-0.134*	-0.155**	-0.153**
$return_{t-1}$	-0.221***	-0.163**	-0.172**	-0.084	-0.012	-0.017	-0.298***	-0.269***	-0.282***
$return_{t-2}$	-0.076	-0.066	-0.071	0.054	0.062	0.059	-0.166	-0.164*	-0.171*
$value_{t-1}$	-2.105*	-1.885**	-1.964**	0.192	-1.698**	-1.712**	-6.564**	-5.499**	-5.586**
gdp_{t-1}	0.364**	0.336	0.332	0.386	0.739***	0.731***	0.588**	0.480**	0.484**
$gdpdeflator_{t-1}$	0.060	0.057	0.056	-0.113	0.214	0.211	0.418*	0.342*	0.350*
cp_{t-1}	0.124**	0.044	0.042	0.102	0.057	0.055	-0.084	-0.060	-0.062
dcp_{t-1}	0.070	0.038	0.041	-0.021	0.018	0.014	0.143	0.072	0.082
$rate_{t-1}$	0.531***	0.575***	0.573***	0.651***	0.196	0.195	1.933***	2.082***	2.060***
$dexch_{t-1}$	-0.305**	-0.318**	-0.313**	-0.059	-0.095	-0.090	-0.442**	-0.371*	-0.370*
pre	-1.259	-1.335	-1.333	-0.004	1.313	1.371	-0.624	-0.828	-0.867
$post$	0.166	-0.225	-0.162	-0.620	0.175	0.258	0.991	0.076	0.163
$entropy_{t-1}$		-5.269	-5.015		34.498***	33.656***		-22.328**	-21.475**
$artcount_{t-1}$		2.791***	2.804***		3.908***	3.930***		2.487**	2.420**
$sMkt_{t-1}$		-1.787	-1.391		0.331	0.822		-0.933	-0.589
$fMkt_{t-1}$		0.066	0.143		4.248**	4.404**		-5.188***	-5.206***
$sGovt_{t-1}$		2.277**	1.995**		0.175	0.233		2.439	2.455*
$fGovt_{t-1}$		3.787***	3.561***		3.534**	3.525**		0.057	0.024
$sCorp_{t-1}$		2.872**	2.564**		-0.541	-0.735		3.309***	2.854**
$fCorp_{t-1}$		3.011***	2.906***		3.278***	3.143***		-1.427	-1.567
$sComms_{t-1}$		-1.422*	-1.309*		-1.628***	-1.707***		-1.407	-0.994
$fComms_{t-1}$		-0.421	-0.379		-0.699	-0.767		-1.182	-1.015
$sMacro_{t-1}$		-1.283	-1.132		-0.235	-0.284		-0.571	-0.848
$fMacro_{t-1}$		-1.941**	-1.922**		0.155	0.146		-2.195**	-2.356**
$R2$	0.0616	0.121	0.118	0.0798	0.221	0.223	0.173	0.252	0.249
$start$	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
$stderr$	both	both	both	both	both	both	both	both	both

DM drawdown residuals



Empirical results

EM drawdown residuals



Empirical results

Out-of-sample testing

- Do we have too many explanatory variables?
- What about regime shifts?
- Check out-of-sample forecasting performance
- Run rolling 5-year regressions in *t-60,…,t-1* for forecasting month *t* outcomes

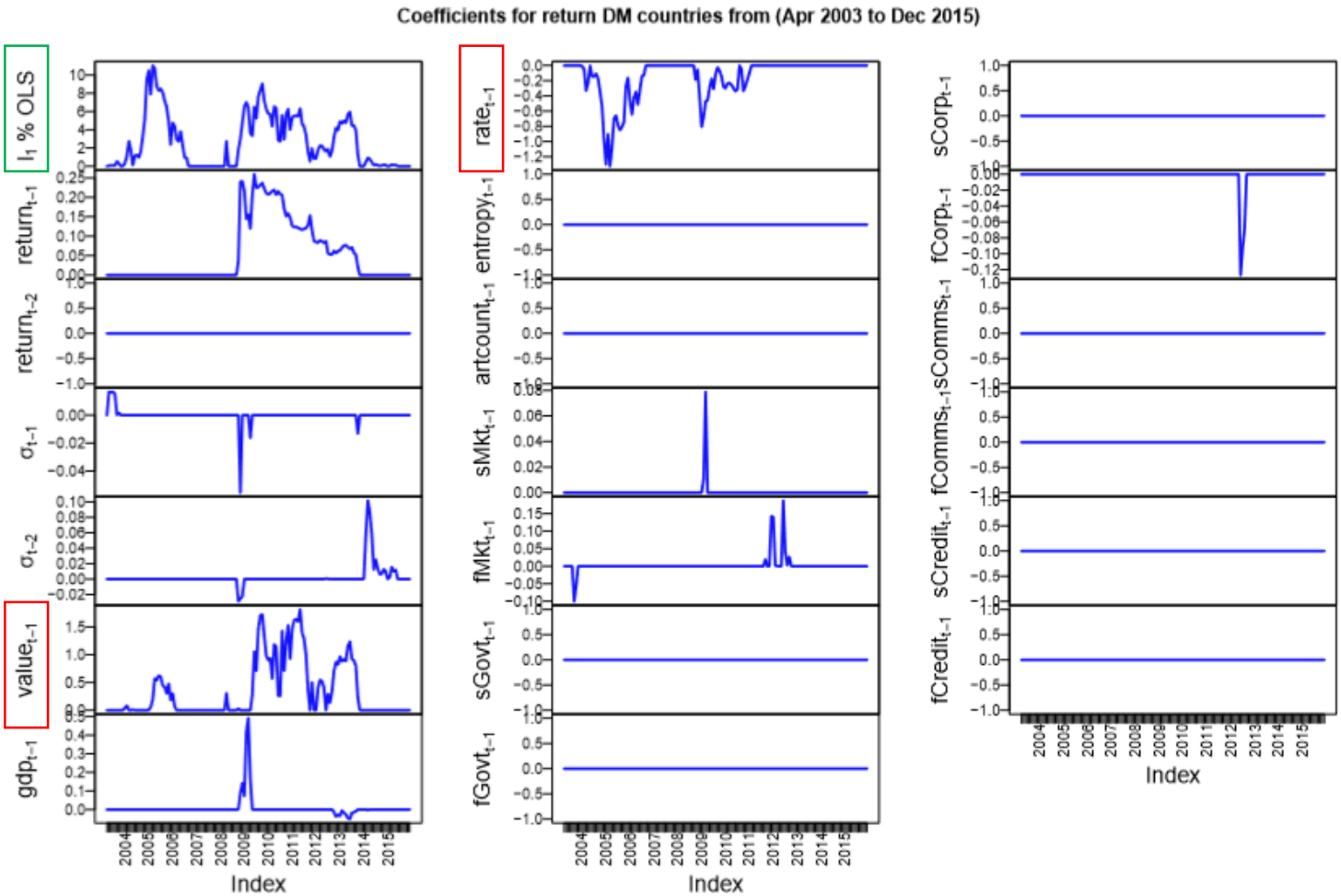
<u>*Lasso*</u> (least absolute shrinkage and selection operator)

$$\min_{\beta} \frac{1}{2N} \sum_{1,t} \left(y_{i,t} - x'_{i,t-1}\beta\right)^2 + \lambda \|\beta\|_1$$

- Lasso does shrinkage and model selection
- Very good at identifying correlated variables
- Amount of shrinkage given by $\|\beta\|_1 / \|\beta^{OLS}\|_1$

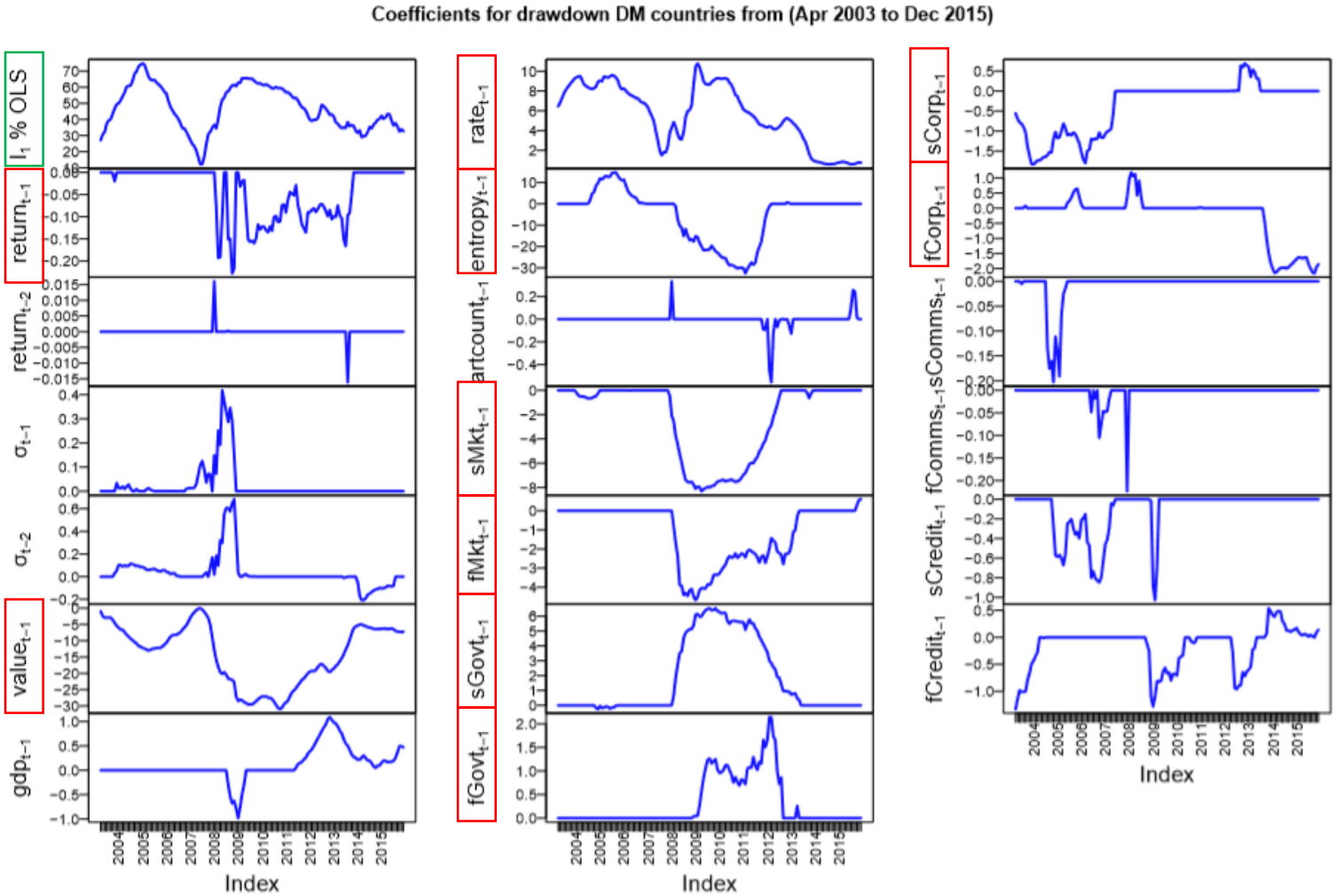
When the model has little to say

Coefficient time series from lasso for monthly returns for DM countries



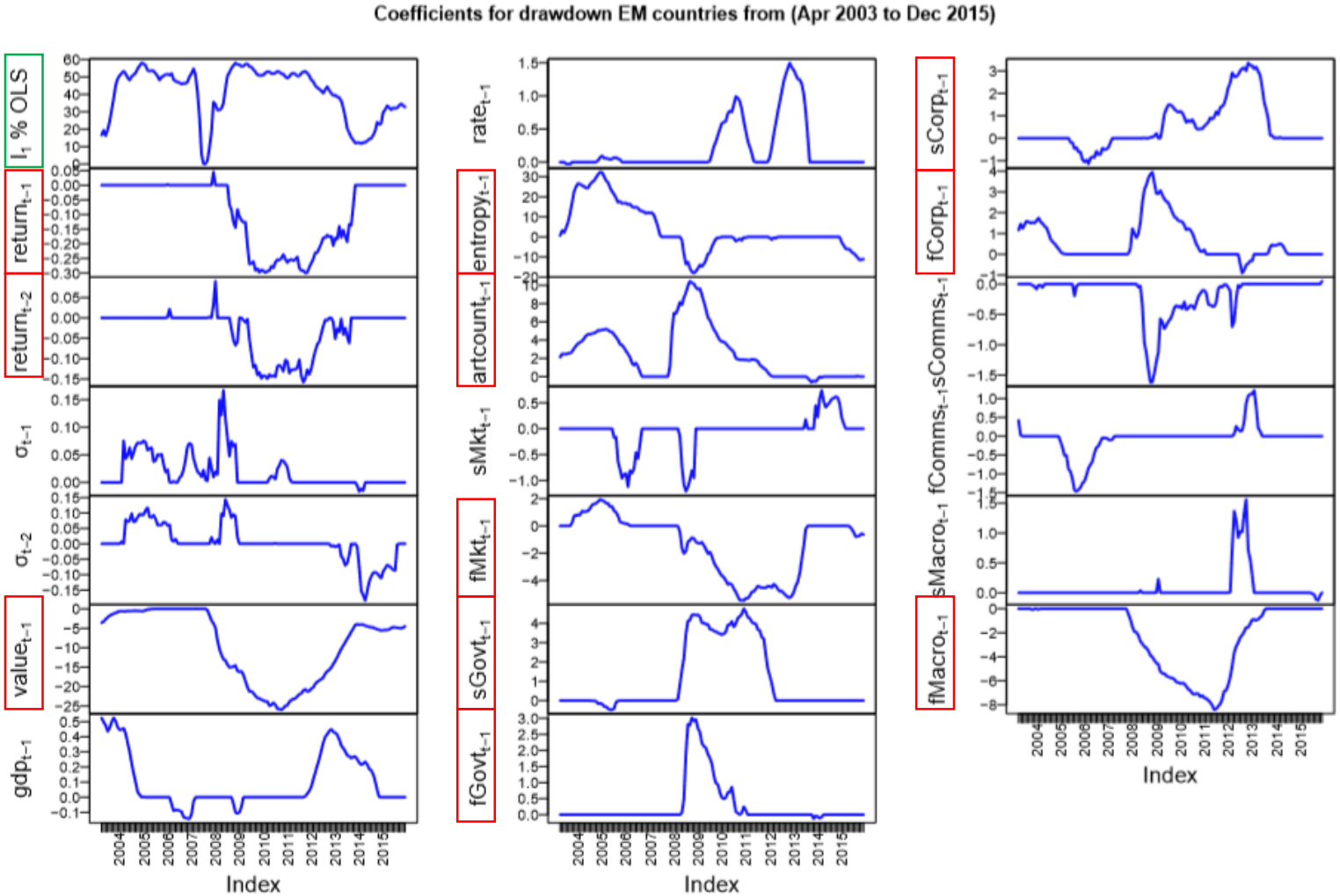
Rolling lasso for DM *drawdown*

Coefficient time series from lasso for next 12 month drawdown for DM countries



Rolling lasso for EM *drawdown*

Coefficient time series from lasso for next 12 month drawdown for EM countries



Out-of-sample performance

Out of sample forecasting tests for developed markets sample

LHS	AvgNumObs	Naive	Base	CM
$return_t$	1272.38	6.67	6.78	6.78
$return_t^{12}$	1271.59	28.74	24.99	24.37
$sigma_t$	1272.38	11.54	8.89	8.89
$drawdown_t$	1272.38	17.57	14.92	14.32

Out of sample forecasting tests for emerging markets sample

LHS	AvgNumObs	Naive	Base	CM
$return_t$	1396.29	8.48	8.54	8.51
$return_t^{12}$	1396.29	42.17	41.59	39.45
$sigma_t$	1393.25	12.05	10.20	10.18
$drawdown_t$	1396.29	18.29	17.66	16.78

- Look at RMSE for out-of-sample forecasting
- *Naïve* model forecasts using country fixed effects
- *Base* model includes only the non-text variables
- *CM* includes all text measures
- All models estimated using lasso

Out-of-sample comparisons to *EPU*

Out of sample forecasting tests for developed markets sample – Baker-Bloom-Davis subsample

LHS	AvgNumObs	Naive	Base	EPU	CM	Both
$return_t$	594.35	6.21	6.32	6.33	6.33	6.34
$return_t^{12}$	594.35	25.30	22.41	22.20	21.38	21.31
$sigma_t$	594.35	11.73	8.98	8.97	8.95	8.94
$drawdown_t$	594.35	16.63	14.12	13.88	12.86	12.88

Out of sample forecasting tests for emerging markets sample – Baker-Bloom-Davis subsample

LHS	AvgNumObs	Naive	Base	EPU	CM	Both
$return_t$	337.25	8.57	8.70	8.70	8.71	8.71
$return_t^{12}$	337.25	43.05	42.59	42.58	40.35	40.38
$sigma_t$	337.25	13.53	10.75	10.75	10.69	10.69
$drawdown_t$	337.25	17.94	17.53	17.57	16.49	16.51

- *EPU* counts articles from 10 major papers that contain triplets from

  uncertainty x

  economic x

  {policy terms}

- For 5 EM and 11 DM countries where we have EPU data, compare out-of-sample performance of *Base* vs *Base* + alternative text measures