Essays on Novel Methods in Empirical Finance

by

Vitaly Meursault

Submitted In Partial Fulfillment of the

Requirements for the Degree

Doctor of Philosophy

at the

Carnegie Mellon University

David A. Tepper School of Business

Pittsburgh, Pennsylvania

Dissertation Committee:

Professor Bryan Routledge (Chair)

Professor Chester Spatt

Professor Stephen Karolyi

Professor Dokyun Lee

March, 17, 2020

Contents

Acknowledgements **7**

1 The language of earnings announcements **9**

1.1 Introduction 9

1.2 Data 13

1.2.1 Earnings announcements corpus 13

1.2.2 Compustat and CRSP data 13

1.2.3 Data split 14

1.3 Statistical model of language 14

1.3.1 Model overview 15

1.3.2 Model interpretation 17

1.4 Information content of language 20

1.5 What makes earnings announcements informative? 27

1.5.1 Earnings, earnings components, and loss 28

1.5.2 Sentiment 31

1.5.3 Forward-looking statements 37

1.5.4 Quotes . . . 41
1.5.5 Non-GAAP measures . . . 45
1.5.6 Comparing different themes . . . 49
1.6 Discussion . . . 51
1.7 Conclusion . . . 52
Appendix A: Earnings announcement dataset construction . . . 54
Appendix B: CNN-GRU . . . 55
Document . . . 55
Embedding . . . 55
Convolution . . . 56
Max Pooling . . . 57
Repeating Convolutions and Max Pooling . . . 58
Gated Recurrent Unit . . . 58
Repeat GRU . . . 59
Linear layer . . . 60
Integrated Gradients . . . 60

2 Investment fund letters and the economy 63

2.1 Introduction . . . 63
2.2 Data . . . 66
2.2.1 Fund shareholder letters . . . 66
2.2.2 Outcome variables . . . 68
2.3 Hierarchical Attention Network (HAN) . . . 69

2.4 HAN vs. Elastic Net . . . 71
2.5 Comparing HAN and Elastic Net predictions . . . 73
2.6 Summarizing the narratives with HAN . . . 75
2.6.1 Integrated Gradients . . . 75
2.6.2 Integrated Gradients example . . . 76
2.6.3 Word lists . . . 77
2.6.4 Important words for different targets . . . 80
2.7 Conclusion . . . 82
Appendix A . . . 84
HAN vs Elastic Net prediction example 1 . . . 84
HAN vs Elastic Net prediction example 2 . . . 85

3 Bank credit supply and shadow mortgage lending 101

3.1 Introduction . . . 101
3.2 The rise of the shadow mortgage lending and the originate-to-sell model . . . 106
3.3 Supply-demand decomposition using HMDA dataset . . . 110
3.3.1 Geographic distribution of the shocks . . . 118
3.3.2 Validity of the shocks . . . 119
3.4 Bank credit supply and shadow lending growth . . . 121
3.4.1 More granular lender types . . . 128
3.5 Shadow lending growth and bank credit supply with split samples . . . 129
3.6 Conclusion . . . 135
References . . . 136

Acknowledgements

My advisor, Bryan Routledge, backed the ideas behind this thesis from the very beginning. His support gave me the courage to pursue the research I wanted to do. I am also thankful to my committee members, Chester Spatt, Steve Karolyi and Dokyun Lee, for helping me to shape my work and guide it throughout graduate school.

Research like mine is unimaginable without the open-source software community. I am grateful to the communities behind R, RStudio, Python, Anaconda and many others. I wouldn't have made much progress without the vast knowledge shared on Stack Overflow. I am also thankful to the Pittsburgh Supercomputing Center for being able to run my code promptly.

Tepper School of Business Ph.D. students are lucky to have administrative assistance from Lawrence Rapp and Laila Lee. Their tireless work was fundamental for my progress.

Getting a Ph.D. is about managing one's mental health as much as it about research. I am very thankful to mental health professionals of CMU and Pittsburgh in general for supporting me through the many trials of graduate school.

I am thankful to all my friends in Pittsburgh and St. Petersburg who greatly supported me through the years, especially Saryuna Lambayeva, Nadya Romanova, Ksenia Sautina and Madeline Scanlon. As a Russian proverb goes, don't have a hundred roubles but a hundred friends.

I wouldn't have gotten my Ph.D. if I didn't get my Master's degree at the European University at St. Petersburg. I thank Maxim Bouev, Yulia Vymyatnina, Kirill Borissov and everyone else at the department of Economics for always believing in me.

Most of all, I am thankful to my wife Casey Roark, the best partner anyone could wish for. I am very happy to have an amazing family in America – Robert and Janita Roark, Samantha, Nathan, and Caleen Geerdes. I dedicate this dissertation to my grandparents, babushka Masha and dedushka Gena, who wanted me to be a doctor since I was a little boy.

Chapter 1

The language of earnings announcements

1.1 Introduction

Earnings announcements contain important information for investors. Absolute price changes on announcement days are more than twice the size of changes on other days (Beaver, McNichols, and Wang 2018). Various aspects of earnings press releases are shown to be informative for the markets, including the presence of operational details in a release (Francis, Schipper, and Vincent 2002a) and forward-lookingness (Bozanic, Roulstone, and Van Buskirk 2018). Interestingly, the content of earnings announcements is varied. For example, companies can mention topics such as their segment sales ("Meanwhile, sales for our disk drive analyzer products reached their highest level in more than two years"; LeCroy, 2005, Q3), or relay their

excitement about the potential of their products ("We believe that these new and enhanced products represent great potential"; Mobius Management System, 2006, Q1). This paper shows that we can systematically understand what language is important.

To quantify the information content of announcement language, we use a flexible statistical model. Specifically, the deep neural network approach employed here allows extracting information from language (words in order) without prejudging the value of content. The model estimates a series of latent factors based on all document content. These factors reflect word usage, local context, and word order. Here, the latent factors are used to predict absolute returns around the time of an earnings announcement.

Earnings announcement language has high information content. By our estimate, the statistical model of language explains 11% of the variation in absolute announcement returns. Regression analysis of the same returns data with standard financial variables has comparable explanatory power. Combining language with the numerical variables consistently improves the explanatory power of the model. About 13% of the information contained in the language is incremental to my full set of numbers. Overall, language both substitutes for and complements the numbers.

Since the model explains the absolute price changes well, we use it to understand what features of the language of earnings announcements are valuable to investors. To do that, we project the latent factors onto the document space by quantifying, for every sentence, how much the prediction of the model would change if the sentence were absent. This process enables a variety of thought experiments about how language conveys information.

For example, we estimate what happens with removal of all sentences mentioning earnings components, forward-looking sentences, sentences with tone markers, quotes, and so on.

The magnitude of absolute price changes depends on the content of earnings announcements. When a document mentions earnings components, the price changes tend to be larger than when it mentions bottom-line net income. Mentions of losses are associated with especially large swings in prices.

It is interesting to see that the way earnings announcements say something can be as important as what they say. Our model encapsulates interactions between the content groups and language features. Sentiment and forward-lookingness amplify the information content of all content groups. Quotes from managers are informative when they talk about earnings components and operations. Markets also respond to discussions of non-GAAP measures.

Starting with Beaver (1968), many studies examine the importance of various items contained in earnings announcements. Examples include earnings value (Ball and Brown 1968), earnings components (Lipe 1986), recurrent and nonrecurrent components of earnings (Fairfield, Sweeney, and Yohn 1996), non-GAAP earnings (Bradshaw and Sloan 2002), and balance sheet items (Collins et al. 1997).

A series of papers discusses specific textual features of earnings announcements. Examples include the provision of operational details (Francis, Schipper, and Vincent 2002b), an emphasis on non-GAAP measures (Bowen, Davis, and Matsumoto 2005), sentiment (Davis, Piger, and Sedor 2012), and prevalence of forward-looking statements (Bozanic, Roulstone, and Van Buskirk 2018). An extensive literature studies financial texts other than earnings

announcements. Popular collections of texts include forms 10Q/10K, earnings calls, and financial media. Many papers focus on extracting specific features of text like sentiment (Tetlock 2007; Li 2010), readability (Loughran and McDonald 2014), or prevalence of forward-looking statements (Muslu et al. 2015). It is common to extract textual features either by hand or using rule-based algorithms.

Another literature applies statistical models to financial disclosure texts. In Finance, Kogan et al. (2009) pioneered the use of text regression, in which word and phrase counts are used in predictive tasks. Routledge, Sacchetto, and Smith (2017) and Chebonenko, Gu, and Muravyev (2018) take a similar approach. The current study uses a deep neural network that can model language in a more flexible way.[1]

Statistical models of language can help link disclosure theory to textual data. In models, scholars tend to look at signals with low dimensional value. For example, in Kim and Verrecchia (1991), Kim and Verrecchia (1994), and in Tetlock (2010), the signal is univariate normal. Statistical models of text provide a tool for processing raw text data into a signal. The mapping of text to signal can be a source of disagreement and lead to market frictions.

Statistical models of language can be of practical use to investors. Earnings announcements are informative, but investor attention is constrained. Our model can ease the constraints by extracting the most informative text.

Earnings announcements also attract attention from regulators, often because of the possibly

[1]Topic modeling is a statistical approach alternative to text regression. It involves extracting latent factors (topics) based on word concurrence across documents. The topics can then become inputs to regressions. Hoberg and Lewis (2017) is an example of this approach.

misleading non-GAAP reporting. Our model can estimate the magnitude of price changes associated with non-GAAP measures. This feature can be useful in the regulatory process.

1.2 Data

Data are the texts of earnings announcements from SEC's EDGAR database combined with information from Compustat and CRSP. The dataset contains 128,317 observations at the firm-quarter level and covers the period from 2005 to 2017.

1.2.1 Earnings announcements corpus

The corpus includes forms 8-K that contain a "Results of Operations and Financial Condition" item according to their metadata and that are filed within five days of an earnings release date as recorded in Compustat. This paper refers to the whole text of the 8-K as the earnings announcement, even if it contains other items besides the earnings press release. The median number of words, punctuation marks and special symbols (collectively referred to as tokens) in an earnings announcement in this corpus is 3,278. Across the sample years, the median size of a document grows from 2,282 to 3,816 tokens.

The Appendix A presents details of corpus construction.

1.2.2 Compustat and CRSP data

The usual numerical variables related to the firms' performance come from Compustat and CRSP. Earnings announcements are matched to Compustat observations corresponding to the quarter discussed in the earnings announcement. For example, an earnings announcement

released during the second quarter that discusses the first quarter is matched with the Compustat observation for the first quarter. This way earnings announcements are matched to accounting numbers that are public at the time of the release or at least soon after.

Calculation of the variables derived from CRSP data, such as absolute announcement returns and stock volatility, is based on the earnings announcement date from Compustat.

1.2.3 Data split

We split the data into training, validation, and test sets to measure information content. Since the models used in this paper involve large numbers of parameters and have the potential for overfitting the data, the performance of the models was evaluated on the test set, a subset of data not seen by the models during estimation.

The total number of firms represented in the dataset is 6,202. We randomly split the firms into training (4,462 firms and 91,914 observations), validation (882 firms and 18,431 observations), and test sets (858 firms and 17,972 observations). The training and validation sets are used to estimate the model, and the test set is used for the model evaluation. All the results presented in this paper are obtained using only the test set—that is, the observations generated by the firms whose data were not used for model estimation.

1.3 Statistical model of language

The statistical model of language used here is a CNN-GRU (Convolutional Neural Net with Gated Recurrent Units) that operates by sequentially creating a set of latent factors. These

factors reflect word usage, local context, and word order. Minimization of the prediction error drives factor extraction.

1.3.1 Model overview

CNN-GRU is a composition of several functions:

$$f(X_{f,t}; w) = \mathrm{L} \circ \mathrm{GRU}_2 \circ \mathrm{GRU}_1 \circ \mathrm{Max}_3 \circ \mathrm{Conv}_3 \circ \mathrm{Max}_2 \circ \mathrm{Conv}_2 \circ \mathrm{Max}_1 \circ \mathrm{Conv}_1 \circ \mathrm{Emb}(X_{f,t}),$$

where $X_{f,t}$ is the document, w is the set of model parameters, Emb is the embedding, Conv is the convolution, Max is max pooling, GRU is the Gated Recurrent Unit, and L is linear layer.

The document is represented as a vector $X = [x_1 \cdots x_n]$, where x_j is the index of the j'th word in the vocabulary, and n is the maximum document length. If a document is shorter than the maximum document length, the extra space is filled with special padding words.

Embedding (Emb) assigns embedding vectors to individual words. The starting values for word embeddings are obtained using the word2vec algorithm (Mikolov et al. 2013) applied to the training set. The embedding vectors are then updated during estimation. An important property of embeddings is that words that likely to be used interchangeably tend to cluster in the embedding space. Figure 1.1 shows several such clusters, including names of the months, corporate titles, and the words "earnings," "income," and "loss." Thus, embeddings reflect word usage.

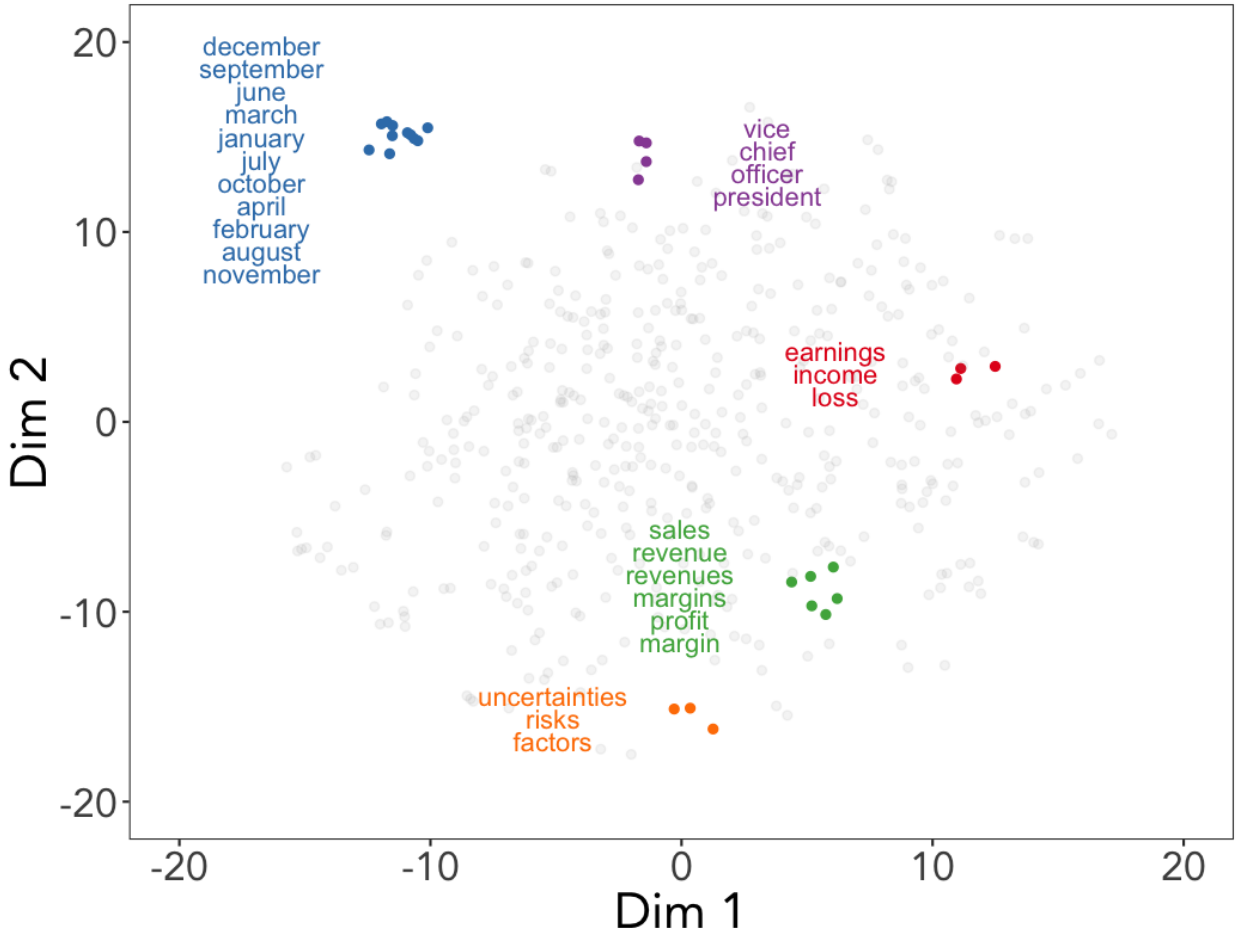

Figure 1.1: Embeddings example. Words that tend to be used interchangeably cluster together in the embedding space. We perform Principal Component Analysis on 200-dimensional word embeddings and visualize the first two principal components. The clusters are picked manually. The embeddings are intially obtained using the word2vec algorithm (Mikolov, 2013) on earnings announcements. They are further updated during the estimation of CNN-GRU.

The model proceeds by creating a series of latent factors. Convolutions create latent factors that reflect local context. For example, a latent factor might learn to distinguish "net income" from "operating income." Max pooling only prunes the factors, keeping the ones most associated with absolute announcement returns. At the next stage, gated recurrent units create latent factors that take into account word order. For example, the model can learn to distinguish between the phrases "the board of directors dismissed the CEO" and "the CEO dismissed the board of directors." Finally, the linear layer takes the final set of latent factors and runs them through a linear regression to produce the prediction.[2]

All layers of CNN-GRU are estimated simultaneously. The objective function is the sum of squared errors:

$$\min_{w} \sum_{f,t} \left[y - f(X; w) \right]^2 ,$$

where $f(X, w)$ is the prediction of the model. The weights are estimated using an iterative procedure.[3]

1.3.2 Model interpretation

We want to use the model to know what is important in earnings announcements. Intuitively, if a passage is important, removing it should change the model's prediction. CNN-GRU is a nonlinear model containing more than 600,000 parameters. To interpret it, one computes impact scores, which represent the change in model prediction resulting from moving from

[2]Convolutional neural nets were originally used for computer vision tasks, but quickly made headway into natural language processing (Collobert et al. 2011; Kim 2014). Along with other word-order-aware models, convolutional neural networks are becoming one of the standard choices in sentence and document classification tasks (including sentiment analysis) and make some headway into the social sciences, where they are used to, for example, classify political discourse (Bilbao-Jayo and Almeida 2018).

[3]We use adaptive moment estimation (Adam) algorithm, which is a form of stochastic gradient descent.

an empty space to the observed word.

Impact scores were computed using integrated gradient method (Sundararajan, Taly, and Yan 2017).[4] Integrated gradients correspond to Aumann-Shapley values from Game Theory (Aumann and Shapley 1974). They are computed at the word level by propagating the derivatives associated with individual words through CNN-GRU. The same word can have different impact scores depending on its context. Within every document the scores sum to the model prediction. The Appendix B provides details.

Figure 1.2 provides a sample paragraph with words highlighted according to their impact scores. Red represents wirds with an impact larger than zero. Removing them would drive the model's prediction downwards. Blue words have an impact smaller than zero. Removing them would result in higher $|\hat{R}^a|$. Color density represents the magnitude. In this example, the word "earnings" and the phrase "strategic decision" are associated with low absolute announcement returns, while the phrase "Mike Brooks, chief executive officer, commented" is associated with more repricing activity.

The Appendix presents details of the model's architecture and training. Table 1.1 presents an overview of the model, including the number of parameters.

[4]Other methods of computing attribution scores include Shapley additive values (SHAP) (Lundberg, Allen, and Lee 2017), DeepLift (Shrikumar, Greenside, and Kundaje 2017), and Relevance Propagation (Arras, Horn, et al. 2017; Arras, Montavon, et al. 2017)

Figure 1.2: Integrated gradients example. The model associates the words in red with high absolute announcements returns, and the words in blue with low absolute announcement returns. The color density corresponds to the magnitude of the association. The weights are computed using the integrated gradients method (Sundararajan et al., 2017), which is one way to attribute model predictions to input words.

Table 1.1: Dimensions of CNN-GRU. This tables presents the operations (layers) of CNN-GRU. For each layer, the associated parameters and hyperparameters are listed. Parameters are estimated, while hyperparameters are picked using the validation set.

Layer	Notation	Output dim		Num of par.	Par. Notation	Hyperparameters
Input	X	12,000		0		
Embedding	Emb	12,000	200	6,000,400	$e_{i;j}$	d(E) = 200
Convolution	$Conv^1$	11996	128	128,128	$w^c; b^c$	$h^c = 5$; d(C) = 128
Max Pooling	Max^1	2,399	128	0		$h^m = 5$
Convolution	$Conv^2$	2,395	128	82,048	$w^c; b^c$	$h^c = 5$; d(C) = 128
Max Pooling	Max^2	479	128	0		$h^m = 5$
Convolution	$Conv^3$	475	128	82,048	$w^c; b^c$	$h^c = 5$; d(C) = 128
Max Pooling	Max^2	13	128	0		$h^m = 35$
GRU (Bidir)	GRU^1	13	256	198,144	$f W; U; B g^{f u;r;h g}$	128
GRU (Bidir)	GRU^2	13	128	123,648	$f W; U; B g^{f u;r;h g}$	64
Linear	L	1		1665	$w^l; b^l$	

Total parameters: 6,616,081

Parameters excluding embedding: 615,681

Bank supply shocks appear to be more persistent geographically than shadow supply shocks, even as they change significantly in magnitude over time. This is consistent with the fact that compared to shadow firms banks hold a higher percentage of originated loans in their portfolio, which can lead to poor performance of past loans influencing the origination of new ones. The weaker geographic persistence of the shadow supply shocks can reflect the higher versatility of these firms.

3.3.2 Validity of the shocks

To support the validity of shocks, I examine the correlations between the shocks and variables that are plausibly linked to changes in credit supply or demand expansions. The results are presented in Table 3.4. Columns (1) to (3) are related to firm supply shocks, and Columns (4) and (5) – to county demand shocks.

The crisis exposure variable used in the regression represented in Column (1) is computed following Bord, Ivashina, and Taliaferro (2017). The paper argues that bank holding companies most exposed to the decline in housing prices during the crisis constrained their lending in all counties, including the ones not affected by the crisis. Following their methodology, I first determine which counties belong to the bottom quartile on the measure of declining house prices between 2007 and 2006. Then I use the Summary of Deposits data to determine which bank holding companies belong to the top quartile of exposure to the counties most affected by the crisis. My firm level crisis exposure variable is a dummy equal to one for the banks belonging to bank holding companies in the top quartile. The regression is run using the loan growth data of 2007–2008. Statistically significant negative correlation between crisis

Table 3.4: Validity of supply and demand shocks. The regression results presented in this table show that the shocks obtained using supply-demand decomposition are correlated with variables that can reasonably affect firm-level supply and county level demand. Columns (1) to (3) correspond to regressions with firm supply shocks on the left-hand side, and Columns (4) and (5) correspond to regression with county demand shocks on the left-hand side. Crisis exposure variable in Column (1) is calculated following Bord et al. (2017) and is a dummy that is equal to one for firms most exposed to counties with biggest declines in house prices. Following Bord et al. (2017), the regression in Column (1) is for 2008 only. Columns (2) and (3) include Troubled Asset Relief Program (TARP) recipient dummy and TARP amount as the right-hand side variables. Top 250 TARP recipients are considered. County-level household number growth and household income growth are the U.S. Census Bureau estimates.

	Firm supply shock			County demand shock	
	(1)	(2)	(3)	(4)	(5)
Crisis exposure	-0.231^{***}				
	(0.083)				
TARP recipient		0.272^{***}			
		(0.067)			
TARP amount			0.016^{***}		
			(0.004)		
Household number growth				0.0004^{***}	
				(0.0002)	
Household income growth					0.044^{***}
					(0.005)
Constant	0.468^{***}	0.219^{***}	0.219^{***}	-0.006^{***}	0.004^{**}
	(0.070)	(0.011)	(0.011)	(0.002)	(0.001)
Observations	2,678	16,384	16,384	8,574	27,319
Adjusted R^2	0.003	0.001	0.001	0.001	0.003
Note:					*p<0.1; **p<0.05; ***p<0.01

exposure variable and the lender supply shocks suggests that my decomposition correctly captures exposure to the housing crisis as a negative lender supply shock.

To provide further evidence about the validity of lender supply shocks, I examine the correlation between my shocks and participation in the Troubled Asset Relief Program (TARP). TARP recipient variable is a dummy equal to one for banks that were among top 250 of TARP recipients by dollar value. TARP amount variable is the logarithm of the dollar amount of relief received by top 250 recipients. The regression uses 2009–2010 data. The positive correlation between the relief and the lender supply shocks suggests that my shock variable correctly captures the credit expansion following participation in the TARP.

The last two columns present regression of the county demand shocks on the county-level household number growth and household income growth. The positive correlations suggest that my measure of demand shocks correctly captures the growing need to acquire housing credit, and higher ability to pay for it.

3.4 Bank credit supply and shadow lending growth

I analyze the relationship between bank mortgage supply and shadow lending growth using the following equation:

$$\begin{aligned} \text{shadow\_lending\_growth}_{ct} = \\ \beta_1 \cdot \text{bank\_supply}_{ct} + \beta_2 \cdot \text{shadow\_supply}_{ct} + \beta_3 \cdot \text{sub\_supply}_{ct} + \\ \beta_4 \cdot \text{demand}_{ct} + \text{interaction\_terms}_{ct} \cdot \Gamma + \text{year}_t + \text{county}_c, \end{aligned} \tag{3.5}$$

where c and t are the county and year indices, shadow_lending_growth$_{ct}$ – shadow lending growth from period $t-1$ to period t, demand$_{ct}$ – county demand shock, bank_supply$_{ct}$, shadow_supply$_{ct}$ and sub_supply$_{ct}$ – county specific supply shocks associated with the corresponding lender types. The interaction_terms$_{ct}$ vector includes interactions between county demand shocks and the supply shocks of each lender type, and year$_t$ and county$_c$ are fixed effects.

The sign of the coefficient associated with bank supply shock can help us characterize the relationship between bank and shadow mortgage lending. A negative coefficient would imply substitution. We would observe it, for example, if shadow lenders specifically targeted areas most exposed to decreases in bank supply, and avoided areas where bank supply is increasing. A negative coefficient would be the most consistent with shadow lenders engaging in regulatory arbitrage. In contrast, a positive coefficient would imply complementarity between bank and shadow lending. We can expect to see a positive coefficient if shadow mortgage firms strongly rely on bank warehouse lending for short-term funding, and if the bank warehouse lending is positively correlated with bank mortgage lending.

The model also includes an interaction between demand shock and the bank supply shock. A significant coefficient corresponding to this interaction would indicate that the relationship between bank and shadow mortgage lending can change in the presence of demand pressure. For example, the borrowers in a county dominated by banks might be more likely to engage with shadow lenders if the demand is high and bank supply is low at the same time.

Table 3.5 presents the results of regressions with gradually increasing complexity of specifi-

cations. The observations are at the county-year level. Column (1) presents the results of univariate regression with shadow lending growth on the left-hand side and county specific bank supply on the right-hand side. The negative coefficient means that unconditionally shadow mortgage lending grew faster in areas exposed to decreases in bank credit supply. However, the sign of the coefficient flips when demand shocks and year fixed effects are included, as shown in Column (2).[3] The flipping sign suggests that shadow lending grew faster in years when bank supply declined across the board and that decreases in bank supply were stronger in areas with positive demand pressure. Once demand and time trend are controlled for, the positive bank supply coefficient shows that shadow lending grows faster in areas exposed to increases in bank supply. Column (3) shows that inclusion of other lender type supply variables does not significantly change the bank supply coefficient.

The complete specification in Column (4) includes interactions between lender type specific supply variables and county demand. Bank supply shocks enter the equation with a positive coefficient, but the interaction between bank supply and county demand – with a negative one. This suggests that bank and shadow mortgages are complements on average, but can become substitutes in the presence of a positive demand pressure.

Table 3.6 presents the main result with standardized variables. The biggest point estimate of coefficients corresponds to demand shocks – a one standard deviation change in demand corresponds to 0.49 standard deviation change in shadow lending growth. The coefficients for shadow supply (0.25) and bank supply (0.16) follow. The relatively large positive coefficient

[3]Including a median shock leads to the same results as including a year fixed effects, which reflects the fact that these variables are constructed in a very similar way.

Table 3.5: Shadow lending growth and supply-demand decomposition. This table demostrates the relationships between the elements of the supply-demand decomposition and shadow lending growth. For every year, county level percentage loan growth is decomposed into *median shock* that affects every firm-county pair equally, county *demand shock* that affects all firms present in a county proportionally to their last year's market shares, county specific *bank supply shock* that is the sum of all bank firm level supply shocks weighted by the firms' last year's market shares, county specific *shadow supply shock* that is the sum of all shadow firm level supply shocks weighted by their last year's market shares, county specific *subsidiary supply shock* that is the sum of all subsidiary firm level supply shocks weighted by their last year's market shares. The table also includes interaction terms between lender type supply and demand shocks. Errors are clustered at county level. Regressions include county and year fixed effects.

	Shadow lending growth			
	(1)	(2)	(3)	(4)
Demand shock		0.767***	0.819***	0.826***
		(0.023)	(0.022)	(0.025)
Bank supply shock	−0.583***	0.534***	0.522***	0.373***
	(0.016)	(0.029)	(0.028)	(0.031)
Shadow supply shock			2.339***	2.520***
			(0.072)	(0.073)
Sub. supply shock			−0.077	−0.128
			(0.082)	(0.083)
Bank supply × demand shock				−0.304***
				(0.030)
Shadow supply × demand shock				1.297***
				(0.243)
Sub. supply × demand shock				−0.736**
				(0.289)
FE	No	Yes	Yes	Yes
Observations	25,816	25,816	25,816	25,816
Adjusted R^2	0.092	0.436	0.478	0.488
Note:				*p<0.1; **p<0.05; ***p<0.01

of bank supply suggests that the complementarity relationship between bank and shadow mortgages can be quite strong. At the same time, the interaction of demand shock and bank supply shock is smaller in magnitude (-0.061), suggesting that substitution effects are more limited in size. This result confirms

Table 3.7 presents the analysis of variable importance that follows Lindeman, Merenda, and Gold (1980). The procedure computes the relative contributions of variables to the explained variance of the model. I consider a regression with shadow lending growth on the left-hand side, and median, demand and lender type specific supply shocks, as well as the interaction of demand and lender type specific supply shocks on the right-hand side. The results show that median shock accounts for approximately 49% of explained variance, and shadow supply shock – for 21%. Demand and bank supply are also relatively important in explaining the variance, contributing about 10% each. The interaction terms are less important, but demand – bank supply is the most important of them, accounting for approximately 5% of explained variance.

The results in this section suggest that shadow lenders' own supply shocks are more important for explaining the variation in shadow lending growth than bank supply shocks. This is consistent with a view that shadow lenders' operational advantages over banks are an important driver of their growth. At the same time, the importance of complementarity effect between bank and shadow lending supports the view that different lender types should be viewed not as opposing but reinforcing each other, for example through warehouse funding ties.

Table 3.6: Shadow lending growth and supply-demand decomposition with normalized variables. This table demostrates the relationships between the elements of the supply-demand decomposition and shadow lending growth. For every year, county level percentage loan growth is decomposed into *median shock* that affects every firm-county pair equally, county *demand shock* that affects all firms present in a county proportionally to their last year's market shares, county specific *bank supply shock* that is the sum of all bank firm level supply shocks weighted by the firms' last year's market shares, county specific *shadow supply shock* that is the sum of all shadow firm level supply shocks weighted by their last year's market shares, county specific *subsidiary supply shock* that is the sum of all subsidiary firm level supply shocks weighted by their last year's market shares. The table also includes interaction terms between lender type supply and demand shocks. All variables are normalized by subtracting the mean and dividing by standard deviation. Errors are clustered at county level. Regressions include county and year fixed effects.

	Shadow lending growth				
	(1)	(2)	(3)	(4)	(5)
Demand shock			0.439^{***}	0.468^{***}	0.485^{***}
			(0.011)	(0.011)	(0.011)
Bank supply shock	-0.018^{**}	-0.029^{***}	0.161^{***}	0.150^{***}	0.113^{***}
	(0.007)	(0.009)	(0.010)	(0.009)	(0.010)
Shadow supply shock				0.231^{***}	0.243^{***}
				(0.007)	(0.007)
Sub. supply shock				0.012^{*}	0.013^{*}
				(0.007)	(0.007)
Bank supply × demand shock					-0.063^{***}
					(0.006)
Shadow supply × demand shock					0.038^{***}
					(0.006)
Sub. supply × demand shock					-0.010
					(0.008)
FE	No	Yes	Yes	Yes	Yes
Observations	25,816	25,816	25,816	25,816	25,816
Adjusted R^2	0.0003	-0.034	0.121	0.172	0.183

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.7: Relative importance of different shock variables in explaining county level shadow lending growth. The metric is based on Lindeman et al. (1980). The numbers represent relative contributions of variables to the model's R^2 normalized to sum to one.

Shock	Relative importance
Median	0.497
Demand	0.104
Bank supply	0.111
Shadow supply	0.206
Sub. supply	0.027
Bank supply × demand	0.050
Shadow supply × demand	0.002
Sub supply × demand	0.002

The high explanatory power of shadow lenders' own supply shocks helps us explain the recent success of shadow originators. First, it can be due to the net effect of variables defined at the institution-year level. For example, shadow lenders can possess better origination technology than the banks, or have a lower cost of funding. Second, it is possible that shadow lenders can better respond to aggregate shocks, for example, by quickly adjusting prices after the interest rate movements. Third, the shadow lenders may be able to tailor their supply to the counties they operate in better than the banks, for example, by offering better loans to specific demographics.

The complementarity result informs us about the relationship between lenders in the mortgage origination markets. The regression of shadow lending growth on bank credit supply could have had three possible outcomes corresponding to kinds of market structure. A zero coefficient would suggest that shadow lenders act independently of banks. We would expect this in a market where banks and shadow lenders are competitors, but demand is not saturated. A negative coefficient would suggest substitution effects resulting from shadow lenders increasing

the supply in areas exposed to weak banks. In contrast to that, the positive coefficient that we observe suggests positive supply spillovers between different lenders.

The positive spillover can be explained by the warehouse funding ties between traditional banks and shadow lenders. Warehouse funding allows banks to participate in mortgage origination market without underwriting the mortgages or holding them on the balance sheet. The incentives for banks include lower compliance costs and increased ability to originate risky assets. It is possible that warehouse funding of shadow mortgage companies represents regulatory arbitrage by the banks with the help of shadow lenders, as opposed to regulatory arbitrage by the shadow lenders at the expense of the banks.

The results above make a twofold contribution to the literature. First, they suggest that factors related to shadow lenders themselves, like technology (Buchak et al. 2017) or cost of funding (Ganduri 2016) contributed more to their rise compared to exploitation of decreasing bank supply via regulatory arbitrage (Buchak et al. 2017). Second, they suggest that despite the presence of a large number of lenders with varying business models, the mortgage origination market involves positive spillovers between lenders rather than marked competitiveness (Stanton, Walden, and Wallace 2014). A potential explanation for the interconnectedness, warehouse lending, is an important topic for future research.

3.4.1 More granular lender types

The definition of bank employed in my main analysis covers all depository institutions. Bundling all banks in one category can mask heterogeneities between financial institutions

of very different sizes and business models. To explore the heterogeneity, I repeat the main analysis using more granular definitions of lender types. Using the HMDA Lender File provided by Robert Avery and FDIC Call Reports, I assign all mortgage lenders into one of the following categories: top 5 banks[4], community banks[5], mortgage banks[6], middle-tier banks[7], thrifts[8], credit unions, independent shadow firms[9], shadow firm affiliates of traditional banks.[10]

To test whether my results are robust to using more granular lender types, I run a regression with a specification similar to equation 3.5, that uses lender type classification described above. The results are presented in Table 3.8. The results are qualitatively similar to the main analysis in Table 3.5. The coefficients are positive for the top 5 banks, mid-tier and community banks. The coefficient for mortgage banks, thifts and credit unions is insignificant. The insignificant coefficients are consistent with warehouse lending story, since these lender types are less likely (or unable) to provide funding to shadow mortgage companies.

3.5 Shadow lending growth and bank credit supply with split samples

In this section, I explore whether the results are driven by the proportionality assumption. The supply-demand decomposition requires that supply shocks vary at the firm-year level

[4]By total assets in 2015
[5]Using Call Reports classification.
[6]Using Call Reports classification, banks that have more than 50% of their total assets in mortgage loans.
[7]Banks that belong to neither of the previous three categories
[8]Excluding the ones that are classified as mortgage banks.
[9]Independent mortgage lenders in HMDA Lender File.
[10]Affiliated mortgage lenders in HMDA Lender File.

Table 3.8: Shadow lending growth and supply-demand decomposition with granular lender types. This table demostrates the relationships between the elements of the supply-demand decomposition and shadow lending growth. For every year, county level percentage loan growth is decomposed into *median shock* that affects every firm-county pair equally, county *demand shock* that affects all firms present in a county proportionally to their last year's market shares, and county level lender type specific supply shocks that represent the sum of the supply shocks of financial institutions belonging to a specific lender type weighted by their last year's market shares. The table also includes interaction terms between lender type specific supply and demand shocks. Column (2) presents results of a regression where all variables are normalized by subtracting the mean and dividing by standard deviation. Errors are clustered at county level. Regressions include county and year fixed effects.

	Independent shadow lending growth	
	Raw variables	Normalized variables
	(1)	(2)
Demand shock	1.126*** (0.141)	0.371*** (0.026)
Top bank supply shock	0.783*** (0.132)	0.023** (0.011)
Mid-tier bank supply shock	0.866*** (0.117)	0.086*** (0.011)
Community bank supply shock	0.927*** (0.138)	0.050*** (0.014)
Mortgage bank supply shock	0.957*** (0.278)	0.009 (0.007)
Thrift supply shock	0.186 (0.696)	0.010 (0.009)
Credit union supply shock	0.227 (0.188)	−0.005 (0.009)
Independent shadow supply shock	1.628*** (0.175)	0.171*** (0.010)
Affiliated shadow supply shock	−0.673 (1.784)	−1.263 (1.263)
Top bank supply × demand shock	−0.057*** (0.021)	−0.025*** (0.005)
Mid-tier bank supply × demand shock	0.012* (0.007)	−0.003 (0.002)
Community bank supply × demand shock	−0.006*** (0.002)	−0.013*** (0.002)
Mortgage bank supply × demand shock	−0.016 (0.027)	0.015*** (0.006)
Thrift supply × demand shock	−0.703 (0.504)	−0.012*** (0.004)
Credit union supply × demand shock	−0.129 (0.111)	−0.008*** (0.003)
Independent shadow supply × demand shock	0.042* (0.022)	−0.001 (0.002)
Affiliated shadow supply × demand shock	−8.896 (8.712)	0.518 (3.600)
FE	Yes	Yes
Observations	28,068	28,068
Adjusted R^2	0.105	0.086
Note:		*p<0.1; **p<0.05; ***p<0.01

while demand shocks vary at the county-year level. However, it is possible that some shocks to lender credit supply can affect lending to different kinds of consumers differently. For, example, loans to lower-income borrowers might attract more regulatory scrutiny, incentivizing the lender to focus on higher income borrowers. Since income levels vary across counties, a shock to lending to lower-income borrowers will affect counties differently.

To account for some of the possible heterogeneities, I allow the shocks to vary at the firm–product–year and demand–product–year level. Specifically, I separate the lending to low and high-income borrowers, and FHA and non–FHA lending. To do that, I split every firm-county-year pair in two, separating firm lending into different products. Each firm–product and county–product pair gets its own identifier, and the shocks are calculated at the firm–product–year and the county–product–year levels.

This procedure allows every firm to have a separate supply shock for each product, and each county to have a separate demand shock for each product, relaxing the proportionality assumption. This allows to capture, for example, the effects of regulation that affect firm's lending to borrowers from some demographics, but not others.

I use two regression specifications. First one is similar to equation 3.5:

$$\begin{aligned} &\text{shadow\_lending\_growth}_{cpt} = \\ &\beta_1 \cdot \text{bank\_supply}_{cpt} + \beta_2 \cdot \text{shadow\_supply}_{cpt} + \beta_3 \cdot \text{sub\_supply}_{cpt} + \\ &\beta_4 \cdot \text{demand}_{cpt} + \text{interaction\_terms}_{cpt} \cdot \Gamma + \text{year}_t + \text{county-product}_{cp}. \end{aligned} \tag{3.6}$$

The only difference between this and main regression is the level at which the variables are defined. Instead of county-year level loan growth and shocks, I use county-year-product variables, where the product is either loans to lower/higher income borrowers, or FHA/non-FHA loans. This specification allows to estimate coefficients that characterize the relationship between shocks and shadow lending growth on average.

The second specification involves estimating the coefficients for the subsamples separately:

$$\text{shadow\_lending\_growth}_{cpt} = \tag{3.7}$$

$$\sum_{p\in 1,2} \beta_{1p} \cdot \text{bank\_supply}_{cpt} \cdot \text{product}_p + \sum_{p\in 1,2} \beta_{2p} \cdot \text{shadow\_supply}_{cpt} \cdot \text{product}_p +$$

$$\sum_{p\in 1,2} \beta_{3p} \cdot \text{sub\_supply}_{cpt} \cdot \text{product}_p + \sum_{p\in 1,2} \beta_4 \cdot \text{demand}_{cpt} \cdot \text{product}_p +$$

$$\sum_{p\in 1,2} \text{interaction\_terms}_{cpt} \cdot \Gamma \cdot \text{product}_p + \text{year}_t + \text{county-product}_{cp}.$$

Table 3.9 presents results for the sample that is split according to whether the income of the borrower is below or above the local median, and Table 3.10 corresponds to the split of FHA vs. non-FHA loans. Column 1 in both tables corresponds to the specification from equation 3.6, and Column 2 – to the specification in equation 3.7.

The results based on the sample with income split are largely similar to the ones produced using the full sample: shadow and bank lending are complements on average, but become substitutes in the presence of demand shocks. At the same time, the substitution effects differ in magnitude between the subsamples. In the presence of positive demand shocks lower-income consumers substitute more relative to their higher-income counterparts (the

Table 3.9: Shadow lending growth and supply-demand decomposition with separate shocks for lower and higher income borrowers. This table demostrates the relationships between the elements of the supply-demand decomposition and shadow lending growth allowing the shocks to vary at the firm–demographic–year and county–demographic–year level. For every year, county level percentage loan growth is decomposed into *median shock* that affects every firm-county pair equally, county *demand shock* that affects all firms present in a county proportionally to their last year's market shares, county specific *bank supply shock* that is the sum of all bank firm level supply shocks weighted by the firms' last year's market shares, county specific *shadow supply shock* that is the sum of all shadow firm level supply shocks weighted by their last year's market shares, county specific *subsidiary supply shock* that is the sum of all subsidiary firm level supply shocks weighted by their last year's market shares. The table also includes interaction terms between lender type supply and demand shocks. The effects are calculated both for the whole sample and separately for the subsamples of lower and higher income borrowers. Errors are clustered at county level. Regressions include county and year fixed effects.

	Shadow lending growth	
	(1)	(2)
Demand shock	0.816^{***} (0.018)	0.809^{***} (0.018)
Bank supply shock	0.807^{***} (0.019)	0.799^{***} (0.019)
Shadow supply shock	0.895^{***} (0.021)	0.937^{***} (0.021)
Sub. supply shock	0.820^{***} (0.020)	0.813^{***} (0.020)
Bank supply × demand shock	-0.001^{***} (0.0002)	
Shadow supply × demand shock	0.013^{***} (0.001)	
Sub. supply × demand shock	-0.003^{***} (0.001)	
Bank × demand × high income		-0.001^{***} (0.0002)
Shadow × shock × high income		0.008^{***} (0.001)
Sub. × demand × high income		-0.003^{***} (0.001)
Bank × demand × low income		-0.002^{***} (0.001)
Shadow × demand × low income		0.022^{***} (0.001)
Sub. × demand × low income		-0.008^{***} (0.001)
Observations	48,445	48,445
Adjusted R^2	0.327	0.328
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 3.10: Shadow lending growth and supply-demand decomposition with separate shocks for FHA and non-FHA loans. This table demostrates the relationships between the elements of the supply-demand decomposition and shadow lending growth allowing the shocks to vary at the firm–product–year and county–product–year level. For every year, county level percentage loan growth is decomposed into *median shock* that affects every firm-county pair equally, county *demand shock* that affects all firms present in a county proportionally to their last year's market shares, county specific *bank supply shock* that is the sum of all bank firm level supply shocks weighted by the firms' last year's market shares, county specific *shadow supply shock* that is the sum of all shadow firm level supply shocks weighted by their last year's market shares, county specific *subsidiary supply shock* that is the sum of all subsidiary firm level supply shocks weighted by their last year's market shares. The effects are calculated both for the whole sample and separately for the subsamples of FHA and non-FHA loans. Errors are clustered at county level. Regressions include county and year fixed effects.

	Shadow lending growth	
	(1)	(2)
Demand shock	0.759^{***} (0.013)	0.762^{***} (0.013)
Bank supply shock	0.752^{***} (0.013)	0.755^{***} (0.013)
Shadow supply shock	0.764^{***} (0.013)	0.766^{***} (0.013)
Sub. supply shock	0.757^{***} (0.013)	0.760^{***} (0.013)
Bank supply × demand shock	-0.0001^{***} (0.00004)	
Shadow supply × demand shock	0.0001^{**} (0.00003)	
Sub. supply × demand shock	-0.0003^{***} (0.0001)	
Bank × demand × non fha		-0.045^{***} (0.008)
Shadow × shock × non fha		0.101^{***} (0.018)
Sub. × demand × non fha		-0.040^{***} (0.015)
Bank × demand × fha		-0.0001^{**} (0.00004)
Shadow × demand × fha		0.0001^{**} (0.00003)
Sub. × demand × fha		-0.0002^{***} (0.0001)
Observations	41,440	41,440
Adjusted R^2	0.466	0.473
Note:		*p<0.1; **p<0.05; ***p<0.01

Wald test rejects the hypothesis of the coefficients being equal at 1% level). This result suggests that shadow companies focus on capturing the lower income borrowers that become underserved by the banks when they constrain their credit.

The results for the FHA/non-FHA split are similar. The substitution effects are stronger in the non-FHA sample. Since FHA borrowers tend to be lower income, this result together with the results in Table 3.9 means that the substitution is stronger among lower-income borrowers who do not participate in FHA programs. This result suggests that the rise of shadow lending in FHA market is due to specific advantages of shadow firms rather than a substitution of constrained bank lending.

The most important implication of this analysis is that the main result of the paper holds when the shock proportionality assumption is relaxed. This increases the confidence that the conclusion about the importance of complementarity between bank and shadow mortgage lending, potentially due to funding ties, was not driven by the assumption behind the supply-demand decomposition.

3.6 Conclusion

Shadow mortgage companies more than doubled their market share after 2008, accounting for 42% of originations in 2015. Some of the aspects of shadow firms' business model, like reliance on short-term funding and propensity to sell the vast majority of originated mortgages, can have an impact on the overall economy. Previous literature provided insights about the funding and operational advantages of shadow lenders, as well as their ability to engage in

regulatory arbitrage.

The contribution of this paper is twofold. First, I show that shadow lenders' own supply shocks explain 50% more variation in shadow lending shocks than bank supply shocks. This suggests that shadow lenders came to prominence thanks to factors like operational advantages, rather than by exploiting decreases in bank supply, perhaps through regulatory arbitrage.

Second, I show that on average shadow lending grows *faster* in areas exposed to *increases* of the bank credit supply. The complementarity is not surprising given that shadow lenders receive a large part of their short-term funding from traditional banks. These results of this paper suggest that the effects of the warehouse funding can be quite large. Considering different ways in which banks and shadow mortgage firms are connected is an important direction for future research.

References

Aiyar, Shekhar, Charles W. Calomiris, and Tomasz Wieladek. 2014. "Does Macro-Prudential Regulation Leak? Evidence from a UK Policy Experiment." *Journal of Money, Credit and Banking* 46 (s1): 181–214.

Amiti, Mary, and David E Weinstein. 2016. "How much do idiosyncratic bank shocks affect investment? Evidence from matched bank-firm loan Data." *Journal of Political Economy*.

Arras, Leila, Franziska Horn, Grégoire Montavon, Klaus-Robert Müller, and Wojciech Samek. 2017. ""What is relevant in a text document?": An interpretable machine learning approach."

PLOS ONE 12 (8): e0181142.

Arras, Leila, Grégoire Montavon, Klaus-Robert Uller, and Wojciech Samek. 2017. "Explaining recurrent neural network predictions in sentiment analysis." *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 159–68.

Aumann, Robert J., and Lloyd S. Shapley. 1974. *Values of non-atomic games*.

Ball, Ray, and Philip Brown. 1968. "An empirical evaluation of accounting income numbers." *Journal of Accounting Research* 6 (2): 159–78.

Beaver, William H. 1968. "The information content of annual earnings announcements." *Empirical Research in Accounting: Selected Studies* 6: 67–92.

Beaver, William H., Maureen F. McNichols, and Zach Z. Wang. 2018. "The information content of earnings announcements: new insights from intertemporal and cross-sectional behavior." *Review of Accounting Studies* 23 (1): 95–135.

Bhattacharya, Nilabhra, Ervin L. Black, Theodore E. Christensen, and Chad R. Larson. 2003. "Assessing the relative informativeness and permanence of pro forma earnings and GAAP operating earnings." *Journal of Accounting and Economics* 36: 285–319.

Bhattacharya, Nilabhra, Ervin L. Black, Theodore E. Christensen, and Richard D. Mergenthaler. 2007. "Who trades on pro forma earnings information?" *The Accounting Review* 82 (3): 581–619.

Bilbao-Jayo, Aritz, and Aitor Almeida. 2018. "Political discourse classification in social

networks using context sensitive convolutional neural networks." *Proceedings of the Sixth International Workshop on Natural Language Processing for Social Media*, 76–85.

Bord, Vitaly M, Victoria Ivashina, and Ryan D Taliaferro. 2017. "Large banks and the transmission of financial shocks." *Working Paper*.

Bowen, Robert M., Angela K. Davis, and Dawn A. Matsumoto. 2005. "Emphasis on pro forma versus GAAP earnings in quarterly press releases: determinants, SEC intervention, and market reactions." *The Accounting Review* 80 (4): 1011–38.

Bozanic, Zahn, Darren T. Roulstone, and Andrew Van Buskirk. 2018. "Management earnings forecasts and other forward-looking statements." *Journal of Accounting and Economics* 65 (1): 1–20.

Bradshaw, Mark T., and Richard G. Sloan. 2002. "GAAP versus the street: an empirical assessment of two alternative definitions of earnings." *Journal of Accounting Research* 40 (1): 41–66.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru. 2017. "Fintech, regulatory arbitrage, and the rise of shadow banks." *Working Paper*.

Chebonenko, Tatiana, Lifeng Gu, and Dmitriy Muravyev. 2018. "Text sentiment's ability to capture information: evidence from earnings calls." *Working Paper*.

Collins, Daniel W, Edward L Maydew, Ira S Weiss, and Ira S Weiss. 1997. "Changes in the value-relevance of earnings and book values over the past forty years." *Journal of Accounting*

and Economics 24: 67.

Collobert, Ronan, Jason Weston, Leon Bottou, Koray Kavukcuoglu, and Pavel Kuksa. 2011. "Natural language processing (almost) from scratch." *Journal of Machine Learning Research* 12: 2493–2537.

Davis, Angela K., Jeremy M. Piger, and Lisa M. Sedor. 2012. "Beyond the numbers: measuring the information content of earnings press release language." *Contemporary Accounting Research* 29 (3): 845–68.

Doyle, Jeffrey T, and Russell J Lundholm. 2003. "The predictive value of expenses excluded from pro forma earnings." *Review of Accounting Studies* 8: 145–74.

Fairfield, Patricia M, Richard J Sweeney, and Teri Lombardi Yohn. 1996. "Accounting classification and the predictive content of earnings." *The Accounting Review* 71 (3): 337–55.

Francis, Jennifer, Katherine Schipper, and Linda Vincent. 2002a. "Earnings announcements and competing information." *Journal of Accounting and Economics* 33 (3): 313–42.

———. 2002b. "Expanded disclosures and the increased usefulness of earnings announcements." *The Accounting Review* 77 (3): 515–46.

Ganduri, Rohan. 2016. "Repo regret?" *Working Paper*.

Gete, Pedro, and Michael Reher. 2017. "Mortgage supply and housing rents." *Working Paper*.

Gornicka, Lucyna A. 2016. "Banks and shadow banks: competitors or complements?" *Journal of Financial Intermediation* 27: 118–31.

Hoberg, Gerard, and Craig Lewis. 2017. "Do fraudulent firms produce abnormal disclosure?" *Journal of Corporate Finance* 43: 58–85.

Khwaja, Asim Ijaz, and Atif Mian. 2008. "Tracing the impact of bank liquidity shocks: evidence from an emerging market." *American Economic Review* 98 (4): 1413–42.

Kim, Oliver, and Robert E Verrecchia. 1991. "Trading volume and price reactions to public announcements." *Journal of Accounting Research* 29 (2): 302–21.

———. 1994. "Market liquidity and volume around earnings announcements." *Journal of Accounting and Economics* 17: 41–67.

Kim, Yoon. 2014. "Convolutional neural networks for sentence classification." *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1746–51.

Kogan, Shimon, Dimitry Levin, Bryan R Routledge, Jacob S Sagi, and Noah Smith. 2009. "Predicting risk from financial reports with regression." *Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the ACL*, 272–80.

Lee, Chen-Yu, Saining Xie, Patrick Gallagher, Zhengyou Zhang, and Zhuowen Tu. 2015. "Deeply-supervised nets." *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, Proceedings of machine learning research, 38: 562–70.

Li, Feng. 2010. "The information content of forward-looking statements in corporate filings—a naïve bayesian machine learning approach." *Journal of Accounting Research* 48 (5): 1049–1102.

Lindeman, Richard Harold, Peter Francis Merenda, and Ruth Z. Gold. 1980. *Introduction to bivariate and multivariate analysis.*

Lipe, Robert C. 1986. "The information contained in the components of earnings." *Journal of Accounting Research* 24: 37–64.

Loughran, Tim, and Bill McDonald. 2014. "Measuring readability in financial disclosures." *The Journal of Finance* 69 (4): 1643–71.

Lundberg, Scott M, Paul G Allen, and Su-In Lee. 2017. "A unified approach to interpreting model predictions." *Advances in Neural Information Processing Systems 30*, 4765——4774.

Manning, Christopher D, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J Bethard, and David Mcclosky. 2014. "The Stanford CoreNLP natural language processing toolkit." *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations.*

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Distributed representations of words and phrases and their compositionality." *NIPS'13 Proceedings of the 26th International Conference on Neural Information Processing Systems* 2: 3111–9.

Muslu, Volkan, Suresh Radhakrishnan, K R Subramanyam, and Dongkuk Lim. 2015. "Forward-looking MD&A disclosures and the information environment." *Management Science* 61 (5).

Nystrom, Nicholas A., Michael J. Levine, Ralph Z. Roskies, and J. Ray Scott. 2015. "Bridges." *Proceedings of the 2015 XSEDE Conference on Scientific Advancements Enabled by Enhanced Cyberinfrastructure - XSEDE '15*, 1–8.

Pennington, Jeffrey, Richard Socher, and Christopher D. Manning. 2014. "Glove: global vectors for word representation." *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, 1532–43.

Purnanandam, Amiyatosh. 2011. "Originate-to-distribute model and the subprime mortgage crisis." *Review of Financial Studies* 24 (6): 1881–1915.

Routledge, Bryan R, Stefano Sacchetto, and Noah A Smith. 2017. "Predicting merger targets and acquirers from text." *Working Paper*.

Shao, Shuai, Robert Stoumbos, and X Frank Zhang. 2018. "The power of accounting information in explaining stock returns." *Working Paper*.

Shiller, Robert J. 2017. "Narrative economics." *American Economic Review* 107 (44).

Shrikumar, Avanti, Peyton Greenside, and Anshul Kundaje. 2017. "Learning important features through propagating activation differences." *Proceedings of Machine Learning Research* 70: 3145–53.

Stanton, Richard, Johan Walden, and Nancy Wallace. 2014. "The industrial organization of the US residential mortgage market." *Annual Review of Financial Economics* 6 (1). Annual Reviews: 259–88.

Sundararajan, Mukund, Ankur Taly, and Qiqi Yan. 2017. "Axiomatic attribution for deep networks." *ICML'17 Proceedings of the 34th International Conference on Machine Learning* 70: 3319–28.

Tetlock, Paul C. 2007. "Giving content to investor sentiment: the role of media in the stock market." *The Journal of Finance* 62 (3): 1139–68.

———. 2010. "Does public financial news resolve asymmetric information?" *Review of Financial Studies* 23 (9): 3520–57.

Yang, Zichao, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. 2016. "Hierarchical Attention Networks for document classification." *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1480–9.

Zou, Hui, and Trevor Hastie. 2005. "Regularization and variable selection via the elastic net." *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 67 (2): 301–20.