Economic Quarterly—Volume 105, Number 3—Third Quarter 2019—Pages 133–164

Sentiment Analysis of the Fifth District Manufacturing and Service Surveys

Santiago M. Pinto

The Richmond Fed conducts monthly surveys of business conditions in the Fifth Federal Reserve District in order to obtain timely information about economic conditions and to provide context to data obtained from other sources. The survey instruments allow respondents to enter free-form comments. This article employs basic text analytic techniques to quantify the sentiment embodied in those survey comments.

An important portion of the information collected and received by regional Reserve Banks is communicated in an unstructured or textual form. The qualitative data conveyed through surveys, or gathered at roundtable meetings with business firms or Bank directors, are very valuable pieces of information for the Banks. This information is generally used to corroborate and provide context to other sources of data. However, the data also reflect sentiment or attitudes derived from economic conditions, a perspective that constitutes a key determinant of firms' and households' economic decisions as supported by an extensive academic literature.[1]

Quantifying and measuring sentiment is not straightforward. Recent development of text analytic tools, however, could be useful. Dif-

Federal Reserve Bank of Richmond; santiago.pinto@rich.frb.org. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] See, for instance, Bram and Ludvigson (1998), Souleles (2004), Barsky and Sims (2012), among many others. The importance of gauging sentiment has recently been highlighted in a speech by Richmond Fed President Thomas Barkin (see Barkin 2019). Barkin not only describes his view on how confidence affects investment decisions by businesses and consumers' expenditures on big-ticket items, but he claims that these reactions have become a lot more sensitive over time.

ferent applications of text analytic techniques are becoming widespread in government agencies, academia, and the private sector as a way to uncover some of the information hidden in unstructured and textual resources. These techniques are useful not only because they could potentially quantify qualitative data, but also because they could uncover novel information hidden in unstructured and textual resources. The availability of relevant and timely economic indicators at the local level is generally limited. The development of a systematic approach that uses text analytic tools to examine and evaluate the information content of a variety of sources, including media and qualitative surveys, could offer new opportunities to better understand changes in local economic conditions and predict economic sentiment.

Using very simple text analytic tools, this article extracts and analyzes the sentiment expressed in comments provided by participants in two surveys conducted by the Richmond Fed: the Fifth District Survey of Manufacturing Activity and the Fifth District Survey of Service Sector Activity. Specifically, the article first develops a set of sentiment indicators that intend to capture the "emotions" reflected in the open-ended comments. The indicators are intended to track three categories of sentiments: negative, positive, and uncertain. Second, to evaluate the information content of the indicators, the article contrasts the sentiment measures against responses to other questions included in the surveys. This kind of exercise is meaningful because these other questions are supposed to specifically inquire about monthly changes in business conditions experienced by survey participants. Third, the article examines the evolution of the sentiment indicators over time and compares their behavior to an indicator of economic activity reported by the Richmond Fed, the manufacturing composite diffusion indices (DIs). Fourth, the article also shows that this methodology can be employed to identify the extent to which responses by individual survey participants show a systematic pattern. For instance, based on the sentiment implicit in the their written comments, this approach can identify those respondents who are systematically positive, negative, or uncertain.

This approach, of course, has its limitations. As with any other method, it is subject to bias and misinterpretation, and the results should always be contrasted against other methods and data. However, the analysis of qualitative data may help enhance the predictive accuracy and corroborate the information provided by other more traditional sources.

The article is organized as follows. Section 2 briefly reviews the text analytic methodology and its application, focusing on sentiment analysis. Section 3 applies these techniques to the survey comments of the Richmond Fed Surveys and discusses the main findings. Finally, Section 4 summarizes the conclusions of the analysis and highlights other potential applications of the present approach.

1. QUALITATIVE DATA COLLECTED BY THE FEDERAL RESERVE BANK OF RICHMOND

All twelve Reserve Banks have regional economics departments that collect, analyze, and publish regional and national data. These data are both quantitative (such as the unemployment rate, employment growth rate, housing prices, etc.) and qualitative, from conversations with representatives from different sectors of the local economy and from surveys. The most visible use of the analysis is to give the president of each Reserve Bank a summary of regional economic conditions, information that is later shared at the Federal Open Market Committee (FOMC) meetings and made available to policymakers, consumers, and businesses. The information collected and disseminated in this way constitutes an additional instrument to evaluate economic conditions: it not only provides context for data obtained from other sources, it is useful to confirm developing trends and understand their effect on the broader economy.[2]

As part of these efforts, the Richmond Fed conducts several monthly surveys that collect qualitative information on business activity. The two largest ones in terms of number of participants are the Fifth District Survey of Manufacturing Activity (the "Manufacturing Survey") and the Fifth District Survey of Service Sector Activity (the "Service Survey"). In order to identify the factors that drive current and expected business conditions in real time, the surveys ask participants a number of questions concerning changes in various measures of activity. Most of the questions in the surveys are qualitative in nature, since respondents are only required to report whether they experienced an increase, decrease, or no change in each economic variable from the preceding month or if they expect to observe similar changes six months ahead.[3]

For example, participants in the Manufacturing Survey are asked, among other questions, whether employment, orders, or shipments decreased, did not change, or increased from the previous month and

[2] The article by Macheras et al. (2015) explains in more detail how and why regional economic conditions may help policymakers understand economic changes observed at the macro level.

[3] The use of the term "qualitative" is common in the literature to refer to directional changes rather than quantitative changes in a specific variable. The term "qualitative" is also used in the present article to refer to textual data.

how they expect those variables to change in the next six months. The Service Survey includes questions that overlap with those asked in the Manufacturing Survey (such as changes in employment, wages, and local economic conditions), in addition to a few other specific questions (such as changes in revenue and product demand). The qualitative information collected through these surveys is later aggregated and combined into several DIs.[4] For the Manufacturing Survey, the Richmond Fed also reports a composite DI defined as the weighted sum of three individual DIs: employment, shipments, and orders.[5]

The survey also allows participants to provide feedback through open-ended textual comments. The comments are not only valuable because they offer information about emerging topics and trends, but they also indicate the respondents' perceptions or sentiment regarding the surrounding economic environment during a given time period.

The present analysis uses basic text mining techniques to closely examine the survey comments submitted by the surveys' participants during the period April 2002 to December 2018.[6] The analysis intends to evaluate the sentiment implicit in those comments, examine how sentiment changes over time, and evaluate the connection between sentiment and participant responses to the other questions included in the survey.

2. TEXT ANALYTICS

What can text analytics do and how does it work?

Text analytics has several different uses and applications. For instance, it can be used to find hidden connections, patterns, and models in plain language narratives or unstructured data. It might be useful to detect emerging areas of concern or interest in specific target groups. Alternatively, it could be used to find trending themes by identifying topic

---

[4] DIs are used and reported by various agencies and organizations, such as the BLS, the Institute of Supply Management (ISM), and the University of Michigan Surveys of Consumers. The diffusion index calculated by the Richmond Fed is simply the difference between the proportion of those that report an increase and those that report a decrease. For additional background information on the structure and information content of the surveys and DIs, see Price and Watson (2014), Waddell (2015), Pinto et al. (2015), and Lazaryan and Pinto (2017).

[5] The panel is unbalanced. The subset of respondents may change from one period to the next. Approximately 45 percent of 200 contacts respond to the Manufacturing Survey in a typical month. The numbers are similar for the Service Survey. Also, panel members may drop out of the survey or they may be removed because they have not responded for an extended period of time.

[6] The input, or corpus, to be analyzed is the entire database of comments from these two surveys.

areas that are either novel or are growing in importance, or it could be used to consistently track concepts that are generally difficult to quantify (such as risk or uncertainty). Specific text analytic tools include text clustering (the classification and grouping of documents according to similarity measures), content categorization (assignment of text documents into predefined categories and building models), concept extraction, entity extraction (identifying named text features, such as people, organizations, places, etc.), entity relation modeling (learning relations between named entities), text summarization, and sentiment analysis.

Text analytic methods generally involve four steps. The first step consists of selecting the input or sources to be analyzed, usually referred to as "corpus." The input can be any textual data, such as open-ended questions in surveys, a collection of documents, or transcribed minutes from a meeting. The second is a preprocessing step that involves the implementation of several methods and techniques to simplify the data. The process includes the extraction and identification of individual words (usually referred to as "tokenization" of the textual document), word stemming and lemmatization, the recognition of names, entities, places, and dates, and the removal of common or "stop" words that do not provide any meaning to the text (e.g.: "the," "at," "in," and "with"). Stemming is a process through which words are reduced to their roots or stems. For example, the words "fox" and "foxes" may be reduced to the root "fox." Lemmatization also tries to group words, but the process is somewhat more complicated because it attempts to associate words according to their meanings. For example, the lemma of the words "paying," "paid," and "pay" is "pay." The objective of both stemming and lemmatization is to match and group words in order to reduce the size of the data and, consequently, reduce processing time and memory. The third step is the analysis. At this stage, the goal is to extract features from the documents, define a model based on those features, and train the model with a subsample of the data. Lastly, the fourth step consists of the validation of the results from the analysis. The validation is both internal, i.e., using available data not employed to construct the model, and external, i.e., using other available data sources and methods.

Sentiment analysis

Sentiment analysis uses the tools of text analytics to measure and classify the emotional content of unstructured textual data. This classification has typically been used to analyze opinions and product ratings, to inform political strategy, and used in research methods to quantify qualitative data.[7] The goal of this approach is essentially to map a piece of text to a specific sentiment category, such as positive, negative, or uncertain. Different techniques are generally employed to construct this mapping. Some of them are based on predefined dictionaries (the lexical or "bag of words" approach), while others rely on machine learning algorithms. See, for example, Hansen et al. (2018). They all, however, share the general principles.

The lexical or "bag of words" approach assigns textual data to each sentiment category using a predefined dictionary or list of words typically associated with those categories. Sentiment is then determined by the frequency of words in each category found in the text. However, relying exclusively on these kinds of dictionaries may lead to errors and misinterpretations. In general, the task of classifying text according to its sentiment is a lot more complicated because the meaning of a word may depend on the context and the specific combination of words found in an expression. For instance, if a word that reflects a positive sentiment is combined with a word that has a negative connotation, then the overall sentiment becomes negative. Other factors, such as sarcasm or slang, may complicate even more the analysis based on dictionaries. The approach followed later in the article (explained in Section 3) extends the "bag of words" approach by incorporating short expressions associated with different tonalities and by implementing general linguistic rules to deal with some of the problems described above.[8]

[7] Recent work in economics and finance has used text analytic tools to develop various indicators of economic activity. See, among others, Nyman et al. (2018), Thorsrud (2018), and Calomiris and Mamaysky (2019).

[8] Provalis Research, vendor for QDA Miner and WordStat text analytic software, provides a general sentiment dictionary in a website download. The WordStat Sentiment Dictionary was created by combining negative and positive words from three dictionaries: the Harvard IV TagNeg dictionary of negative words, the Martindale Regressive Imagery dictionary, and the Pennebaker Linguistic and Word Count dictionary. The dictionary building utility program in WordStat was then used to expand the word list, generating over 9,500 negative and nearly 4,700 positive word patterns. The word lists themselves do not measure sentiment; rather, sentiment is determined by applying two linguistic rules. Negative sentiment is measured by "negative words not preceded by a negation (no, not, never) within four words in the same sentence" and "positive words preceded by a negation within four words in the same sentence." Positive sentiment can be measured similarly but is not as predictive. Improving the accuracy of a sentiment dictionary requires additional "training" of the generic dictionary to customize for a particular domain or body of content. For additional information, see "Sentiment Dictionaries" (Provalis Research) at https://www.provalisresearch.com/products/content-analysis-software/wordstat-dictionary/sentiment-dictionaries (accessed November 1, 2018). Two examples of dictionaries customized for specific domains include the Loughran and McDonald financial sentiment dictionary (for more information, see Loughran and McDonald [2011] and Loughran and McDonald [2015]) and the Lexicoder Sentiment Dictionary (see Young and Soroka [2012]) for the analysis of political news. The developers of the two dictionaries took different approaches toward achieving a greater accuracy of sentiment analysis.

The exercise developed in the present article is closely related to the work by Shapiro et al. (2018), in which the authors examine sentiment embodied in the news media. They use text analytic techniques to construct sentiment indices intended to capture the opinions expressed in economic and financial newspaper articles and to determine the writer's attitude toward certain issues. To develop their indices, the paper uses a proprietary machine learning predictive model developed by a company called Kanjoya.[9] They next analyze the information content of these measures by examining their correlation with different indicators of business economic conditions and their predictive accuracy. They find not only a strong contemporaneous correlation between sentiment and key business cycle variables, but also that sentiment helps in forecasting inflation and the federal funds rate.

3. SENTIMENT ANALYSIS OF THE FIFTH DISTRICT SURVEYS OF ECONOMIC ACTIVITY

The main objective of the exercise is to construct different measures that capture the sentiment and opinions embodied in the open-ended comments offered by survey participants and to examine how sentiment changes over time. To assess the information content of these measures, I compare them to the participants' responses to other questions in the survey that are supposed to track monthly changes in economic activity.[10]

Preliminary analysis: views of participants who write comments

Before proceeding with the textual analysis, and to understand the limitations and scope of the methodology described in the next section, it should be noted that not all respondents choose to write comments. In fact, during the period under consideration, on average, 26 percent of survey participants in the Service Survey and 30 percent in the Manufacturing Survey offer written comments. To draw meaningful conclusions from the textual analysis of comments, it is important to understand the behavior of participants who take the time and effort

[9] See Shapiro et al. (2018) for a thorough description of their methodology.

[10] The present analysis should simply be regarded as an exercise that shows the potential use of text mining techniques. Applying these techniques would probably make more sense when dealing with large bodies of text rather than with the surveys mentioned above, since they they only target a limited number of participants. However, even for small samples, it is still valuable to develop a methodology, using some of these techniques, that systematically and consistently examines the qualitative data collected by the Richmond Fed.

to offer such information. Specifically, is the group of participants who write comments biased in a particular direction or can this subset of participants be regarded as a representative subsample?

One way of examining the differential behavior across groups is by determining the extent to which writing comments covaries with responses to the other questions included in the surveys. To do this, I compare the behavior of the two groups by evaluating how they respond to the question on changes in current employment.[11] Figure 1 shows the monthly difference between the employment DIs calculated using responses from each group of survey participants (i.e., [$DI_{\text{no comments}} - DI_{\text{comments}}$]) along with its HP-filtered trend (solid line). The values reported in the figure combine responses from the two surveys: Service and Manufacturing. The series do not seem to indicate systematically different behavior between the two groups until approximately October 2014. While until October 2014 the difference between DIs indicated that those who write comments assessed economic conditions more negatively than those who do not write comments (i.e., the employment DI calculated using responses from the group of participants who write comments is lower than the employment DI calculated using responses from the group who don't write comments), the difference has become negative from that time period onward.

In order to examine the extent to which this kind of behavior differs across the service and manufacturing sectors, I perform the same exercise using data from each survey separately. The results are plotted in Figure 3. The figure shows periods in which the series move together (from April 2007 until April 2012) and periods in which they behave differently (from the beginning of the sample until April 2007, and from April 2012 until the end of the sample). In those periods when the series do not coincide, the survey participants who write comments in the Service Survey tend to be relatively less optimistic about economic conditions than those who write comments in the Manufacturing Survey. However, beginning in September 2017, the pattern has changed: those who write comments in the Service Survey become increasingly pessimistic, while survey participants in the Manufacturing Survey tend to show the opposite behavior.

Overall, the latter exercises suggest that the conclusions obtained from the sentiment analysis performed on survey comments should be interpreted with caution. Specifically, the conclusions could be biased because the analysis relies on information provided by a subsample of survey participants, whose incentives to report written comments might

[11] Only the analysis that considers the employment question is reported here. Similar conclusions can be drawn by comparing responses to other survey questions.

be effectively driven by their own perceptions of economic conditions (as indicated by their responses to other questions in the survey).

Sentiment analysis: methodology

The first step of the analysis is to preprocess the textual survey data following the steps described in Section 2.[12] Next, I construct different sentiment indicators by extending the lexical or "bag-of-words" methodology discussed previously. The approach involves the following steps. First, I define the set of sentiment categories $I = \{negative, positive, uncertain\}$, where $i \in I$ is a representative element of this set. Second, I analyze the text and detect the list of words that belong to each of the categories based on a predefined dictionary.[13] Third, in addition to identifying such words, I categorize text according to the use of different short expressions that commonly reflect certain types of emotions.[14]

Fourth, I define several linguistic rules that take into account the context of words to assess sentiment. The idea is that sentiment is not simply determined by the frequency of words present in the dictionary. For instance, positive words are assumed to reflect positive sentiment if their meanings are not modified by the presence of other words. Specifically, positive sentiment is captured by positive words "not near" a negation (such as "no," "not," and "never") or "not near" a negative word and by negative words "near" a negation or "near" another negative word.[15] Negative sentiment can be measured using similar rules. In this way, negative sentiment would be described by the presence of negative words not near negations or other negative words and by positive words near negations or negative words. According to this approach, comments like "*We* ***do not*** *think it is a cause for* ***concern,***" "*We* ***have not*** *had a* ***problem*** *hiring entry level staff,*" or "*Customer traffic is* ***not bad***" would be classified as positive, and comments like "*From September 10 to mid-October, business* ***was not*** *at all* ***good,***"

[12] These steps are common to most every analysis performed on textual data. This stage essentially entails the identification and removal of frequently used words that appear in a content set and do not have sentiment connotations. The removal of words with many occurrences reduces the "noise" in the subsequent sentiment analysis.

[13] The methodology uses the dictionary constructed by Loughran and McDonald (2011) as the starting point. The dictionary is modified and trained for the specific corpus under study.

[14] An explanation of the methodology, including examples from the survey comments, is described in the Appendix (see Section B).

[15] In the present exercise, a word is defined to be "near" another word if they are within five words of each other (before or after), in the same sentence. Some of these rules are variations of those suggested by Provalis Research.

"*The market is* ***not*** *as* ***strong*** *at retail as last fall,*" or "*things are* ***not*** *as* ***good*** *as everyone thinks*" would be categorized as negative (i.e., "not good"). Uncertainty is simply assessed by determining the presence of words or expressions generally associated with this sentiment.

Finally, I analyze the mix of positive, negative, and uncertain "words" and assess the overall sentiment embodied in the text by calculating three types of indicators. To calculate the first sentiment measure, I sum the number of survey comments (or occurrences) assigned to each sentiment category. To do this, I define a case-specific indicator function that is equal to one when a comment from a survey participant (i.e., a case) contains at least one expression that belongs to the previously defined categories (negative, positive, or uncertain), and I then sum over all the indicator functions.[16]

The second measure of sentiment is based on the number of words in each category showing up in the comments. According to this indicator, the sentiment of a comment would depend on the relative frequency of words. Compared to the previous measure, this one reflects more accurately differences in the intensity of each sentiment expressed in the textual data. However, it does not contain information about the spread of the sentiment among respondents. In other words, it could be possible for a few comments to drive the sentiment in a specific time period if those comments include many words associated with the respective categories.

For the third measure, I construct an indicator that also uses the frequency of words in each category but normalized by the total number of words (the values are expressed as a rate every 10,000 words). The results of the analysis are summarized in the following sections.

Results of the analysis

Sentiment and responses to questions on business activity

In this section, I determine the extent to which the sentiment embodied in the written comments is associated with changes in business conditions. To establish this relationship, I compare the sentiment of the comments to the responses offered by survey participants to other questions included in the Fifth District Surveys. As mentioned earlier, these questions ask participants to determine if a specific variable has changed from the previous month (current changes) or is expected to change in the next six months (expected changes). The set of possible

[16] A specific comment may be assigned to more than one category.

responses $J = \{1, 2, 3\}$, where "(1)" is decrease, "(2)" is remain unchanged, and "(3)" is increase, and $j \in J$ is a representative response from set J.

In the first place, I evaluate the sentiment of survey comments in conjunction with the responses to the question on current changes in employment.[17] Figure 2 summarizes the association between sentiment categories and responses to the employment question. The tables on the left (top and bottom) are constructed by counting the cases (or respondents) in each sentiment group i who respond j to changes in employment. The table on the top left reports the column percentages, i.e., the percentage of cases in each category i as a proportion of those who respond j. The table on the bottom left shows the row percentages, i.e., the number of cases that respond $j = 1, 2, 3$, as a proportion of cases in each category i. The tables in the middle show similar percentages constructed using the frequency of words, and the table on the right uses the frequency of words as a proportion of total words.

Consider the top left table. The largest percentage of cases in the negative category is observed when the response is (1) or "decrease," with 59 percent, and then smallest when the response is (3) or "increased," with 47 percent. When the response is (2) or "remain unchanged," the percentage is 51, in between the other two. For the positive category, response (3) has the largest percentage and response (1) has the smallest. For the uncertainty category, the maximum is reached when the response is (2). The table on the bottom left shows that the percentage of those who report (1) is highest for the negative category (19 percent), the percentage of those who report (2) is highest for the uncertainty category (69 percent), and the percentage who report (3) is highest for the positive category (22 percent).

The tables constructed using word frequencies, both tables in the middle and the table on the right, show identical results.[18] Finally, Figure 4 in Appendix A shows the results from a similar analysis performed separately for the Manufacturing and Service Surveys.[19] From that table, it can be concluded that the sentiment indicators accurately reflect the opinion of those participating in the two separate surveys.

In general, the tables suggest that the sentiment indicators based on textual data accurately reflect participants' perceptions about eco-

[17] The question about current changes in employment is common to both the Manufacturing and Service Surveys. The present analysis combines the information from both surveys in order to work with a larger sample size.

[18] Similar conclusions are obtained using expected changes in employment.

[19] Only the tables using frequency of words (rate per 10,000) are reported in Figure 4 in Appendix A. The tables calculated using case occurrences and frequency of words show the same conclusions.

nomic conditions. In other words, the information offered by these indicators seems to be consistent with other information conveyed by survey participants, in this case, the information revealed by their responses to the question that asks about changes in employment.

Finally, I examine the correspondence between the sentiment categories and other questions included in the surveys, such as current changes in: (i) local economic conditions (Figure 5 in Appendix A; the tables are constructed using data from the Manufacturing and Service Surveys since this question is common to both), (ii) shipments (Figure 6; data from the Manufacturing Survey), (iii) orders (Figure 7; data from the Manufacturing Survey), (iv) demand (Figure 8; data from the Service Survey), and (v) revenues (Figure 9; data from the Service Survey). The results confirm the conclusions from the previous analysis that compares sentiment and employment changes and further validate the measures of sentiment introduced earlier.

Changes in sentiment by month

Figures 10, 11, and 12 display the monthly evolution of the measures of sentiment. Figure 10 shows the changes in the negative, positive, and uncertain indicators, calculated as the number of cases or respondents assigned to each sentiment category. The series reported in the graph are simply the percentage of cases in each category. Figure 11 shows the evolution of sentiment indicators that include the frequency of words in each category. The series, as before, are expressed as the percentage of words in each category at each period of time. Finally, Figure 12 shows the frequency of words in each category, normalized by the total number of words in each period (the numbers are expressed as a rate per 10,000 words). All figures include the series' twelve-month moving averages (solid lines).

The following observations are worth pointing out from the graphs. First, the behavior of all the sentiment indicators is similar in all three figures. Moreover, the category representing negative sentiment is relatively more important than the other two categories. However, the value of the information offered by these sentiment indicators is not determined by the level of such measures but by how these measures change in time, reflecting changing views and perceptions about the evolution of the economy.

Second, in all cases, the negative sentiment indicator reaches its maximum (within the sample considered in the analysis) at the end of 2008, and declines thereafter until April 2010. This series reaches a new peak in the second half of 2013 and later steadily declines until August 2017. Since August 2017, negative sentiment has been increasing.

Third, the series that reflect positive sentiment evolve in the opposite way. In fact, the correlations between the negative and positive sentiment indicators are -0.71, -0.89, and -0.53 in Figures 10, 11, and 12, respectively.[20]

Fourth, the indicator of uncertainty shows a somewhat different behavior. The uncertainty measure rises prior to the Great Recession, reaching a peak in the middle of 2007. After a brief decline it rises again reaching another peak at the end of 2012. Since then, the indicator has been declining, except for a short period of time from mid-2016 to approximately September 2017 in which it slightly increased.[21]

Next, I construct a sentiment indicator that aggregates the individual information described above. Specifically, the sentiment indicator is defined as the difference between negative and positive sentiment, i.e., $[Negative - Positive]$. This means that higher values of this indicator would be associated with higher overall negative perceptions and views about the economy. The evolution of this indicator is depicted in Figure 15.[22] A striking feature of this series is that negative sentiment has been steadily increasing since mid 2017, reaching in December 2018 similar levels as those observed during late 2012 and the beginning of 2013.

It is likely that certain factors affect and drive sentiment differently in the manufacturing and service sectors. I therefore evaluate the extent to which the sentiments associated with the comments included in the two surveys, the Manufacturing (M) and Service (S) Surveys, differ. Figures 17, 18, and 19 display the evolution of the sentiment indicators constructed using frequency of words (rate per 10,000 words) for each survey. The correlation between each sentiment indicator across surveys is positive but low (0.06 for negative sentiment, 0.04 for positive sentiment, and 0.17 for the uncertainty category), suggesting there could be factors affecting sentiment in each sector differently.

Finally, I calculate the sentiment indicator introduced earlier ($[Negative - Positive]$ using frequency of words normalized by the total number of words in each period), but only for the Manufacturing Survey, and I compare the evolution of this indicator to the composite DI described in Section 1. The series are plotted in Figure 16. The

[20] The entire correlation matrix is shown in Figure 14 in Appendix A.

[21] An enlarged version of the series showing the behavior of the uncertainty indicator is shown in Figure 13.

[22] It should be considered that, as mentioned earlier, negative words tend to be more preponderant in comments than positive words. Also, changes in the positive and negative sentiment indicators may individually offer valuable information, each one correlated with different set of variables. Future work will evaluate the information content of each one of the sentiment series.

left axis indicates the units of the manufacturing sentiment indicator, and the right axis the units of the composite index.[23] The series, as expected, have a negative correlation (the correlation between the two (smoothed) series is -0.51). However, it is interesting to note that since approximately October 2017 both series have been increasing.[24] This means that during this period both negative sentiment, and favorable business conditions, captured by the level of the composite DI, have been rising. A similar behavior is only briefly observed in 2004, at least during the sample period considered in the present analysis.

Understanding the factors driving the behavior of the series is, of course, crucial in order to make sense of the conveyed information. A complete investigation is relegated for future research. However, by performing a very preliminary analysis, I was able to identify a positive association between stock market volatility and our indicator of negative sentiment.[25] Specifically, the correlation between the Chicago Board Options Exchange Market Volatility Index (VIX) and the negative sentiment indicator (smoothed) series is 0.44 during the sample period considered in the analysis.[26] The series are plotted in Figure 20.[27]

Now, a final comment regarding the extent to which a methodology like the one developed in this paper could help Reserve Banks in their efforts to evaluate economic conditions. The alignment of the information provided by the sentiment indicator with other qualitative measures, such as the composite DI, would help confirm the Banks' view about economic conditions. It should not be interpreted, how-

[23] The range of the DI is $[-100, 100]$.

[24] The correlation between the (smoothed) series is 0.84 during the period October 2017 to December 2018.

[25] Note that, in principle, the series are supposed to capture changes in sentiment and economic conditions in the Fifth District. A thorough analysis would require the identification of regional and national factors associated with the evolution of those variables. The work by Lazaryan and Pinto (2017), for instance, studies the extent to which the composite DI is associated with regional and national economic variables. A similar analysis could be performed using the negative sentiment indicator developed in this paper.

[26] The VIX indicator is constructed using a number of options included in the S&P 500 index and is supposed to capture the stock market's expectation of volatility over the next thirty days. While the correlation between the VIX and composite diffusion index (smoothed) series during the sample period under consideration is -0.68, the correlation has become positive since the beginning of 2017.

[27] In Pinto et al. (forthcoming), we construct a measure of uncertainty and apply the methodology using data from the Survey of Consumers conducted by the University of Michigan. While the correlation between our measure of uncertainty and consumer confidence (measured by the Index of Consumer Sentiment) is generally negative, as expected, they both tend to rise during the period 2009–14. To some extent, such behavior is similar to the one highlighted above when comparing the evolution of the negative sentiment indicator and the composite DI (even though such behavior is observed at different periods).

ever, that when these indicators move in opposite directions (providing perhaps conflicting evidence about the state of the economy) that the methodology is flawed. In fact, these kinds of scenarios could simply reveal the fact that sentiment gives us different information, not captured by other data, and further exploration would be necessary. The sentiment indicator, as a result, is used in this context as a way to corroborate information obtained from other qualitative assessments.

Sentiment and survey respondents

A similar analysis can be carried out to identify respondents who systematically show a negative, positive, or uncertain sentiment. Note that the surveys conducted by the Richmond Fed have a panel structure. A list of contacts, developed throughout the years and representative of the Fifth District industry composition, receives online surveys every month. Using this panel of respondents, the methodology can determine the extent to which some contacts are systematically more pessimistic or optimistic than others. Understanding the systematic behavior of individual participants and identifying those contacts who consistently express a specific sentiment (positive, negative, or uncertain) would provide a much more accurate assessment and interpretation of the monthly responses by correcting any bias in the results due to sample selection. As an illustration, Figures 21, 22, and 23 list contacts, in decreasing order, according to the sentiment generally communicated through their survey comments.[28]

4. CONCLUSIONS

The present article illustrates the use of basic text analytic tools by evaluating the sentiment of survey comments collected by two surveys conducted by the Richmond Fed: the Fifth District Manufacturing and Service Surveys. First, the article constructs several indicators that intend to capture the sentiment embodied in the open-ended comments written by survey participants. Second, in order to evaluate the information content of these indicators, the article contrasts the sentiment measures against responses to other survey questions. This exercise is useful since the other survey questions are meant to specifically track monthly changes in business conditions experienced by survey participants. Finally, the article analyzes the evolution of the sentiment indi-

[28] I have carried out similar sentiment analysis by industry NAICS code and state. However, due to small sample sizes, the conclusions tend to be very imprecise.

cators and compares their behavior to an indicator of economic activity reported by the Richmond Fed, the composite DI.

Sentiment as measured in the paper (defined as the difference between negative and positive sentiment) generally aligns well with other assessment measures of qualitative data, such as the composite DI. However, there are instances in which these measures convey conflicting information. For example, the sentiment indicator and the composite DI have both been increasing since approximately October 2017. Such behavior has only been briefly observed in 2004.

The fact that sentiment might not fully align with other assessments does not necessarily imply that the methodology is flawed. It could simply mean that sentiment is capturing different information. In this way, the sentiment indicator could be used as a tool to corroborate other information collected by the Bank. When sentiment and diffusion indices head in opposite directions, for example, we would be less confident about what the qualitative surveys are telling us, requiring further exploration.

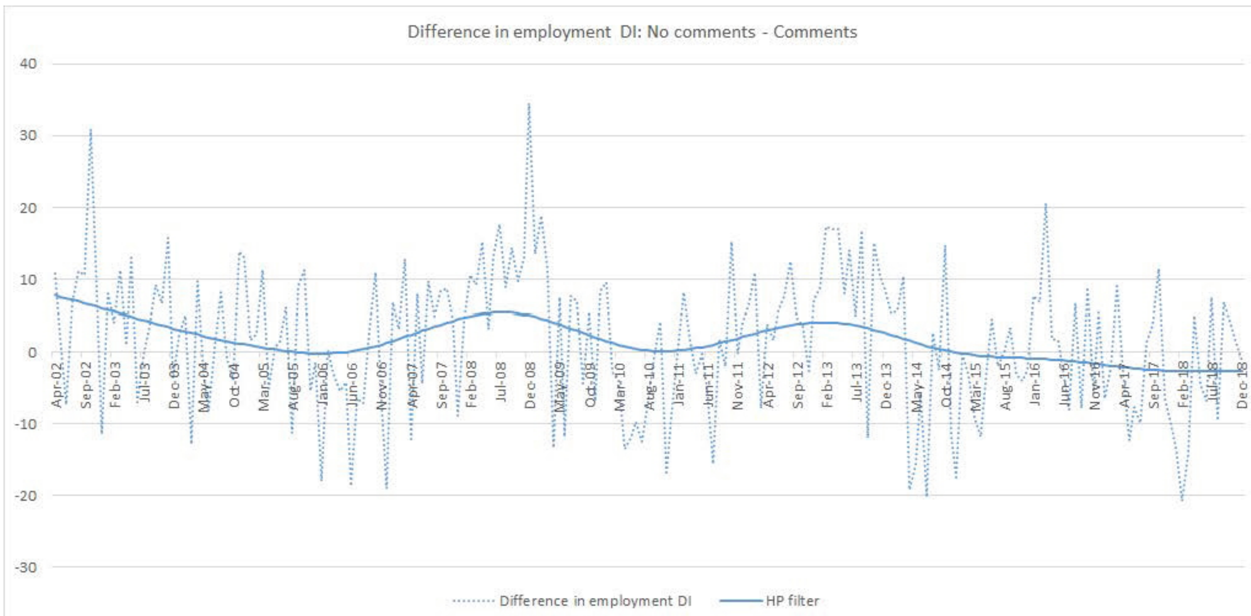
Different factors could potentially play a role in explaining the behavior of sentiment. While a thorough investigation of such determinants is beyond the scope of the present paper, a preliminary analysis allows us to identify a positive correlation between stock market volatility and the negative sentiment indicator.

It should be emphasized that the present exercise is simply a first attempt to evaluate sentiment in survey comments. A more rigorous analysis is definitely required in order to apply this method for other purposes, such as assessing the level of uncertainty in the economy or drawing conclusions about individuals' expectations. However, the preliminary results indicate that this kind of analysis is promising.

There are many other potential applications of text analytics. Some of these applications are meaningful not only to extract information from the surveys conducted by the Richmond Fed, but also to gain insights from the rest of the qualitative data communicated to the Bank. For instance, these tools could be used to uncover recurrent and emerging issues, identify trends, or consistently track the evolution of certain topics (such as "tariffs," "labor market," "inflation," etc.). The use of text mining techniques by regional Reserve Banks is not as widespread as in other sectors of the economy. However, regional Reserve Banks can definitely benefit from these methods both in academic research and policymaking. Unstructured data provide an additional source of information that, jointly with other data collected by the Banks, could offer a more complete description and understanding of the changes taking place in the economy.

APPENDIX: APPENDIX A

Figure 1 Difference in Employment DI: No Comments vs. Comments



Notes: The figure shows the monthly difference between the employment DI calculated for those who don't submit written comments and those who submit written comments ([$\text{DI}_{nocomments}$ - $\text{DI}_{comments}$]). The information used to calculate the DIs includes responses from both the Manufacturing and Service Surveys.

Figure 2 Sentiment and Changes in Current Employment

Case occurrence (column)			
SENTIMENT	1	2	3
NEGATIVE	59%	51%	47%
POSITIVE	30%	34%	41%
UNCERTAINTY	11%	14%	12%

Frequency (column)			
SENTIMENT	1	2	3
NEGATIVE	67%	59%	52%
POSITIVE	26%	32%	40%
UNCERTAINTY	7%	9%	8%

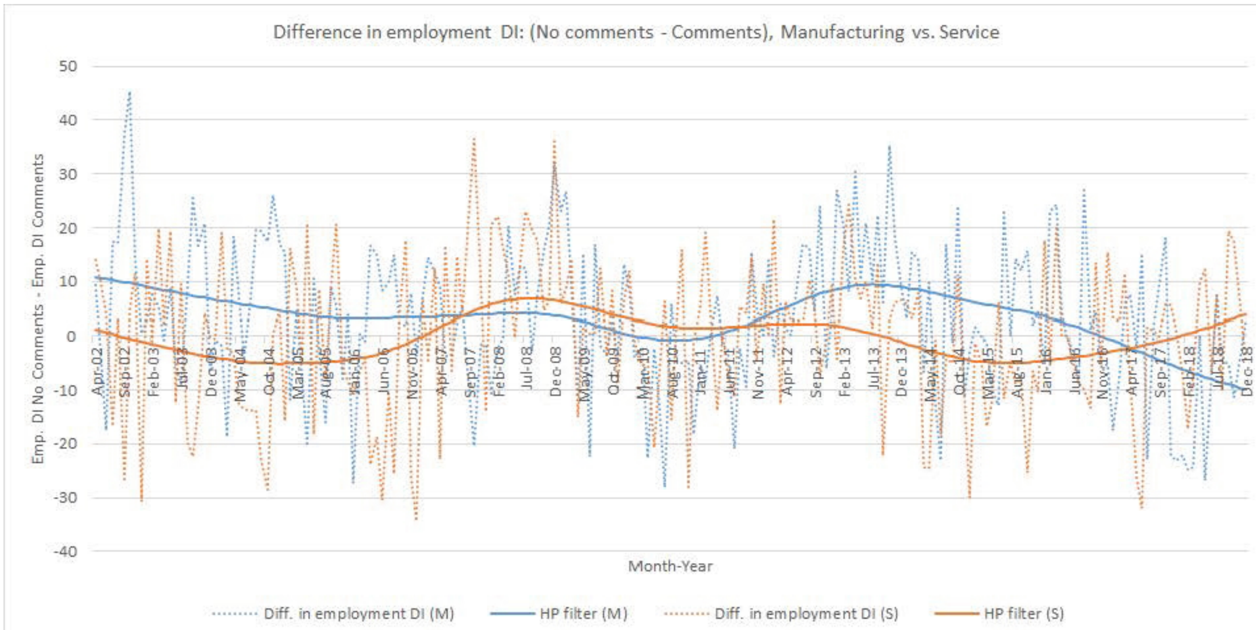
Frequency (rate per 10,000 words)			
SENTIMENT	1	2	3
NEGATIVE	761.01	632.08	537.80
POSITIVE	297.61	342.75	415.66
UNCERTAINTY	81.30	98.45	80.55

Case occurrence (row)			
SENTIMENT	1	2	3
NEGATIVE	19%	64%	17%
POSITIVE	14%	64%	22%
UNCERTAINTY	14%	69%	17%

Frequency (row)			
SENTIMENT	1	2	3
NEGATIVE	19%	65%	16%
POSITIVE	13%	64%	23%
UNCERTAINTY	14%	70%	17%

Notes: (1) decrease, (2) no change, (3) increase. The table is constructed using the combined data from the Manufacturing and Service Surveys.

Figure 3 Difference in Employment DI: No Comments vs. Comments – Manufacturing and Service Surveys



Notes: The figure shows the monthly difference between the employment DI calculated for those who don't submit written comments and those who do submit written comments ([$\text{DI}_{nocomments}$- $\text{DI}_{comments}$]). M: Manufacturing, S: Service.

Figure 4 Sentiment and Changes in Current Employment: Manufacturing and Service Surveys

Frequency (rate per 10,000 words)						
	Service			Manufacturing		
SENTIMENT	**1**	**2**	**3**	**1**	**2**	**3**
NEGATIVE	498.94	420.77	347.33	562.27	491.96	404.08
POSITIVE	275.48	312.78	361.36	239.47	254.05	294.35
UNCERTAINTY	20.57	32.87	32.74	22.62	26.98	24.44

Notes: (1) decrease, (2) no change, (3) increase.

Figure 5 Sentiment and Changes in Current Local Economic Conditions

Case occurrence (column)				Frequency (column)				Frequency (rate per 10,000 words)			
SENTIMENT	**1**	**2**	**3**	**SENTIMENT**	**1**	**2**	**3**	**SENTIMENT**	**1**	**2**	**3**
NEGATIVE	55%	51%	44%	**NEGATIVE**	66%	59%	48%	**NEGATIVE**	727.61	640.56	481.92
POSITIVE	29%	35%	42%	**POSITIVE**	24%	33%	43%	**POSITIVE**	267.82	361.80	437.34
UNCERTAINTY	16%	14%	13%	**UNCERTAINTY**	10%	8%	9%	**UNCERTAINTY**	107.86	89.00	87.28

Case occurrence (row)				Frequency (row)			
SENTIMENT	**1**	**2**	**3**	**SENTIMENT**	**1**	**2**	**3**
NEGATIVE	26%	49%	26%	**NEGATIVE**	28%	49%	24%
POSITIVE	19%	47%	34%	**POSITIVE**	17%	46%	36%
UNCERTAINTY	26%	46%	28%	**UNCERTAINTY**	27%	45%	28%

Notes: (1) decrease, (2) no change, (3) increase. The values are calculated using the combined data from the Manufacturing and Service Surveys.

Figure 6 Sentiment and Changes in Current Shipments

Case occurrence (column)				Frequency (column)				Frequency (rate per 10,000 words)			
SENTIMENT	**1**	**2**	**3**	**SENTIMENT**	**1**	**2**	**3**	**SENTIMENT**	**1**	**2**	**3**
NEGATIVE	56%	55%	48%	**NEGATIVE**	67%	64%	56%	**NEGATIVE**	719.60	724.79	571.44
POSITIVE	30%	32%	39%	**POSITIVE**	25%	28%	36%	**POSITIVE**	267.07	321.24	370.54
UNCERTAINTY	14%	12%	13%	**UNCERTAINTY**	9%	8%	8%	**UNCERTAINTY**	94.65	91.08	82.64

Case occurrence (row)				Frequency (row)			
SENTIMENT	**1**	**2**	**3**	**SENTIMENT**	**1**	**2**	**3**
NEGATIVE	36%	38%	27%	**NEGATIVE**	36%	36%	27%
POSITIVE	30%	35%	34%	**POSITIVE**	29%	34%	37%
UNCERTAINTY	36%	34%	29%	**UNCERTAINTY**	36%	34%	30%

Notes: (1) decrease, (2) no change, (3) increase. The values are calculated using data from the Manufacturing Survey.

Figure 7 Sentiment and Changes in Current Orders

Case occurrence (column)				Frequency (column)				Frequency (rate per 10,000 words)			
SENTIMENT	**1**	**2**	**3**	**SENTIMENT**	**1**	**2**	**3**	**SENTIMENT**	**1**	**2**	**3**
NEGATIVE	56%	56%	48%	**NEGATIVE**	67%	65%	55%	**NEGATIVE**	718.98	743.42	557.75
POSITIVE	30%	32%	39%	**POSITIVE**	24%	28%	37%	**POSITIVE**	261.47	318.02	382.36
UNCERTAINTY	14%	12%	13%	**UNCERTAINTY**	9%	8%	8%	**UNCERTAINTY**	98.02	87.69	82.26

Case occurrence (row)				Frequency (row)			
SENTIMENT	**1**	**2**	**3**	**SENTIMENT**	**1**	**2**	**3**
NEGATIVE	38%	35%	27%	**NEGATIVE**	38%	35%	27%
POSITIVE	32%	33%	36%	**POSITIVE**	29%	31%	39%
UNCERTAINTY	38%	31%	31%	**UNCERTAINTY**	39%	31%	30%

Notes: (1) decrease, (2) no change, (3) increase. The values are calculated using data from the Manufacturing Survey.

Figure 8 Sentiment and Changes in Current Demand

Case occurrence (column)				Frequency (column)				Frequency (rate per 10,000 words)			
SENTIMENT	**1**	**2**	**3**	**SENTIMENT**	**1**	**2**	**3**	**SENTIMENT**	**1**	**2**	**3**
NEGATIVE	54%	50%	44%	**NEGATIVE**	61%	58%	47%	**NEGATIVE**	713.74	629.65	475.69
POSITIVE	32%	36%	43%	**POSITIVE**	30%	33%	45%	**POSITIVE**	348.18	360.65	459.93
UNCERTAINTY	15%	14%	13%	**UNCERTAINTY**	9%	9%	8%	**UNCERTAINTY**	108.42	96.97	85.82

Case occurrence (row)				Frequency (row)			
SENTIMENT	**1**	**2**	**3**	**SENTIMENT**	**1**	**2**	**3**
NEGATIVE	23%	43%	34%	**NEGATIVE**	24%	44%	32%
POSITIVE	18%	40%	43%	**POSITIVE**	17%	37%	45%
UNCERTAINTY	22%	43%	36%	**UNCERTAINTY**	23%	42%	35%

Notes: (1) decrease, (2) no change, (3) increase. The values are calculated using data from the Service Survey.

Figure 9 Sentiment and Changes in Current Revenues

Case occurrence (column)			
SENTIMENT	1	2	3
NEGATIVE	57%	51%	46%
POSITIVE	31%	35%	41%
UNCERTAINTY	12%	14%	13%

Frequency (column)			
SENTIMENT	1	2	3
NEGATIVE	63%	57%	49%
POSITIVE	29%	34%	43%
UNCERTAINTY	8%	10%	9%

Frequency (rate per 10,000 words)			
SENTIMENT	1	2	3
NEGATIVE	704.22	622.60	505.76
POSITIVE	323.86	367.74	444.66
UNCERTAINTY	87.10	105.24	91.65

Case occurrence (row)			
SENTIMENT	1	2	3
NEGATIVE	31%	38%	31%
POSITIVE	24%	36%	40%
UNCERTAINTY	25%	41%	34%

Frequency (row)			
SENTIMENT	1	2	3
NEGATIVE	31%	39%	30%
POSITIVE	23%	36%	41%
UNCERTAINTY	25%	42%	34%

Notes: (1) decrease, (2) no change, (3) increase. The values are calculated using data from the Service Survey.

Figure 10 Sentiment by Month: Case Occurrence

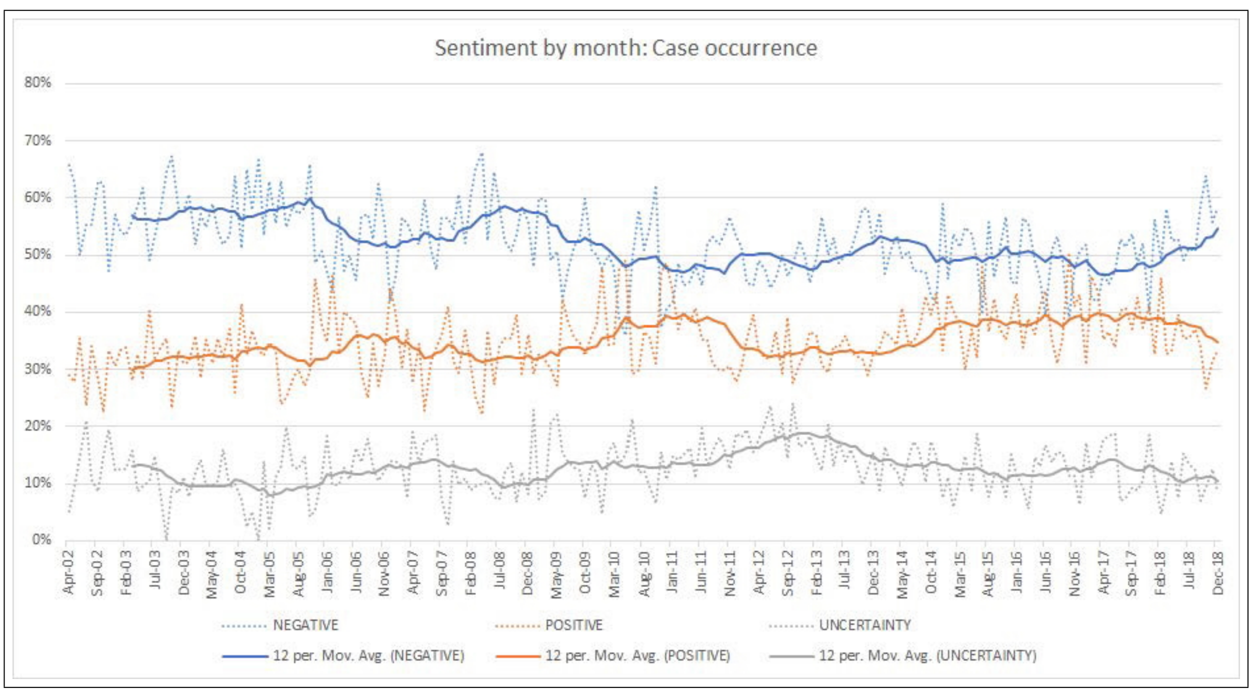


Figure 11 Sentiment by Month: Word Frequency

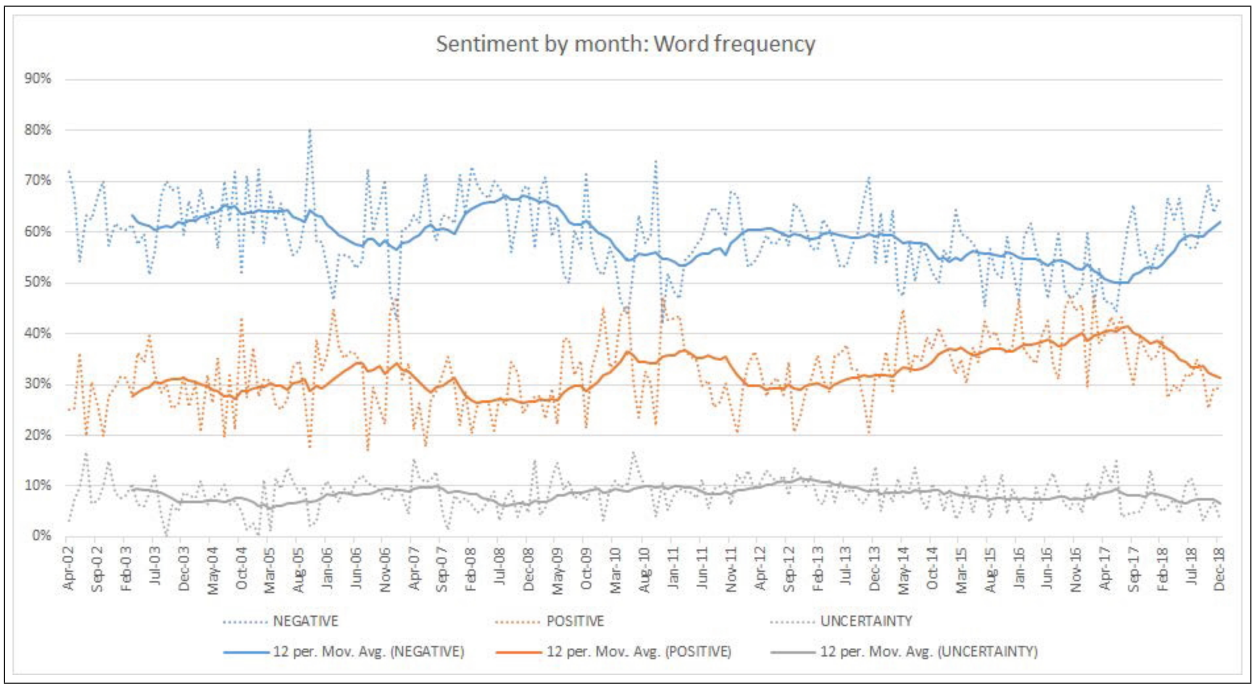
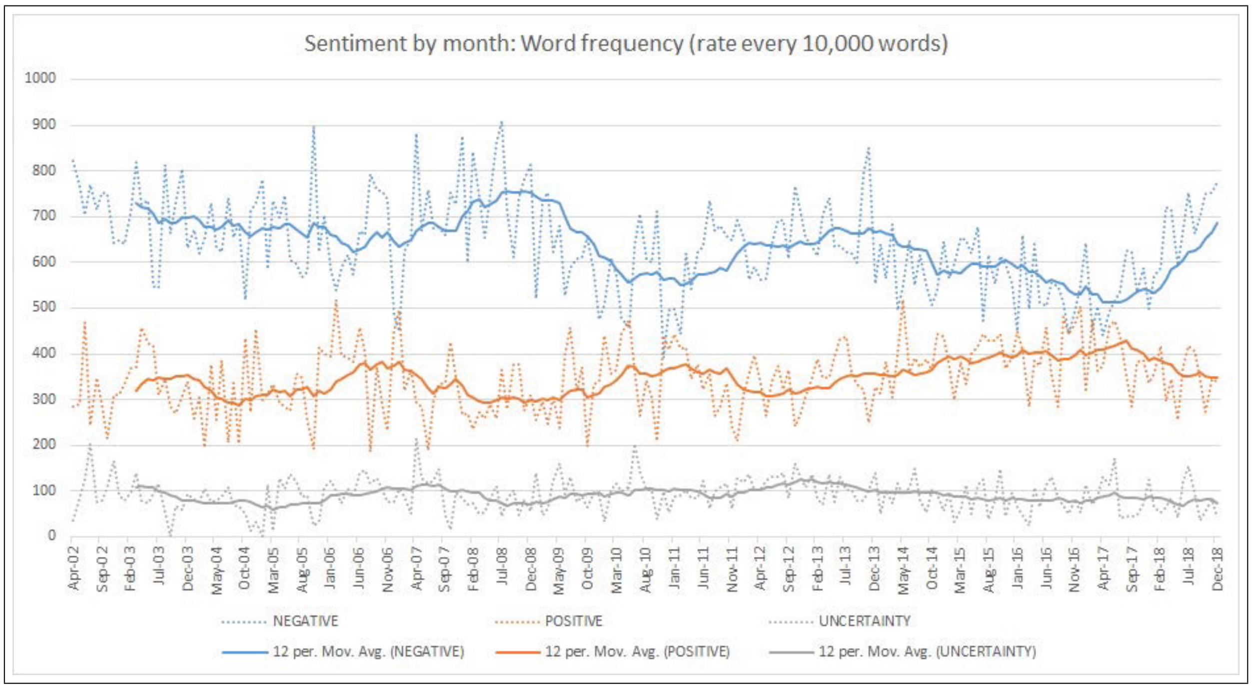
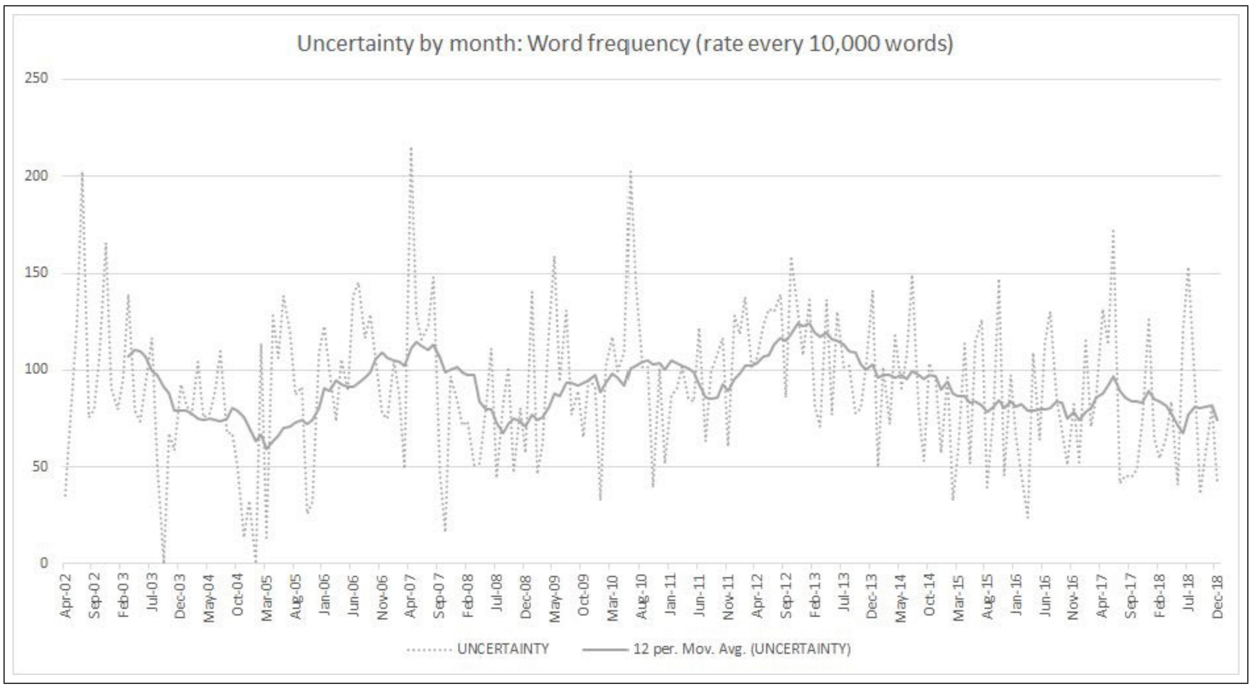


Figure 12 Sentiment by Month: Word Frequency (Rate per 10,000 Words)



Notes: The indicators shown in Figures 10, 11, and 12 are calculated using the combined data from the Manufacturing and Service Surveys.

Figure 13 Uncertainty by Month: Word Frequency (Rate per 10,000 Words)

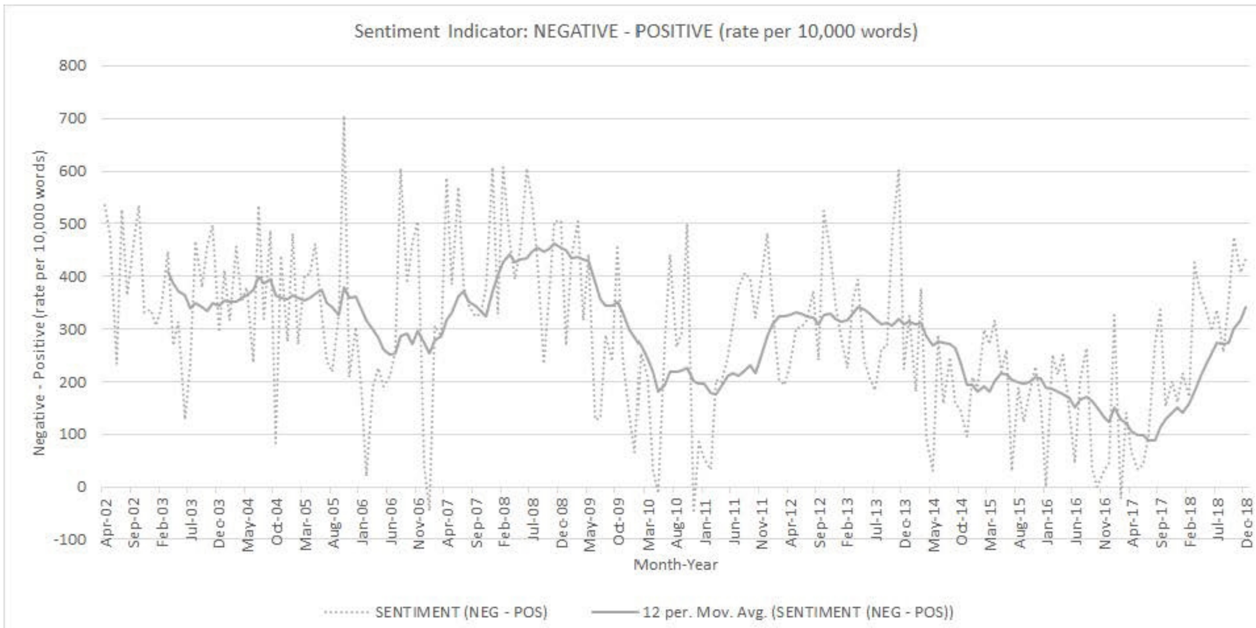


Notes: The indicators are calculated using the combined data from the Manufacturing and Service Surveys.

Figure 14 Correlation between Sentiment Indicators

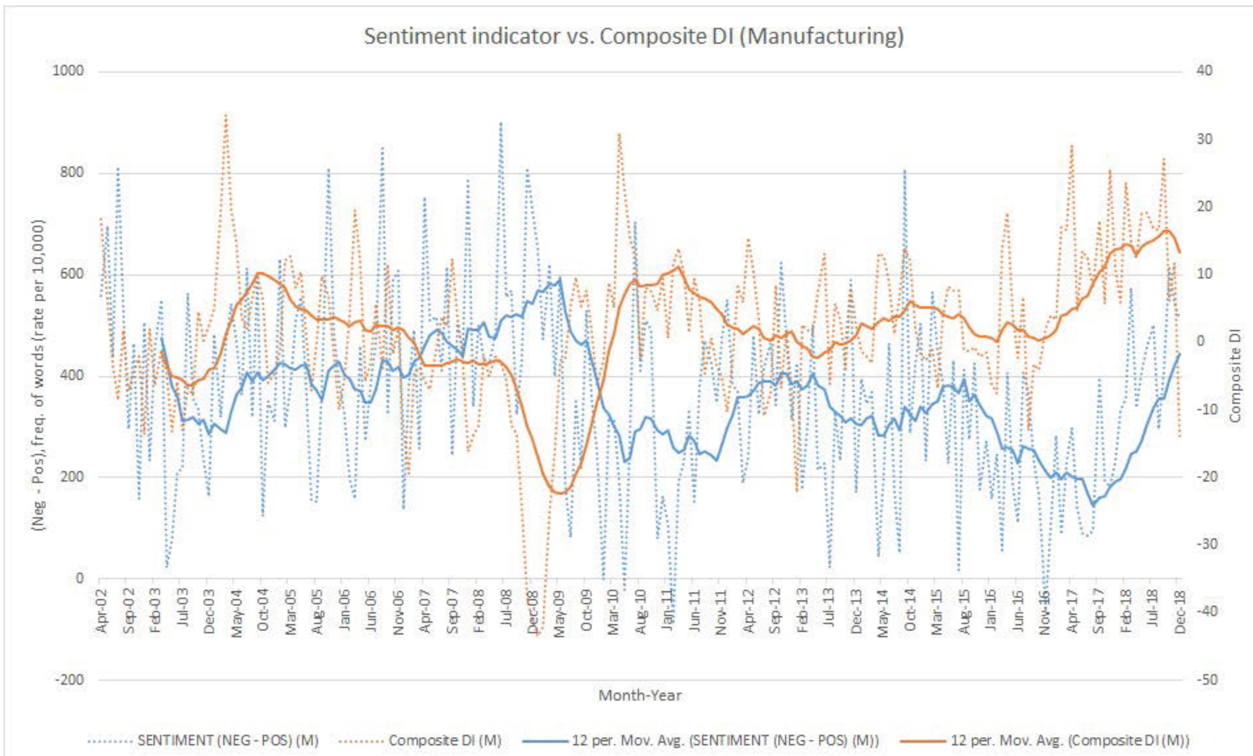
	Case Neg.	Case Pos.	Case Unc.	Freq. Neg.	Freq. Pos.	Freq. Unc.	Freq. rate Neg.	Freq. rate Pos.	Freq. rate Unc.
Case Neg.	1.00								
Case Pos.	-0.71	1.00							
Case Unc.	-0.53	-0.22	1.00						
Freq. Neg.	0.81	-0.65	-0.33	1.00					
Freq. Pos.	-0.65	0.80	-0.07	-0.89	1.00				
Freq. Unc.	-0.43	-0.22	0.87	-0.36	-0.09	1.00			
Freq. rate Neg.	0.70	-0.63	-0.21	0.83	-0.79	-0.21	1.00		
Freq. rate Pos.	-0.57	0.69	-0.04	-0.84	0.92	-0.05	-0.53	1.00	
Freq. rate Unc.	-0.38	-0.26	0.84	-0.31	-0.14	0.97	-0.05	-0.01	1.00

Figure 15 Sentiment Indicator



Notes: The indicators are calculated using the combined data from the Manufacturing and Service Surveys.

Figure 16 Sentiment Indicator vs. Manufacturing Composite DI



Notes: The sentiment indicator and the composite DI are calculated using data from the Manufacturing Survey.

Figure 17 Sentiment by Month: Negative, Manufacturing vs. Service (rate per 10,000 words)

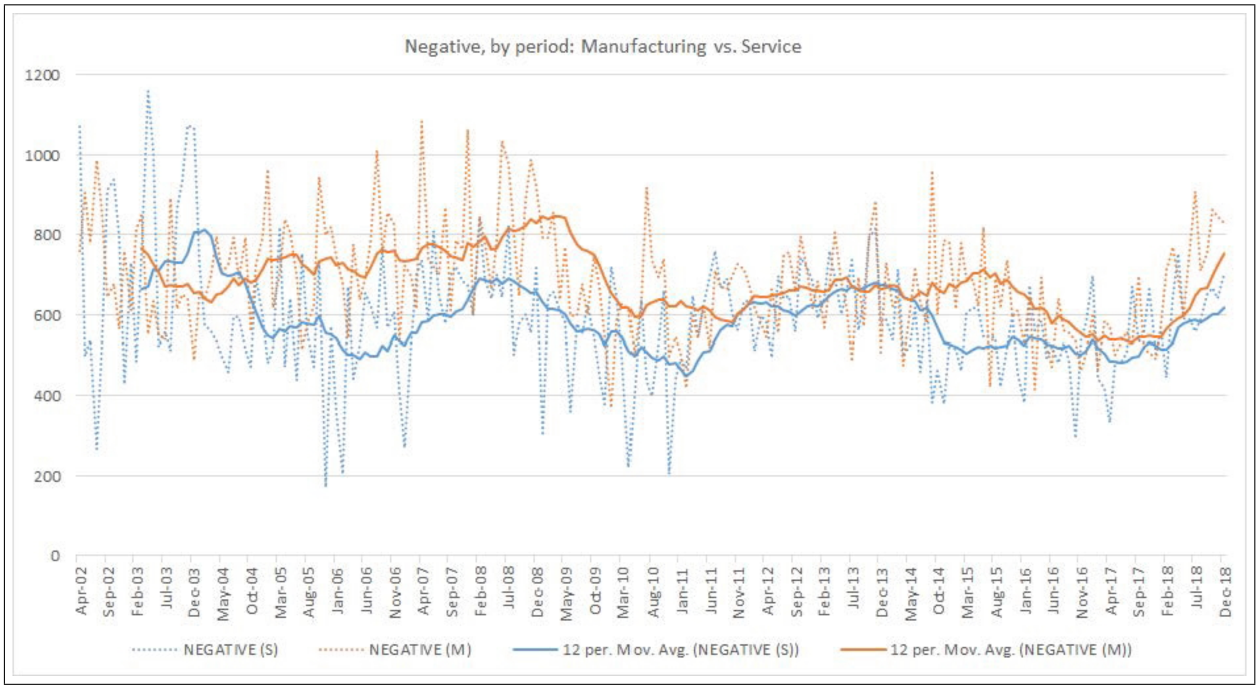


Figure 18 Sentiment by Month: Positive, Manufacturing vs. Service (rate per 10,000 words)

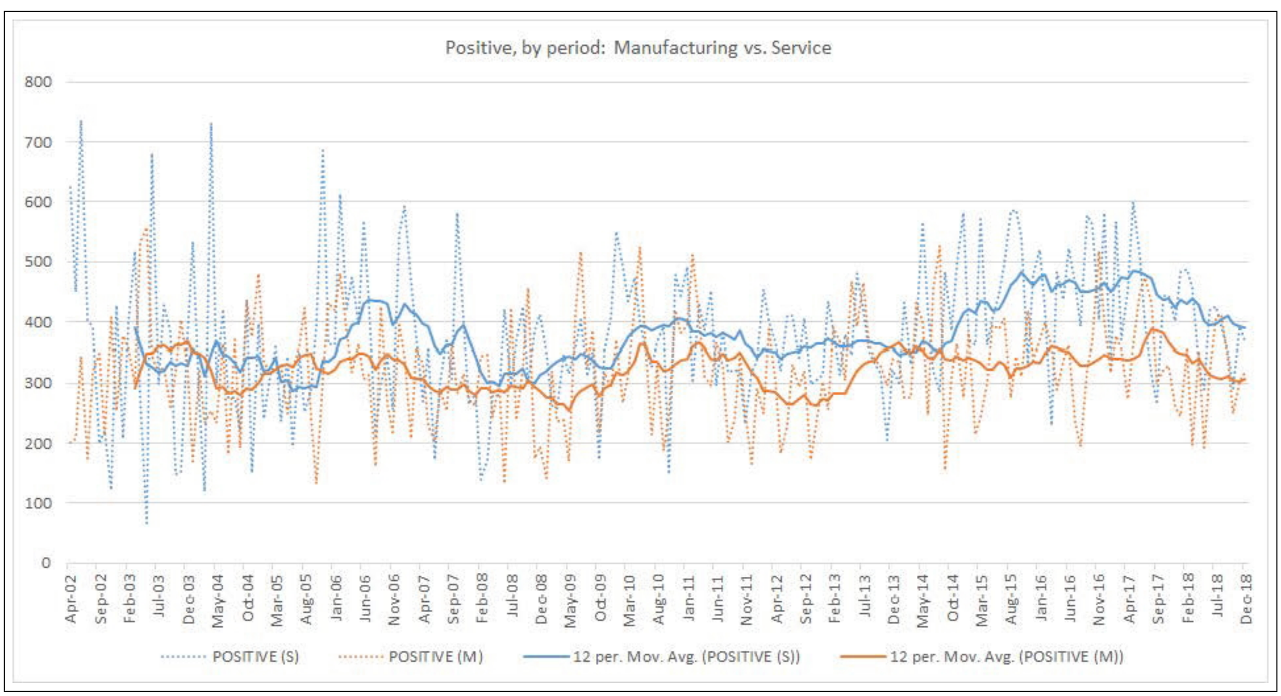


Figure 19 Sentiment by Month: Uncertainty, Manufacturing vs. Service (rate per 10,000 words)

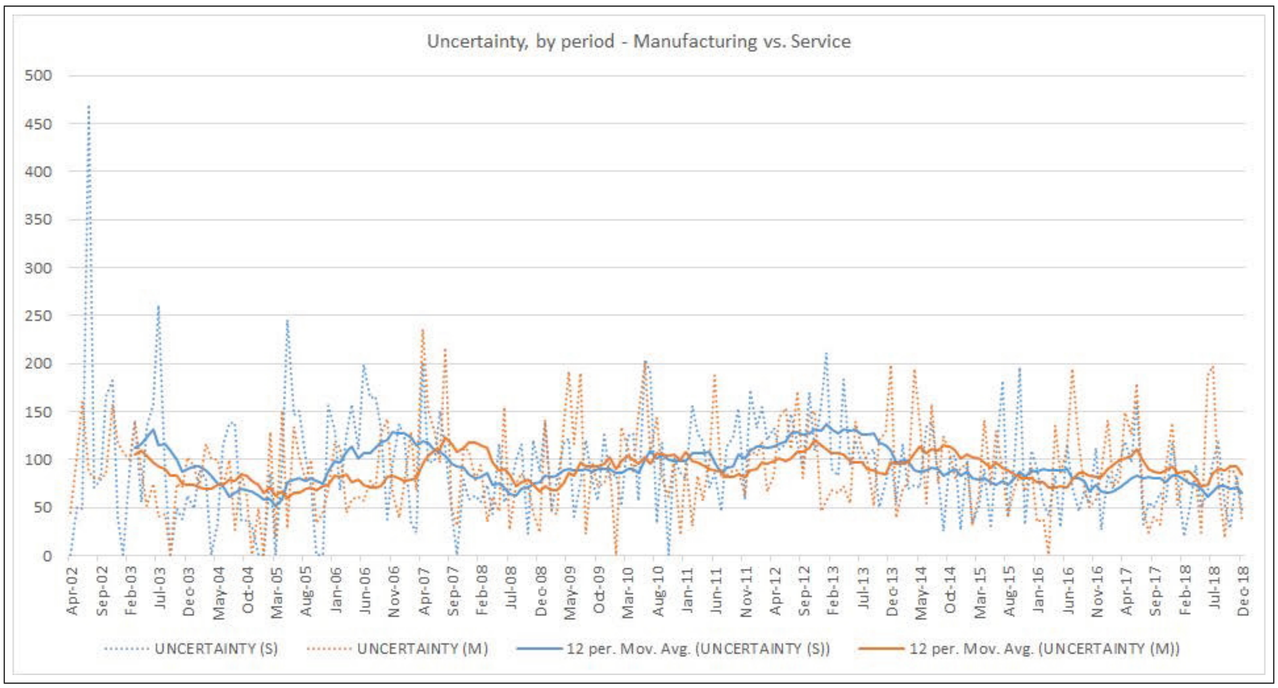
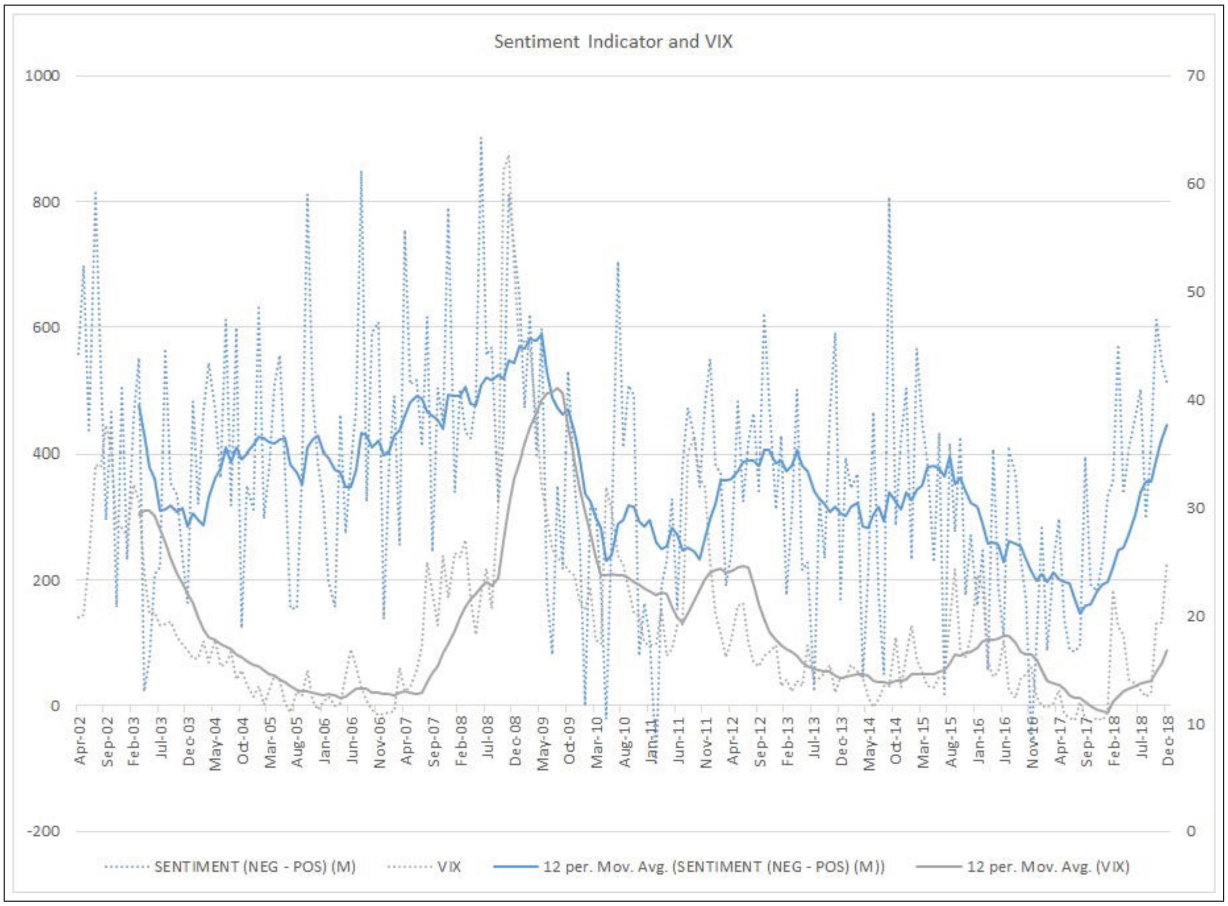


Figure 20 Sentiment Indicator vs. VIX



Notes: Left axis: Sentiment indicator. Right axis: VIX.

Figure 21 Sentiment by ID: Negative

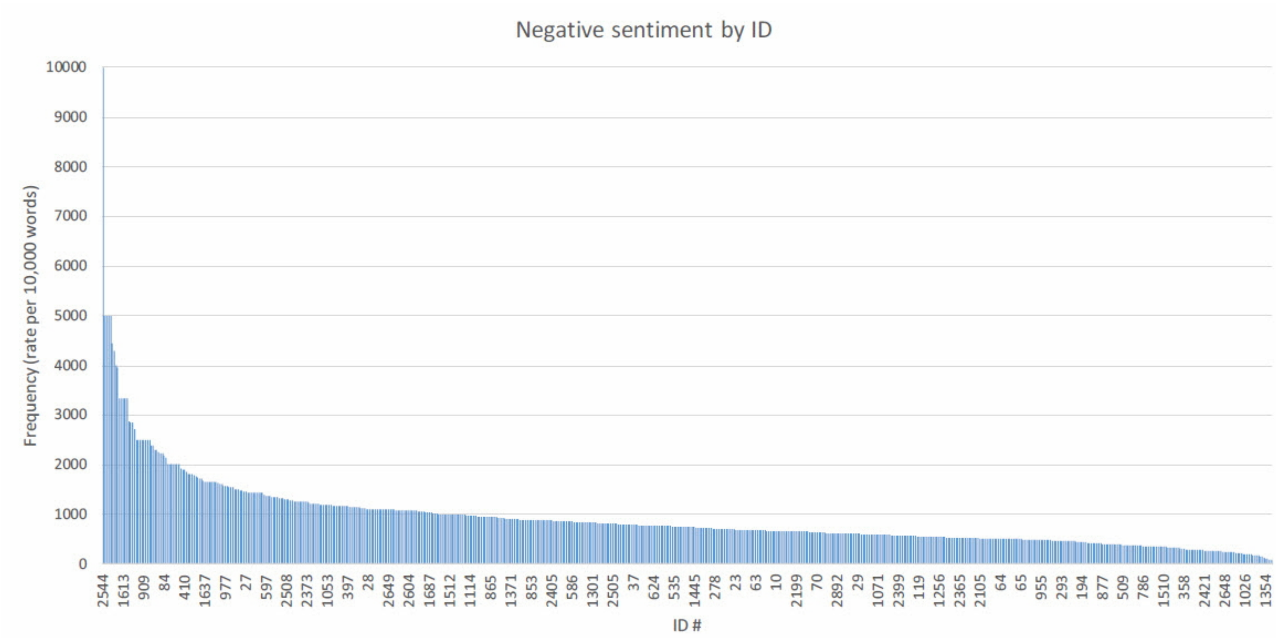


Figure 22 Sentiment by ID: Uncertainty

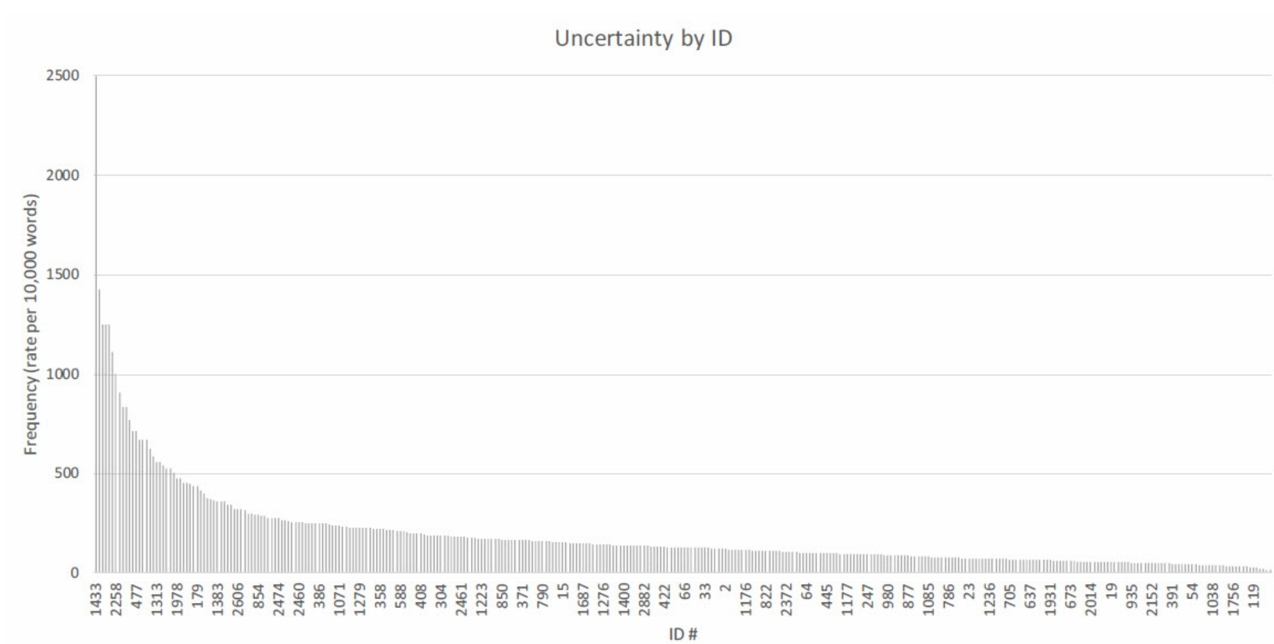
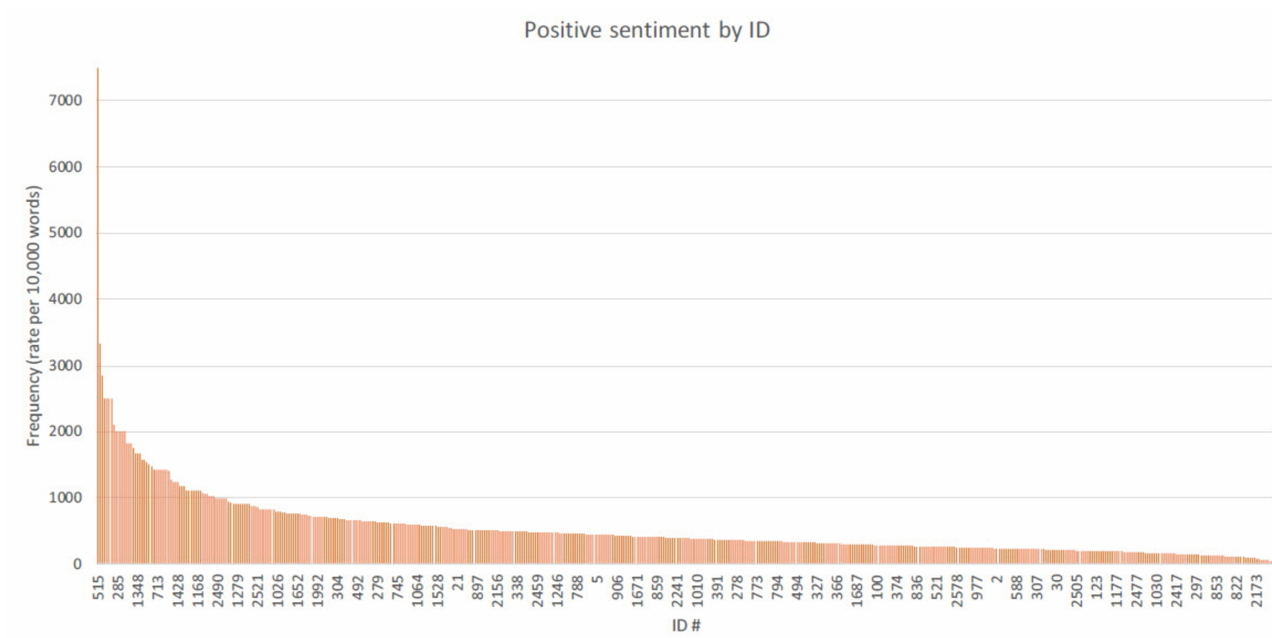


Figure 23 Sentiment by ID: Positive Sentiment



APPENDIX: APPENDIX B

1. METHODOLOGY

The tables in Figures 24 and 25 show a sample of the words included in the dictionary used in the analysis. Note that the words themselves are not directly associated with the respective sentiment category. By using predefined rules, I assume that a sentence expresses a specific sentiment depending on how the words are combined. Figure 26 shows different examples of sentences categorized as "negative," "positive," or "uncertain" using the rules described in the text.

Figure 24 Words

TYPE OF WORD	WORD	FREQUENCY	% TOTAL	NO. CASES	% CASES	TF • ID
DOUBLE NEGATION	A_FEW	70	0.03%	66	0.14%	200.1
NEGATIONS	CAN'T	79	0.03%	74	0.16%	221.9
NEGATIONS	CANNOT	92	0.04%	89	0.19%	251.0
NEGATIONS	DIDN'T	16	0.01%	16	0.03%	55.6
NEGATIONS	DOESN'T	26	0.01%	26	0.05%	84.8
NEGATIONS	DON'T	110	0.05%	103	0.22%	293.1
NEGATIONS	FEW	90	0.04%	88	0.18%	246.0
NEGATIONS	ISN'T	11	0.00%	11	0.02%	40.0
NEGATIONS	LOW	330	0.15%	317	0.67%	718.2
NEGATIONS	NEVER	21	0.01%	20	0.04%	70.9
NEGATIONS	NO	584	0.26%	524	1.10%	1143.6
NEGATIONS	NONE	22	0.01%	22	0.05%	73.4
NEGATIONS	NOR	11	0.00%	11	0.02%	40.0
NEGATIONS	NOT	1228	0.54%	1058	2.22%	2029.9
NEGATIONS	WITHOUT	57	0.03%	55	0.12%	167.4
NEGATIONS	WON'T	12	0.01%	12	0.03%	43.2
NEGATIONS	ZERO	13	0.01%	13	0.03%	46.3
NEGATIVE WORDS	ADVERS*	21	0.01%	21	0.04%	70.5
NEGATIVE WORDS	AFRAID*	15	0.01%	12	0.03%	54.0
NEGATIVE WORDS	AGAINST	27	0.01%	27	0.06%	87.6
NEGATIVE WORDS	AGGRESS*	22	0.01%	20	0.04%	74.3
NEGATIVE WORDS	ANTI	16	0.01%	16	0.03%	55.6
NEGATIVE WORDS	ATTEMPT	11	0.00%	11	0.02%	40.0
NEGATIVE WORDS	AVOID	14	0.01%	14	0.03%	49.4
NEGATIVE WORDS	BAD	128	0.06%	123	0.26%	331.2
NEGATIVE WORDS	BREAK	65	0.03%	64	0.13%	186.6
NEGATIVE WORDS	BROK*	35	0.02%	33	0.07%	110.6
NEGATIVE WORDS	BURDEN	25	0.01%	25	0.05%	82.0
NEGATIVE WORDS	CANCEL	29	0.01%	27	0.06%	94.1
NEGATIVE WORDS	CAUTION	11	0.00%	11	0.02%	40.0
NEGATIVE WORDS	CAUTIOUS	13	0.01%	13	0.03%	46.3
NEGATIVE WORDS	CHEAP	20	0.01%	20	0.04%	67.5
NEGATIVE WORDS	CLAIM	10	0.00%	9	0.02%	37.2
NEGATIVE WORDS	CLOSE	259	0.11%	245	0.51%	592.7
NEGATIVE WORDS	CONCERN	180	0.08%	178	0.37%	436.9
NEGATIVE WORDS	CONFUS*	16	0.01%	15	0.03%	56.0
NEGATIVE WORDS	CONSTRAIN	14	0.01%	14	0.03%	49.4
NEGATIVE WORDS	CONSTRAINT	17	0.01%	16	0.03%	59.0
NEGATIVE WORDS	CRISI*	13	0.01%	13	0.03%	46.3
NEGATIVE WORDS	CRITIC*	16	0.01%	16	0.03%	55.6
NEGATIVE WORDS	CURTAIL	19	0.01%	19	0.04%	64.6
NEGATIVE WORDS	CUTBACK	12	0.01%	12	0.03%	43.2
NEGATIVE WORDS	CUT_BACK	28	0.01%	28	0.06%	90.5
NEGATIVE WORDS	DAMAG*	18	0.01%	18	0.04%	61.6
NEGATIVE WORDS	DEAD	18	0.01%	18	0.04%	61.6
NEGATIVE WORDS	DEBT	34	0.01%	33	0.07%	107.4
NEGATIVE WORDS	DEFENS*	92	0.04%	89	0.19%	251.0
NEGATIVE WORDS	DEFICIT	23	0.01%	23	0.05%	76.3
NEGATIVE WORDS	DEPEND	83	0.04%	83	0.17%	229.0
NEGATIVE WORDS	DEPRES*	72	0.03%	71	0.15%	203.5
NEGATIVE WORDS	DIFFICULT	150	0.07%	146	0.31%	377.0
NEGATIVE WORDS	DIFFICULT*	43	0.02%	42	0.09%	131.3
NEGATIVE WORDS	DIMINISH	31	0.01%	29	0.06%	99.7
NEGATIVE WORDS	DISAPPOINT*	17	0.01%	16	0.03%	59.0
NEGATIVE WORDS	DISMAL*	11	0.00%	11	0.02%	40.0
NEGATIVE WORDS	DISRUPT	15	0.01%	15	0.03%	52.5
NEGATIVE WORDS	DRAG	30	0.01%	30	0.06%	96.0
NEGATIVE WORDS	DRASTIC	17	0.01%	16	0.03%	59.0
NEGATIVE WORDS	DROP	176	0.08%	157	0.33%	436.8
NEGATIVE WORDS	DROP_OFF	29	0.01%	29	0.06%	93.2
NEGATIVE WORDS	DRY	15	0.01%	15	0.03%	52.5
NEGATIVE WORDS	EMPT*	10	0.00%	9	0.02%	37.2
NEGATIVE WORDS	EXCESS	31	0.01%	30	0.06%	99.2
NEGATIVE WORDS	EXCISE	26	0.01%	25	0.05%	85.3
NEGATIVE WORDS	FAIL*	12	0.01%	12	0.03%	43.2
NEGATIVE WORDS	FALL	142	0.06%	138	0.29%	360.3
NEGATIVE WORDS	FALLEN	17	0.01%	17	0.04%	58.6
NEGATIVE WORDS	FALL_OFF	12	0.01%	12	0.03%	43.2
NEGATIVE WORDS	FEAR	41	0.02%	40	0.08%	126.1
NEGATIVE WORDS	FED	53	0.02%	52	0.11%	157.0
NEGATIVE WORDS	FEE	64	0.03%	55	0.12%	188.0
NEGATIVE WORDS	FELL	32	0.01%	29	0.06%	102.9
NEGATIVE WORDS	FIGHT	11	0.00%	11	0.02%	40.0
NEGATIVE WORDS	FIRE	51	0.02%	50	0.11%	151.9
NEGATIVE WORDS	FIX	29	0.01%	27	0.06%	94.1
NEGATIVE WORDS	FLOOD	25	0.01%	25	0.05%	82.0
NEGATIVE WORDS	FROZEN	14	0.01%	13	0.03%	49.9
NEGATIVE WORDS	FRUSTRAT*	17	0.01%	16	0.03%	59.0
NEGATIVE WORDS	GROSS	13	0.01%	13	0.03%	46.3
NEGATIVE WORDS	HAMPER	14	0.01%	14	0.03%	49.4
NEGATIVE WORDS	HARD	121	0.05%	116	0.24%	316.2
NEGATIVE WORDS	HOLD	132	0.06%	130	0.27%	338.4
NEGATIVE WORDS	HOLD_BACK	13	0.01%	13	0.03%	46.3
NEGATIVE WORDS	HOT	22	0.01%	21	0.04%	73.8
NEGATIVE WORDS	HURT	122	0.05%	116	0.24%	318.8
NEGATIVE WORDS	INFLATION	42	0.02%	42	0.09%	128.3
NEGATIVE WORDS	KICK*	15	0.01%	15	0.03%	52.5
NEGATIVE WORDS	KILL*	50	0.02%	50	0.11%	148.9
NEGATIVE WORDS	LACK	179	0.08%	175	0.37%	435.8
NEGATIVE WORDS	LAG	14	0.01%	14	0.03%	49.4
NEGATIVE WORDS	LAID-OFF	16	0.01%	16	0.03%	55.6
NEGATIVE WORDS	LATE	64	0.03%	61	0.13%	185.1
NEGATIVE WORDS	LAYOFF	61	0.03%	58	0.12%	177.8
NEGATIVE WORDS	LET	31	0.01%	29	0.06%	99.7
NEGATIVE WORDS	LIMIT	67	0.03%	67	0.14%	191.0
NEGATIVE WORDS	LOAD	25	0.01%	25	0.05%	82.0
NEGATIVE WORDS	LOOM	11	0.00%	11	0.02%	40.0
NEGATIVE WORDS	LOSE	41	0.02%	40	0.08%	126.1
NEGATIVE WORDS	LOSS	68	0.03%	62	0.13%	196.2
NEGATIVE WORDS	LOST	84	0.04%	82	0.17%	232.2

TYPE OF WORD	WORD	FREQUENCY	% TOTAL	NO. CASES	% CASES	TF • IDF
NEGATIVE WORDS	LOWER	263	0.12%	233	0.49%	607.6
NEGATIVE WORDS	LOWEST	19	0.01%	19	0.04%	64.6
NEGATIVE WORDS	MAR	13	0.01%	12	0.03%	46.8
NEGATIVE WORDS	MEAN	35	0.02%	34	0.07%	110.1
NEGATIVE WORDS	MESS	18	0.01%	16	0.03%	62.5
NEGATIVE WORDS	MINE	75	0.03%	62	0.13%	216.4
NEGATIVE WORDS	MINIMUM	57	0.03%	57	0.12%	166.5
NEGATIVE WORDS	NERVOU*	17	0.01%	17	0.04%	58.6
NEGATIVE WORDS	NOT_*_ENOUGH	12	0.01%	12	0.03%	43.2
NEGATIVE WORDS	NOT_GOOD	21	0.01%	21	0.04%	70.5
NEGATIVE WORDS	NUMB*	186	0.08%	175	0.37%	452.8
NEGATIVE WORDS	OLD	34	0.01%	33	0.07%	107.4
NEGATIVE WORDS	OLDER	14	0.01%	14	0.03%	49.4
NEGATIVE WORDS	OUT	415	0.18%	374	0.79%	873.4
NEGATIVE WORDS	PAIN	12	0.01%	12	0.03%	43.2
NEGATIVE WORDS	PAIN*	12	0.01%	12	0.03%	43.2
NEGATIVE WORDS	PLEAS	34	0.01%	34	0.07%	107.0
NEGATIVE WORDS	POOR	88	0.04%	85	0.18%	241.8
NEGATIVE WORDS	PRESSUR*	235	0.10%	226	0.47%	546.0
NEGATIVE WORDS	PREVENT	15	0.01%	15	0.03%	52.5
NEGATIVE WORDS	PROBLEM	146	0.06%	141	0.30%	369.1
NEGATIVE WORDS	PUSH	55	0.02%	54	0.11%	162.0
NEGATIVE WORDS	QUIT	38	0.02%	37	0.08%	118.2
NEGATIVE WORDS	RAW	566	0.25%	530	1.11%	1105.5
NEGATIVE WORDS	REBOUND	23	0.01%	23	0.05%	76.3
NEGATIVE WORDS	RESTRICT	48	0.02%	48	0.10%	143.8
NEGATIVE WORDS	RISK	21	0.01%	21	0.04%	70.5
NEGATIVE WORDS	ROUGH	11	0.00%	11	0.02%	40.0
NEGATIVE WORDS	SCARE*	23	0.01%	23	0.05%	76.3
NEGATIVE WORDS	SCRAP	44	0.02%	41	0.09%	134.8
NEGATIVE WORDS	SEVER*	150	0.07%	146	0.31%	377.0
NEGATIVE WORDS	SHORT	104	0.05%	104	0.22%	276.7
NEGATIVE WORDS	SHORTER	11	0.00%	11	0.02%	40.0
NEGATIVE WORDS	SHOULD	149	0.07%	144	0.30%	375.4
NEGATIVE WORDS	SHOW	109	0.05%	105	0.22%	289.5
NEGATIVE WORDS	SHRINK	32	0.01%	31	0.07%	102.0
NEGATIVE WORDS	SHUT	46	0.02%	43	0.09%	140.0
NEGATIVE WORDS	SHUTDOWN	37	0.02%	34	0.07%	116.4
NEGATIVE WORDS	SLOWDOWN	129	0.06%	126	0.26%	332.5
NEGATIVE WORDS	SLOWER	106	0.05%	100	0.21%	283.8
NEGATIVE WORDS	SLOWEST	19	0.01%	19	0.04%	64.6
NEGATIVE WORDS	SLOW_DOWN	182	0.08%	177	0.37%	442.2
NEGATIVE WORDS	SLUGGISH*	52	0.02%	52	0.11%	154.0
NEGATIVE WORDS	SLUMP	14	0.01%	14	0.03%	49.4
NEGATIVE WORDS	SMOKE	33	0.01%	33	0.07%	104.2
NEGATIVE WORDS	STAB*	132	0.06%	129	0.27%	338.8
NEGATIVE WORDS	STAGNANT	37	0.02%	37	0.08%	115.0
NEGATIVE WORDS	STALL	16	0.01%	15	0.03%	56.0
NEGATIVE WORDS	STOP	66	0.03%	61	0.13%	190.9
NEGATIVE WORDS	STORM	36	0.02%	35	0.07%	112.8
NEGATIVE WORDS	STRANG*	15	0.01%	15	0.03%	52.5
NEGATIVE WORDS	STRESS	19	0.01%	19	0.04%	64.6
NEGATIVE WORDS	STRIKE	12	0.01%	10	0.02%	44.1
NEGATIVE WORDS	STRUGGL*	51	0.02%	50	0.11%	151.9
NEGATIVE WORDS	SUFFER	30	0.01%	29	0.06%	96.5
NEGATIVE WORDS	TARIFF	151	0.07%	124	0.26%	390.2
NEGATIVE WORDS	TAX	357	0.16%	316	0.66%	777.5
NEGATIVE WORDS	TEMPER*	12	0.01%	11	0.02%	43.6
NEGATIVE WORDS	TERRIBL*	23	0.01%	23	0.05%	76.3
NEGATIVE WORDS	TERROR	23	0.01%	23	0.05%	76.3
NEGATIVE WORDS	THREAT*	31	0.01%	30	0.06%	99.2
NEGATIVE WORDS	TIRE	13	0.01%	12	0.03%	46.8
NEGATIVE WORDS	TOO_MUCH	34	0.01%	31	0.07%	108.3
NEGATIVE WORDS	TOUGH	47	0.02%	47	0.10%	141.3
NEGATIVE WORDS	TROUBL*	23	0.01%	22	0.05%	76.7
NEGATIVE WORDS	TRY	68	0.03%	67	0.14%	193.9
NEGATIVE WORDS	TURMOIL	14	0.01%	14	0.03%	49.4
NEGATIVE WORDS	TURN	72	0.03%	69	0.14%	204.4
NEGATIVE WORDS	UNCERTAIN*	170	0.07%	168	0.35%	416.9
NEGATIVE WORDS	UNFAIR	15	0.01%	15	0.03%	52.5
NEGATIVE WORDS	UNKNOWN	29	0.01%	29	0.06%	93.2
NEGATIVE WORDS	WAIT	54	0.02%	50	0.11%	160.8
NEGATIVE WORDS	WAR	77	0.03%	75	0.16%	215.8
NEGATIVE WORDS	WEAK	155	0.07%	150	0.32%	387.7
NEGATIVE WORDS	WEAK*	29	0.01%	29	0.06%	93.2
NEGATIVE WORDS	WEAKEN	21	0.01%	21	0.04%	70.5
NEGATIVE WORDS	WORR*	27	0.01%	27	0.06%	87.6
NEGATIVE WORDS	WORST	36	0.02%	34	0.07%	113.3
POSITIVE WORDS	ADVANTAG*	18	0.01%	18	0.04%	61.6
POSITIVE WORDS	AFFORD	31	0.01%	27	0.06%	100.6
POSITIVE WORDS	AGREEMENT*	36	0.02%	36	0.08%	112.4
POSITIVE WORDS	ALLOW	55	0.02%	54	0.11%	162.0
POSITIVE WORDS	AMAZ*	12	0.01%	12	0.03%	43.2
POSITIVE WORDS	ASSET	19	0.01%	17	0.04%	65.5
POSITIVE WORDS	ATTRACT*	27	0.01%	25	0.05%	88.5
POSITIVE WORDS	AWARD*	15	0.01%	14	0.03%	53.0
POSITIVE WORDS	BENEFIT	64	0.03%	59	0.12%	186.0
POSITIVE WORDS	BEST	115	0.05%	106	0.22%	305.0
POSITIVE WORDS	BETTER	292	0.13%	270	0.57%	655.9
POSITIVE WORDS	BONUS*	14	0.01%	14	0.03%	49.4
POSITIVE WORDS	BOOM	13	0.01%	13	0.03%	46.3
POSITIVE WORDS	BOOST	15	0.01%	15	0.03%	52.5
POSITIVE WORDS	BRIGHT*	15	0.01%	15	0.03%	52.5
POSITIVE WORDS	BRISK	15	0.01%	15	0.03%	52.5
POSITIVE WORDS	CANDID	41	0.02%	36	0.08%	128.0
POSITIVE WORDS	CARE	155	0.07%	146	0.31%	389.5
POSITIVE WORDS	CHALLENG*	61	0.03%	58	0.12%	177.8
POSITIVE WORDS	CLASS	14	0.01%	14	0.03%	49.4
POSITIVE WORDS	CLEAN	11	0.00%	10	0.02%	40.5
POSITIVE WORDS	CLEAR	22	0.01%	22	0.05%	73.4

Figure 25 Words (continued)

TYPE OF WORD	WORD	FREQUENCY	% TOTAL	NO. CASES	% CASES	TF • IDF
POSITIVE WORDS	CONSTRUCT	397	0.17%	367	0.77%	838.8
POSITIVE WORDS	COOL	22	0.01%	21	0.04%	73.8
POSITIVE WORDS	CORRECT	35	0.02%	34	0.07%	110.1
POSITIVE WORDS	DEAL	51	0.02%	49	0.10%	152.4
POSITIVE WORDS	DECENT	16	0.01%	16	0.03%	55.6
POSITIVE WORDS	EARN	19	0.01%	18	0.04%	65.0
POSITIVE WORDS	ENCOURAG*	18	0.01%	18	0.04%	61.6
POSITIVE WORDS	ENERG*	272	0.12%	260	0.55%	615.4
POSITIVE WORDS	ENOUGH	58	0.03%	55	0.12%	170.4
POSITIVE WORDS	EXCEL*	19	0.01%	17	0.04%	65.5
POSITIVE WORDS	EXPAND	68	0.03%	67	0.14%	193.9
POSITIVE WORDS	EXTEND	31	0.01%	31	0.07%	98.8
POSITIVE WORDS	EXTREM*	74	0.03%	73	0.15%	208.3
POSITIVE WORDS	FAIR	15	0.01%	15	0.03%	52.5
POSITIVE WORDS	FAST	22	0.01%	22	0.05%	73.4
POSITIVE WORDS	FASTER	13	0.01%	13	0.03%	46.3
POSITIVE WORDS	FAVOR*	30	0.01%	29	0.06%	96.5
POSITIVE WORDS	FREE	21	0.01%	21	0.04%	70.5
POSITIVE WORDS	FRESH	12	0.01%	11	0.02%	43.6
POSITIVE WORDS	GAIN	51	0.02%	49	0.10%	152.4
POSITIVE WORDS	GOOD	601	0.26%	529	1.11%	1174.4
POSITIVE WORDS	GREAT*	137	0.06%	133	0.28%	349.9
POSITIVE WORDS	HA	1209	0.53%	1081	2.27%	1987.2
POSITIVE WORDS	HAPPI*	12	0.01%	12	0.03%	43.2
POSITIVE WORDS	HELP	194	0.09%	187	0.39%	466.7
POSITIVE WORDS	HIGHER	287	0.13%	244	0.51%	657.3
POSITIVE WORDS	HIGHEST	35	0.02%	31	0.07%	111.5
POSITIVE WORDS	HOME	260	0.11%	218	0.46%	608.2
POSITIVE WORDS	HOPE	169	0.07%	162	0.34%	417.1
POSITIVE WORDS	IMPROVE*	33	0.01%	32	0.07%	104.7
POSITIVE WORDS	INTEREST*	287	0.13%	282	0.59%	639.2
POSITIVE WORDS	LARGER	58	0.03%	57	0.12%	169.5
POSITIVE WORDS	LEADERSHIP	33	0.01%	30	0.06%	105.6
POSITIVE WORDS	MAINTAIN	47	0.02%	47	0.10%	141.3
POSITIVE WORDS	MIND	13	0.01%	12	0.03%	46.8
POSITIVE WORDS	MODEST	18	0.01%	18	0.04%	61.6
POSITIVE WORDS	NICE*	38	0.02%	37	0.08%	118.2
POSITIVE WORDS	OPPORT*	48	0.02%	46	0.10%	144.7
POSITIVE WORDS	OPTIMI*	88	0.04%	87	0.18%	240.9
POSITIVE WORDS	PARTNER	14	0.01%	10	0.02%	51.5
POSITIVE WORDS	PATIENT	56	0.02%	52	0.11%	165.8
POSITIVE WORDS	PERMIT	12	0.01%	12	0.03%	43.2
POSITIVE WORDS	PREMIUM	31	0.01%	31	0.07%	98.8
POSITIVE WORDS	PROFIT*	198	0.09%	188	0.40%	475.9
POSITIVE WORDS	PROGRESS	16	0.01%	16	0.03%	55.6
POSITIVE WORDS	PROMIS*	12	0.01%	11	0.02%	43.6
POSITIVE WORDS	PROTECT	14	0.01%	14	0.03%	49.4
POSITIVE WORDS	RELIEF*	21	0.01%	20	0.04%	70.9
POSITIVE WORDS	REMODEL	18	0.01%	18	0.04%	61.6
POSITIVE WORDS	REPAIR	14	0.01%	13	0.03%	49.9
POSITIVE WORDS	RICH*	29	0.01%	28	0.06%	93.7
POSITIVE WORDS	RISEN	20	0.01%	20	0.04%	67.5
POSITIVE WORDS	ROBUST	27	0.01%	27	0.06%	87.6
POSITIVE WORDS	SAFE*	33	0.01%	33	0.07%	104.2
POSITIVE WORDS	SECUR*	41	0.02%	39	0.08%	126.5
POSITIVE WORDS	SHARE	51	0.02%	49	0.10%	152.4
POSITIVE WORDS	SKILL	178	0.08%	162	0.34%	439.3
POSITIVE WORDS	SOFT	114	0.05%	111	0.23%	300.1
POSITIVE WORDS	SOFTEN	47	0.02%	46	0.10%	141.7
POSITIVE WORDS	SOFTER	13	0.01%	13	0.03%	46.3
POSITIVE WORDS	SOLID	18	0.01%	18	0.04%	61.6
POSITIVE WORDS	SPECIAL	16	0.01%	16	0.03%	55.6
POSITIVE WORDS	STAINLESS	15	0.01%	15	0.03%	52.5
POSITIVE WORDS	STRENGTH*	24	0.01%	24	0.05%	79.1
POSITIVE WORDS	STRONG*	457	0.20%	418	0.88%	939.8
POSITIVE WORDS	SUCCESS*	22	0.01%	22	0.05%	73.4
POSITIVE WORDS	SUPPORT	47	0.02%	47	0.10%	141.3
POSITIVE WORDS	SURE*	73	0.03%	71	0.15%	206.3
POSITIVE WORDS	SURPRIS*	34	0.01%	34	0.07%	107.0
POSITIVE WORDS	TALENT*	27	0.01%	27	0.06%	87.6
POSITIVE WORDS	THANK	26	0.01%	26	0.05%	84.8
POSITIVE WORDS	THE_REAL	31	0.01%	30	0.06%	99.2
POSITIVE WORDS	TRADITION	24	0.01%	24	0.05%	79.1
POSITIVE WORDS	TRAIN	51	0.02%	45	0.09%	154.2
POSITIVE WORDS	TRAVEL	50	0.02%	42	0.09%	152.7
POSITIVE WORDS	TRUE	22	0.01%	22	0.05%	73.4
POSITIVE WORDS	UNDERSTAND	15	0.01%	14	0.03%	53.0
POSITIVE WORDS	WARM*	127	0.06%	124	0.26%	328.2
POSITIVE WORDS	WELL	154	0.07%	148	0.31%	386.1
POSITIVE WORDS	WIN	19	0.01%	18	0.04%	65.0
POSITIVE WORDS	WISE*	11	0.00%	9	0.02%	41.0
POSITIVE WORDS	WORK	566	0.25%	485	1.02%	1127.4
UNCERTAINTY	ALMOST	74	0.03%	74	0.16%	207.8
UNCERTAINTY	APPEAR	102	0.04%	99	0.21%	273.6
UNCERTAINTY	CAUTIOUS	13	0.01%	13	0.03%	46.3
UNCERTAINTY	DEPEND	83	0.04%	83	0.17%	229.0
UNCERTAINTY	DIFFER	35	0.02%	35	0.07%	109.7
UNCERTAINTY	GUESS	22	0.01%	22	0.05%	73.4
UNCERTAINTY	MIGHT	19	0.01%	19	0.04%	64.6
UNCERTAINTY	PREDICT	28	0.01%	28	0.06%	90.5
UNCERTAINTY	RISK	21	0.01%	21	0.04%	70.5
UNCERTAINTY	SOMEWHAT	67	0.03%	67	0.14%	191.0
UNCERTAINTY	UNCERTAIN	47	0.02%	47	0.10%	141.3
UNCERTAINTY	UNKNOWN	29	0.01%	29	0.06%	93.2

Notes: The symbol "*" is a wildcard that substitutes for several alternative forms of a word or expression. TF IDF: term frequency x inverse document frequency

Figure 26 Examples

Negative	
Not good	**Bad**
Floods in the region won't help either	Apparel industry continues to suffer
govt furlough does not help consumer confidence problem	We have had several plants and retail stores to close in the past few weeks
One local car dealership closed - another not well	December sales were off over 15% as bad as I have seen
Seasonal--economy fair-not robust	Winter has been extremely tough on our business
From Sept. 10 to mid-October business was not at all good	lack of consumer conficence
Positive	
Not bad	**Good**
There should be no further decline	demand for our services continues to be strong
High gas and raw goods prices have not dampened the buying spirit	Business is very good.
No adverse signs yet to our business	Better than it's been since 2006
No noticeable loss of business because of gas prices	Weather has improved and we are very busy
inflation and rising raw material prices are not too bad	Our property management business is growing market share and doing well
Uncertainty	
Appear	
Prices appeared fairly stable in April/sales continued strong	
Demand appears to be returning for permanent placement activity	
Energy costs appear stable	
Richmond economy appears steady	
Office leasing appears to be picking up slightly	
Depend	
We are very dependent on Medicaid funidng.	
Several capital projects are being planned at this time dependent on the market and economy continuing to remain strong - include lumber storage building	
Increased profits depend largely on increased volume of service	
Depends on gas prices	
Increase in reimbursements depends on cotnracts	
Somewhat	
Business slows down somewhat for hot summer months	
This month somewhat reflects last month	
Recovering oil prices are helping us somewhat	
Holidays should help somewhat	
Business somewhat better	

REFERENCES

Barkin, Thomas I. 2019. "Confidence, Expectations and Implications for Monetary Policy." Speech at the Global Interdependence Center's Rocky Mountain Economic Summit, Victor, Idaho, July 11.

Barsky, Robert B., and Eric R. Sims. 2012. "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence." *American Economic Review* 102 (June): 1343–77.

Bram, Jason, and Sydney Ludvigson. 1998. "Does Consumer Confidence Forecast Household Expenditure? A Sentiment Index Horse Race." Federal Reserve Bank of New York *Economic Policy Review* 4 (June): 59–78.

Calomiris, Charles W., and Harry Mamaysky. 2019. "How News and Its Context Drive Risk and Returns around the World." *Journal of Financial Economics* 133 (August): 299–336.

Hansen, Stephen, Michael McMahon, and Andrea Prat. 2018. "Transparency and Deliberation within the FOMC: a Computational Linguistics Approach." *Quarterly Journal of Economics* 133 (May): 801–70.

Lazaryan, Nika, and Santiago Pinto. 2017. "Using the Richmond Fed Manufacturing Survey to Gauge National and Regional Economic Conditions." Federal Reserve Bank of Richmond *Economic Quarterly* 103 (First-Fourth Quarter): 81–136.

Loughran, Tim, and Bill McDonald. 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." *Journal of Finance* 66 (February): 35–65.

Loughran, Tim, and Bill McDonald. 2015. "The Use of Word Lists in Textual Analysis." *Journal of Behavioral Finance* 16 (March): 1–11.

Macheras, Ann, Santiago Pinto, Jessie Romero, and Pierre-Daniel G. Sarte. 2015. "Why Does the Fed Study Regional Economics?" Federal Reserve Bank of Richmond *Economic Brief* 15-01 (January.)

Nyman, Rickard, Sujit Kapadia, David Tuckett, David Gregory, Paul Ormerod, and Robert Smith. 2018. "News and Narratives in Financial Systems: Exploiting Big Data for Systemic Risk Assessment." Bank of England Staff Working Paper 704 (January).

Pinto, Santiago, Sonya Waddell, and Pierre-Daniel G. Sarte. 2015. "Monitoring Economic Activity in Real Time Using Diffusion Indices: Evidence from the Fifth District." Federal Reserve Bank of Richmond *Economic Quarterly* 101 (Fourth Quarter): 275–301.

Pinto, Santiago, Pierre-Daniel G. Sarte, and Robert Sharp. Forthcoming. "The Information Content and Statistical Properties of Diffusion Indices." *International Journal of Central Banking.*

Price, David A., and Aileen Watson. 2014. "The Richmond Fed Manufacturing and Service Sector Surveys: A User's Guide." Federal Reserve Bank of Richmond *Economic Brief* 14-03 (March).

Shapiro, Adam Hale, Moritz Sudhof, and Daniel Wilson. 2018. "Measuring News Sentiment." Federal Reserve Bank of San Francisco Working Paper 2017-01 (June).

Souleles, Nicholas S. 2004. "Expectations, Heterogeneous Forecast Errors, and Consumption: Micro Evidence from the Michigan Consumer Sentiment Surveys." *Journal of Money, Credit and Banking* 36 (February): 39–72.

Thorsrud, Leif Anders. 2018. "Words Are the New Numbers: A Newsy Coincident Index of the Business Cycle." *Journal of Business & Economic Statistics*, https://doi.org/10.1080/07350015.2018.1506344.

Waddell, Sonya. 2015. "Predicting Economic Activity through Richmond Fed Surveys." Federal Reserve Bank of Richmond *Econ Focus* (Fourth Quarter): 36–39.

Young, Lori, and Stuart Soroka. 2012. "Affective News: The Automated Coding of Sentiment in Political Texts." *Journal of Political Communication* 29 (April): 205–31.

Economic Quarterly—Volume 105, Number 3—Third Quarter 2019—Pages 173–200

Wealth Effects with Endogenous Retirement

Borys Grochulski and Yuzhe Zhang

In this article, we discuss the so-called wealth effects: the response of aggregate consumption to exogenous movements in wealth. Wealth effects are of interest to market participants and policymakers, as they can be informative about expected GDP growth given observed movements in asset prices.

Estimated in a standard way, the wealth effect in aggregate data amounts to about 2.4 cents on the dollar. From the point of view of the simple permanent income theory, as in Friedman (1957), Bewley (1977), or Hall (1978), these estimates are surprisingly low. With changes in asset prices unpredictable, an exogenous increase in wealth of \$1 increases the agent's permanent income by r dollars, where r is the riskless rate of interest or the agent's rate of time preference. With the standard estimate of $r = 5$ percent, permanent income theory predicts the wealth effect of 5 cents on the dollar, which is about twice the effect we observe in the data.[1]

Poterba (2000) reviews the main explanations of weak wealth effects that have been proposed in the literature. At the aggregate level, the response of consumption to wealth changes may be low because wealth concentration is high and consumption of high net worth households may be relatively inelastic. Cagetti and De Nardi (2008) and Saez and Zucman (2016) document a recent further increase in wealth

The authors thank Caroline Davis, Felipe Schwartzman, Bruno Sultanum, and John Weinberg for their helpful comments and Emma Yeager for excellent research assistance. The views expressed in this article are those of the authors and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.
Email: borys.grochulski@rich.frb.org; yuzhe-zhang@econmail.tamu.edu.

[1] King and Low (2014) provide estimates of the average world real interest rate of near 5 percent in the 1980s and 1990s, declining after 2000 and strongly so since the financial crisis. The sample we use to estimate wealth effects in Section 1 covers 1958-2018, which does include the period of high real rates identified by King and Low (2014).

inequality in the US. Using state-level data, Calomiris, Longhofer, and Miles (2012) find a large variation in wealth effects correlated with the dispersion of wealth and age distributions across US states.

At the individual level, the response of consumption to changes in wealth may be weak because wealth is allocated to illiquid assets. Kaplan and Violante (2014) and Saez and Zucman (2016) document that about 80 percent of wealth is held in illiquid assets like housing, retirement accounts, and closely held businesses.

In this article, we discuss another reason why wealth effects may be weak: an endogenous reaction of labor supply along the extensive margin, i.e., retirement. When households save for retirement, their optimal retirement timing decision depends on their wealth. A positive wealth shock can make a household adjust their planned retirement date forward, i.e., shorten the remainder of their work career. But a shorter work career means the present value of all future labor income goes down, which partially offsets the impact of the positive wealth shock on consumption.

Zhao (2018) provides direct evidence on the response of the retirement timing decision to wealth shocks. Using data from the Health and Retirement Study, a panel survey of individuals age fifty and older, he shows that declines in housing prices are positively correlated with a drop in retirement probability for homeowners, while no such correlation exists for renters.

To evaluate quantitatively the impact of the endogenous retirement timing decision on the standard wealth effect, in this article we build a simple model in which the retirement decision is optimally taken according to a threshold policy: the agent retires when her financial wealth, W_t, reaches a particular, optimally chosen target level, W^*. A positive wealth shock, specifically, a positive shock to the rate of return on her financial assets, brings the agent closer to retirement. Correspondingly, the monetary value of her human capital, i.e., the present value of the labor income she expects to earn in the remainder of her career, decreases. The agent's optimal consumption decision, naturally, takes into account her total wealth: the sum of her financial wealth and the monetary value of her human capital. The model captures the effect of the offsetting movement of the value of human capital on the response of consumption to financial wealth shocks. Despite a few strong simplifying assumptions we make to keep the model tractable, the model is able to generate a weak wealth effect, which helps explain why the response of consumption to wealth shocks is weak in the data.

We start in Section 1 by presenting a standard estimate of wealth effects in aggregate consumption and wealth data. Using the quarterly US data on wealth and consumption from 1958 to 2018, we estimate the average wealth effect of 2.4 cents on the dollar. This number is in line with standard estimates obtained in the literature, as summarized in Poterba (2000).

In Section 2, we lay out a stylized model of optimal consumption, saving, and retirement decisions, which essentially is a simplified version of Kingston (2000) and Farhi and Panageas (2007). We follow these studies, in particular, in assuming that the retirement decision is irreversible.[2]

In Section 3, as a baseline, we show that when the retirement decision margin is shut down, our model predicts a wealth effect of r cents on the dollar, in line with the simple permanent income theory. In the baseline case, in particular, the agent does not have an active retirement margin because she is already retired. The value of her human capital is therefore nil and all her wealth is financial. With a logarithmic utility function and Brownian motion wealth return shocks, we can solve the baseline case in closed form.

In Sections 4 and 5, we solve the model with an endogenous retirement decision. We show that if the rate of return on wealth is sufficiently high, the agent prefers to retire if and when her wealth reaches a target level, W^*. We discuss the Hamilton-Jacobi-Bellman (HJB) equation for the agent's lifetime utility value function, along with a procedure for finding appropriate boundary conditions.

In Section 6, we discuss the dynamics of the monetary value of the agent's human capital in the solution to her optimal consumption, saving, and retirement problem. The main observation there is that, with the retirement decision taken according to a wealth-threshold policy, the value of human capital decreases when the agent's wealth increases. The key part of the computation of the value of human capital at any point in time is the expected remaining duration of the agent's career, i.e., the amount of time left until retirement. We present a useful lemma that allows for computation of this object in our model.

In Section 7, we analyze a special case of our model in which the expected growth rate of financial wealth is equal to the agent's rate of time preference, r. We use this special case to show clearly the intuition for our main result: the value of human capital responds negatively to wealth shocks, making the response of consumption weaker. In the special case, in particular, the endogenous response of the value of human capital perfectly offsets all shocks to financial wealth prior to retirement, making the agent's total wealth and consumption constant up until the retirement date. With constant consumption, clearly, the

[2] In Section 9, we comment on how our results would change if our model allowed for unretirement.

Table 1 Sample Properties of Aggregate Real Consumption and Net Wealth

	Average growth rate	St. dev. of growth rate
Consumption	2.9%	2.4%
Net Wealth	3.4%	4.4%

offsetting response of the value of human capital is strong enough to make the wealth effect nil at all times during the agent's work career.

Our main results are presented in Section 8, where we calibrate the model to match the wealth effect of 2.4 cents on the dollar, as in the data. The model is capable of generating realistic wealth effects under a reasonable parametrization. The untargeted wealth threshold W^* associated with the desired wealth effect is close to sixteen times annual income. We discuss the key intuition of our model showing that the endogenous response of human capital dampens the response of consumption to wealth shocks in the model.

Section 9 concludes with a discussion of the robustness of our results to several of our simplifying assumptions. There, also, we discuss some further related literature.

1. MEASUREMENT

In this section, we present briefly the data and conduct a simple estimation of the average wealth effect, similar to Iacoviello (2011). Our point estimate is 2.4 cents on the dollar.

The data consist of the series of quarterly aggregate net wealth and quarterly aggregate consumption expenditure. Net wealth data come from the Flow of Funds. Our sample covers the period of 1952:Q1 though 2018:Q3.

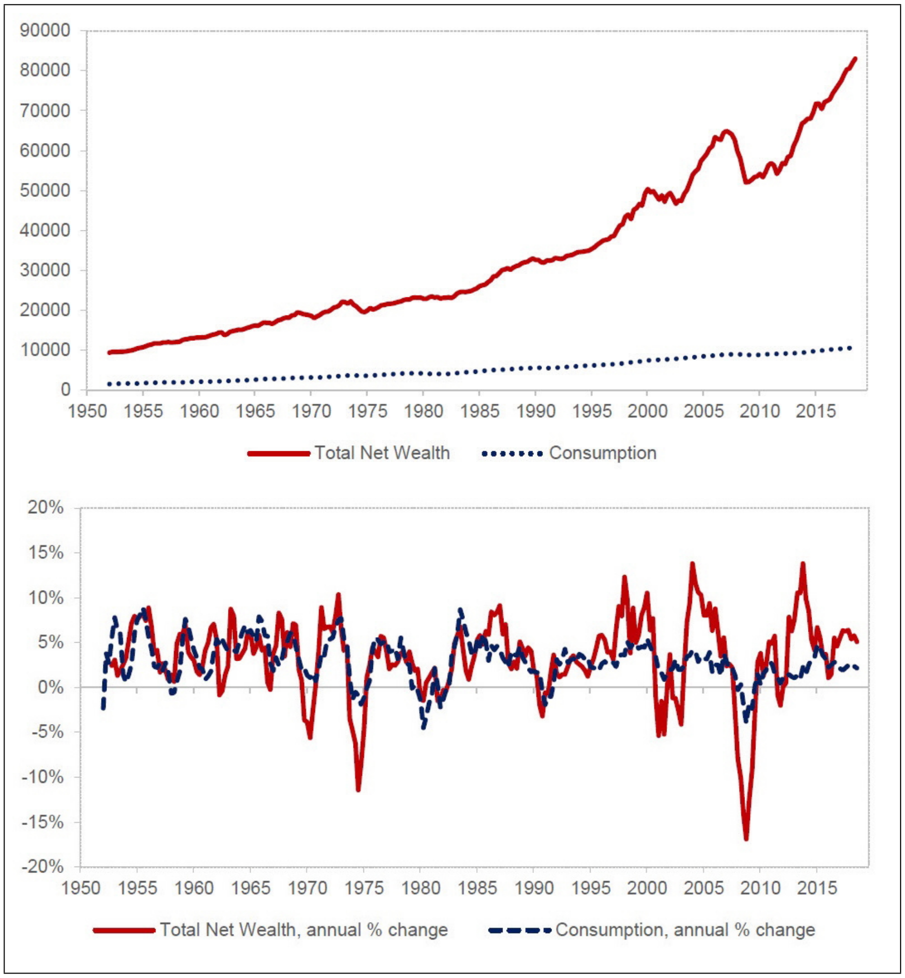
The two variables used in the estimation of wealth effects are as follows:

NW_t = Households and Nonprofit Organizations total assets less total liabilities, constant 2005 dollars (CPI deflated).

C_t = Personal Consumption Expenditure, constant 2005 dollars (CPI deflated).

Figure 1 plots these two series in levels (millions of constant 2005 dollars) and in year-over-year growth rates. In levels, the ratio of sample average net wealth to consumption is 6.1. The summary statistics for the growth rates are in Table 1.

Figure 1 Aggregate Real Consumption and Net Wealth



Notes: Top panel: millions of 2005 dollars. Bottom panel: annualized growth rates.

Following Iacoviello (2011), we obtain the average wealth effect by estimating the following regression equation:

$$\triangle \ln(C_t) = \underset{(.0005)}{.0060} + \underset{(.0286)}{.1475} \triangle \ln(NW_{t-1}),$$

where $\triangle$ denotes the first difference operator. With the average net wealth to consumption ratio of 6.1, the estimated elasticity of consumption to net wealth of 0.1475 gives the wealth effect of 2.4 cents on the dollar.[3]

Several other ways of measuring wealth effects have been considered in the literature. In a sample that ends in 2008:Q4, Iacoviello (2011) estimates a similar wealth effects regression after splitting net wealth into unencumbered housing wealth and net financial wealth. In that regression, he finds overall wealth effects of similar total magnitude, with the effects of changes in the unencumbered housing wealth component being stronger than those estimated for net nonhousing wealth. Piazzesi and Schneider (2016) provide additional discussion of the effects of housing price changes on consumption. The studies reviewed in Poterba (2000) suggest that a \$1 increase in stock market equity values raises consumption in the next quarter by 2 cents, while an analogous increase in non-stock-market wealth raises next-quarter consumption by 1.4 cents.[4]

In our analysis, we will take the average wealth effect of 2.4 cents on the dollar as our target. Using a simple model of optimal consumption, saving, and retirement decisions, we will show how an endogenous response of labor supply along the extensive margin can bring the wealth effect from the baseline level of 5 cents on the dollar down to the estimated value of 2.4 cents.

2. MODEL

Consider the following optimal consumption and saving problem. The retirement decision will be added in Section 3. The agent has initial financial wealth W_0. For simplicity, we will abstract from the agent's portfolio decision of allocating her wealth between different asset classes and instead treat financial wealth as a single, aggregated asset to which

[3] Following Iacoviello (2011), we calculate the average dollar-over-dollar wealth effect from the average elasticity of consumption with respect to net wealth and the average ratio of net wealth to consumption:

$$\frac{\triangle C}{\triangle NW} = \frac{\frac{\triangle C}{C}}{\frac{\triangle NW}{NW}} \frac{C}{NW} = \frac{\triangle \ln(C)}{\triangle \ln(NW)} \div \frac{NW}{C}.$$

[4] See, however, Lettau and Ludvigson (2004), who argue that net worth changes have a significant transitory component, which makes identification of wealth effects not straightforward.

all wealth is allocated.[5] We will refer to this asset simply as financial wealth and denote the amount held by the agent at date t by W_t.[6]

We assume that the expected growth rate and volatility of financial wealth are constant and denote them, respectively, by μ and σ. That is, absent any new investments or withdrawals, the agent's financial wealth follows geometric Brownian motion

$$dW_t = \mu W_t dt + \sigma W_t dZ_t, \tag{1}$$

where Z_t is a cumulative growth rate shock process modeled as standard Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$.

The agent draws utility from two sources: consumption of a single consumption good and leisure that the agent enjoys in retirement. The agent's flow of utility from consumption is $u(c_t)$, where u is her utility function. To facilitate analytical solutions, we will often assume logarithmic utility: $u = \ln$. The agent's flow utility from leisure in retirement is denoted by $\psi > 0$. If not retired, the agent does not receive ψ but earns labor income $y > 0$, which, again for simplicity, we will take to be constant.[7] The agent discounts future payoffs at a constant rate of time preference $r > 0$.

Our simplifying assumption of a single financial asset will force the agent to take on risk as she saves. In particular, our model does not allow the agent to save by investing in a riskless asset. To ensure that this feature of the model does not drive our results, we will assume that the expected growth rate of financial wealth is sufficiently positive.

Assumption 1 $\mu - r > \frac{1}{2}\sigma^2$.

3. WEALTH EFFECTS WITHOUT SAVING FOR RETIREMENT

In this section, we use our model of optimal consumption and saving decisions to derive optimal wealth effects in the absence of an endogenous retirement decision. In particular, for simplicity and also as a building block for the foregoing analysis, we assume in this section that the agent is already retired.

[5] See Kingston (2000) and Farhi and Panageas (2007) for related models with a portfolio choice.

[6] Alternatively, we can interpret W_t as physical capital that can be converted to and from consumption without any transaction costs, so the price of capital in terms of current consumption is always 1.

[7] This assumption is not essential for the main mechanism of our model to work. We discuss this point briefly in the concluding Section 9.

In retirement, as the agent does not earn any labor income, her financial wealth W_t evolves according to

$$dW_t = (\mu W_t - c_t)dt + \sigma W_t dZ_t, \tag{2}$$

where c_t is the agent's consumption flow at date t. The agent also receives the flow of leisure utility ψ.

Let us denote by $V(W)$ the maximal value that the agent can attain in retirement given that her current financial wealth is W. That is, the value function V is defined as

$$V(W) := \max_{\{c_t;t\geq 0\}} \mathbb{E}\{ \left[\int_0^\infty e^{-rt}(u(c_t)+\psi)dt \right] \text{ s.t. (2) and } W_0 = W,$$

where expectation is taken over the realizations of the financial return shock Z_t. Note that the expression for the agent's total discounted expected utility takes into account the flow of leisure utility in retirement ψ.

In the remainder of this section, we will characterize the agent's optimal consumption plan and compute the wealth effect, which shows how the agent's consumption responds to changes in wealth.

HJB equation for the value of retirement

To find the agent's optimal consumption policy, we will use dynamic programming.[8] The intuition behind this approach is as follows. If the agent optimally chooses her consumption, then the flow of value she receives out of financial wealth W_t consists of the current utility flow from consumption and leisure plus the increase in the value she expects going forward. This intuition is succinctly expressed in the following equation:

$$rV(W_t)dt = \max_c \left\{ (u(c)+\psi)\, dt + \mathbb{E}[dV(W_t)] \right\} [dV(W_t)] \} . \tag{3}$$

Equivalently, we can divide both sides by $V(W_t)$ and note that this condition implies that the rate of return the agent earns in value terms is equal to her rate of time preference rdt. The agent's rate of return consists of the "dividend yield" component $(u(c)+\psi)dt/V(W_t)$ and the expected "capital gain" component $\mathbb{E}\left[dV(W_t)\right]/V(W_t)$.

Assuming that V is twice continuously differentiable, we can use Ito's lemma (see, e.g., Karatzas and Shreve 1998) to compute the expected time change in the value V, given that financial wealth follows

[8] For a standard textbook exposition of dynamic programming and Bellman equations, see Dixit (1990) or Kamien and Schwartz (1991). For additional details on the derivation of HJB equations in a related setting, see Grochulski and Zhang (2013).

(2) and the agent consumes at rate $c\,dt$:

$$\mathbb{E}[dV(W_t)] = \left((\mu W_t - c)V'(W_t) + \frac{1}{2}\sigma^2 W_t^2 V''(W_t)\right) dt. \tag{4}$$

Clearly, a higher consumption rate, $c\,dt$, implies lower wealth tomorrow, which, given that the agent's value function is increasing in financial wealth, $V' > 0$, implies a lower change in the agent's value going forward. On the other hand, as we see in (3), the agent's current utility flow is higher when c is higher. This trade-off is captured by the HJB equation for the value function V, which we obtain by substituting (4) into (3):

$$rV(W_t) = \max_c \left\{u(c) + \psi + (\mu W_t - c)V'(W_t) + \frac{1}{2}\sigma^2 W_t^2 V''(W_t)\right\}. \tag{5}$$

We will use this equation to find the value function V next.

Solution to the optimal consumption problem in retirement

With logarithmic utility of consumption, we can solve the retirement value problem in closed form.

Proposition 1 *Suppose $u(c) = \ln(c)$, then the retirement value function V is given by*

$$rV(W_t) = \ln(rW_t) + \psi + r^{-1}\left(\mu - r - \frac{1}{2}\sigma^2\right). \tag{6}$$

To verify the solution in (6), we compute its first and second derivative as

$$V'(W_t) = \frac{1}{rW_t} \quad \text{and} \quad V''(W_t) = \frac{-1}{rW_t^2}. \tag{7}$$

We then use $u = \ln$ and take the first-order condition with respect to c in (5):

$$\frac{1}{c} = V'(W_t), \tag{8}$$

which implies $u(c) = -\ln(V'(W_t))$. We can then write (5) as

$$rV(W_t) = -\ln(V'(W_t)) + \psi + \mu W_t V'(W_t) - 1 + \frac{1}{2}\sigma^2 W_t^2 V''(W_t).$$

When we substitute the derivatives in (7) to the right-hand side of the above equation, we verify (6).

If we suppose that financial wealth is a safe asset with the riskless growth rate equal to the agent's rate of time preference, i.e., if $\mu = r$

and $\sigma = 0$, then, due to the desire for consumption smoothing, the agent's optimal consumption policy would be to consume a constant amount at all dates in retirement. With wealth W, the maximal constant consumption the agent can afford is $c = rW$, which keeps the agent's wealth constant. The value the agent would attain in these conditions thus satisfies $rV(W) = \ln(rW) + \psi$.

Comparing this value with (6), we see that the constant $r^{-1}(\mu - r - \frac{1}{2}\sigma^2)$ represents the impact that the excess expected return $\mu - r$ and volatility σ have on the value the agent attains when the growth rate of wealth is risky, as in (2). Assumption 1 ensures that, despite the risk in the growth rate of the only asset available to the agent in retirement, she is able to attain a higher value than she would with a riskless asset. This attenuates the concern that our results are driven by our simplifying assumption of no riskless asset in the model.

Wealth effects in retirement

We can now use the closed-form solution for V to compute the agent's optimal consumption and the associated wealth effect in retirement. Using the first-order condition in (8) and the marginal value of wealth in (7), we obtain that the agent's optimal consumption satisfies

$$c_t = rW_t. \tag{9}$$

That is, the agent consumes a constant fraction of her wealth at all times. While this policy is the same as the one the agent would choose if the growth rate of wealth was riskless, the value the agent attains, of course, is lower due to risk.

With the agent's consumption function given in closed form in (9), we immediately obtain the wealth effect, or the marginal propensity to consume out of wealth, given as

$$\frac{dc_t}{dW_t} = r.$$

In a typical calibration, we would have $r = 0.05$, which implies a wealth effect of 5 cents on the dollar. This number is about twice what we estimated in Section 1. In the remainder of this article, we will argue that this number is made lower, and thus closer to the wealth effects observed in the data, when an endogenous labor supply decision is taken into account. In particular, we focus on the extensive margin of labor supply, i.e., work versus retirement.

4. WORKING FOREVER

Before we consider, in the next section, the wealth effects while the agent saves for retirement, in this section we discuss the option of never retiring, i.e., working forever. We argue that the plan to never retire is not optimal for the agent. In particular, we show that when the agent's wealth is high enough, being permanently retired is preferable to permanently working. This means that the agent's optimal retirement plan is a threshold policy, where the agent retires as soon as her wealth reaches a certain level, which we analyze in the next section.

While working, the agent earns labor income y, assumed here to be constant, and decides at each point in time how much to consume out of y and out of her stock of financial wealth W_t. The law of motion for the agent's financial wealth W_t is as follows

$$dW_t = (\mu W_t + y - c_t)dt + \sigma W_t dZ_t. \tag{10}$$

If, for example, the agent were to consume exactly her labor income at all times, i.e., if $c_t = y$, then her financial wealth would follow simply geometric Brownian motion (1).

Denote by $F(W)$ the maximal value that the agent whose wealth is W can obtain by never retiring, i.e., working and earning the flow of income y forever, while saving and consuming optimally. That is

$$F(W) := \max_{(c_t;t\geq 0)} \mathbb{E}\left[\int_0^\infty e^{-rt}u(c_t)dt\right] \text{ s.t. (10) and } W_0 = W.$$

The main result of this section is the following

Proposition 2 *There exists $\bar{W}$ such that $F(W) < V(W)$ for all $W \geq \bar{W}$.*

In words, when financial wealth is sufficiently high, the agent would rather be permanently retired than permanently working.

We argue this result by making use of an auxiliary result. Let us define and denote by V_0 the value of being retired but without the utility flow of leisure:

$$V_0(W) := \max_{(c_t;t\geq 0)} \mathbb{E}\left[\int_0^\infty e^{-rt}u(c_t)dt\right] \text{ s.t. (2) and } W_0 = W.$$

Because this function is a special case of the retirement value function V, one in which $\psi = 0$, we know that $V_0(W) = V(W) - \frac{\psi}{r}$. Note that this means that $V_0'(W) = \frac{1}{rW}$, which is the same as $V'(W)$ because the flow utility of leisure ψ enters V additively.

Our auxiliary result is as follows:

Lemma 1 $\lim_{W\to\infty} F(W) = V_0(W)$.

The following sketch of the formal argument for why this is true captures the intuition. Directly from the definition of function F and the function V_0, we see that these two values come out of maximizing the same objective subject to two different laws of motion for financial wealth. In the case of F, the law of motion includes labor income y. Like we did earlier for V, we can write an HJB equation for F as follows

$$rF(W_t) = \max_c \left\{ u(c) + F'(W_t)(\mu W_t + y - c) + \frac{1}{2}\sigma^2 W_t^2 F''(W_t) \right\}.$$

We can see that the contribution of y to F is $F'(W)y$, where $F'(W)$ represents the marginal valuation of y at wealth W. Next, we observe that this marginal valuation must be smaller than $V_0'(W)$, which is how income y would be valued by an agent who does not have it. Therefore, we must have $F'(W)y < V_0'(W)y = \frac{y}{rW}$, which goes to zero as W goes to infinity. Thus, the difference between F and V_0 must go to zero as wealth approaches infinity, which proves the lemma. Intuitively, when the stock of her financial wealth becomes larger and larger, whether or not the agent earns some constant labor income y matters less and less and becomes immaterial when financial wealth is sufficiently large.

Lemma 1 implies Proposition 2 because

$$V_0(W) + \frac{\psi}{r} = V(W).$$

Clearly, since $F(W)$ coverages to $V_0(W)$ as W becomes large, $F(W)$ must fall below $V_0(W) + \frac{\psi}{r}$ at some point. This point is represented by $\bar{W}$ in the statement of Proposition 2.

In sum, we have argued in this section that when the agent's financial wealth is high enough, she prefers being permanently retired to working forever. In the next section, we will use this fact when we define and solve the agent's problem of optimal saving for retirement.

5. SAVING FOR RETIREMENT

In this section, we define the consumption and saving problem with endogenous retirement and describe the method for solving it. Since retirement is irreversible in our model, the choice of the optimal timing of retirement is a real-option exercise problem similar to the exercise problems studied in the investment literature.[9]

Let us define and denote by $J(W)$ the maximum lifetime utility an agent with financial wealth W can attain by working and retiring at

[9] See, for example, Pindyck (1991).

some point in the future. That is,

$$J(W) := \max_{(c_t;t\geq 0),\tau} \mathbb{E}\left[\int_0^\tau e^{-rt}u(c_t)dt + e^{-r\tau}V(W_\tau)\right] \text{ s.t. (10) and } W_0 = W,$$

where τ is the agent's preferred retirement time, $V(W_t)$ is the value of retiring with wealth W_t, and the expectation $\mathbb{E}$ is taken over the realizations of the wealth growth shock process Z_t. As we see, the agent's utility flow before retirement does not include the value of leisure, ψ, and the law of motion for financial wealth accounts for the agent's labor income y.

As we did for V and F, we can use dynamic programing to obtain the following HJB equation for J

$$rJ(W_t) = \max_c \left\{u(c) + J'(W_t)(\mu W_t + y - c) + \frac{1}{2}\sigma^2 W_t^2 J''(W_t)\right\}. \tag{11}$$

Since HJB equations only account for the local flows of utility and changes in wealth, the HJB equations F and J are the same. The difference between the optimal value functions F and J comes from their boundary conditions, which we discuss next.

Boundary conditions and existence of optimal solution

At $W_t = 0$, the HJB equation simplifies to $rJ(0) = \max_c \{u(c) + J'(0)(y-c)\}$. The first-order condition for consumption with $u = \ln$ gives us, as before, $c = 1/J'(0)$ and $u(c) = -\ln(J'(0))$. Substituting to the HJB, we obtain a boundary condition for J:

$$rJ(0) = -\ln(J'(0)) + J'(0)y - 1. \tag{12}$$

Since we do not allow negative financial wealth, drift of W_t cannot be negative when $W_t = 0$. That is, the agent must be saving a part of her labor income when $W_t = 0$, i.e., we must have $c \leq y$, which gives us a restriction on the slope of J at zero: $J'(0) = \frac{1}{c} \geq \frac{1}{y}$.[10]

Using the boundary condition (12), we can solve the HJB equation forward from the boundary point $W = 0$.[11] This gives us a unidimensional family of solution curves $J(W)$, indexed by $J'(0) \geq \frac{1}{y}$.

[10] As a side note, we also know that $rJ(0) \geq u(y)$ because the agent with zero financial wealth has the option to consume $c = y$ and work forever.

[11] Because volatility of financial wealth is zero when wealth itself is zero, advancing the solution out of $W = 0$ presents a challenge for numerical ODE solvers. This challenge can be quite easily overcome by using local or global numerical approximation methods.

The analysis of the value of working forever, discussed in the previous section, implies here that solution curves $J(W)$ that remain strictly above the curve $V(W)$ for all W can be classified as inadmissible. Indeed, since V represents the value of being retired, if a solution curve J such that $J(W) > V(W)$ for all W were to represent the maximum utility value the agent could obtain, then it would be optimal for the agent to never retire, i.e., to set $\tau = \infty$. But we know from the previous section that the value of working forever, denoted by $F(W)$, is not above $V(W)$ for all W, which gives us a contradiction. Thus, admissible solutions J are those that satisfy the boundary condition (12) at $W = 0$ and satisfy $J(W) = V(W)$ for some W. The highest of the admissible solutions is the optimal one.

It can be shown that the maximum admissible J exists, that it satisfies $J(W) \geq V(W)$ for all W, and that the set of W on which $J(W) = V(W)$ consists of just a single point. Intuitively, if there is more than one point in this set, we can shift the curve J up to obtain a better admissible solution. We can continue this process until a single point remains in the set on which J overlaps with V. That point, which we will denote by W^*, is when the agent optimally chooses to retire.

The agent's optimal retirement time τ is thus a stopping time. That is, the agent works until her financial wealth attains the threshold W^* for the first time, and she retires at that time. The retirement date, thus, satisfies

$$\tau = \min\{t : W_t = W^*\}.$$

Geometrically, since W^* is the single point on which J and V overlap, in addition to the values $J(W^*)$ and $V(W^*)$ being the same, the first derivatives must match, $J'^*) = V'^*)$, and J must be less concave than V at W^*, i.e., J must have a less negative second derivative than V at the retirement threshold.[12]

In sum, the procedure for finding the optimal solution J is as follows. Start with some $J'(0)$ close to $1/y$ and use the boundary condition (12) to determine $J(0)$. Solve the HJB equation forward from zero. If $J'(0)$ is close enough to $1/y$, this solution will at some W cross the value of being retired, V. Increase $J'(0)$ to obtain a new solution curve, which is everywhere above the one obtained from the initial guess. Repeat until the solution J becomes tangent to V. The point of tangency becomes the optimal wealth retirement threshold W^*.

[12] Our construction of the optimal solution curve J provides a simple example of how the so-called smooth-pasting optimality conditions arise in many optimal control problems. For a much more extensive discussion of these conditions, see Dixit (1993).

6. THE ENDOGENOUS VALUE OF HUMAN CAPITAL

In this section, we discuss the impact of wealth shocks on the value of the agent's human capital when retirement is endogenous.

Generally, the value of one's human capital is defined as the expected present value of all of one's future labor earnings.[13] Following this definition and denoting the value of the agent's human capital at time t by H_t, we have

$$H_t := \mathbb{E}_t \left[\int_0^\tau e^{-rs} y ds \right],$$

where $\tau = \min\{t : W_t = W^*\}$ is the agent's retirement date. Because income y is constant in our model as long as the agent is working, we have

$$H_t = \frac{y}{r} \mathbb{E}_t \left[\int_0^\tau r e^{-rs} ds \right] = \frac{y}{r} (1 - \mathbb{E}_t \left[e^{-r\tau} \right]) = \frac{y}{r} (1 - G(W_t)), \quad (13)$$

where

$$G(W_t) := \mathbb{E}_t \left[e^{-r\tau} \right]$$

is the expected discount factor from the present until retirement.

Clearly, the value of the agent's human capital is endogenous with respect to her retirement decision. When her financial wealth hits W^*, the agent retires immediately, which means $\tau = 0$ and $G(W^*) = E_t \left[e^0 \right] = 1$. Thus, $H_t = 0$ when $W_t = W^*$. At the other extreme, suppose, as we did when we discussed the value function F, that the agent chooses to never retire, i.e., sets $\tau = \infty$. In this case, the value of the agent's human capital is constant at all dates and states: $H_t = \frac{y}{r}$.

As we see, the value of the agent's human capital depends on how soon she expects to retire. This time horizon, in turn, depends on the agent's saving rate and her current financial wealth. But, since savings are also determined by the financial wealth position, we can express the value of human capital as a function of the agent's financial wealth alone: $H_t = H(W_t)$. Note that the function H is decreasing: higher W_t means nearer retirement, i.e., a shorter remaining career, and thus a lower monetary value of human capital.

As we see in (13), the key part in the computation of $H(W)$ is the function G. We have the following

Lemma 2 *G satisfies $G(W^*) = 1$ and*

$$rG(W) = (\mu W + y - c(W))G'(W) + \frac{1}{2}\sigma^2 W^2 G''(W), \quad (14)$$

[13] See, e.g., the review article by Benzoni and Chyruk (2015).

where $c(W)$ is the agent's optimal consumption policy before retirement.

The differential equation in this lemma accounts for the agent's remaining time before retirement, taking into account the dynamics of her financial wealth position W_t. In order to capture these dynamics correctly, we need to know the agent's optimal consumption policy $c(W)$, which comes from the value function J characterized in the previous section. We provide a short formal proof of this lemma in the Appendix.[14]

7. SPECIAL CASE WITH NO RISK PREMIUM

In this section, we discuss briefly a special case in which $\mu = r$, i.e., the case in which aggregate wealth offers no risk premium. This case admits a closed-form solution to the problem of optimal consumption and saving for retirement and provides a clear illustration of our point that wealth effects are weaker when retirement is endogenous.

However, this special case violates Assumption 1. The results of this section, therefore, need to be taken with caution. In particular, the agent's only vehicle for saving in this section is an asset that offers zero excess return and positive risk. Such an asset would clearly be dominated by a riskless asset with the expected rate of return equal to the agent's rate of time preference. In practice, even if such a completely safe asset is not available, close substitutes are. To ensure that, despite this low risk-adjusted return of the financial asset, the agent wants to save for retirement, in this section we impose a stronger assumption on the value of leisure in retirement.

Assumption 2 $\psi > 1 + \frac{\sigma^2}{2r}$.

Under this assumption, as we show next, the agent wants to save for retirement even though the asset she can use is risky and offers no excess return.

Proposition 3 *Under Assumption 2, if $u = \ln$ and $\mu = r$, then J is linear:*

$$J(W_t) = V(W^*) + V'(W^*)(W_t - W^*), \tag{15}$$

where

$$W^* = \frac{y}{r\psi - \frac{1}{2}\sigma^2}.$$

[14] For a textbook treatment of expected hitting times in Brownian models, see Stokey (2009).

The agent's optimal consumption and total wealth are constant at all times prior to retirement:

$$c_t = rW^* \quad \text{and} \quad W_t + H(W_t) = W^* \quad \text{at all } t \leq \tau.$$

To verify the value function in (15), we can differentiate it and substitute to the HJB equation (11) using the retirement value function V given in (6).

The first-order condition for consumption, $u'(c_t) = 1/c_t = J'(W_t) = V'(W^*) = rW^*$, confirms that the agent consumes a constant amount prior to retirement. The terms in the expression for W^* show how the agent's retirement wealth threshold depends on the parameters. Assumption 2 implies that the agent saves: $c_t < y$ at all dates prior to retirement. Indeed, $c_t = rW^* = \frac{y}{\psi - \frac{\sigma^2}{2r}} < y$.

Proposition 3 also shows that the value of the agent's human capital offsets perfectly any shocks to financial wealth as the agent saves for retirement, leaving the agent's total wealth, $W_t + H(W_t)$, constant. To verify that indeed $H(W_t) = W^* - W_t$ at all t, we use Lemma 2 to compute the expected discount factor until retirement, $G(W)$. In particular, we can guess and verify by differentiation and substitution to (14) that

$$G(W) = 1 - \frac{r}{y}(W^* - W). \tag{16}$$

Substituting this expression for $G(W)$ into (13) yields the result.

This result helps us understand why consumption prior to retirement is constant: the agent's total wealth is constant. The agent desires to smooth her consumption of the consumption good but not necessarily her consumption of leisure. For this reason, the shocks to her financial wealth position are absorbed by her supply of labor, on the extensive margin, which allows her to keep total wealth constant and her consumption perfectly smoothed. Clearly, negative shocks to the return on financial wealth still hurt the agent by making her postpone retirement but do not affect her consumption. Likewise, positive shocks improve the agent's value, which we see in (15), where the value function J is strictly increasing in financial wealth.

We conclude this section by noting that constant consumption $c_t = rW^*$ for any $W_t < W^*$ implies the wealth effect is nil:

$$\frac{dc_t}{dW_t} = 0.$$

With no consumption response, the whole financial return risk is absorbed by the agent's present value of future utility from leisure. Indeed, we can substitute (16) and (6) into (15) to express the agent's value function prior to retirement as

$$rJ(W) = u(rW^*) + G(W)\left(\psi - \frac{\sigma^2}{2r}\right).$$

This formula shows that the expected present value of the agent's consumption, $u(rW^*)$, does not at all respond to changes in financial wealth, and J depends on W only through the expected present value of the utility from leisure in retirement, $G(W)\psi$, and the penalty for risky consumption in retirement, $-G(W)\sigma^2/2r$.

The zero wealth effect result, although too extreme relative to the wealth effects measured in Section 1, highlights our main point: the endogenous response of labor supply to realized wealth shocks attenuates the impact of these shocks on consumption. In the next section, we calibrate our model to obtain realistic wealth effects consistent with those measured in Section 1.

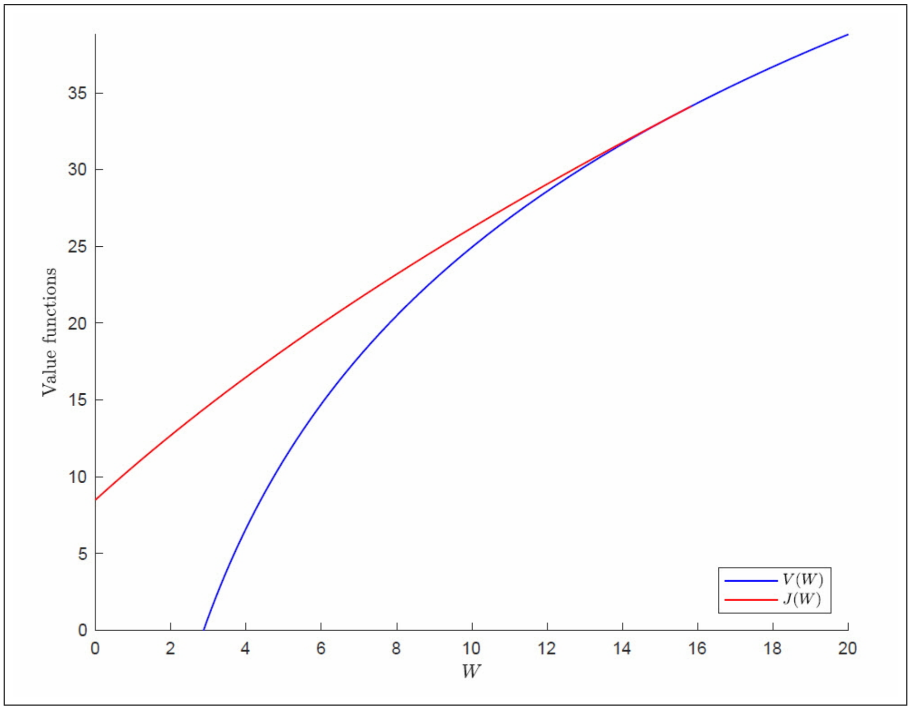
8. CALIBRATION FOR REALISTIC WEALTH EFFECTS

In this section, we show that this model can generate realistic wealth effects of about 2.4 cents on the dollar, as measured in Section 1. We calibrate the model allowing for positive risk premium $\mu > r$ and solve it numerically. We present optimal value functions, consumption policies, and wealth effects.

We normalize the labor income flow $y = 1$ and use the standard rate of time preference of $r = 0.05$. In order to pick parameters μ and σ, we use the average growth rate and the standard deviation of the growth rate of aggregate net wealth presented in Table 1. We will not ask this stylized model to also match the aggregate consumption statistics given in Table 1. Instead, we approximate consumption by using the simple dynamics our model predicts in retirement. That is, we approximate consumption as $c_t = rW_t$ and substitute it to the law of motion for wealth in retirement (2), which implies the average growth rate of wealth after consumption of $\mu - r$. We calibrate this to match the 3.4 percent reported in Table 1, which gives us $\mu = 0.084$. The volatility parameter for wealth after consumption we calibrate to match the standard deviation of net wealth reported in Table 1, which gives us $\sigma = 0.044$. Finally, the parameter ψ is chosen to match our target of the average wealth effects of 2.4 cents on a dollar. The calibrated value of ψ is 1.28.

Under this parametrization, we solve the model numerically and obtain the optimal value function J using the procedure described in Section 5. The solution, along with the retirement value function V characterized in Proposition 1, is plotted in Figure 2. Clearly, as shown

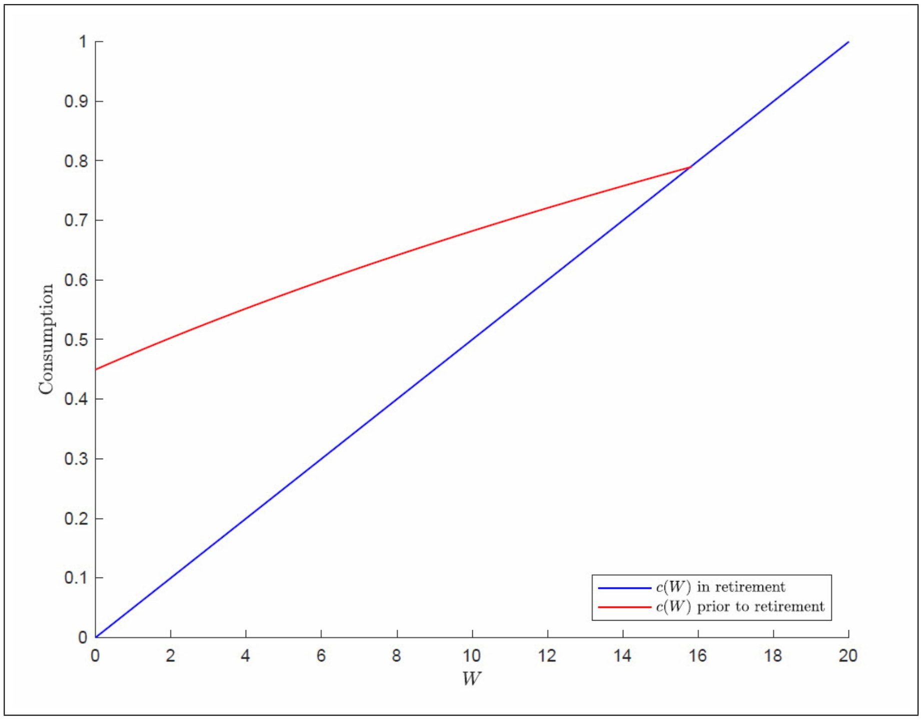
Figure 2 Pre- and Postretirement Value Functions



by the vertical difference between J and V, the agent's labor income y is most valuable to her when her financial wealth is low. Positive labor income allows the agent to better insulate her consumption from the wealth return shocks Z_t, consistent with J being everywhere flatter and less concave than V. The two value functions paste together at the retirement threshold of $W^* = 15.8$, which means the agent retires when her financial wealth reaches roughly sixteen times annual income. This number, which did not target in the calibration, may seem high, but this is to be expected of our infinitely lived agent model.

The optimal consumption functions prior to and after retirement are plotted in Figure 3. Consistent with previous observations, consumption is flatter, i.e., less responsive to wealth return shocks, prior to retirement. This shape of the optimal consumption function implies lower wealth effects prior to retirement. In retirement, $c(W)$ has a constant slope equal to r. Prior to retirement, this slope is less positive and not exactly constant.

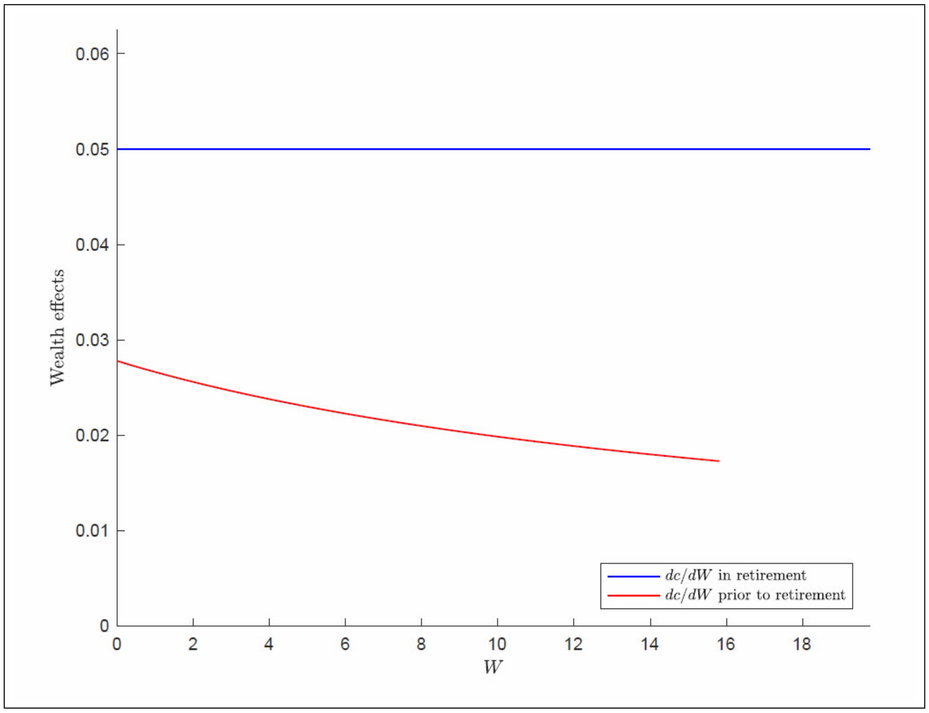
Figure 3 Optimal Consumption



The exact wealth effects, i.e., the slopes of the two consumption functions, are plotted in Figure 4. While constant in retirement, the wealth effects are decreasing in financial wealth prior to retirement, ranging from almost 3 cents on the dollar when wealth is low to a bit below 2 cents when wealth is high. The average of these effects matches the number estimated in Section 1 of 2.4 cents on the dollar. When the agent retires, the wealth effect jumps to 5 cents. This discontinuity comes from the irreversible switching of the agent's state from working to retired as wealth reaches the threshold W^*.

The intuition for why wealth effects are weaker prior to retirement than after is similar to that of the special case discussed in the previous section: when retirement is endogenous, changes in the value of human capital offset the shocks to the agent's financial wealth, keeping her total wealth relatively more stable, which dampens the response of consumption. In the special case discussed earlier, the offsetting response of the value of human capital was one-to-one: total wealth was constant, the wealth effect for consumption was nil, and the financial

Figure 4 Wealth Effects

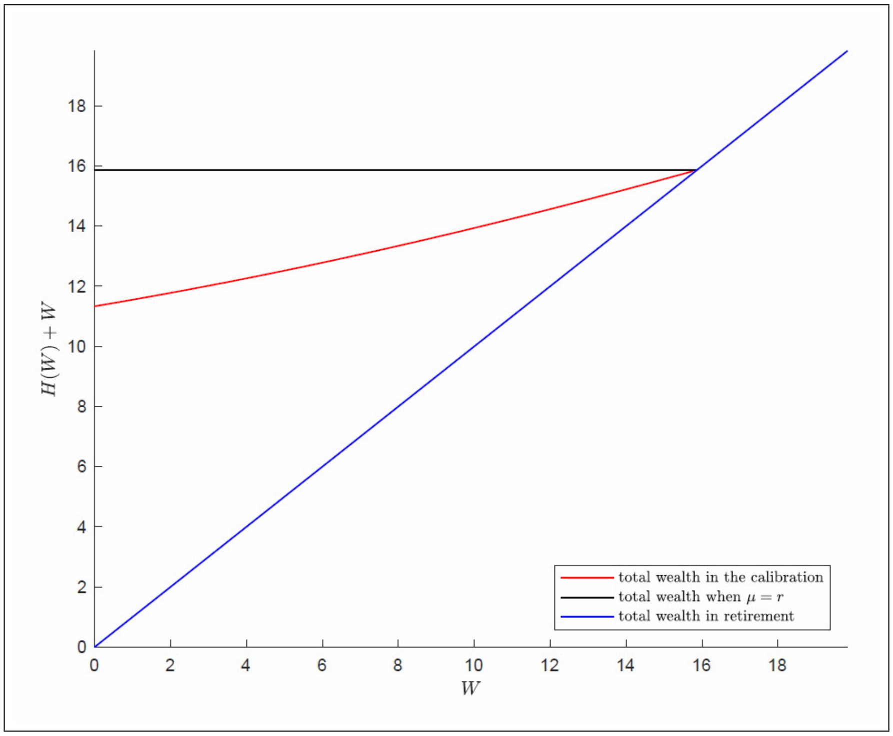


return shocks were fully absorbed by the changes in expected utility from leisure. Here, the offsetting effect is present but smaller than one-to-one, which gives rise to wealth effects for consumption that are positive but smaller than r.

The decreasing pattern of wealth effects prior to retirement, observed in Figure 4, can be understood in terms of the less than one-to-one response of the value of human capital and the associated change in the expected present value of leisure in retirement. When the agent's financial wealth is low, she plans to work for a long duration, and the present value of her retirement leisure is low. By adjusting her already far-off retirement date, the agent can change the value of her human capital by a little and thus is able to transmit only a small portion of the financial return risk to the value of her retirement leisure, which leaves her consumption relatively more exposed and makes wealth effects relatively strong.

When the agent's wealth is close to the retirement threshold, she plans to work for just a short time, and the present value of her retire-

Figure 5 Total Wealth



ment leisure is high. By adjusting her retirement date in response to realized financial return shocks, the agent can change the value of her human capital drastically and thus is able transmit a lot of the financial return risk to the value of her retirement leisure, which insulates her consumption better, yielding a weaker wealth effect. Intuitively, the value of human capital is "less endogenous" when retirement is far off and "more endogenous" when retirement is near.

As the agent retires, she loses completely the ability to adjust the value of her human capital or the present value of her retirement leisure. This means all the financial return risk must now be absorbed by consumption, leading to a jump in the wealth effects at retirement.

Figure 5 plots the agent's total wealth, $H(W_t)+W_t$. In the baseline example with no risk premium, as we saw in Section 7, the endogenous adjustment of the value of human capital perfectly offsets the shocks to financial wealth, so the agent's total wealth is constant, equal to W^*, at all times prior to retirement, and, consequently, the wealth effect is nil. In retirement, the value of human capital is fixed at zero forever, i.e., it cannot respond at all to shocks to financial wealth, and thus the effect of these shocks on consumption is strong. In the calibrated model, prior to retirement, we see that the agent's total wealth is a less steep function of W than it is in retirement, but it is more steep than prior to retirement in the baseline case of $\mu = r$. Correspondingly, the wealth effect in the calibration is not as strong as in retirement but remains positive.

The same intuition also helps us understand comparative statics with respect to the retirement leisure utility flow ψ. The comparative statics result, which we can show numerically, is that higher ψ leads to a lower average wealth effect. Intuitively, high ψ means a lower threshold W^*, faster desired retirement, and thus a "more endogenous" value of human capital. With the value of human capital reacting to the wealth shocks more strongly, the agent can stabilize her consumption better, which explains a weaker average wealth effect of the financial return shock on consumption.

9. CONCLUSION AND FURTHER READING

In this article, we study a simple model of optimal consumption, saving, and retirement decisions. We use this model to show that the endogenous response of the agent's planned career length and the associated change in the value of her human capital dampen the response of consumption to financial return shocks. In a reasonable parametrization, the model is able to generate wealth effects of the magnitude observed in the data.

We make several simplifying assumptions that do not affect our main result. First, we abstract from portfolio choice, which makes the consumption-saving-retirement problem very easily tractable. Farhi and Panageas (2007) study a more general problem with an endogenous portfolio choice as in Merton (1971) and show that wealth effects are weakened by the endogeneity of the retirement decision also in their model. In addition, they show that the portfolio weight of risky assets is increased by the inclusion of this margin.

Second, we abstract from shocks to labor income and labor supply decisions on the intensive margin. It is easy to see that exogenous labor income risk would change the retirement threshold without affecting the result that the endogeneity of the value of human capital moderates wealth effects. Liu and Neis (2002) and Farhi and Panageas (2007) discuss the case with endogenous response of labor supply along the intensive margin. They show that wealth effects are lower than in retirement but not necessarily decreasing in wealth, depending on assumptions.

Third, we follow Kingston (2000) and Farhi and Panageas (2007), among many others, in modeling retirement as an absorbing state with zero hours worked, which is a strong assumption. Ruhm (1990) and Rust and Phelan (1997) show that less than 40 percent of workers retire from full-time career jobs by exiting the labor force completely. About half of the workers transition from full-time career jobs to part-time retirement "bridge" jobs. About 25 percent of workers experience episodes of reemployment after retirement.[15] The assumption of a single transition into full retirement, clearly, helps keep our analysis simple. The intuition that arises out of our model strongly suggests, however, that more flexible retirement would help us generate low wealth effects. In particular, an option to unretire would make the value of the agent's human capital "more endogenous," allowing it to better absorb the financial return shocks and thus helping the agent stabilize her consumption, which would dampen the wealth effects of the financial return shock on consumption.

Further, we abstract from mortality risk, the bequest motive, and the possible nonseparability of utility between consumption and leisure. Dybvig and Liu (2010) include these features and show that the agent's consumption and portfolio allocation jump at retirement. Farhi and Panageas (2007) and Dybvig and Liu (2010) also allow for borrowing, i.e., negative wealth, subject to appropriate borrowing constraints. Each of these assumptions changes the results quantitatively but not qualitatively.

In other areas of economics, the endogeneity of the retirement decision has been shown to be important when assessing magnitudes of various economic forces. For example, Rogerson and Wallenius (2013) show its importance for measurement of the intertemporal elasticity of substitution for labor supply. Ndiaye (2018) shows its importance for calibrating optimal labor income tax rates and Social Security benefits.

[15] However, based on evidence from the Panel Study of Income Dynamics and the Current Population Survey, Rogerson and Wallenius (2013) argue that an abrupt transition from full time to little or no work approximates well the process of retirement for male heads of households in the US.

APPENDIX: PROOF OF LEMMA 2

Define a bounded random variable $M := \int_0^\infty re^{-rt}1_{t<\tau}dt$ and a martingale $M_t := \mathbb{E}_t\,[M]$. We have $M_t = \int_0^t re^{-rs}ds + e_t^{-rt}\mathbb{E}_t[\int_0^\infty re^{-rs}1_{t<\tau}ds] = 1 - e^{-rt}G(W_t)$. The drift of M_t is $re^{-rt}G(W_t) - e^{-rt}G'(W_t)(\mu W_t + y - c(W_t)) - \frac{1}{2}\sigma^2 W^2 e^{-rt}G''(W_t)$, which must be zero. QED

REFERENCES

Benzoni, Luca, and Olena Chyruk. 2015. "The Value and Risk of Human Capital." *Annual Review of Financial Economics* 7: 179–200.

Bewley, Truman. 1977. "The Permanent Income Hypothesis: A Theoretical Formulation." *Journal of Economic Theory* 16 (December): 252–92.

Cagetti, Mariacristina, and M. De Nardi. 2008. "Wealth Inequality: Data and Models." *Macroeconomic Dynamics* 12 (September): 285–313.

Calomiris, Charles W., Stanley D. Longhofer, and William Miles. 2012. "The Housing Wealth Effect: The Crucial Roles of Demographics, Wealth Distribution and Wealth Shares." Working Paper 17740. Cambridge, Mass.: National Bureau of Economic Research (January).

Dixit, Avinash K. 1990. *Optimization in Economic Theory, Second Edition*. Oxford: Oxford University Press, 164.

Dixit, Avinash K. 1993. *The Art of Smooth Pasting*. New York: Gordon and Breach Publishing Group, 88.

Dybvig, Philip H., and Hong Liu. 2010. "Lifetime Consumption and Investment: Retirement and Constrained Borrowing." *Journal of Economic Theory* 145 (May): 885–907.

Farhi, Emmanuel, and Stavros Panageas. 2007. "Saving and Investing for Early Retirement: A Theoretical Analysis." *Journal of Financial Economics* 83 (January): 87–121.

Friedman, Milton. 1957. *A Theory of the Consumption Function*. Princeton: Princeton University Press.

Grochulski, Borys, and Yuzhe Zhang. 2013. "Saving for Retirement with Job Loss Risk." Federal Reserve Bank of Richmond *Economic Quarterly* 99 (First Quarter): 45–81.

Hall, Robert E. 1978. "Stochastic Implications of the Life Cycle–Permanent Income Hypothesis: Theory and Evidence." *Journal of Political Economy* 86 (December): 971–87.

Iacoviello, Matteo M. 2011. "Housing Wealth and Consumption." Board of Governors International Finance Discussion Paper 1027 (August).

Kamien, Morton I., Nancy L. Schwartz. 1991. *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management, Second Edition.* Amsterdam: Elsevier, 261.

Kaplan, Greg, and Giovanni Violante. 2014. "A Model of the Consumption Response to Fiscal Stimulus Payments." *Econometrica* 82 (July): 1199–239.

Karatzas, Ioannis, and Steven E. Shreve. 1998. *Brownian Motion and Stochastic Calculus, Second Edition.* New York: Springer Science+Business Media, Inc.

King, Mervyn, and David Low. 2014. "Measuring the 'World' Real Interest Rate." Working Paper 19887. Cambridge, Mass.: National Bureau of Economic Research (February).

Kingston, Geoffrey H. 2000. "Efficient Timing of Retirement." *Review of Economic Dynamics* 3 (October): 831–40.

Lettau, Martin, and Sydney Ludvigson. 2004. "Understanding Trend and Cycle in Asset Values: Reevaluating the Wealth Effect on Consumption." *American Economic Review* 94 (March): 276–99.

Liu, Jun, and Eric Neis. 2002. "Endogenous Retirement and Portfolio Choice." UCLA Working Paper.

Merton, Robert C. 1971. "Optimum Consumption and Portfolio Rules in a Continuous-Time Model." *Journal of Economic Theory* 3 (December): 373–413.

Ndiaye, Abdoulaye. 2018. "Flexible Retirement and Optimal Taxation." Federal Reserve Bank of Chicago Working Paper 2018-18 (November).

Piazzesi, M., and M. Schneider. 2016. "Housing and Macroeconomics." In *Handbook of Macroeconomics*, Vol. 2, edited by John B. Taylor and Harald Uhlig. Amsterdam: Elsevier, 1547–640.

Pindyck, Robert S. 1991. "Irreversibility, Uncertainty, and Investment." *Journal of Economic Literature* 29 (September): 1110–48.

Poterba, James M. 2000. "Stock Market Wealth and Consumption." *Journal of Economic Perspectives* 14 (Spring): 99–118.

Rogerson, Richard, and Johanna Wallenius. 2013. "Nonconvexities, Retirement, and the Elasticity of Labor Supply." *American Economic Review* 103 (June): 1445–62.

Ruhm, Christopher J. 1990. "Bridge Jobs and Partial Retirement." *Journal of Labor Economics* 8 (October): 482–501.

Rust, John, and Christopher Phelan. 1997. "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets." *Econometrica* 65 (July): 781–831.

Saez, Emmanuel, and Gabriel Zucman. 2016. "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data." *Quarterly Journal of Economics* 131 (May): 519–78.

Stokey, Nancy L. 2009. *The Economics of Inaction*. Princeton: Princeton University Press.

Zhao, Bo. 2018. "Too Poor to Retire? Housing Prices and Retirement." *Review of Economic Dynamics* 27 (January): 27–47.