The University of Osaka
Institutional Knowledge Archive

Title	Sentiment Analysis for the Financial Market
Author(s)	數見, 拓朗
Citation	大阪大学, 2017, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/61470
rights	
Note	やむを得ない事由があると学位審査研究科が承認したため、全文に代えてその内容の要約を公開しています。全文のご利用をご希望の場合は、大阪大学の博士論文についてをご参照ください。

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

様式３

論　文　内　容　の　要　旨

氏　名　（　　數見拓朗　　）	
論文題名	Sentiment Analysis for the Financial Market （金融資産市場のセンチメント分析）

論文内容の要旨

The field of sentiment analysis is gaining increasing interest in both academia and business. As sentiment reflects the atmosphere of economic activities and the psychology of economic agents, analyzing sentiment can help us to understand the economy and securities markets in a more sophisticated manner. This thesis aims to quantify the market sentiment as four indexes and examine whether they can help predict stock prices and real estate prices in the Japanese market. We construct four sentiment indexes that quantify the positive or negative emotion in articles from the *Nikkei*, the most popular business newspaper in Japan. We found that the sentiment indexes significantly predict asset pricing in the Japanese stock and real estate markets.

The chapters of this thesis are organized as follows.

Chapter 1 describes the background and aims of our thesis. We show some theories attempt to support econometric analysis for evidence of market anomalies. We clarify the importance of sentiment analysis in econometric analysis.

Chapter 2 introduces some semantic dictionaries essential to construct the sentiment indexes. These dictionaries are well-known in the field of the sentiment analysis and Natural Language Processing. Natural Language Processing, a field of computer science, is concerned with enabling computers to derive meaning from natural languages. We survey the methodologies used to extract sentiment from text and some well-known semantic dictionaries in sentiment analysis.

Chapter 2 elaborates on how to construct market sentiment indexes and provides the summary statistics. There is no universal viewpoint on what data set should be employed and how sentiment index should be generated. We construct real-valued and integer-valued sentiment index, and compare the predictive power of each. In addition, as we will discuss in Chapter 5, we create a new sentiment index only reflecting economics-and-business-related news. We employ the *Nikkei* as a news source reflecting Japanese sentiment.

Chapter 4 describes several methods often employed in the field of time-series analysis and our experimental designs to examine the predictability of the Japanese stock and real estate markets when using the sentiment indexes. More specifically, we test for the *Granger causality*, impulse responses, and out-of-sample predictions to validate the predictability of the markets.

Chapter 5 demonstrates the experimental results we achieved using the methodologies described in the previous chapter 4. First, we build a single 31-year VAR model with the SIs constructed from general news, and clarify long-term correlations between the SIs and the Nikkei 225. Then, we build VAR models year-by-year, and interpret the estimates of VAR models using Japanese business cycles. Second, we explore whether the SIs can help us predict real estate prices as well as stock prices. Third, we employ a SI constructed only from economics-and-business-related news using a topic model which can discover hidden topics in each document. We conclude that our SIs can be used to predict asset pricing in Japanese markets.

Chapter 6 concludes this thesis.

様式７

論文審査の結果の要旨及び担当者

氏　名　（　敷見　拓朗　）		
論文審査担当者	（職）	氏　名
	主 査　教 授	大西　匡光
	副 査　教 授	太田　亘
	副 査　准教授	松村　真宏

論文審査の結果の要旨

［論文内容の要旨］

本論文の研究は，研究の背景・動機・目的を述べ，そして過去の研究のサーベイを行った，導入部分となる第１章，本論となる第２，３，４章，そして，結論と今後の研究課題を述べた第５章の計５章から構成される．

以下，本論にあたる第２，３，４章についての要旨を述べる．

本論文では，先ずは，我が国の代表的な経済新聞である日本経済新聞に掲載された見出しと記事に基づいて，日本経済のセンチメントを反映したセンチメント・インデックス（Sentiment Index，以下，SI）の作成を提案している．SIの構築手順は，１．形態素解析，２．単語を抽出する対象範囲，２．日本語極性辞書，３．ネガティブ・ポジティブな感情を計量化する方法，からなるが，本研究では，以下の観点・範囲から，複数種のSIの作成を試みている：

- Takamura (2005) の評価極性辞書（単語感情極性対応表）を使用
- 1984年から2014年までの31年間の日次・月次SI
- センチメント・スコアリングと対象範囲に応じたインデックス
- 読者が想起する感情の大きさを反映
- 新聞の見出しと記事本文ごとのSIの作成
- 世の中の全体の雰囲気を表すSIとトピックごとのSIの作成

次いで，それら各種SIを時系列モデルのVAR (Vector Auto-Regressive model) の変数に組み入れ，下記の統計分析を通じて，我が国の株価（株価指数日経225，TOPIX）と不動産価格（不動産投資信託J-REIT，実物不動産価格）の予測可能性について，実証的に検討している：

- VARモデルのパラメータ推定
- 頑健推定量を用いたGrangerの因果性検定
- インパルス応答関数
- アウト・オブ・サンプルにおける予測精度

より具体的には，下記の実証分析を行っている：

- 日経225に関する日次分析
- 日経225に関する年次分析
- TOPIXとJ-REITに関する日次分析
- 日本の資産市場における月次分析
- 記事のトピックごとのセンチメント・インデックスの日本株式市場への影響

それらの実証分析の結果，日次と月次の両観測頻度において，我が国の株価指数（日経225，TOPIX），J-REIT，実物不動産価格のいずれに対しても，トータル・スコアに基づく実数値型の見出しSIと記事SIは，資産価格に対して頑健なGranger因果性を持つことを結論している．また，経済ニュースに限定したSIは，リターン・リバーサルが観測されないこと，トピックごとのSIは我が国の株価に対する影響が異なること，事後的にはビジネス・サイクルが因果性に影響を与えていること，などの知見を得ている．

［審査結果の要旨］

本論文の研究は，我が国の代表的な経済新聞である日本経済新聞に掲載された見出しと記事に基づいて，日本経済のセンチメントを反映した独自のSIを複数種作成した上で，それらを変数に組み入れた時系列モデルのVARを構成して，日経225，TOPIX，J-REIT，等の資産価格の予測可能性についての実証分析を行った意欲的な研究であり，多くの経済学的に興味深く，新しい知見を得ていて，高く評価できる．

経済新聞に代表される伝統的なメディアと各種 SNS（Social Networking Service）との差違，SIの精緻化，分野特有の日本語極性辞書の選択，資産価格のボラティリティ変動との関連，ビジネス・サイクルの因果性への影響についての経済学的解釈，等々，今後に解決・吟味・検討すべき多くの問題が残されているものの，審査担当者らは，上記の貢献から，本論文を博士（経済学）の学位を授与するに値するものと判断する．