



日本銀行ワーキングペーパーシリーズ

通貨オプション市場における投資家センチメントの要因分析：機械学習アプローチ

鷺見和昭*

kazuaki.washimi@boj.or.jp

No.20-J-8
2020年10月

日本銀行
〒103-8660 日本郵便（株）日本橋郵便局私書箱30号

* 金融市場局

日本銀行ワーキングペーパーシリーズは、日本銀行員および外部研究者の研究成果をとりまとめたもので、内外の研究機関、研究者等の有識者から幅広くコメントを頂戴することを意図しています。ただし、論文の中で示された内容や意見は、日本銀行の公式見解を示すものではありません。

なお、ワーキングペーパーシリーズに対するご意見・ご質問や、掲載ファイルに関するお問い合わせは、執筆者までお寄せ下さい。

商用目的で転載・複製を行う場合は、予め日本銀行情報サービス局(post.prd8@boj.or.jp)までご相談下さい。転載・複製を行う場合は、出所を明記して下さい。

通貨オプション市場における投資家センチメントの要因分析： 機械学習アプローチ*

鷺見 和昭†

2020年10月

【要旨】

本稿では、ドル／円に関する主体別にみた投資家センチメントを分析するため、日本における取引情報蓄積機関や金融機関からの通貨オプション取引報告データを用いて、ランダムフォレストおよび勾配ブースティング (XGBoost) による要因分析を行った。投資家センチメント指数は、月内の新規取引 (想定元本ベース) のうち「コールオプションの買い・プットオプションの売り (ドル／円の上昇を予測)」と「プットオプションの買い・コールオプションの売り (ドル／円の下落を予測)」のポジションの差分によって構築した。

分析結果をみると、米国のイールドカーブは事業法人・機関投資家双方のセンチメントに影響する一方、米国の通商政策に関する不確実性は事業法人のセンチメントへの影響が大きいことが窺われる。さらに、米国のイールドカーブと米国の通商政策に関する不確実性の変数重要度や他の変数との交互作用は、VIX 指数以上に大きく、モニタリングの観点からは、これらの様々な不確実性指標を点検する必要があることが示唆された。今後の課題としては、データの蓄積およびクレンジング精度の向上が挙げられる。

JEL 分類記号 : F31、G11、G23

キーワード : 機械学習、通貨オプション、高粒度データ、投資家センチメント

* 本稿は、2020年人工知能学会全国大会・2020年統計関連学会連合大会において報告された論文を改訂したものである。本稿の作成にあたり、上記の学会参加者および日本銀行スタッフより、有益なコメントを頂戴した。記して感謝の意を表したい。本稿のあり得べき誤りは筆者個人に帰する。なお、本稿の内容や意見は、筆者個人に属するものであり、日本銀行の公式見解を示すものではない。

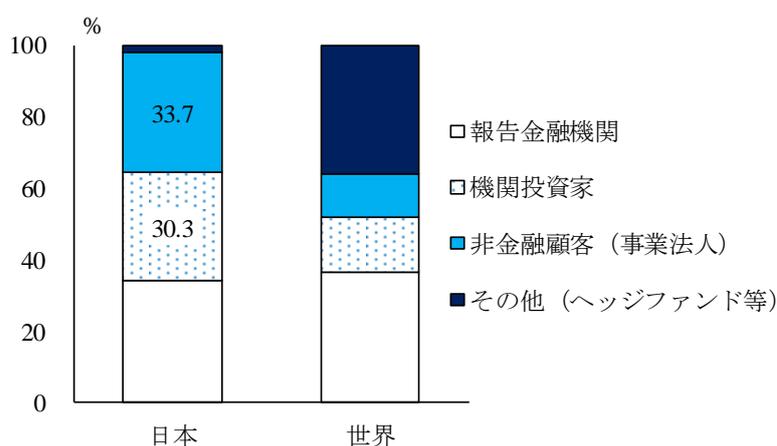
† 日本銀行金融市場局 (Email: kazuaki.washimi@boj.or.jp)

1. はじめに

通貨オプション取引は、リスク・リバーサルに代表されるように市場参加者のリスク認識（通貨の上昇／下落に関する予測）を抽出できることが知られている¹。もっとも、店頭デリバティブ取引については、取引所を通さない相対取引であるため、カウンターパーティー情報の取得が困難であり、これまで日本の通貨オプション取引について、主体別にみたリスク認識（以下、投資家センチメント）を分析した事例は、著者の知る限り存在しない。

例えば、通貨オプション市場に関する代表的な統計として知られる「外国為替およびデリバティブに関する中央銀行サーベイ」²によれば、ドル／円に関しては、日本の通貨オプション取引高は世界の約4分の1を占めている（2019年4月時点）。また、カウンターパーティー別にみると、日本では事業法人・機関投資家がそれぞれ取引高のうち約3分の1を占めている一方、グローバルではヘッジファンド等のプレゼンスが大きいなどの差異が窺われる（図表1）。このように、一口に投資家センチメントと言っても、地域ごとに投資家特性が異なっており、投資家行動の把握にあたっては、主体別にみた分析が重要と考えられる。本稿では、こうした問題意識のもと、ドル／円の通貨オプション取引データを用いて、主体別にみた投資家センチメントの構築および要因分析を行った。

（図表1） 報告金融機関からみた通貨オプション取引（ドル円）の取引相手（19/4月時点）



（資料） BIS 「外国為替およびデリバティブに関する中央銀行サーベイ」

¹ リスク・リバーサルの解説については、加藤・福永・山田（2012）や ECB（2003）を参照。

² 国際決済銀行（BIS）では3年に一度、世界の中央銀行の協力を得て、「外国為替およびデリバティブに関する中央銀行サーベイ」を行っている。同調査は世界の主要な金融機関を対象とし、包括的な項目を、国際的に整合性のとれた形で調査したものである。

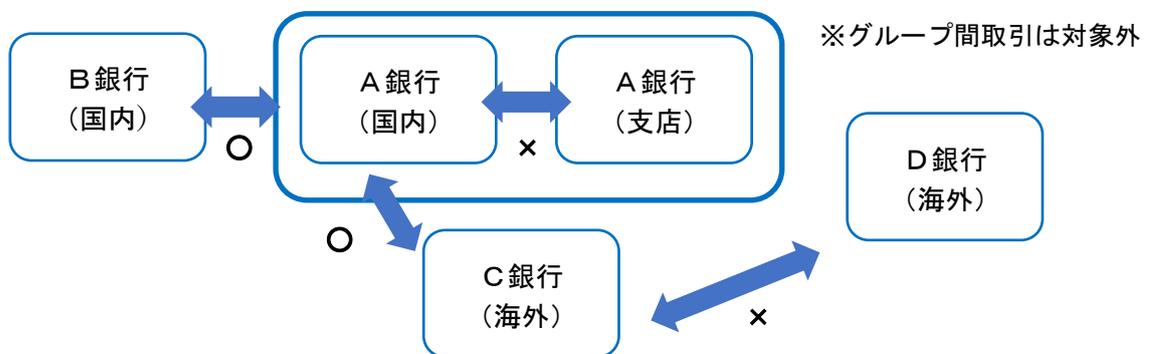
本稿の構成は、次のとおりである。2節では、分析対象データを整理する。3節では、分析手法について紹介する。4節では、実証分析の結果を説明したうえで、解釈を提示する。5節はまとめである。

2. データ

(1) 店頭デリバティブ取引データ

利用するデータは、取引情報蓄積機関（Trade Repository）および金融機関からの報告に基づく日本における店頭デリバティブ取引データを用いている。取引の対象範囲は、取引者のうち少なくとも一方が本邦居住者の取引となっている（図表 2）。本稿では、取引の匿名性を確保するため、カウンターパーティー情報に基づき、マーケットメーカーを除く事業法人・機関投資家の2分類を対象とし、高粒度の日次取引データを基に月次データを作成している。通貨ペアはドル／円とし、データクレンジングの観点から、カウンターパーティー情報の欠落および行使価格の異常値を判定し、当該取引を取り除いた上で、サンプル期間は2015年4月から2020年6月までとした。

（図表 2）店頭デリバティブ取引データの対象範囲（A銀行からみた場合）



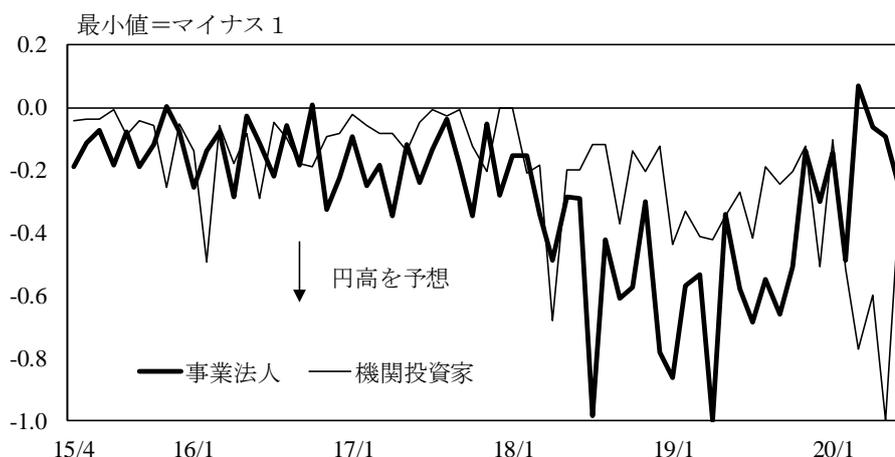
店頭デリバティブ取引データは、リーマン破綻後、G20 ピッツバーグサミットにおいて、その重要性が再認識され、各国当局（日本の場合、金融庁）がデータ収集を始めている。もっとも、データ量が膨大であり、データクレンジングに相応の時間とコストがかかるほか、個別取引の秘匿性確保が必要とされている（BIS（2018）、IAG（2017））。このため、分析結果の公表は海外では徐々に広が

りつつあるものの³、日本では限定的となっている。この点、本研究は日本の店頭デリバティブ取引データを用いた分析のパイロットケースと考えられる。

(2) センチメント指数の構築

通貨オプションには、コールオプションとプットオプションが存在するほか、取引主体からみた買いと売りを区別する必要がある。すなわち、投資家がドル／円の上昇を予想している場合、コールオプションの購入とプットオプションの売却が考えられる。このため、センチメント指数は、月内の新規取引（想定元本ベース）のうち「コールの買い・プットの売り（ドル／円の上昇を予想）」から「プットの買い・コールの売り（ドル／円の下落を予想）」を差し引くことで、構築している⁴。同センチメント指数をみると、事業法人・機関投資家ともに、基本的にはリスク・リバーサルと同様に、ドル／円の下落を警戒したポジションに偏っている傾向が窺われる（図表3）。

(図表3) センチメント指数



³ 例えば、各国・地域の店頭デリバティブ市場に関しては、HKMA (2015)、RBA (2018)、ESMA (2018) が市場構造を紹介しているほか、Abad et al. (2016)、Cielinska et al. (2017)、D'Errico et al. (2018) 等の分析事例が存在する。

⁴ 同指数は、リスク・リバーサル戦略（権利行使価格の異なるコールオプションとプットオプションを同時に反対売買するような取引）の場合、通貨の上昇／下落のどちらに備えたポジションかが反映されるものの、ストラドル戦略（同じ権利行使価格のコールオプションとプットオプションを組み合わせることで得られるポジション）の場合、ポジションが相殺されてしまう。もっとも、ストラドル戦略のように、相場変動の大小にベットするポジションは短期筋が構築するケースが多く、本稿の分析への影響は少ないとみられる。

なお、海外では、ECB (2019) が店頭デリバティブ取引データのうち金利スワップのポジションデータをもとに、同様に市場参加者の先行きの金利観を示す指数を作成している。

3. 分析手法

前節で構築したセンチメント指数がどのような要因によって影響を受けるかを分析するため、ランダムフォレスト⁵および勾配ブースティング (XGBoost⁶) を用いる (図表 4)。分析の主眼は予測精度ではなく、変数の関係性の探索であるため、変数重要度や部分従属プロットによる直感的な解釈可能性を重視している⁷。また、ランダムフォレストは、通常の回帰分析に比べ、因果関係は推論できないものの、モデルの前提条件が少ないいうえ、非線形な関係や変数間の交互作用を分析できるといった利点がある⁸。特に、今回のように互いに相関関係がある多くの説明変数を扱う場合、通常の回帰分析では、多重共線性の問題が生じやすく、パラメータが安定的でなくなる可能性があると考えられる⁹。

(図表 4) 使用する機械学習モデルの概要

モデル	特徴
ランダムフォレスト	決定木を構築する際、サンプルをランダムに抽出し、多数の決定木を独立に構築するモデル
勾配ブースティング (XGBoost)	決定木を構築する際、それまでに作られたすべての決定木の予測結果を利用し、逐次的に決定木を構築していくモデル

ソフトウェアは R を利用し、iml (Interpretable Machine Learning) パッケージを活用している。生成する決定木の数は 500 本とし、毎回抽出される説明変数は 4 個とした。勾配ブースティングについては、デフォルト設定を用いている¹⁰。

⁵ 詳細は Breiman (2001) を参照。

⁶ XGBoost (eXtreme Gradient Boosting) は勾配ブースティングツリーのアルゴリズムが実装されているモデル・ライブラリ。詳細は Chen and Guestrin (2016) を参照。

⁷ 日本銀行 (2017) は、同様のアプローチを用いて、企業のインフレ予想がどのように形成されるかを分析している。機械学習に関する解釈可能性については、原 (2018) を参照。

⁸ 例えば、López de Prado (2018) は、ファイナンスの計量分析において、重回帰分析では変数間の非線形な関係性を捉えにくい点や外れ値による影響を受けやすい点を指摘しており、機械学習によるアプローチを推奨している。重回帰を用いた場合との比較は、補論を参照。

⁹ 通常、先行研究や統計量に基づいて、変数の取捨選択を行う必要があるが、Stepwise 法等は局所探索型のプロセスのため、必ずしも最適な説明変数が選択されない可能性がある。また、変数選択を行う場合、LASSO 回帰を用いることも考えられるが、同モデルでは非線形な関係や変数間の交互作用を捉えられないため、今回はランダムフォレストを用いている。なお、説明変数間の相関関係については、参考図表 1 を参照。

¹⁰ ランダムフォレストおよび勾配ブースティングとも、ハイパーパラメータのチューニング

センチメントの決定要因を探るため、主として、様々なマクロ経済・金融市場の不確実指標を標準化した上で¹¹、説明変数に用いた（図表 5）。マクロ経済の不確実性に関しては、景気後退確率との関連性が知られている米国イールドカーブ（Estrella and Trubin（2006））や、ニュースや記事ベースでのキーワード等に基づく政策不確実性指数（日本・米国）を用いている（Baker et al.（2015））。金融市場の不確実性については、株式市場の指標として知られている VIX 指数・日経 VI（米国 S&P500・日経平均のインプライド・ボラティリティ）に加え、VIX 先物カーブ（期近一期先のスプレッド）¹²や S&P500・日経平均の実現ボラティリティおよびスキュー指数も補完的な指標として用いている。加えて、より包括的な金融市場の不確実性を測る指標として、米国の Office for Financial Research による金融ストレス指数も用いている（Monin（2017））。同指数は、5つのカテゴリー（クレジット、株価バリュエーション、資金調達環境、安全資産需要、ボラティリティ）に基づいた 33 の金融変数から構成されている。

（図表 5）説明変数リスト

説明変数	出所
米国イールドカーブ（10年国債金利－2年国債金利）	Bloomberg
ドル／円の前月差	Bloomberg
日経VI	日本経済新聞社
日経平均の実現ボラティリティ（RV）	Oxford Man Institute
VIX指数	Bloomberg
VIX先物カーブ（1か月物－6か月物）	Bloomberg
S&P500の実現ボラティリティ（RV）	Oxford Man Institute
S&P500のSKEW指数	Bloomberg
金融ストレス指数	Office for Financial Research
米国の政策不確実性指数（経済、通商）	Economic Policy Uncertainty
日本の政策不確実性指数（経済、通商）	Economic Policy Uncertainty

グを行った場合でも、変数重要度の結果に大きな変化はみられない。

¹¹ ランダムフォレストや勾配ブースティングの場合、説明変数の標準化は必ずしも必要ないが、後述の部分従属プロット等の図示において、見やすさの観点から単位を揃えている。

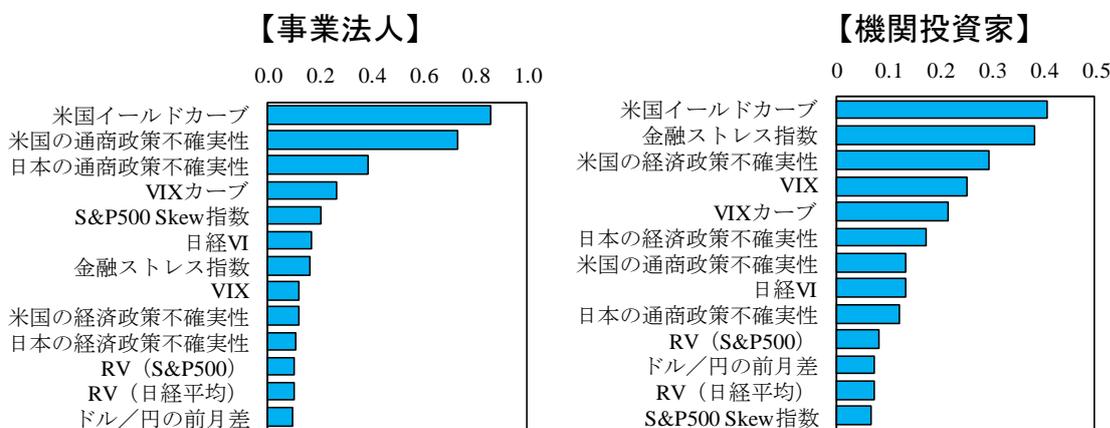
¹² 金融市場の不確実性が高まる局面では、期近物が期先物を上回るバックワーデーション（逆ザヤ）となることが知られている。

4. 分析結果

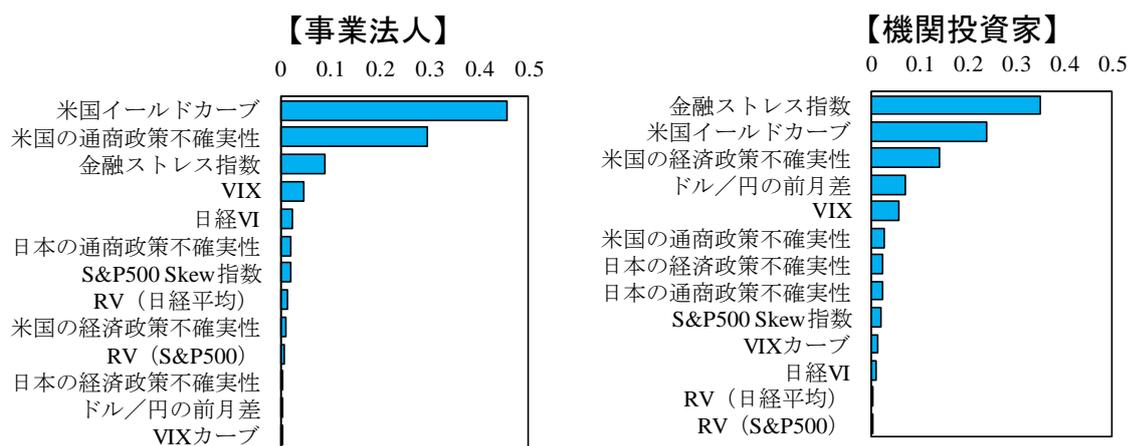
(1) 変数重要度¹³

変数重要度をみると（図表 6、7）、米国のイールドカーブが事業法人・機関投資家双方のセンチメントに影響していることが窺われる。また、事業法人のセンチメントに対しては、米国の通商政策に関する不確実性の影響が大きい一方、機関投資家のセンチメントに対しては、金融ストレス指数の影響が大きいことが分かる。一方、金融市場の不確実性指標として知られている VIX 指数や実現ボラティリティ（RV）の寄与は相対的に小さいことが分かる¹⁴。

（図表 6）変数重要度＜ランダムフォレスト＞



（図表 7）変数重要度＜勾配ブースティング＞



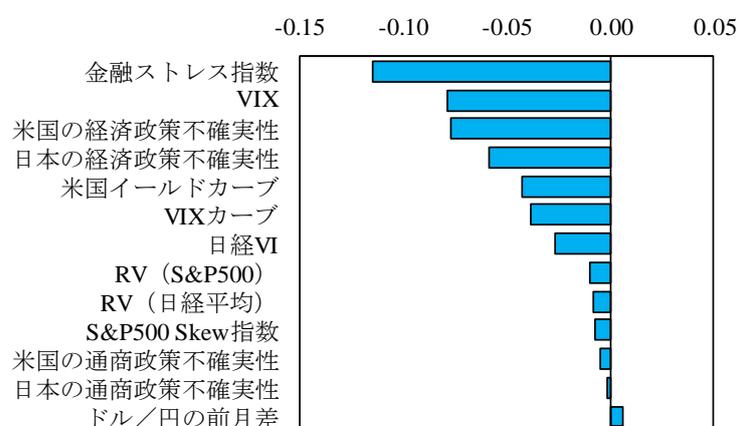
¹³ 変数重要度は、説明変数ごとに観測値をランダムに並び替えた場合、予測誤差がどの程度大きくなるかを計測したもので、予測誤差の大きい説明変数ほど重要度が高いと評価できる。XGBoost については、Gain を用いている。

¹⁴ 今回、コロナショック前後の影響を除くため、サンプル期間を 2015 年 4 月～2019 年 10 月とした場合でも、概ね同様の結果を得た。

一つの解釈としては、2018年および2019年には、VIX指数は過去のショック時に比べ低位で推移したものの、米中の通商問題に関する対立や米国の逆イールドが話題に挙がるなど、グローバル経済に関する不確実性が意識された可能性がある。加えて、事業法人は通常輸出入にかかる通貨エクスポージャーのヘッジに3か月から6か月程度の期間を伴うこと¹⁵等から、ごく短期的な金融市場のボラティリティの高まりに対する反応が限定的であった可能性が考えられる。他方、機関投資家については、金融ストレス指数がセンチメントに影響を与えている様子が窺われ、2020年3月のコロナ危機時など金融市場にストレスが掛かる局面では、円高に備えたポジションを造成している可能性が示唆される¹⁶。

実際、予測値に対する各説明変数の寄与度を示す Shapley Value フレームワーク¹⁷を用いて、2020年3月時の機関投資家のセンチメントに対する寄与度をみると、金融ストレス指数の急速な悪化が円高に備えたポジション増加に寄与した可能性が窺われる（図表8）。ただし、サンプル期間が短いこともあり、結果については幅を持ってみる必要がある。

（図表8）Shapley Value フレームワークによる寄与度
 <ランダムフォレスト>



（注）2020年3月時点における機関投資家のセンチメント指数に対する各説明変数の限界的な貢献度（φ）を示している。

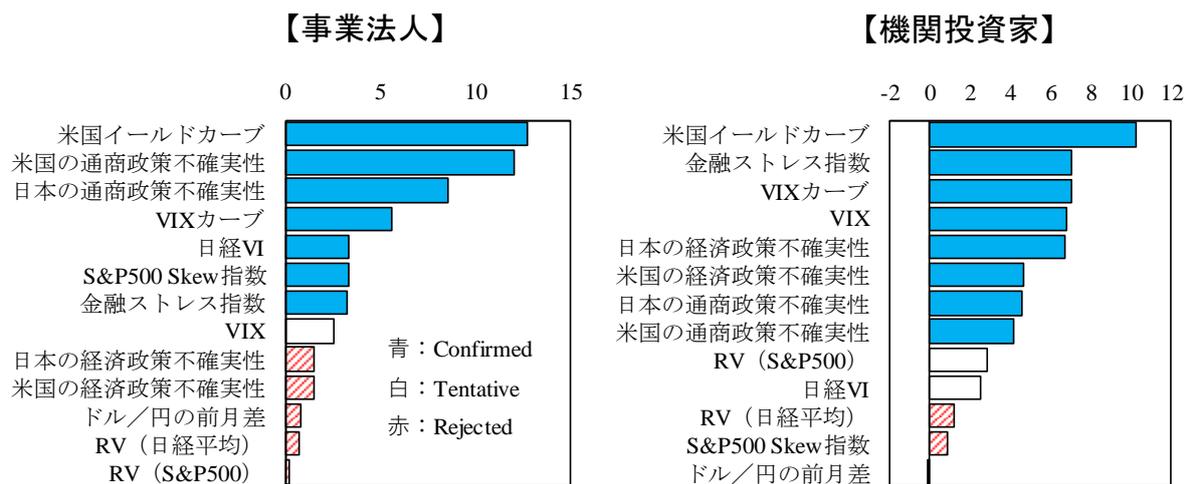
¹⁵ 事業法人および機関投資家の通貨オプションにかかる平均的な残存期間は、巻末の参考図表2を参照。

¹⁶ 2020年3月にかけて、事業法人の円高警戒感は一見弱まっているように見える（前掲図表3）。これは、米国の通商政策に関する不確実性が和らいだこと等に加え、輸出・輸入業者が先行きの為替エクスポージャーに対する不透明感の高まりや在宅勤務の拡がり等を受けて、一時的に取引を縮小させたことが影響している可能性がある。

¹⁷ 元々、ゲーム理論において協力によって得られた利得を各プレイヤーへ公正に分配する方法の一つ。詳細は、Strumbelj and Kononenko (2010)を参照。

さらに、Boruta (Kursa and Rudnicki (2010)) と呼ばれるランダムフォレストと検定を用いた変数選択手法¹⁸を利用した場合、事業法人・機関投資家のどちらについても、米国イールドカーブ、日米の通商政策不確実性指数、金融ストレス指数、VIXカーブの5つの変数が重要な変数として選択された (図表9)。

(図表9) Borutaによる変数選択結果



(注) 各変数の重要度について、Confirmedは有意である、Rejectedは有意であると言えない、Tentativeはどちらとも言えないことを示す。試行回数は100回。

(2) 部分従属プロット¹⁹

次に、変数重要度の高い項目について、部分従属プロットをみると(図表10)、直感的な解釈と整合的な関係が確認できる²⁰。すなわち、米国のイールドカーブのフラット化ないし逆転は、ドルの下落・円の上昇に偏ったポジションを増加させるほか、米国の通商政策に関する不確実性は、事業法人に関して、ドルの下落・円の上昇に偏ったポジションを増加させるとみられる。こうした点は、実体経済の減速や米中通商摩擦によって、事業法人のビジネスが影響を受けるため、オプションによる為替ヘッジを増やした可能性を示唆している。また、米国の通商政

¹⁸ 同手法は、ランダムフォレストの変数重要度に関して、あえて目的変数と関係のないダミーの説明変数(オリジナルの説明変数の値をシャッフルすることで作成)を追加し、ダミーの説明変数の変数重要度と、オリジナルの説明変数の変数重要度を比較する。有意水準(今回は、デフォルト値の5%)を設定した上で、比較をもとに変数選択を行う手法。

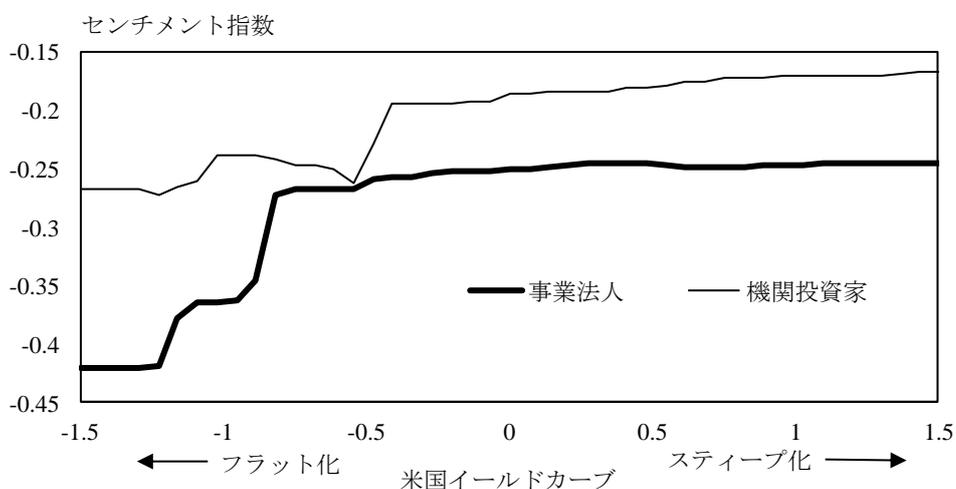
¹⁹ 部分従属プロット(Partial Dependence Plots)は、学習済みのモデルに対して、どの特徴量が予測に影響しているかを視覚化する方法。予測変数と説明変数との関係をその他の説明変数を周辺化(average out)し、交差項を捨象することで求める。

²⁰ その他の説明変数に関する部分従属プロットについては、巻末の参考図表3を参照。

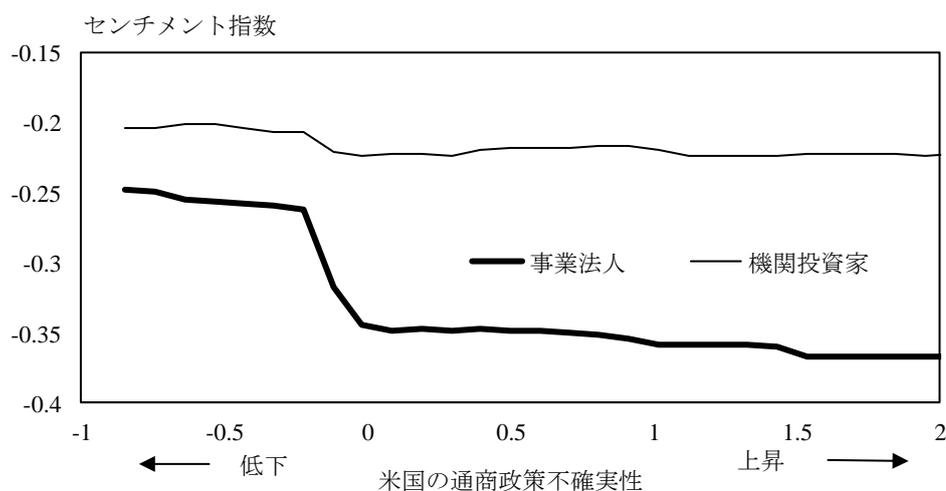
策に関する不確実性については、非線形な関係が顕著にみられており、一定の水
準を超えた場合に事業法人がドル／円の下落に備えたポジションを造成した可
能性が示唆される。他方、金融ストレス指数や VIX 指数などの金融市場の不確
実性指標については、事業法人のポジションには影響は少ないものの、不確実性
が急速に高まる局面では、機関投資家は円高に備えたポジションを増加させる
傾向が窺われる。

(図表 10) 部分従属プロット<ランダムフォレスト>

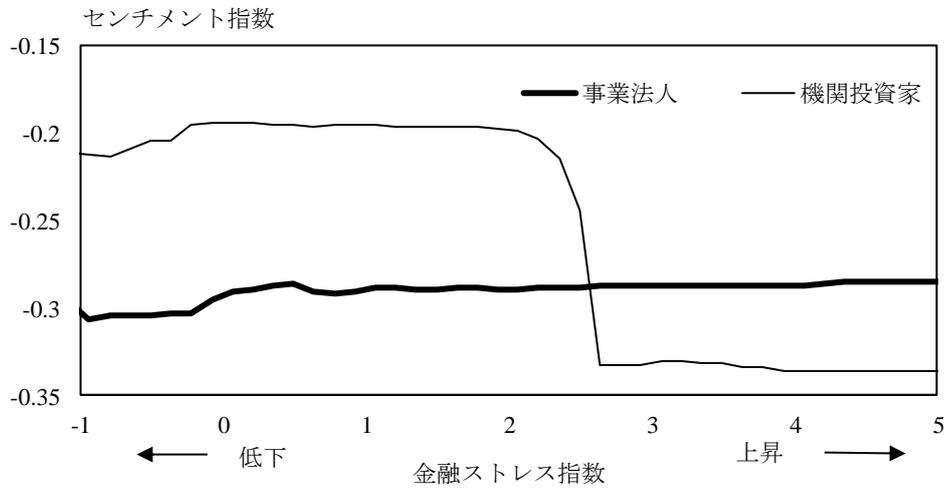
<米国イールドカーブ>



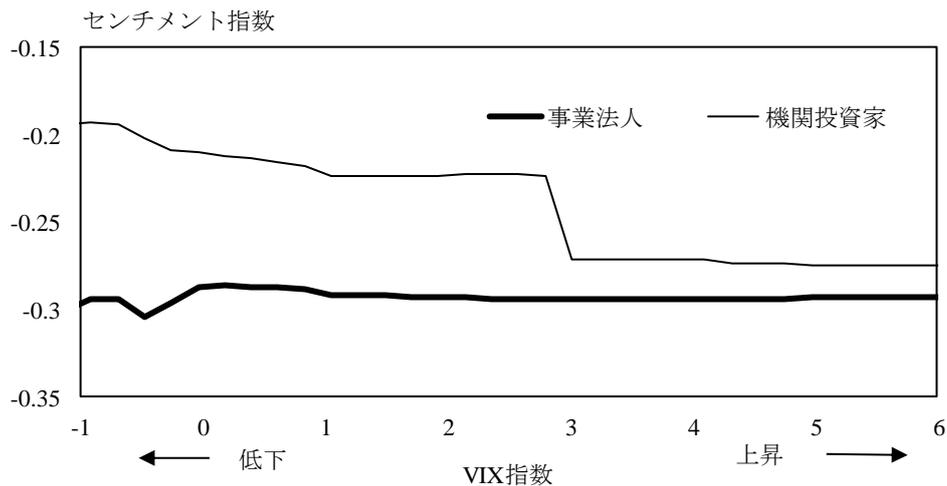
<米国の通商政策不確実性>



<金融ストレス指数>



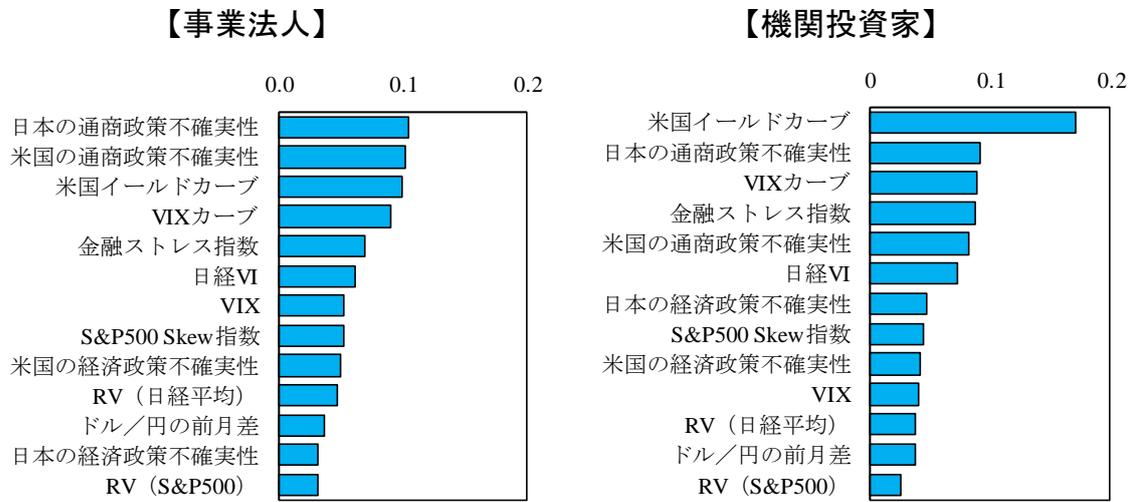
<VIX指数>



(3) 交互作用

上述の通り、VIX 指数自体はセンチメントへの影響度合いは大きくないと考えられるが、他の変数との交互作用をみることは一定の意義があると考えられる。先行研究 (Friedman and Popescu (2008)) に基づき、交互作用の相対的な強さを H 統計量によって計測した。結果をみると (図表 11)、変数重要度と概ね整合的な結果となっており、事業法人・機関投資家とも、VIX 指数の他の変数との交互作用は小さくなっている。他方、米国イールドカーブや金融ストレス指数の方が、他の変数との交互作用が相対的に強い様子が窺われる。これは、投資家行動を分析する場合には、VIX 指数のみならず、その他の不確実性指標を考慮することが重要な点を示唆していると考えられる。

(図表 11) 交互作用の強さ<ランダムフォレスト>



(注) H 統計量は、0~1 の値をとり、数値が大きいほど他の変数との交互作用が大きいことを示す。

5. まとめ

本研究では、店頭デリバティブ取引データからドル/円に関する主体別センチメント指数を構築し、その要因分析を行った。サンプル期間の短さには留意する必要があるものの、米国イールドカーブや米国の通商政策に関する不確実性、金融ストレス指数が影響を与えている可能性が示唆された。また、同分析結果に基づくと、投資家行動の把握にあたっては、金融市場に関する不確実性指標として知られる VIX 指数のみならず、幅広い指標を点検することが重要であると考えられる。

今後の課題としては、データの蓄積・クレンジングの精度向上が挙げられるほか、分析対象を他通貨ペア・他のプロダクトに拡張することや他地域を含めることでヘッジファンドなどの投資行動の分析等が可能と考えられる。金融市場のモニタリングやシステミックリスクの把握にとって、こうした店頭デリバティブ取引の分析の蓄積やデータの透明性向上は重要な役割を担っている。

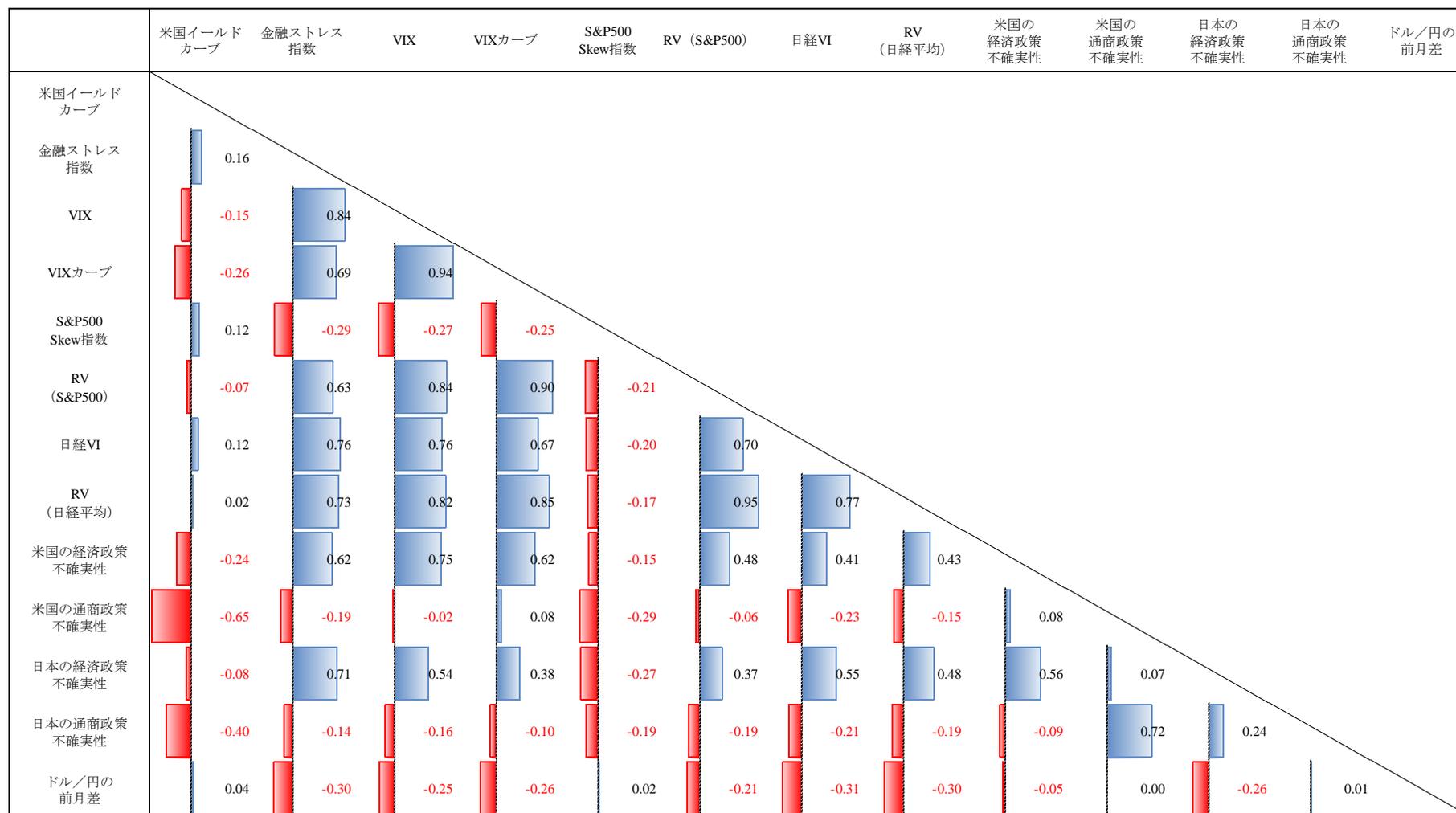
以上

(参考文献)

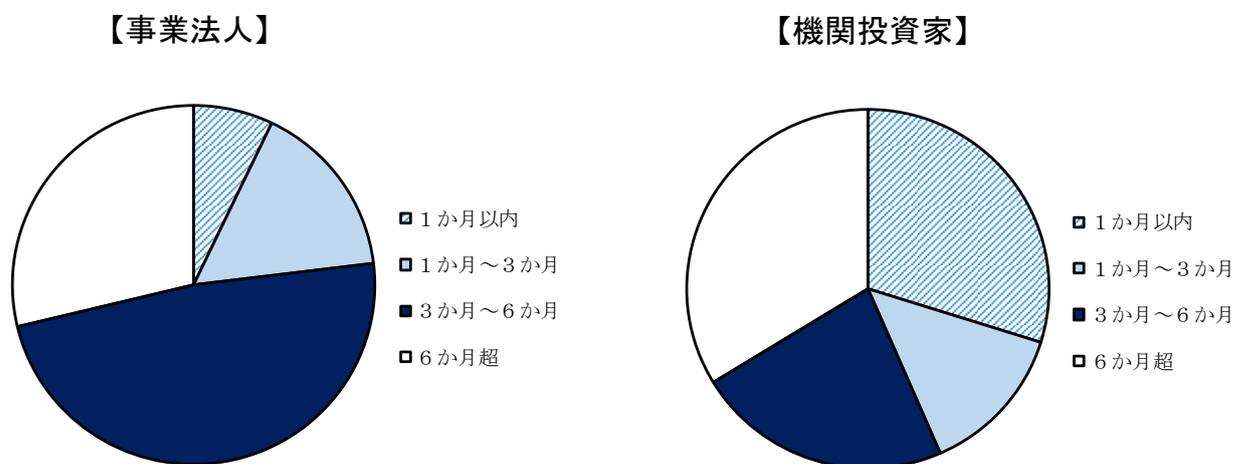
- 加藤晴子・福永一郎・山田健 (2012) 「リスク・リバーサルからみた為替変動へのリスク認識」、日銀レビュー・シリーズ、2012-J-14.
- 日本銀行 (2017) 「企業のインフレ予想形成に関する新事実: Part II—機械学習アプローチ」、日本銀行ワーキングペーパー・シリーズ、2017-J-4.
- 原聡 (2018) 「機械学習における解釈性 (私のブックマーク)」、人工知能、Vol.33、No.3、pages 366-369.
- Arturo Estrella and Mary R. Trubin (2006). “The yield curve as a leading indicator: some practical issues guidelines on how best to construct the yield curve indicator and interpret the measure in real time,” *Current Issues in Economics and Finance* (12) 5.
- Bank for International Settlements (2018). “Central banks and trade repositories derivatives data,” IFC Report.
- Erik Strumbelj and Igor Kononenko (2010). “An efficient explanation of individual classifications using game theory,” *Journal of Machine Learning Research*, Vol. 11, 1–18.
- European Central Bank (2003). “Using currency options-based indicators to assess sentiment in the foreign exchange markets,” *ECB Economic Bulletin*: 42-44.
- European Central Bank (2019). “Derivatives transactions data and their use in central bank analysis,” *ECB Economic Bulletin*, 6.
- European Securities and Markets Authority (2018). “EU derivatives markets 2018,” *ESMA Annual Statistical Report*.
- Hong Kong Monetary Authority (2015). “A first analysis of derivatives data in the Hong Kong trade repository,” *HKMA Quarterly Bulletin*.
- Inter-Agency Group on Economic and Financial Statistics (IAG) (2017). *Update on the Data Gap Initiative and the Outcome of the Workshop on Data Sharing*.
- Jerome H. Friedman and Bogdan E. Popescu (2008). “Predictive learning via rule ensembles,” *The Annals of Applied Statistics* (2): 916–54.
- Jorge Abad, Iñaki Aldasoro, Christoph Aymanns, Marco D’Errico, Linda Fache Rousová, Peter Hoffmann, Sam Langfield, Martin Neychev, and Tarik Roukny (2016). “Shedding light on dark markets: first insights from the new EU-wide OTC

- derivatives dataset,” ESRB Occasional Paper, 11, European Systemic Risk Board.
- Leo Breiman (2001). “Random forests,” *Machine Learning*, 45: 5-32.
- Marco D'Errico, Stefano Battiston, Tuomas Peltonen, and Martin Scheicher (2018). “How does risk flow in the credit default swap market?,” *Journal of Financial Stability*, 35: 53-74.
- Marcos López de Prado (2018). “Advances in financial machine learning,” Wiley.
- Miron B. Kursu, Witold R. Rudnicki (2010). “Feature selection with the boruta package,” *Journal of Statistical Software* (36) 11: 1-13.
- Olga Cielinska, Andreas Joseph, Ujwal Shreyas, John Tanner, and Michalis Vasios (2017). “Gauging market dynamics using trade repository data: the case of the Swiss franc de-pegging,” *Bank of England Financial Stability Papers*, 41.
- Reserve Bank of Australia (2018). “The Australian OTC derivatives market: insights from new trade repository data,” *RBA Bulletin*.
- Phillip Monin (2017). “The OFR financial stress index,” OFR Working Paper, 17-04, Office of Financial Research, Washington, D.C.
- Scott R. Baker, Nicholas Bloom, and Steven J. Davis (2015). “Measuring economic policy uncertainty,” *NBER Working Paper*, No.21633.
- Tianqi Chen and Carlos Guestrin (2016). “XGBoost: A scalable tree boosting system,” In *Proceedings of the 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 785-794.

(参考図表1) 説明変数の相関係数マトリックス



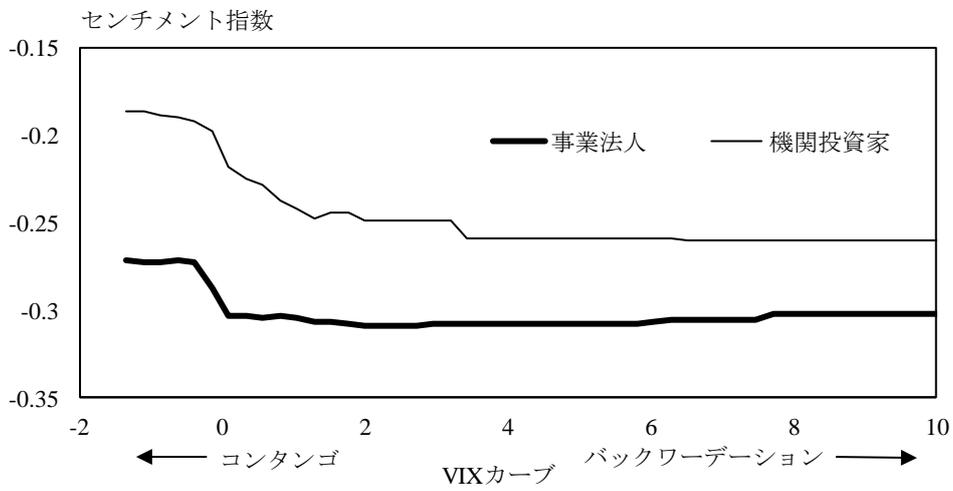
(参考図表 2) 通貨オプションの残存期間別にみた想定元本シェア (2019 年)



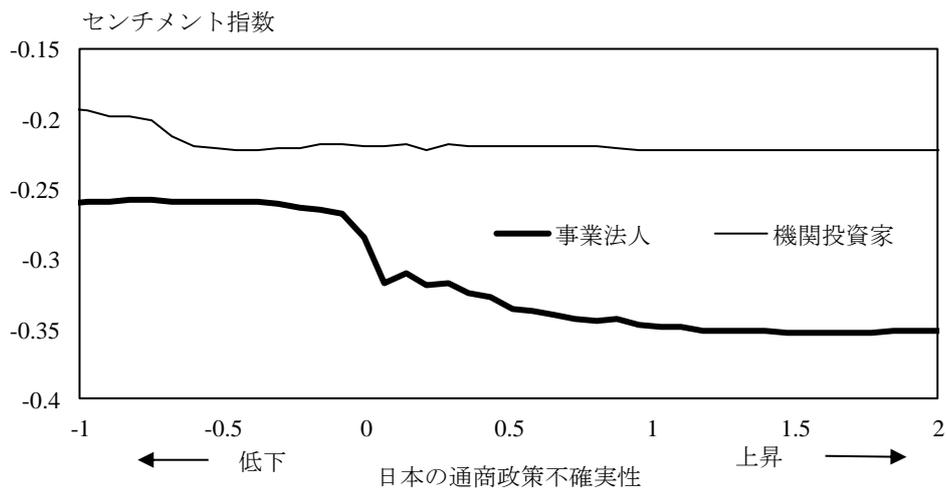
(注) 2019 年中の新規取引 (ドル/円) が対象。もっとも、データのクレンジング次第では、結果が変わりうる点には留意する必要がある。

(参考図表 3) 部分従属プロット

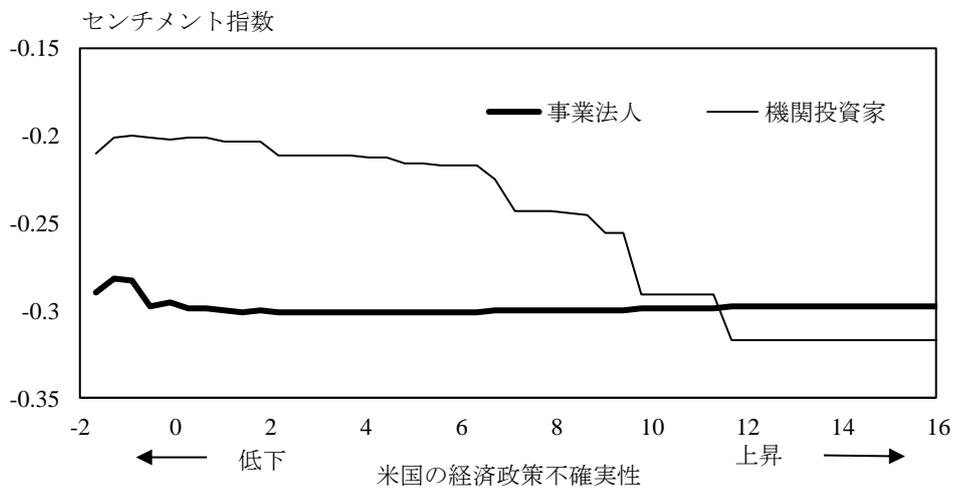
<VIXカーブ>



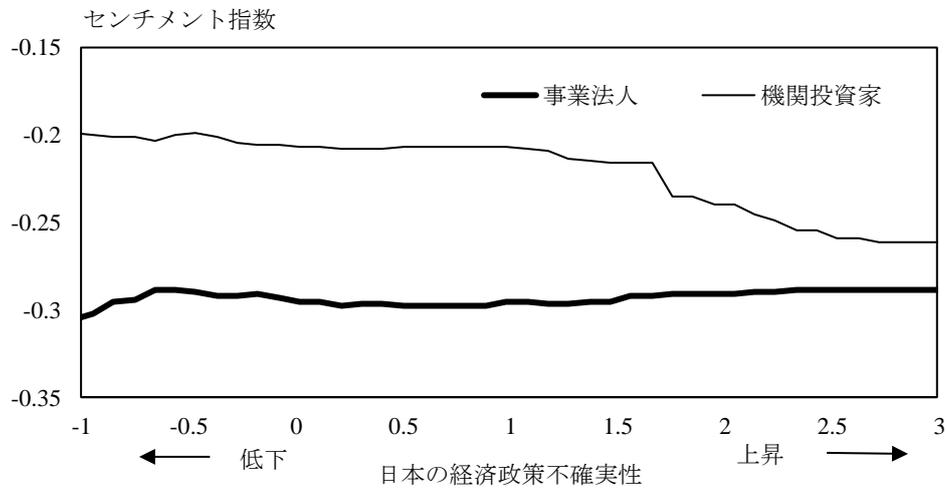
<日本の通商政策不確実性>



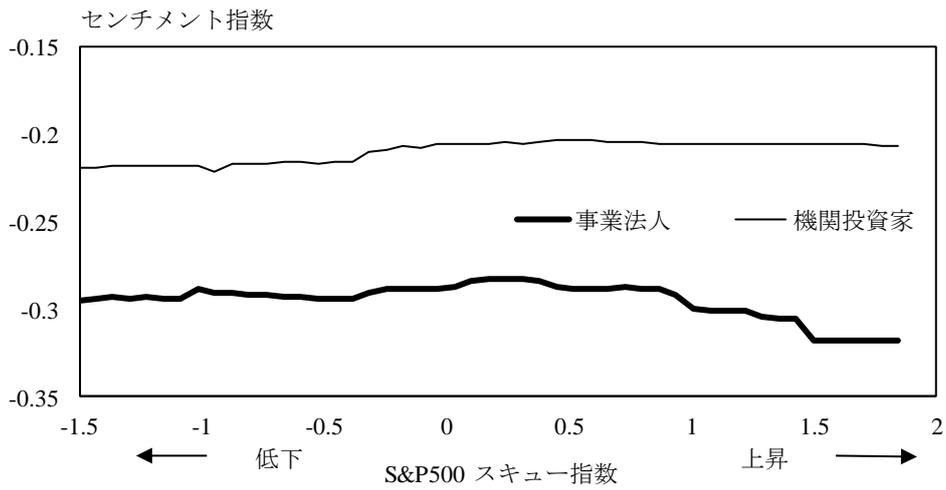
<米国の経済政策不確実性>



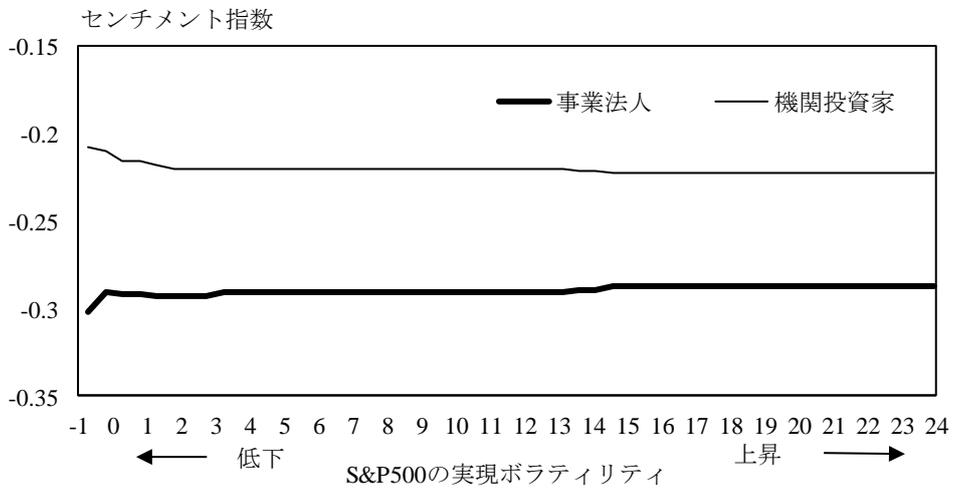
<日本の経済政策不確実性>



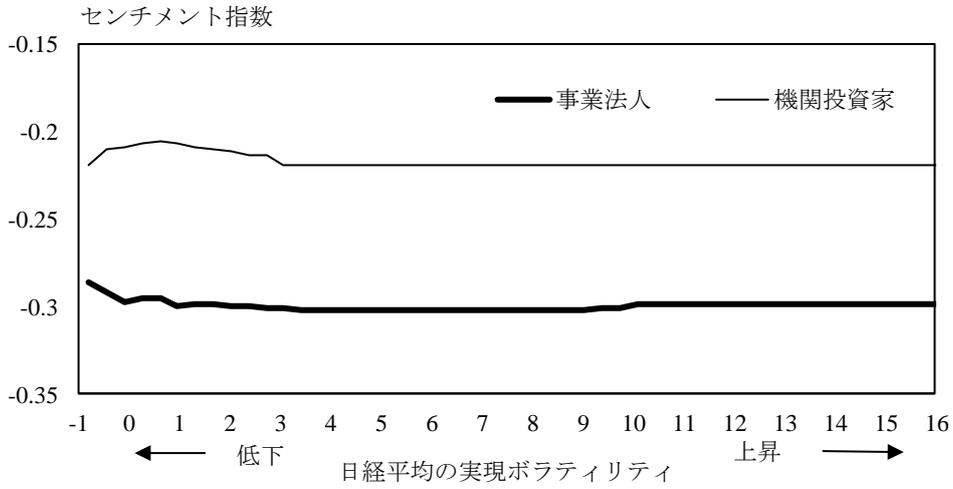
<S&P500のスキュー指数>



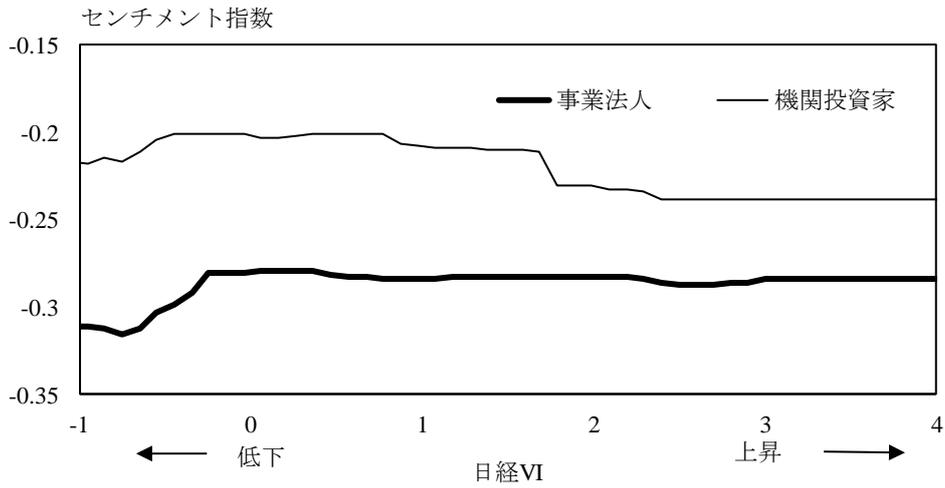
<S&P500の実現ボラティリティ>



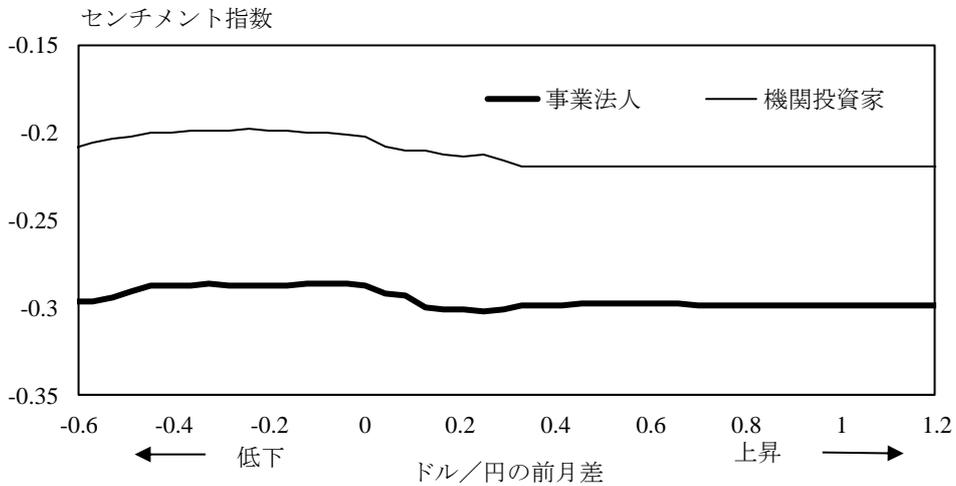
<日経平均の実現ボラティリティ>



<日経VI>



<ドル/円の前月差>



(補論) 重回帰モデルとランダムフォレストの結果比較

機械学習と従来の回帰分析との主な違いをみるために、事業法人・機関投資家のそれぞれについて、重回帰モデルの分析結果を紹介する。なお、説明変数は、ランダムフォレストによる変数重要度に基づき、上位2変数を用いている²¹。すなわち、事業法人については、説明変数は米国イールドカーブと米国の通商政策にかかる不確実性指標を用いているほか、機関投資家については、米国イールドカーブと金融ストレス指数を用いた。

推計結果をみると、いずれの説明変数も有意となっており、符号条件もランダムフォレストと同様の結果を得た(補論図表1)。もともと、線形性を仮定しているため、ランダムフォレストに比べて、各説明変数が被説明変数に与える影響を正しく捉え切れていない可能性も示唆される。補論図表2では、重回帰モデルで推定された係数をもとに、各説明変数の値が変化した場合にどのように被説明変数が変化するかをプロットしている。一例として、米国の通商政策不確実性や金融ストレス指数は、ある閾値を超えた場合にセンチメントに影響する可能性があるため、今回のような場合、線形性を仮定したモデルでは、ダミー変数を利用することも選択肢の一つとして考えられる²²。

(補論図表1) 重回帰モデルによる推計結果

	事業法人	機関投資家
米国イールドカーブ	0.102 ^{***} (0.023)	0.117 ^{***} (0.019)
米国の通商政策不確実性	-0.080 ^{***} (0.031)	
金融ストレス指数		-0.078 ^{***} (0.019)
定数項	-0.282 ^{**} (0.032)	-0.183 ^{***} (0.019)
Adj. R ²	0.425	0.494
S.E. of reg.	0.185	0.145

(注) 推計期間は、2015年4月～2020年6月。
括弧内は標準誤差。***は1%、**は5%水準で有意。

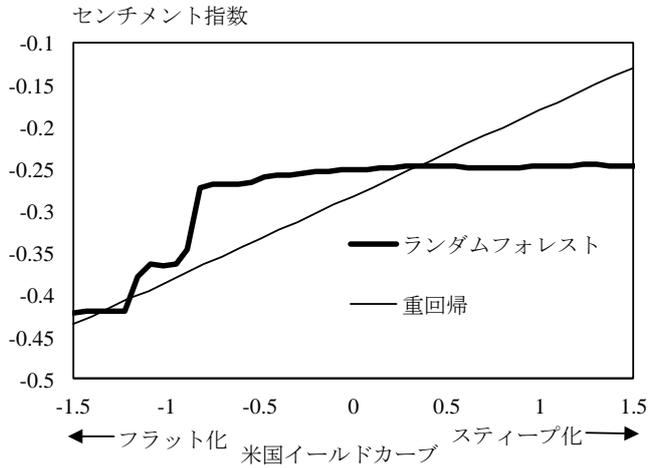
²¹ 前述の通り、説明変数間に相関があるため、多重共線性の問題が生じており、変数減少法に基づいて変数選択を行うと、米国の通商政策不確実性などが取り除かれてしまう。

²² 実際、米国の通商政策不確実性や金融ストレス指数をダミー変数(ある閾値を超えた場合に1、そうでない場合に0)として加えた場合、決定係数の改善がみられる。

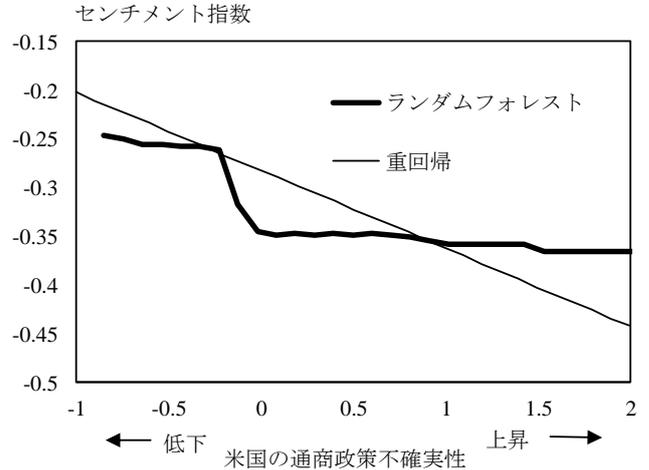
(補論図表 2) ランダムフォレストによる部分従属プロットとの比較

【事業法人】

＜米国イールドカーブ＞

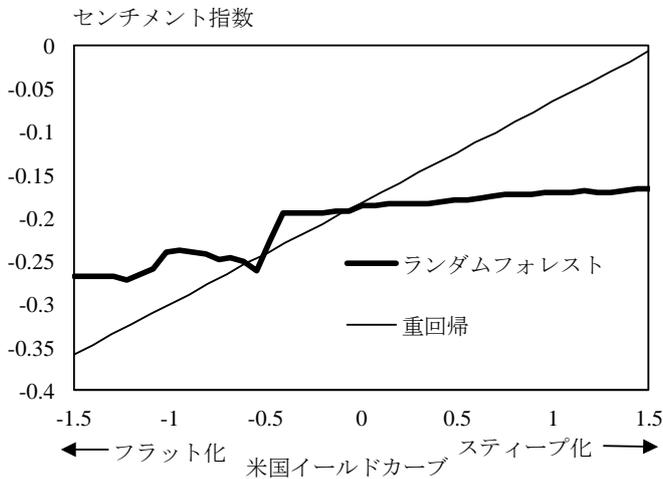


＜米国の通商政策不確実性＞

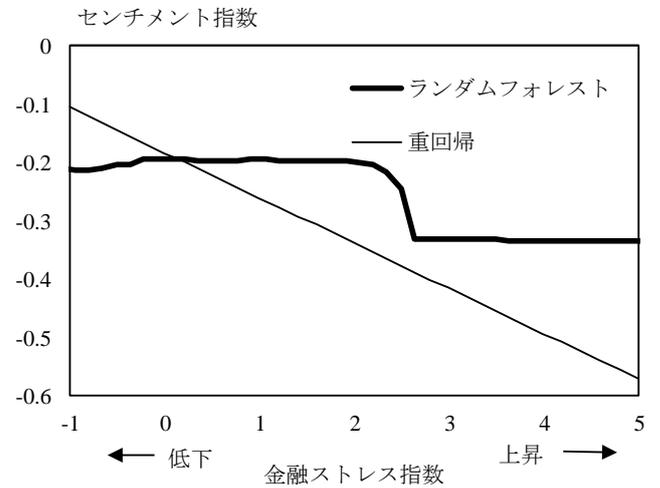


【機関投資家】

＜米国イールドカーブ＞



＜金融ストレス指数＞



(注) 重回帰モデルについては、(各説明変数の値×推定された係数+定数項)をそれぞれ示している。