

経済調査 レポート

コンセンサス予測に勝つ方法はあるのか～月次指標予測を用いた分析

経済調査部門 斎藤 太郎 (さいとう たろう)
(03) 3512-1836 tsaito@nli-research.co.jp

【要旨】

- ・ 本稿では、鉱工業生産、消費者物価など 13 の月次指標を用いて、コンセンサス予測と個別機関（22 機関）予測の特徴を分析した。
- ・ コンセンサス予測は個別機関の予測と比較して相対的に優秀な成績となる。2003 年度から 2009 年度までのデータを用いて指標毎に個別機関の予測誤差、コンセンサスの予測誤差を計算して順位付けしたところ、コンセンサス予測は全ての指標で上位となった。鉱工業生産、経常収支は 1 位であり、最も順位の低い失業率でも 6 位であった。13 指標全体の成績はコンセンサス予測が 1 位となった。
- ・ 個別機関の予測は総じてコンセンサス予測に劣っているが、各機関の予測精度には差が見られた。最も優秀な予測機関は 13 指標中、6 指標でコンセンサス予測を上回っている一方、22 機関の約半数にあたる 10 機関はコンセンサス予測を上回る指標がひとつもなかった。
- ・ 第 3 次産業活動指数など特定の指標については、機関によって予測精度に大きな差があり、単年度の成績が優秀な機関は全期間にわたって優秀な成績となる傾向があることが確認された。
- ・ 個別機関の予測が長期にわたってコンセンサス予測に勝つことは難しいが、比較的パフォーマンスの良い機関の予測値を組み合わせることによって、予測精度を高めることが可能である。単年度の成績が優秀な機関を複数（上位 5 機関、上位 10 機関のケースを試算）取り出し、これらの機関の予測平均値を使うと、全ての指標で全期間（2003～2009 年度）の予測誤差が個別機関全体の平均値よりも小さくなり、鉱工業生産を除くほとんどの指標でコンセンサスを上回るパフォーマンスが得られた。

はじめに

本稿では、鉱工業生産、消費者物価などの月次指標を用いてコンセンサス予測、個別機関予測の評価を行う。

経済予測の評価に関する分析は、年度ベースの実質GDP成長率について行われることが多いが、その際にはいくつかの問題が存在する。

ひとつは、成長率の予測値は頻繁に修正されるため、予測時点の異なる多くの予測値が存在することである。予測機関の多くは、四半期のGDP速報が発表される（GDP速報は1次速報、2次速報がそれぞれ年4回発表される）ごとに予測値の修正を行うが、最近では毎月のように見直しをする機関もある。河越（2007）は、経済企画協会が実施している「ESPフォーキャスト調査」の評価に際して、「予測時点と公表時点が近いほど予測者が直面する不確実性は小さく、精度の高い予測値となるはずである」とし、公表時点に近いほど予測誤差に高いウェイトを与えた上で予測の評価を行っている。しかし、経済予測を利用する立場からすれば、実績値が公表される直前の予測値よりも1年前の予測値のほうにより高い価値を見出すかもしれない。たとえば、成長率の予測値を自社の経営計画を策定する際の前提として利用する場合、実績値が公表される直前の予測が正確であることよりもその年度が始まる前の予測が正確であることのほうがより価値が高いと考えるだろう。

また、GDP統計は一度公表された実績値がその後改定されるため、予測誤差を計算する際にどの実績値を用いるべきかという問題もある。多くの分析では最初に公表された速報値を実績値として用いているが、速報値よりも確報値のほうがより真実に近い値であるとすれば、速報値よりも確報値に対する誤差が小さいほうが優れた予測であるという見方もできる。

この点、月次指標にはこうした問題は基本的に存在しない。月次指標の予測は原則として実績値が公表される前週に1度行われるだけであり、予測値はひとつしかない。また、鉱工業生産や貿易統計などの月次指標もGDP統計と同様に、最初に公表された速報値は確報値、年間補正、基準改定などの際に改定される。しかし、月次指標予測の目的はあくまでも最初に公表される速報値を当てることであることは明白で、対象とすべき実績値がはっきりしている。

月次指標の結果はしばしば金融市場を大きく動かす。6月4日に公表された5月の米国雇用統計で、非農業部門の雇用者数が前月から43.1万人の増加となったにもかかわらず、市場予想を大きく下回ったことから、株価が大幅に下落したことは記憶に新しい。このことは、金融市場の反応が経済指標の実績値そのものよりも実績値が予測値（市場コンセンサス）からどれだけ乖離したかによってもたらされることが多いことを示しており、こうした事象は日本においても確認される。たとえば、草場（2010）は、「長期金利の日次変動にとっては、経済指標そのものの良し悪しよりも事前予想と比べたときの良し悪し（サプライズの有無）の方が重要であり、サプライズの度合いが大きいほど金利変動にもたらす影響も大きくなりやすい」としている。

金融市場が注目するのは、予測値の平均であるコンセンサス予測だが、当然のことながらその裏には個別機関の予測値があり、機関毎の予測誤差にはばらつきがある。仮に、その中にコンセンサス予測よりも精度の高い予測が存在するとすれば、その情報を入手することが金融市場で高いパフォーマンスを得ることにもつながるだろう。今回の分析における最終的な目的は、何らかの方法によってコンセンサス予測を上回る予測を見つけ出すことである。

1. データの解説

1-1. 本稿で取り扱う指標

月次指標の予測は実績値が公表される週の前週に通信各社が調査を行っている。回答の締め切りが設定されている場合は概ね金曜日の午前中となっており、その日のうちに集計結果が公表される。

今回の分析では、各機関の予測値はロイターが集計したものをを用いた。過去に遡って個別機関の予測値が入手できたためである。ただし、第3次産業活動指数については、2009年1月分をもってロイターが集計を取りやめたため、それ以降については共同通信社の集計を用いた。

本稿では以下の13指標を取り上げた。対象期間は2003年度から2009年度までの7年間、84ヵ月である。ただし、ロイターの調査開始が遅かった4指標（国内企業物価：2004年8月～、新設住宅着工戸数、実質消費支出：2005年7月～、マネーストック：2005年8月～）については対象期間が短くなっている。

- ① 鉱工業生産（前月比、%）「鉱工業指数（経済産業省）」
- ② 第3次産業活動指数（前月比、%）「第3次産業活動指数（経済産業省）」
- ③ 機械受注（船舶・電力を除く民需、前月比、%）「機械受注統計（内閣府）」
- ④ 全国・消費者物価（生鮮食品を除く総合、前年比、%）「消費者物価指数（総務省）」
- ⑤ 東京都区部・消費者物価（生鮮食品を除く総合、前年比%）「消費者物価指数（総務省）」
- ⑥ 失業率（季節調整値、%）「労働力調査（総務省）」
- ⑦ 有効求人倍率（季節調整値、倍）「一般職業紹介状況（厚生労働省）」
- ⑧ 貿易収支（原数値、億円）「貿易統計（財務省）」
- ⑨ 経常収支（原数値、億円）「国際収支状況（財務省、日本銀行）」
- ⑩ 国内企業物価（前年比、%）「企業物価指数（日本銀行）」
- ⑪ 実質消費支出（前年比、%）「家計調査（総務省）」
- ⑫ 新設住宅着工戸数（前年比、%）「建築着工統計（国土交通省）」
- ⑬ マネーストック（M2、前年比、%）「マネーストック統計（日本銀行）」（2008年4月まではマネーサプライ統計のM2+CD）

ロイターの集計ではそれぞれの指標について、毎月30前後の機関が回答しているが、本稿では個別機関の分析を行うことも目的としているため、対象は継続的に回答をしている機関に限定した。具体的には対象期間中に調査回数の7割以上回答している22機関を対象とした。全体としては7割以上回答しているが、指標によって回答をしていない、あるいは回答数が少ない機関もあるため、7割以上回答していない指標については、分析の対象から除外した。

なお、分析の対象とした機関が回答しなかった場合（欠損値）については、予測誤差が各機関の平均値になるような予測値を回答したものとして扱った。

1-2. 予測誤差の評価尺度

ある指標について、 t 期の予測誤差 e_t は、予測値を f_t 、実績値を y_t とすると、

$$e_t = y_t - f_t$$

と表すことができる。

コンセンサスといった場合、各機関の平均値を指す場合と中央値を指す場合があるが、本稿では平均値を使った。なお、今回の分析で用いるコンセンサスは、対象とした22機関の予測値をもとに筆者が計算したものである。

予測誤差の評価尺度としては平均絶対誤差 (Mean Absolute Error : MAE) を用いる。

$$\text{平均絶対誤差} = \frac{1}{k} \sum_{t=1}^k |y_t - f_t|$$

となる。本稿では単に予測誤差といった場合、対象期間中の平均絶対誤差を指すこととする。

2. コンセンサス予測の特徴

2-1. 各指標の市場予測の特徴

本節では、まず各指標に関する市場予測の特徴を確認する。

各指標のコンセンサス予測の誤差 (平均絶対誤差) は図表1の通りである。たとえば、鉱工業生産の予測誤差は2003年度から2009年度まで(84ヵ月)の平均で0.70%、機械受注は5.04%、貿易収支は1451億円などとなっている。指標によって単位や実績値の変動の大きさなどが異なるため、予測誤差の単純な比較はできない。たとえば、鉱工業生産に比べ機械受注は月々の変動が大きいため、同じ予測誤差でもその意味が違ってくる。

図表1 指標毎の市場予測の概要

	コンセンサスの 予測誤差	コンセンサスの 的中率	レンジ内確率
鉱工業生産(前月比、%)	0.70	8.3%	73.8%
第3次産業活動指数(前月比、%)	0.51	3.6%	88.1%
機械受注(前月比、%)	5.04	2.4%	52.4%
全国・消費者物価(前年比、%)	0.059	54.8%	96.4%
東京・消費者物価(前年比、%)	0.085	39.3%	89.3%
失業率(%)	0.123	27.4%	64.3%
有効求人倍率(倍)	0.0134	26.2%	67.9%
貿易収支(億円)	1451	0.0%	76.2%
経常収支(億円)	1393	0.0%	78.6%
国内企業物価(前年比、%)	0.22	19.1%	86.8%
実質消費支出(前年比、%)	1.36	0.0%	59.6%
住宅着工戸数(前年比、%)	5.43	0.0%	50.9%
マネーストック(前年比、%)	0.16	23.2%	64.3%

実績値の月々の変動が比較的小さい指標としては、消費者物価、失業率、有効求人倍率が挙げられるが、消費者物価の平均的な予測誤差は公表値の最低単位である0.1%以下となっているのに対し、失業率、有効求人倍率はともに最低単位(それぞれ0.1%、0.01倍)を超える予測誤差が生じている。

同じ消費者物価でも全国のほうが東京都区部よりも誤差が小さい。東京都区部の消費者物価(中甸速

報値)は全国よりも1ヵ月前に公表されるため、予測者の多くが前月の東京都区部の結果を参考にして全国の予測値を作成していることが推察される。

次に指標毎の的中率を見てみよう。ここでは、予測値と実績値が完全に一致した場合を的中とする。的中率が最も高いのは全国・消費者物価の54.8%である。84ヵ月中46ヵ月でコンセンサス予測と実績値が一致している。その他の指標で的中率が高いのは東京・消費者物価(39.3%)、失業率(27.4%)、有効求人倍率(26.2%)、マネーストック(23.2%)で、これらの指標に共通するのは、実績値の変動幅が小さいことである。逆に、対象期間内にコンセンサス予測が実績値と1度も一致しなかった指標は貿易収支、経常収支、実質消費支出、住宅着工戸数の4つである。ただし、貿易収支、経常収支が的中しないのは予測値、実績値の桁数が多いことによる部分が大きく、このことによって予測が外れやすいという判断はできない。

予測の当たり(外れ)やすさを見る指標としては、実績値が市場予測のレンジ内(最大値~最小値)に収まる確率を見る方法がある。各指標が市場予測のレンジ内に収まる確率を求めると、全国・消費者物価が96.4%で最も高い。84ヵ月のうち予測レンジから外れたのは3回だけで、そのうち1回は2006年7月分の公表時に基準改定(2000年基準→2005年基準)が実施され、旧基準から新基準への下方改定幅が各機関の予測を大きく上回ったケースである。

以下、レンジ内に収まる確率が比較的高いのは、順に東京・消費者物価(89.3%)、第3次産業活動指数(88.1%)、国内企業物価(86.8%)となっている。

逆に、レンジ内に収まる確率が最も低いのは住宅着工戸数の50.9%で、約2回に1回は実績値が予測レンジを外れている。以下、機械受注(52.4%)、実質消費支出(59.6%)などがレンジ内に収まる確率が低い指標である。また、失業率、有効求人倍率は月々の変動は小さいが、予測レンジから外れることが比較的多い(レンジ内確率はそれぞれ64.3%、67.9%)指標である。

2-2. コンセンサス予測が優秀となる理由

コンセンサス予測は個別機関の予測と比べて相対的に優秀な成績となることが知られている。経済企画協会が実施している「ESPフォーキャスト調査」では毎年、個々のフォーキャスターの予測評価を行うとともに、予測総平均(コンセンサス)の評価が示されている。総平均の順位は2004年度から2008年度まで順に5位、8位、6位、9位、7位(評価対象人数はそれぞれ37人、37人、33人、29人、27人)といずれも上位にランキングされている。

コンセンサス予測のパフォーマンスは、直感的には予測機関全体の中位程度になるように思われるかもしれないが、実はコンセンサス予測の誤差が各機関の平均的な予測誤差よりも小さくなることは以下の数式から明らかである。

予測機関を*i*とすると、

$$t \text{ 期におけるコンセンサス予測の絶対誤差} = \left| y_t - \frac{1}{n} \sum_{i=1}^n f_{it} \right| \cdots \textcircled{1}$$

$$\text{一方、各機関予測の絶対誤差の平均} = \frac{1}{n} \sum_{i=1}^n |y_t - f_{it}| \cdots \textcircled{2}$$

で表される

①式を変形すると、

$$\frac{1}{n} \left| n \cdot y_t - \sum_{i=1}^n f_{ti} \right| = \frac{1}{n} \left| (y_t - f_{t1}) + (y_t - f_{t2}) + (y_t - f_{t3}) + \dots + (y_t - f_{tm}) \right| \dots \text{①}$$

②式を変形すると、

$$\frac{1}{n} \left(|y_t - f_{t1}| + |y_t - f_{t2}| + |y_t - f_{t3}| + \dots + |y_t - f_{tm}| \right) \dots \text{②}$$

となる。

①式右辺と②式右辺の違いは各項を合計したものに絶対値がついているか、各項それぞれに絶対値がついているかである。各項が全て同じ符号の場合、①式＝②式となり、各項に異なる符号が存在する場合、①式<②式となる。つまり、①式≤②式が必ず成立する。

具体的には、実績値が市場予測のレンジから外れた場合（市場予測の上限値、下限値の場合も含む）、①式右辺の各項は全て同じ符号となり、①式は②式と一致する。すなわち、コンセンサス予測の誤差と個別機関の予測誤差の平均は等しくなる。実績値が市場予測のレンジ内となった場合には、①式右辺の各項はプラスとマイナスが混在し、これらが互いに打ち消し合うことにより、①式は②式よりも小さくなる。すなわち、コンセンサス予測の誤差は個別機関の予測誤差の平均よりも小さくなる。

結局、コンセンサス予測の誤差が個別機関の予測誤差の平均よりも大きくなることは原理的にありえない。このような試行を繰り返すことにより、コンセンサス予測は相対的に優秀な成績をおさめることになるのである。

2-3. コンセンサス予測の順位

実際のデータでこれらのことを確認してみよう。2003年度から2009年度までのコンセンサスの予測誤差は、13指標の全てについて個別機関の予測誤差の平均値よりも小さい。特に、鉱工業生産、経常収支については予測誤差が最も小さい機関よりも誤差が小さくなっている（図表2）。

図表2 コンセンサス予測の特性

	コンセンサスの予測誤差	予測誤差の各機関平均	予測誤差の最小値	予測誤差の最大値
鉱工業生産(前月比、%)	0.70	0.83	0.71	0.99
第3次産業活動指数(前月比、%)	0.51	0.65	0.39	0.82
機械受注(前月比、%)	5.04	5.44	4.99	6.18
全国・消費者物価(前年比、%)	0.059	0.074	0.059	0.119
東京・消費者物価(前年比、%)	0.085	0.098	0.076	0.144
失業率(%)	0.123	0.128	0.113	0.146
有効求人倍率(倍)	0.0134	0.0143	0.0129	0.0165
貿易収支(億円)	1451	1770	1408	2089
経常収支(億円)	1393	1686	1398	2238
国内企業物価(前年比、%)	0.22	0.26	0.21	0.37
実質消費支出(前年比、%)	1.36	1.50	1.17	1.73
住宅着工戸数(前年比、%)	5.43	5.84	5.26	7.37
マネーストック(前年比、%)	0.16	0.17	0.15	0.21

コンセンサス予測の精度が22機関中何位に該当するかを見ると、全ての指標で上位に位置している（図表3）。鉱工業生産、経常収支は1位で、最も順位の低い失業率でも6位である。13指標全体の順位は1位となる。全体の順位は、指標毎に予測誤差の大きさが異なるので、各指標の予測誤差（平均絶

対誤差) を標準化変量で基準化した上で 13 指標を平均することによって求めた。

予測機関*i*の予測誤差 (MAE_i) の標準化変量 (z_i) は

$$z_i = \frac{MAE_i - Mean(MAE)}{Std(MAE)}$$

である。

図表3 コンセンサス予測の順位

	全期間	03年度	04年度	05年度	06年度	07年度	08年度	09年度
鉱工業生産	1	5	8	5	7	4	3	4
第3次産業活動指数	3	7	2	7	8	2	3	6
機械受注	3	7	6	3	8	8	8	6
全国・消費者物価	2	2	5	9	9	10	15	2
東京・消費者物価	4	5	4	8	11	7	6	10
失業率	6	8	11	13	11	12	12	8
有効求人倍率	5	10	8	7	10	10	11	11
貿易収支	2	7	4	8	8	5	3	2
経常収支	1	10	5	3	5	3	6	5
国内企業物価	5	—	7	2	6	5	7	9
実質消費支出	3	—	—	5	4	8	9	4
住宅着工戸数	4	—	—	8	8	6	6	9
マネーストック	4	—	—	7	6	9	2	10
全体	1	2	1	1	3	1	3	2

(注)第3次産業活動指数、住宅着工戸数、マネーストックは18機関、経常収支は20機関、機械受注、実質消費支出は21機関、それ以外は22機関中の順位

失業率 (6 位)、有効求人倍率 (5 位) は比較的順位が低いですが、両指標に共通するのは、個別機関の予測値がコンセンサスと一致するケースが多いこと、実績値が市場予測のレンジから外れることが比較的多いことである。前者はコンセンサス予測と誤差が一致する機関が多いこと、後者はコンセンサスの予測誤差と個別機関の予測誤差の平均値が一致することを意味する。両者ともにコンセンサスの予測誤差と個別機関の平均的な予測誤差の差を小さくする要因として働く。

コンセンサスの順位を指標毎に年度単位で区切って見てみると、ほとんど全ての指標、全ての年度で上位となっているが、全期間を通したものよりは低い順位にとどまっているものが多い。失業率は順位が中位を下回る年度がある (2005 年度の 13 位) が、これは予測誤差の分布に偏りがあったためであり、予測誤差の大きさ自体は各機関の平均値よりも小さい。また、コンセンサス予測の 13 指標全体の成績は年度毎に見ても、1 位が 3 回、2 位が 2 回、3 位が 2 回となっており、非常に良好なパフォーマンスを示している。

前述したとおり、コンセンサスの予測誤差は、各回の調査毎に必ず個別機関の予測誤差の平均と等しいかそれよりも小さくなる。この試行が繰り返されることにより、コンセンサス予測の優位性が増していくため、予測の回数が増えるほどコンセンサス予測の順位が上がりやすくなると考えられる。

3. 個別機関の予測の特徴

3-1. 個別機関の順位、コンセンサスに対する勝率

本節では個別機関の予測に焦点を当てる。図表 4 では指標毎に予測誤差 (2003 年度から 2009 年度までの平均値) に関する各機関の順位を示している。前節でみたとおり、いずれの指標に関してもコンセンサス予測よりも優秀な機関は少数である。全体の順位は各指標の予測誤差を標準化変量で基準化した

上で13指標を平均することによって求めた。

全体の成績が1位の機関は13指標中、6指標でコンセンサス予測を上回るパフォーマンスを示している。上位6機関までは少なくともひとつはコンセンサス予測を上回る指標があるが、22機関のうち約半数にあたる10機関は全ての指標でコンセンサス予測のパフォーマンスを下回っている。なお、前節で見たとおり、コンセンサス予測の13指標全体の成績は1位であるため、全ての機関が全体の成績ではコンセンサス予測に負けているということになる。

図表4 各機関の指標毎の順位

	鉱工業生産	第3次産業活動指数	機械受注	全国・消費者物価	東京・消費者物価	失業率	有効求人倍率	貿易収支	経常収支	国内企業物価	実質消費支出	住宅着工戸数	マネー・ストック	全体
機関A	4	1	10	2	3	3	9	3	4	2	7	2	3	1
機関B	10	3	3	1	10	13	8	1	14	7	1	5	4	2
機関C	1	12	9	8	4	4	3	6	12	4	4	1	13	3
機関D	3	—	5	10	9	7	13	4	3	3	2	11	1	4
機関E	5	2	17	3	1	19	6	2	1	5	5	14	16	5
機関F	12	9	19	12	15	5	10	5	10	6	6	4	5	6
機関G	2	10	13	16	6	6	7	15	8	8	12	—	—	7
機関H	8	4	4	13	14	21	1	21	18	1	10	17	2	8
機関I	19	11	16	18	12	1	4	11	17	11	16	—	9	9
機関J	15	13	2	5	7	12	21	20	15	9	9	3	15	10
機関K	11	5	21	20	17	15	5	18	7	17	3	6	7	11
機関L	14	8	8	4	11	10	12	13	20	10	11	13	10	12
機関M	21	—	—	11	5	14	18	7	5	14	—	—	—	13
機関N	7	7	7	6	16	20	22	8	2	16	21	15	6	14
機関O	20	14	6	14	21	16	15	9	6	12	13	9	12	15
機関P	16	—	1	9	22	8	16	10	11	13	14	16	—	16
機関Q	9	15	15	21	20	2	2	16	16	19	20	10	14	17
機関R	6	6	14	7	2	22	14	19	13	18	18	8	18	18
機関S	13	16	12	15	19	17	19	12	—	21	19	—	11	19
機関T	18	17	20	19	13	9	20	22	9	15	15	12	8	20
機関U	22	—	11	17	8	11	11	14	19	22	17	7	—	21
機関V	17	18	18	22	18	18	17	17	—	20	8	18	17	22
コンセンサス	1	3	3	2	4	6	5	2	1	5	3	4	4	1

(注)シャドーはコンセンサス予測を上回っているもの

図表5 各機関の指標毎の勝率(対コンセンサス)

	鉱工業生産	第3次産業活動指数	機械受注	全国・消費者物価	東京・消費者物価	失業率	有効求人倍率	貿易収支	経常収支	国内企業物価	実質消費支出	住宅着工戸数	マネー・ストック	全体
機関B	41.4%	45.7%	58.2%	38.5%	38.9%	40.0%	42.1%	51.3%	40.5%	39.6%	56.9%	50.0%	48.1%	46.6%
機関D	50.7%	—	55.8%	21.4%	25.0%	43.5%	26.3%	47.4%	43.0%	48.6%	60.4%	41.8%	53.8%	46.4%
機関A	41.4%	69.0%	42.3%	34.8%	43.3%	58.8%	41.7%	41.7%	42.5%	51.2%	45.1%	43.9%	44.4%	46.3%
機関C	53.0%	29.7%	44.9%	32.3%	40.7%	45.1%	51.4%	41.3%	45.1%	50.0%	50.0%	58.9%	34.8%	45.2%
機関E	41.3%	46.3%	40.5%	37.0%	52.9%	45.1%	47.9%	47.6%	48.8%	46.3%	48.1%	45.6%	23.3%	44.8%
機関H	42.3%	49.3%	46.3%	25.0%	35.0%	32.3%	55.2%	36.9%	37.3%	56.3%	49.0%	40.0%	50.0%	42.6%
機関N	41.2%	52.8%	52.4%	37.5%	31.6%	33.3%	26.7%	41.0%	54.4%	37.3%	37.0%	40.4%	44.8%	42.6%
機関L	38.7%	42.9%	44.3%	31.0%	40.0%	46.7%	41.2%	34.6%	32.1%	44.4%	48.1%	52.7%	40.7%	41.0%
機関G	39.3%	38.9%	43.4%	20.8%	38.9%	54.5%	55.6%	33.3%	44.1%	43.2%	42.2%	—	—	40.8%
機関F	27.9%	36.1%	39.0%	27.3%	34.3%	54.5%	38.9%	42.7%	40.7%	43.6%	50.0%	51.9%	42.3%	40.1%
機関K	41.6%	44.3%	34.2%	24.2%	31.7%	37.5%	46.3%	41.3%	42.9%	33.3%	53.8%	39.2%	31.6%	39.7%
機関P	38.5%	—	49.3%	25.8%	17.8%	50.0%	37.8%	40.3%	42.9%	27.3%	44.0%	39.6%	—	38.8%
機関I	27.9%	35.0%	41.3%	19.4%	41.5%	75.0%	54.5%	39.2%	37.0%	41.3%	35.4%	—	—	38.5%
機関J	27.7%	27.3%	52.7%	30.8%	37.1%	41.7%	21.2%	37.2%	47.4%	39.1%	44.9%	48.1%	28.9%	37.1%
機関R	39.4%	16.7%	48.1%	36.6%	46.5%	27.0%	42.2%	31.3%	41.0%	33.3%	32.7%	42.6%	16.7%	37.9%
機関M	28.6%	—	—	33.3%	41.9%	48.5%	38.5%	35.9%	46.6%	26.9%	—	—	—	37.0%
機関O	29.0%	23.3%	48.0%	22.6%	22.7%	42.3%	40.5%	35.7%	43.4%	36.7%	41.8%	51.8%	26.1%	36.8%
機関T	35.5%	24.7%	42.2%	22.2%	38.5%	50.0%	36.6%	24.7%	37.0%	38.1%	35.8%	48.2%	44.4%	35.9%
機関U	37.1%	—	37.0%	16.7%	38.9%	48.1%	44.1%	26.7%	34.2%	20.0%	33.3%	56.9%	0.0%	35.3%
機関Q	42.4%	27.4%	45.2%	5.7%	24.1%	62.5%	51.7%	27.5%	31.9%	27.7%	29.4%	50.9%	34.4%	34.9%
機関S	34.3%	32.9%	39.7%	27.8%	30.2%	40.0%	29.4%	32.5%	—	25.5%	37.3%	—	—	33.2%
機関V	37.1%	—	34.7%	10.5%	35.0%	39.3%	—	36.9%	—	31.0%	41.2%	36.5%	16.7%	31.8%
平均	38.1%	38.7%	44.8%	28.1%	35.7%	45.3%	41.4%	37.7%	41.7%	38.2%	43.2%	46.6%	35.9%	39.9%

(注)シャドーはコンセンサスに対する勝率が50%を超えているもの

さらに、毎回の調査毎に、各機関の予測誤差がコンセンサスの予測誤差よりも小さかった場合に勝ち、大きかった場合に負け、同じだった場合には引き分けとし、引き分けを除いた上で全期間における勝率(勝ち数/(勝ち数+負け数))を計算した(図表5)。

いずれの指標についても、コンセンサス予測に対する勝率が 50%を超える機関は少数である。また、指標毎に全機関の勝率を求めると、全ての指標が 50%を下回っており、最も勝率が高い住宅着工戸数でも 46.6%である。最も勝率が低いのは全国・消費者物価の 28.1%で、全ての機関が 50%を下回っている。13 指標合計の勝率をみると、最も高い機関でも 46.6%と 50%を下回っており、最も低い機関は 31.8%である。全指標、全機関の勝率は 39.9%である。

図表 5 では引き分けを除いた勝率を示したが、指標によっては引き分けの回数が多いことに留意が必要である。図表 6 では引き分けも考慮した勝率、負け率、引き分け率を示した。当然のことながら、実績値の変動が小さい指標では予測値がコンセンサスと一致する機関が多いため、引き分けの回数が多くなる。引き分け率が最も高いのは失業率で 64.3%、以下、全国・消費者物価 (56.4%)、有効求人倍率 (56.3%)、東京・消費者物価 (52.4%) は引き分けの確率が 50%を超えている。引き分けを除いたベースで見ると、失業率は突出して勝率が高い機関が存在する (75.0%、図表 5 参照) が、実際には引き分けの回数も多く、引き分けを分母に含めた勝率は 21.1% (全機関の平均は 16.2%) とそれほど高いわけではない。

図表6 コンセンサスに対する勝率、負け率、引き分け率

	勝率	負け率	引き分け率
鉱工業生産	33.8%	55.1%	11.1%
第3次産業活動指数	33.0%	52.2%	14.9%
機械受注	43.9%	54.2%	1.9%
全国・消費者物価	12.2%	31.3%	56.4%
東京・消費者物価	17.0%	30.6%	52.4%
失業率	16.2%	19.5%	64.3%
有効求人倍率	18.1%	25.6%	56.3%
貿易収支	37.7%	62.3%	0.1%
経常収支	41.7%	58.3%	0.0%
国内企業物価	26.5%	42.9%	30.6%
実質消費支出	40.3%	52.9%	6.8%
住宅着工戸数	45.8%	52.5%	1.7%
マネーストック	18.0%	32.1%	49.9%
全体	28.8%	43.4%	27.9%

(注)全機関の平均

3-2. 異なる指標間の予測誤差の相関

各機関の予測精度には偏りがあるようにも見受けられる。たとえば、22 機関、13 指標 (一部の指標は 22 機関未満) のうち、コンセンサス予測を上回っているものが延べ 30 指標あるが (図表 4 のシャド一部分)、このうち 6 割に当たる 18 指標が全体の成績上位 5 機関に集中している。

異なる指標間の各機関の予測誤差の相関係数をみると、第 3 次産業活動指数、国内企業物価は他の指標との相関が比較的高いことが分かる (図表 7)。第 3 次産業活動指数は他の 4 指標、国内企業物価は他の 3 指標との相関係数が 0.5 を超えている。また、各指標の予測誤差と 13 指標全体の予測誤差の相関をみると、国内企業物価 (0.803)、第 3 次産業活動指数 (0.774)、貿易収支 (0.735) の順に相関係数が高くなっている。特定の指標に関しては、その指標の予測精度が高い機関はそれ以外のいくつかの指標についても予測精度が高く、指標全体の予測精度も高い傾向があるという結果が得られた。

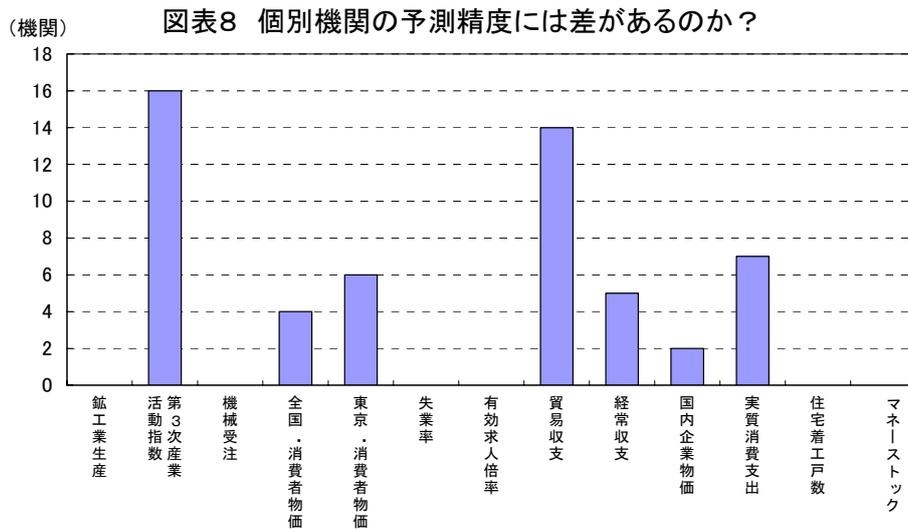
図表7 指標間の予測誤差の相関係数

	鉱工業生産	第3次産業	機械受注	消費者物価(全国)	消費者物価(東京)	失業率	有効求人倍率	貿易収支	経常収支	国内企業物価	実質消費支出	住宅着工戸数	マネーストック	全体
鉱工業生産	1.000													
第3次産業活動指数	0.480	1.000												
機械受注	0.078	0.125	1.000											
消費者物価(全国)	0.268	0.666	0.487	1.000										
消費者物価(東京)	0.281	0.648	-0.197	0.370	1.000									
失業率	-0.034	-0.157	-0.104	-0.159	-0.161	1.000								
有効求人倍率	0.292	0.369	-0.146	-0.064	0.152	0.324	1.000							
貿易収支	0.269	0.559	0.185	0.451	0.199	0.235	0.270	1.000						
経常収支	0.304	-0.178	-0.191	0.157	0.063	-0.160	-0.246	0.380	1.000					
国内企業物価	0.657	0.607	0.285	0.488	0.283	0.140	0.274	0.417	0.299	1.000				
実質消費支出	0.368	0.409	0.097	0.234	0.215	0.080	0.324	0.564	-0.047	0.603	1.000			
住宅着工戸数	0.199	0.372	0.135	0.224	0.326	0.341	0.204	0.281	-0.122	0.304	0.216	1.000		
マネーストック	0.041	0.276	0.270	0.233	-0.164	0.384	0.114	0.263	-0.065	0.449	0.269	0.299	1.000	
全体	0.582	0.774	0.292	0.682	0.419	0.247	0.419	0.735	0.320	0.803	0.648	0.619	0.487	1.000

(注) 全体との相関係数は予測誤差を標準化変量で基準化した上で計測した。シャドーは相関係数が0.5を超えるもの

3-3. 指標毎の予測精度の差

それでは、各指標について個別機関の予測精度には差があるのだろうか。この点を統計的に確かめるために指標毎に各機関の予測誤差について平均値の差の検定を行ったところ、13指標中、第3次産業活動指数、全国・消費者物価、東京・消費者物価、貿易収支、経常収支、国内企業物価、実質消費支出の7指標については、予測精度が最も高い（予測誤差が最も小さい）機関に対して、有意（5%水準）に劣っている機関があった（図表8）。特に、第3次産業活動指数については、成績が最も良い機関を除く17機関のうち16機関の予測誤差が有意に大きいという結果となった。



(注) 指標毎の予測誤差について、成績が1位の機関と5%水準で有意差があった機関の数

さらに、各機関の成績が安定的なものであるかどうかをみるために、各指標について異なる年度間の予測誤差の相関係数を計測した（図表9）。2003年度から2009年度まで7年間のデータがある指標については、21通りの組み合わせ（2003年度－2004年度、2003年度－2005年度、・・・、2008年度－2009年度）についてそれぞれ相関係数を求め平均した。全体としては相関係数がそれほど高くないが、国内企業物価（0.385）、第3次産業活動指数（0.362）、実質消費支出（0.277）などは比較的相関が高くなっている。逆に、鉱工業生産、有効求人倍率は相関係数がほぼゼロで、ほとんど相関がないという結果であった。

図表9 予測誤差の年度間相関係数

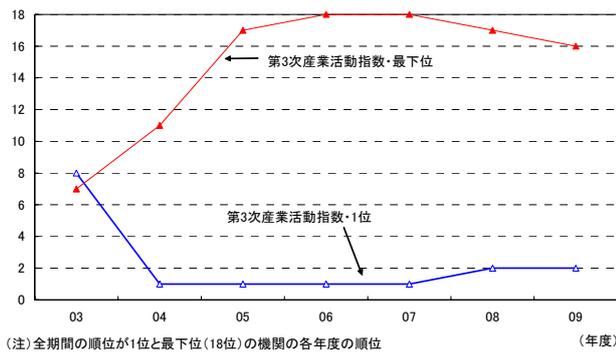
	鉱工業生産	第3次産業	機械受注	消費者物価(全国)	消費者物価(東京)	失業率	有効求人倍率	貿易収支	経常収支	国内企業物価	実質消費支出	住宅着工戸数	マネーストック	全体
年度間	0.000	0.362	0.064	0.146	0.225	0.072	-0.007	0.147	0.229	0.385	0.277	0.186	0.153	0.390
各年度対全期間	0.336	0.671	0.450	0.485	0.565	0.446	0.361	0.510	0.584	0.665	0.655	0.539	0.496	0.669

(注)年度間は2003年度から2009年度までの全ての組み合わせを平均したもので、シャドローは相関係数が0.5を超えるもの

各年度と全期間の相関は比較的高く、13指標のうち7指標は相関係数が0.5以上、国内企業物価、第3次産業活動指数、実質消費支出の3指標は0.6以上となっている。これらの指標については、ある年度に関して成績が優秀だった機関は全期間にわたって優秀な成績を収める確率が高いということであり、比較的成績が安定している指標といえることができるだろう。

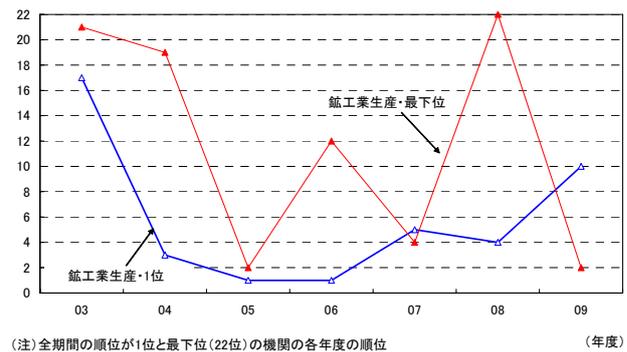
指標によって成績が安定しているかどうかを実際の事例によって確認してみると、第3次産業活動指数は、全期間の成績が1位の機関は7年度中4回が1位、2回が2位、1回が8位と毎年安定的に優秀な成績となっており、全期間の成績が最下位の機関はほぼ毎年悪い成績となっている。一方、年度間の相関係数が低い鉱工業生産については、全期間で1位（最下位）の機関であっても年度毎の順位にはばらつきが見られる（図表10-1、2）。

図表10-1 成績が安定している指標の年度毎の順位



(注)全期間の順位が1位と最下位(18位)の機関の各年度の順位

図表10-2 成績が不安定な指標の年度毎の順位



(注)全期間の順位が1位と最下位(22位)の機関の各年度の順位

4. コンセンサス予測を上回るための試み

4-1. 予測値の組み合わせによる効果

前節まで見てきたように、個別機関の予測が長期にわたってコンセンサス予測を上回ることが非常に困難である。13指標のうち11指標ではコンセンサス予測よりも優秀な成績をおさめている機関が存在するが、あくまでも2003年度～2009年度の7年間のデータに基づいた結果であり、これが今後も続く保証はない。また、仮に個別機関に関する今回の結果が安定的なものだったとしても、予測の利用者（たとえば投資家）が優秀な機関を見抜くことは現実的には極めて困難だろう。

一方、今回の分析で明らかになったことは、指標によっては機関毎の予測精度に差があること、指標毎の各年度の成績と全期間の成績には一定の相関があることである。また、第2節でみたように、コンセンサス予測が個別機関の予測よりも優秀となるのは、複数の機関の予測値を組み合わせることにより、個別機関のプラスの誤差とマイナスの誤差が打ち消されることによる。

こうした事実に基づけば、指標毎にある年度の優秀機関を複数取り出し、その予測値を組み合わせることで、長期にわたり予測誤差の縮小が期待できるのではないかと考えられる。1年間という限られた

期間であれば、個別機関の予測値と実績値を比較することで、投資家などが優秀な機関を見分けることも可能だろう。

ここでは、指標毎に各年度の成績上位5機関（10機関）の予測値を取り出し、その平均を全期間（2003年度～2009年度の84ヵ月）にわたって利用した場合の誤差を求めた。5機関（10機関）の組み合わせは2003年度から2009年度までの7通り（一部の指標は2004年度～2009年度の6通り、2005年度～2009年度の5通り）で、結果はその平均値である。指標間の比較がしやすいように、予測誤差を標準化変量で標準化したものを示した。各機関の予測誤差の平均値はゼロ、標準化変量のマイナスが大きいほど予測誤差が小さいことを意味する。コンセンサス予測に対する勝率も計算した（図表11）。

図表11 組み合わせによる予測精度の改善状況

	予測誤差(標準化変量)			勝率(対コンセンサス)		
	コンセンサス	上位5機関	上位10機関	全機関平均	上位5機関	上位10機関
鉱工業生産	-2.17	△2.08	△2.14	38.1%	48.2%	47.5%
第3次産業活動指数	-1.18	-1.74	-1.58	38.7%	65.3%	68.4%
機械受注	-1.15	-1.29	-1.40	44.8%	50.6%	50.5%
消費者物価(全国)	-1.12	-1.15	-1.17	28.1%	34.8%	33.3%
消費者物価(東京)	-0.98	-1.11	-1.13	35.7%	51.7%	61.4%
失業率	-0.66	-1.09	-0.95	45.3%	62.4%	65.0%
有効求人倍率	-0.91	△0.90	-1.15	41.4%	55.4%	63.2%
貿易収支	-1.87	-1.94	-2.20	37.7%	51.2%	57.5%
経常収支	-1.52	△1.51	-1.65	41.7%	49.5%	52.0%
国内企業物価	-0.98	-1.25	-1.28	38.2%	55.3%	62.8%
実質消費支出	-0.95	-1.47	-1.25	43.2%	62.5%	64.2%
住宅着工戸数	-0.84	-1.07	-1.07	46.6%	52.8%	55.6%
マネーストック	-0.93	-1.30	-1.18	35.9%	55.4%	53.3%
全体	-1.17	-1.38	-1.40	39.9%	53.3%	55.9%

(注)各年度の上位5機関(10機関)の予測平均値を使った場合の誤差、勝率(2003～2009年度)
標準化変量はマイナスが大きいほど予測誤差が小さいことを表す(各機関の平均値はゼロ)
△はコンセンサス予測よりも誤差が大きいことを表す

全ての指標について、予測誤差の標準化変量はマイナスとなり、個別機関の平均値よりも予測誤差が縮小することが確認された。コンセンサス予測との比較では、上位5機関を取り出した場合は、鉱工業生産、有効求人倍率、経常収支を除く10指標、上位10機関を取り出した場合には、鉱工業生産を除く12指標でコンセンサス予測よりも誤差が小さくなった。

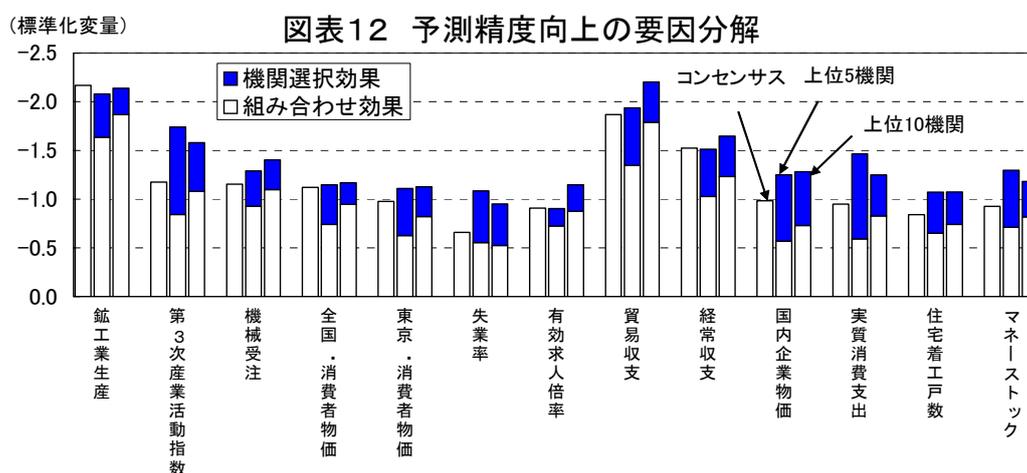
勝率についても、全ての指標について個別機関の平均値よりも高くなり、全体の勝率は5機関、10機関の場合ともに50%を上回った。ただし、経常収支は上位5機関の場合に50%を下回り、鉱工業生産、全国・消費者物価は上位5機関、10機関ともに50%を下回った。

4-2. 予測精度向上の要因分解

図表12は、個別機関全体の平均に比べ予測誤差が縮小した要因を、優秀な機関の予測値を選択した効果（機関選択効果）と予測値の組み合わせによって誤差が縮小した効果（組み合わせ効果）に分けて見たものである。機関選択効果は、選択した各機関の平均誤差と全機関の平均誤差の差、組み合わせ効果はそれ以外の部分とした。

コンセンサスの予測誤差縮小は全て組み合わせ効果によるものであるが、予測値を利用する機関を限定することによって、組み合わせ効果は小さくなる。一方、コンセンサス予測では機関選択効果はゼロ

だが、予測値を利用する機関を比較的優秀な機関に限定することによって、機関選択効果はプラスとなる。今回の例では、組み合わせ効果は5機関を選択した場合よりも10機関を選択した場合のほうが大きくなり、機関選択効果は5機関のほうが10機関よりも大きくなる傾向がある。コンセンサス予測よりも誤差が縮小するかどうかは、組み合わせ効果の縮小分よりも機関選択効果のプラス幅が大きくなるかによって決まる。



鉱工業生産については、上位5機関、10機関ともにコンセンサス予測を上回ることができなかったが、これにはふたつの要因が考えられる。第一に、鉱工業生産はコンセンサス予測の全期間の順位が1位となっていることから分かる通り、コンセンサス予測の組み合わせ効果が非常に大きく、機関を限定することによって組み合わせ効果が大きく縮小してしまう。これは5機関を選択した場合にコンセンサス予測を下回った経常収支（コンセンサス予測の順位1位）と共通の事象である。第二に、鉱工業生産は年度間の予測誤差の相関係数が低い（0.000）ことから分かるように、機関毎の予測精度の差が小さいため、機関選択効果が小さい。これは、5機関を選択した場合にコンセンサス予測を下回った有効求人倍率（相関係数：-0.007）と共通の事象である。

逆に、今回の試行によってコンセンサス予測を大きく上回った指標は、第3次産業活動指数、国内企業物価、実質消費支出で、これらの指標に共通するのは、個別機関の予測誤差の年度間の相関が高いことである。選択された機関が全期間にわたって優秀な成績となる確率が高いため、機関選択効果が大きくなり、このことが予測精度の大幅な改善につながっていると考えられる。

今回の試みからは、一定の条件を満たせば、単年度で優秀な機関の予測値を複数組み合わせることにより、コンセンサス予測を上回る予測値を得ることが可能であることが確認された。

まとめ

本稿では、13の月次指標を用いてコンセンサス予測と個別機関予測の特徴を分析した。コンセンサス予測が個別機関の予測よりも優秀となるのは、複数の機関の予測値を組み合わせることにより、個別機関のプラスの誤差とマイナスの誤差が打ち消されるためである。2003年度から2009年度までのデータ

を用いて指標毎に個別機関の予測誤差、コンセンサスの予測誤差を計算して順位付けしたところ、コンセンサス予測は全ての指標で上位となった。

個別機関の予測は総じてコンセンサス予測に劣っているが、各機関の予測精度には差が見られた。最も優秀な予測機関は 13 指標中、6 指標でコンセンサス予測を上回っている一方、22 機関の約半数にあたる 10 機関はコンセンサス予測を上回る指標がひとつもなかった。

第 3 次産業活動指数など特定の指標については、機関によって予測精度に大きな差が見られ、単年度の成績が優秀な機関は全期間にわたって優秀な成績となる傾向があることが確認された。

比較的パフォーマンスの良い機関の予測値を組み合わせることによって、予測精度を高めることが可能である。指標毎に単年度の成績が優秀な機関を複数（上位 5 機関、上位 10 機関のケースを試算）取り出し、これらの機関の予測平均値を使うと、全ての指標で全期間（2003～2009 年度）の予測誤差が個別機関全体の平均値よりも小さくなり、鉱工業生産を除くほとんどの指標でコンセンサスを上回るパフォーマンスが得られた。

個別機関の予測が長期にわたってコンセンサス予測に勝つことは難しいが、比較的優秀な機関の予測値を組み合わせることにより、コンセンサス予測に勝つ確率は高くなる。このことは投資家など月次指標の予測を利用する者にとって有益な情報と言えるだろう。

今回の分析では、指標毎に単年度の成績上位機関の予測値を組み合わせるという方法で予測精度の向上を試みたが、組み合わせの方法にはさらなる工夫の余地があるだろう。たとえば、今回の分析では指標によっては他の指標との相関が高いという結果が得られたが、こうした情報を活かした組み合わせは試さなかった。また、指標によってはある機関の予測精度が突出して高いことが観察された。このような指標は複数の機関を組み合わせるよりも、最も優秀な機関の予測値を単独で利用したほうが予測精度は高くなるが、その機関の成績が安定的に優れているかどうかを見極めることは必ずしも容易ではない。こうした点につき分析を深めることが今後の課題である。

（参考文献）

- 芦屋政浩（2008）「GDP・CPI・為替レート・失業率の予測成績に関する補正済み予測順位を用いた分析」平成 18 年度内閣府経済社会総合研究所委託調査報告書「平成 19 年度 多数の予測機関、エコノミスト等の経済に関する予想を用いた期待形成の調査研究」第 1 章
- 河越正明（2007）「コンセンサス予測は単なる平均的な予測か？：ESP フォーキャスト調査の評価の再検討」ESRI Discussion Paper Series No. 180 内閣府経済社会総合研究所
- 河越正明（2009）「2008 年度 ESP フォーキャスト調査の評価に関するテクニカルノート」ESP フォーキャスト調査参考資料（2009. 8. 26）
- 草場洋方（2010）「経済指標、国債入札、金融政策、季節性・・・長期金利の日時変動にとって意味のある情報とは？」みずほマーケットインサイト 2010 年 5 月 12 日
- （社）経済企画協会（2009）「ESP フォーキャスト調査（2008 年度）の予測評価について」ESP フォーキャスト調査参考資料（2009. 8. 26）
- 小巻泰之・矢嶋康次（2006）「景気指標の予測精度と予測形成」ニッセイ基礎研究所所報 2006 年 Vol. 42