

法人企業統計の季節調整における新型コロナウイルス
の影響について

高岡 慎*

2021年3月

*琉球大学国際地域創造学部

1 はじめに

2019 年に出現した新型コロナウイルス感染症は、2020 年を通して世界各国で大量の感染者を発生させており、経済に対しても大きな影響を及ぼしている。日本においても 2020 年半ばから主要なマクロ経済データでイレギュラーな変動が生じている。

本報告書は、法人企業統計調査における新型コロナウイルスの影響について、特に季節調整に対する影響という観点から、2020 年 10-12 月期までの実績値に基づいて状況をまとめたものである。

法人企業統計調査では、季節調整プログラム X-12-ARIMA によって季節調整を施した調整済系列に基づく前期比増加率を原数値と合わせて公表している。X-12-ARIMA の運用においては、原系列の統計的性質に合致する時系列モデル (季節 ARIMA モデル) を適切に選択する必要があるが、法人企業統計調査では毎年 6 月に同年 1-3 月期までのデータに基づくモデルの再選択を実施しているが、一般にモデル選択のプロセスにおいては、大きなレベルシフトなどの存在が結果に影響を与えることが多く、注意深い対応が必要となる。

新型コロナウイルスによるデータの変動は、2021 年 3 月時点では収束したかどうかはまだ不明瞭であることに加え、2020 年 6 月の再選択では、調査票回収の遅延のために速報値を利用した検討で代替するなど、季節調整の作業においては様々な点で通常と異なる対応をせざるを得ない状況が生じている。こうしたことから、本報告書では、2021 年 6 月のモデル再選択に向けた予備的な検証として、2020 年 10-12 月期までのデータの変動傾向について分析した結果を示す。

2 2020 年 10-12 月までの季節調整値について

まず、2020 年 10-12 月までの季節調整値の変動を確認しておく。図 1-1 から図 1-10 は、2020 年 1-3 月期までのデータに基づいて選択されたモデルを用いた季節調整系列を表している。

2020 年 1-3 月期までの実績値を確認した時点では、4-6 月期以降にリーマン・ショックの際に生じたような急激な下落が予想されたものの、それがどの程度の期間に渡って継続するかは不明であり、適当なダミー変数の選択は難しかったが、2020 年 10-12 月までの実績値からは以下のような傾向が読み取れる。

- 設備投資関連の4系列を除く6系列では、トレンドおよび不規則成分などから、2020年4-6月を谷とする大きな下落があり、10-12月にかけて回復に向かっている。
- 設備投資関連の4系列については、新型コロナウイルスの感染拡大と連動するような変動はあまり明瞭ではない。
- 設備投資・ソフト除く（製造業）および設備投資・ソフト除く（非製造業）については、コロナ以前から傾向的な低下が継続している。
- 設備投資・ソフト（製造業）および設備投資・ソフト（非製造業）については、2020年中も含めて上昇傾向を示している。

3 2020年の変動の特性について

次に、2020年中に生じた変動がどの程度イレギュラーなものであったかを確認するため、2020年1-3月期までのデータに基づいて選択されたモデルを用いて2020年4-6月期から2020年10-12月期までの3期間を予測し、実績値との比較を行った。結果は図2-1から図2-10に示した。図から観察できる点は以下のようなになる。

- 設備投資・ソフト（製造業）を除く9系列では、2020年4-6月期および7-9月期で実績値が予測値を大きく下回っている。これは2020年1-3月期までのデータの傾向からは予想されない大きな変動であり、新型コロナウイルスによる経済活動への影響が示唆される。
- 設備投資・ソフト除く（製造業）および設備投資・ソフト除く（非製造業）の2系列を除く8系列では、2020年10-12月期には実績値が予測値と同等かそれ以上の水準に回復しつつある。
- 設備投資・ソフト除く（製造業）および設備投資・ソフト除く（非製造業）の2系列では、2020年4-6月以降予測値を下回る傾向が続いている。

4 モデル選択への影響の試算

法人企業統計調査では毎年6月に季節調整モデルの見直しを実施し、その際に選択されたモデルをその後1年間継続して使用するという方針を採用している。モデルの見直

しの際には、前年のモデル設定後に蓄積される4期分のデータを追加した系列を用い、81通りのモデル候補の中からAICを最小にモデルを採用している。ただし、モデルの変更によって生じる過去の季節調整系列の改訂の程度をコントロールするために、各モデルについて「平均乖離幅(SR)」を計算し、AICと併用している。平均乖離幅は、1年前のモデル見直しの際に選択されたモデル(現行モデル)をそのまま使用して計算した場合の季調系列から求めた前期比増加率と、各候補モデルから計算される季調系列による前期比増加率との間の平均的な乖離幅として定義される。計算の詳細については付録を参照されたい。

AICとSRを併用する際には、各系列についてSRの上限値を定めておき、81通りのモデルの中でSRが一定の上限値を下回るモデルのみを候補とし、候補モデルの中でAICが最小になるモデルを最適モデルとして採用するという手順をとっている。これは、モデル変更によって生じる公表系列のギャップを軽減する効果がある。

SRの上限値(以下 a と表記する)については、過去データによる検証結果から、経験的に売上高(製造業): $a = 1.0$ 、売上高(非製造業): $a = 1.0$ 、経常利益(製造業): $a = 2.0$ 、経常利益(非製造業): $a = 1.0$ 、営業利益(製造業): $a = 2.0$ 、営業利益(非製造業): $a = 1.0$ 、設備投資・ソフト除く(製造業): $a = 1.0$ 、設備投資・ソフト除く(非製造業): $a = 1.0$ 、設備投資・ソフト(製造業): $a = 1.0$ 、設備投資・ソフト(非製造業): $a = 1.0$ としている。

こうしたモデル選択法によると、系列の統計的性質が安定している場合、AICの小さいモデルはSRも同様に小さくなる傾向にあるが、直近1年以内の変動に大きな外れ値やレベルシフトが含まれると、81通りのモデル候補のSR値が大きくばらつき、上限値以下となるモデルの数がかなり少なくなる可能性がある。その結果、AICの観点から見てあまり適切ではないモデルが選択される恐れがあるため、直近付近でのイレギュラーな変動が予想される場合は、SR値の分布をモニターしておくことが望ましい。

4.1 SR値の分布

図3-1から図3-10は、現行のモデル選択法に基づくモデル選択を、2019年1-3月期から仮想的に每期行った場合に、81通りの候補モデルのSRがどのように分布するかを示した図である。各図の表示や記号の意味は以下の通りである。

- 各時点の縦のばらつきは、その4期前における最適モデルから測った、各候補モデルのSRを示している。

- 2018年1-3月期から2018年10-12月期の最適モデルは、2018年6月のモデル選択で採用されたモデルを用いている。
- 「AIC改善モデル」は、4期前の最適モデルと比較してAICが減少しているモデルを表す。
- 「AIC悪化モデル」は、4期前の最適モデルと比較してAICが増加しているモデルを表す。
- 「AIC最小モデル」は、81通りのモデルのうちAICが最小になるモデルを表す。
- 「最適モデル」は、 $SR \leq a$ を満たすモデルのうちAICが最小になるモデルを表す。
- 横方向の実線は現行のSR上限値 a を表す。
- 最適モデルのSR値がゼロになっている場合は、1年前の最適モデルが再度選択されていることを表している。
- 最適モデルとAIC最小モデルのSR値が重なっている場合は、AIC最小モデルが選択されている。
- 10本の系列は、モデルの選択結果に大きな変化が無い安定した系列と、全体的にSRのばらつきが大きく上限値の調整を検討すべき系列に分かれている。
- モデルの変化が少ない安定した系列では、ほとんどの候補モデルのSRは現行の上限値以下に収まっており、SRが大きいモデルは概ねAICも大きくなっている。
- 経常利益および営業利益では、2020年4-6月期および2020年7-9月期でSR値のばらつきが急激に増大しており、直近付近の大きなショックがモデル選択に対する影響が予想される。

4.2 AO変数の推定

表4-1は、2020年1-3月期までのデータから選択された最適モデルに対し、2020年4-6月、2020年7-9月、2020年10-12月の3時点にAO変数を設定した場合の係数推定値（括弧内はt値）を表している。また最右列はAO変数の導入によって生じたAIC値の変化を表している。なお、経常利益および営業利益については係数推定値の絶対値が

他の系列と大きく異なっているが、これは経常利益および営業利益では X-12-ARIMA の適用時に対数変換を行っていないことによる。

表 4-1: AO 変数の係数推定値

	ao2020.2	ao2020.3	ao2020.4	AIC 変化
売上高 (製造業)	-0.159 (-5.724)	-0.093 (-2.553)	-0.022 (-0.478)	-34.348
売上高 (非製造業)	-0.097 (-3.891)	-0.033 (-1.027)	0.021 (0.492)	-21.945
経常利益 (製造業)	-22876.184 (-3.507)	-4581.970 (-0.554)	13194.448 (1.375)	-13.159
経常利益 (非製造業)	-51304.193 (-6.430)	-21279.251 (-2.430)	-5587.203 (-0.590)	-30.075
営業利益 (製造業)	-25654.767 (-6.156)	-4459.791 (-0.665)	14698.221 (1.704)	-60.445
営業利益 (非製造業)	-50194.602 (-6.900)	-17478.331 (-2.204)	-4201.251 (-0.466)	-36.124
設備投資・ソフト除く (製造業)	-0.120 (-2.002)	-0.092 (-1.174)	-0.149 (-1.456)	1.427
設備投資・ソフト除く (非製造業)	-0.107 (-1.879)	-0.156 (-2.293)	-0.107 (-1.314)	0.115
設備投資・ソフト (製造業)	-0.005 (-0.061)	-0.083 (-0.951)	0.051 (0.579)	4.436
設備投資・ソフト (非製造業)	-0.172 (-1.592)	-0.193 (-1.657)	0.009 (0.079)	2.514

※括弧内は t 値

表 4-1 から読み取ることができる要点をまとめると以下の通りとなる。

- 売上高、経常利益、営業利益については、2020 年 4-6 月期の AO 係数が概ねマイナスに有意となる。2020 年 7-9 月期は部分的にマイナスに有意となる。
- 売上高、経常利益、営業利益の 2020 年 10-12 月期の AO 係数は、いずれも有意とは判断できない。
- 売上高、経常利益、営業利益については、異常値にダミー変数を導入することにより AIC の改善が見られる。
- 設備投資については、AO 変数の t 値の絶対値が全体的に小さく、AO 変数による AIC の改善も見られない。

5 結論

ここまでの計算を踏まえた結論として、2021 年 6 月に予定されているモデル再選択における検討課題等をまとめておく。

(1) 新型コロナウイルスに対応する変動へのダミーの適用について

- 売上高、経常利益、営業利益の6系列に対しては、AOまたはRAMPを適用することにより適切なモデルが選択されると思われる。
- 現実的にはコロナの影響が進行中の間はAOで対応し、RAMPは変動が一段落した後、事後的に適用を検討することになる。
- 売上高、経常利益、営業利益の6系列については、コロナウイルスによる影響が、4-6月期を谷として、10-12月期には概ね収束する比較的短期間の変動となる可能性があるため、リーマンショックなどとは異なり、2020年4-6月（および2020年7-9月）にAO変数を導入するのみで対応できる可能性がある。
- 設備投資関連の4系列については、ダミー変数を導入するほどの外れ値は発生していないと思われる。

(2) ダミー変数と季節調整系列の関係

- X12では、推定された回帰変数の影響をデータから除去した残差の変動をARIMAモデルで予測する。
- ダミー変数を直近のコロナの変動に適用した場合、直近付近の下落は回帰変数により処理されるので、ARIMAモデルによる残差の予測は、コロナ以前の変動傾向をそのまま伸ばしたような値になる。最終的には、回帰変数により処理されたコロナの影響は、AOの場合は不規則変動に、RAMPの場合はトレンド成分にそれぞれ割り振られる。
- ダミー変数を用いない場合、ARIMAモデルによる短期的な予測は、直近の下落を反映して下降傾向を示すと思われる。コロナの影響は、移動平均による平滑化の過程でトレンド成分と不規則変動の両方に適当に割り振られる。最終的な季節調整系列は、ダミーを用いた場合と比較すると、変化点が不明瞭な系列になる。
- ただし、季節調整系列はトレンドと不規則成分を合成した系列として定義されているので、ダミーを用いるかどうかや、AOとRAMPのどちらを用いるかといったことは、コロナの変動をトレンドと不規則変動のどちらに分類するかという問題には影響するが、季節調整系列そのものに対してはそれほど大きな影響は持たない。

(3) モデル選択法

- ダミー変数をどのように扱うかという点は、最終的な季節調整系列そのものにはそれほど大きくは影響しないが、SR の上限値の設定によっては、AIC があまり小さくない、統計的には適切でないモデルが選択される恐れがある。
- 季節 ARIMA モデルが統計的に適切でないモデルである場合には、点予測の精度が落ちることによって、季節調整系列が事後的な改訂の大きい不安定な系列になる可能性がある。
- 現行のモデル選択法では、過去に選択された最適モデルからの乖離幅 (SR) を考慮するので、ある時点であまり適切でないモデルを選択してしまうと、その影響がそれ以降のモデル選択結果にしばらく残存する可能性がある。

(4) 2021 年 6 月のモデル検討の方針について

- 2021 年 1-3 月までの数値を利用し、コロナの変動へのダミー変数の適用を検討する。
- 2020 年 4-6 月、2020 年 7-9 月、2020 年 10-12 月の 3 時点は異常値になっている可能性があるが、売上高、経常利益、営業利益の 6 系列は 2020 年 4-6 月、2020 年 7-9 月における AO 変数の導入を中心に検討する。変化点の候補が隣接しているため、RAMP 変数より AO 変数の方がより適切と思われる。
- 設備投資・ソフト除く (製造業) および設備投資・ソフト除く (非製造業) については、2018 年または 2019 年から傾向的な下落が続いており、コロナウイルスに関連した特段の対応は不要となる可能性が高い。
- 設備投資・ソフト (製造業) および設備投資・ソフト (非製造業) については、長期的な上昇トレンドが継続しており、コロナウイルスに関連した特段の対応は不要となる可能性が高い。
- 経常利益および営業利益の 4 系列については、SR の上限の設定がモデル選択の幅をどの程度制約しているかを確認する必要がある。
- 2021 年 1-3 月期については、緊急事態宣言の影響等により、再度数字が落ち込む可能性がある点に注意する。

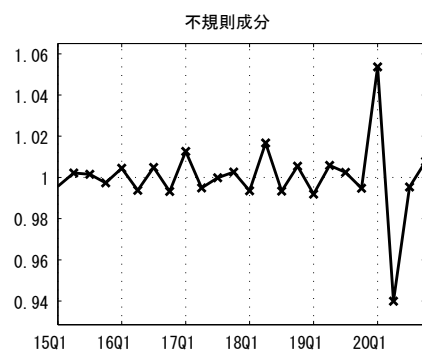
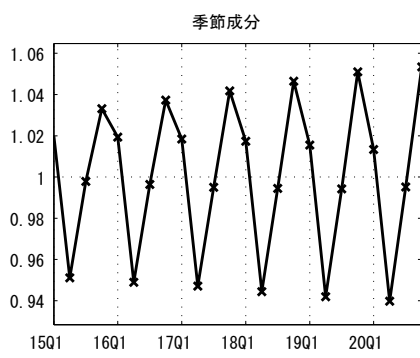
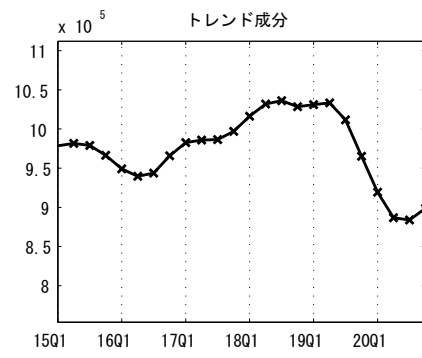
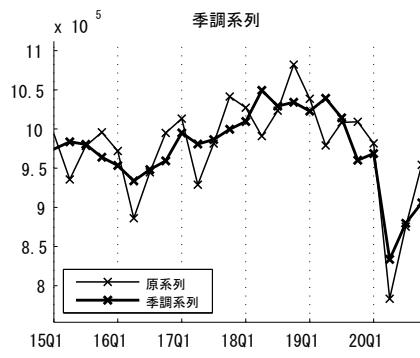


図 1-1: 売上高 (製造業)

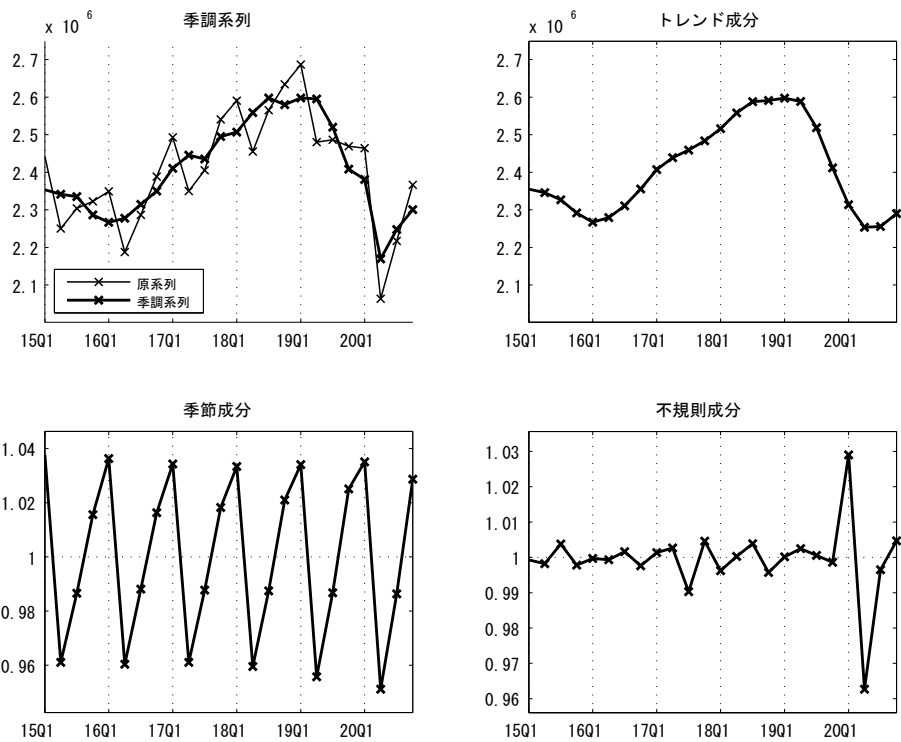


図 1-2: 売上高 (非製造業)

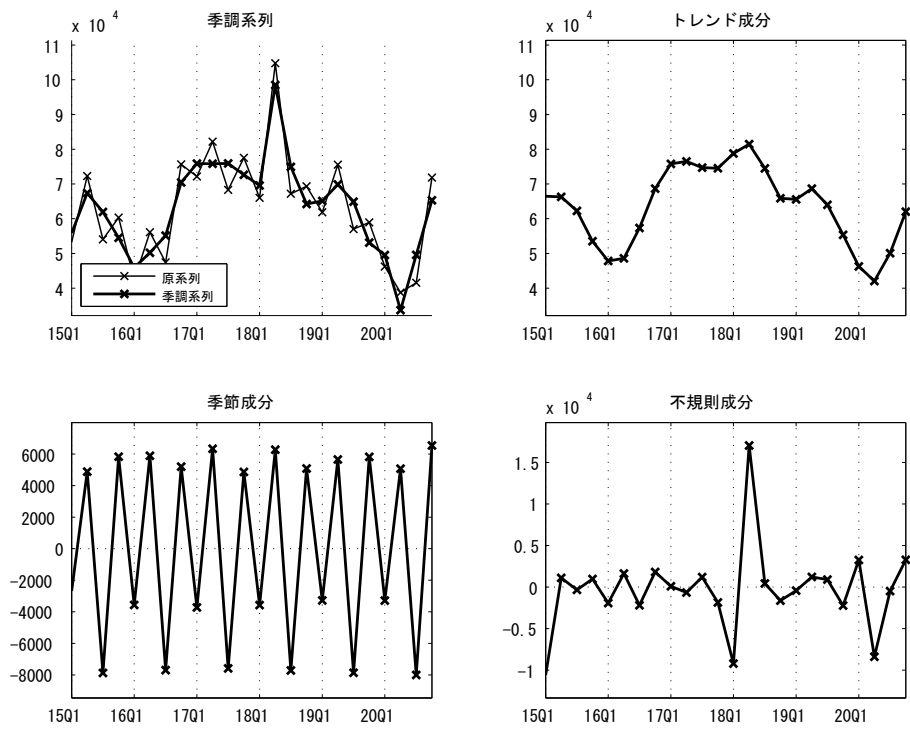


図 1-3: 経常利益 (製造業)

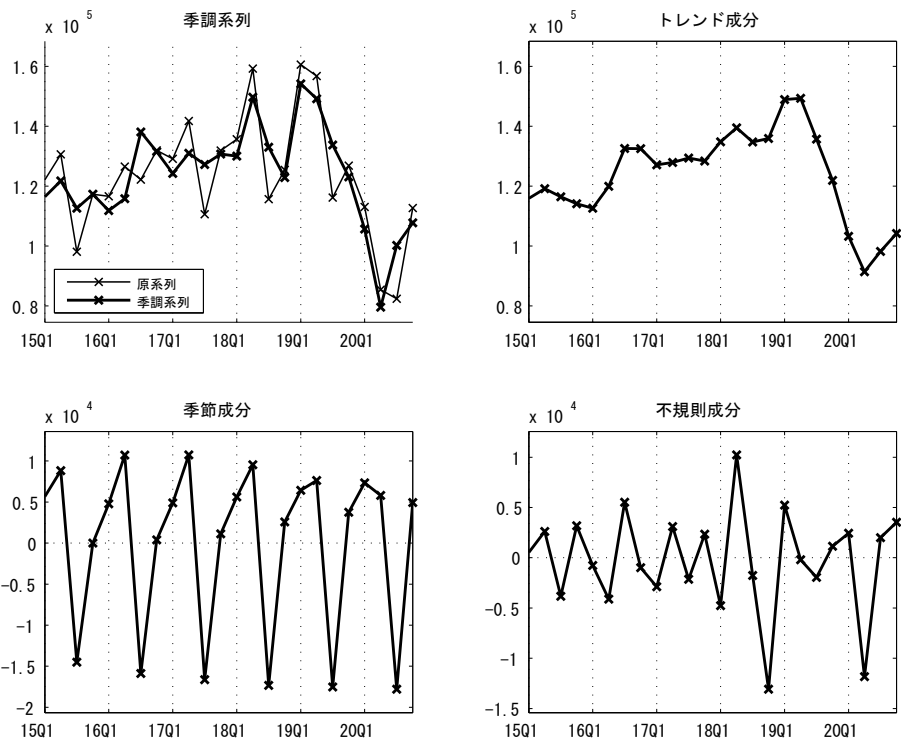


図 1-4: 経常利益 (非製造業)

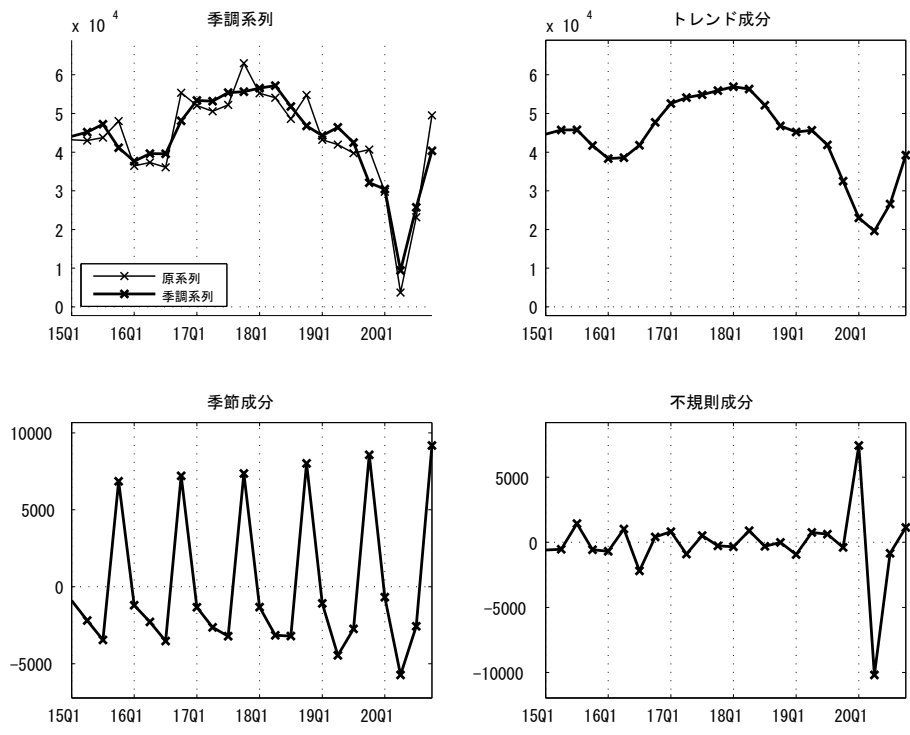


図 1-5: 営業利益 (製造業)

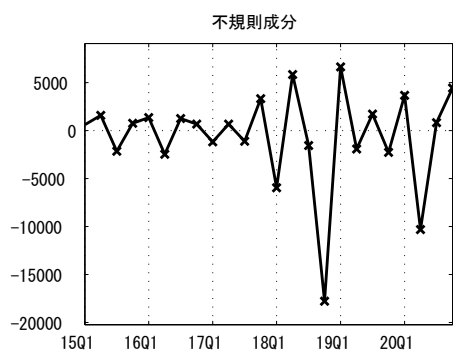
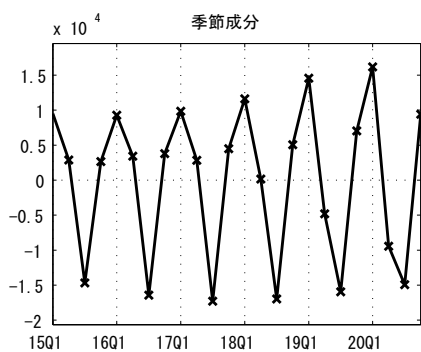
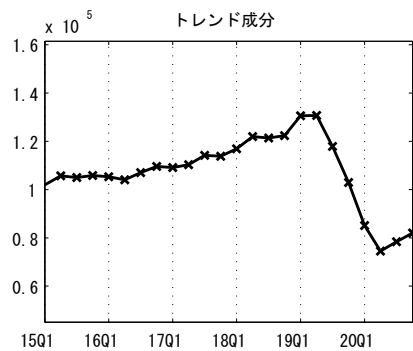
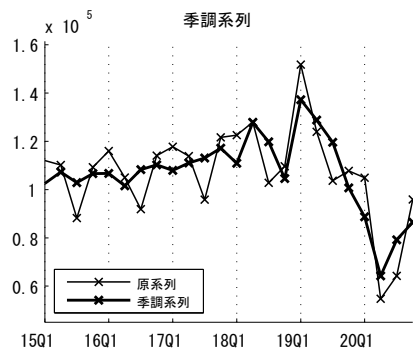


図 1-6: 営業利益 (非製造業)

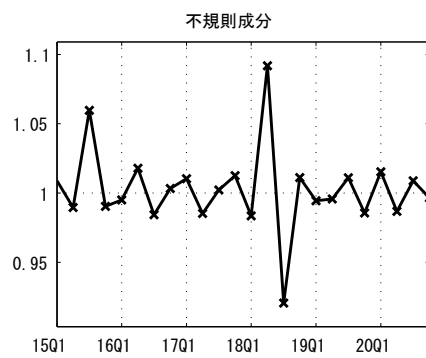
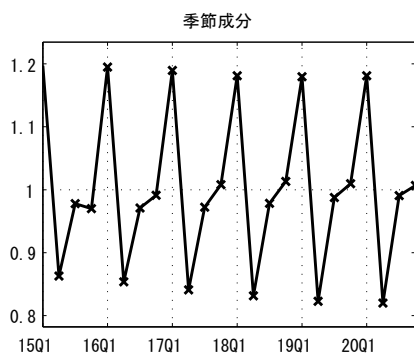
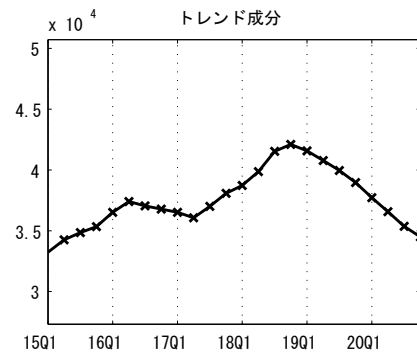
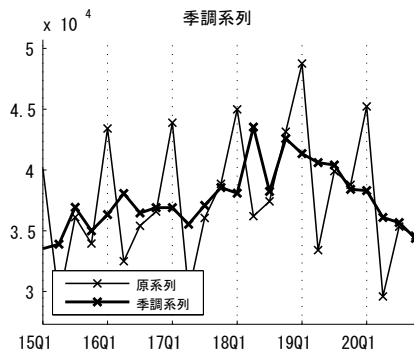


図 1-7: 設備投資・ソフト除く (製造業)

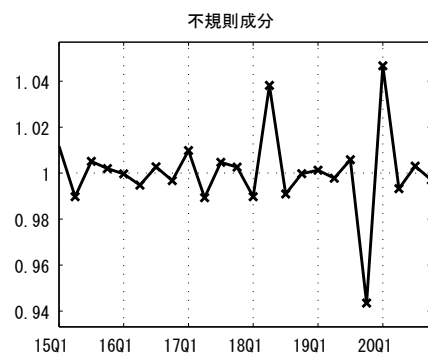
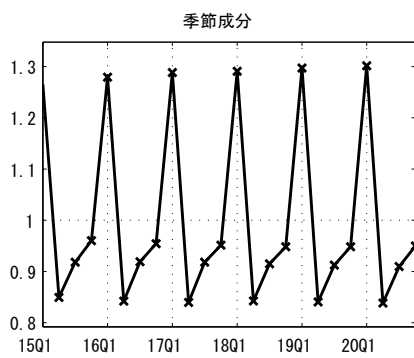
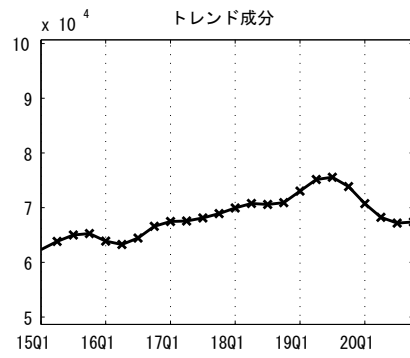
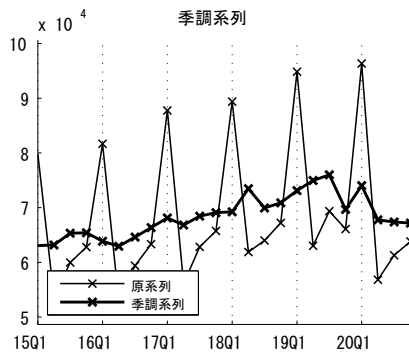


図 1-8: 設備投資・ソフト除く (非製造業)

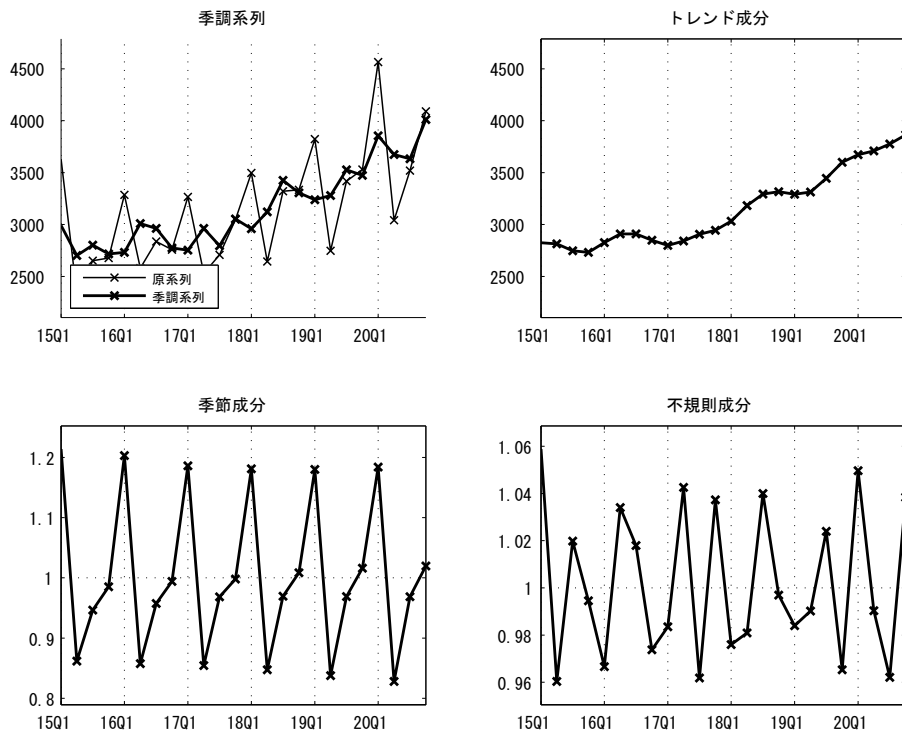


図 1-9: 設備投資・ソフト (製造業)

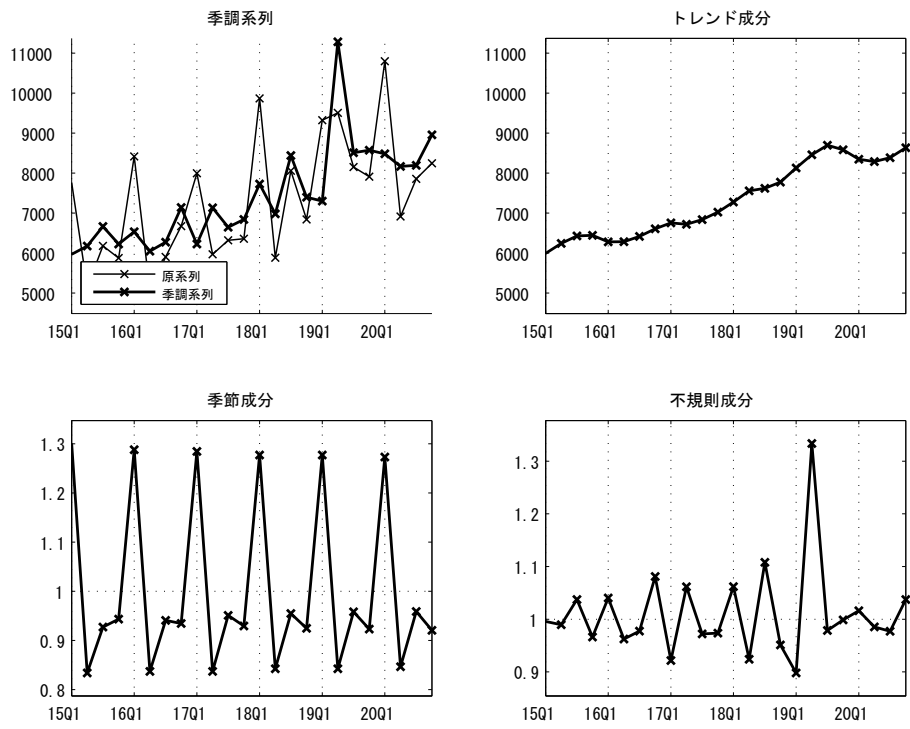


図 1-10: 設備投資・ソフト (非製造業)

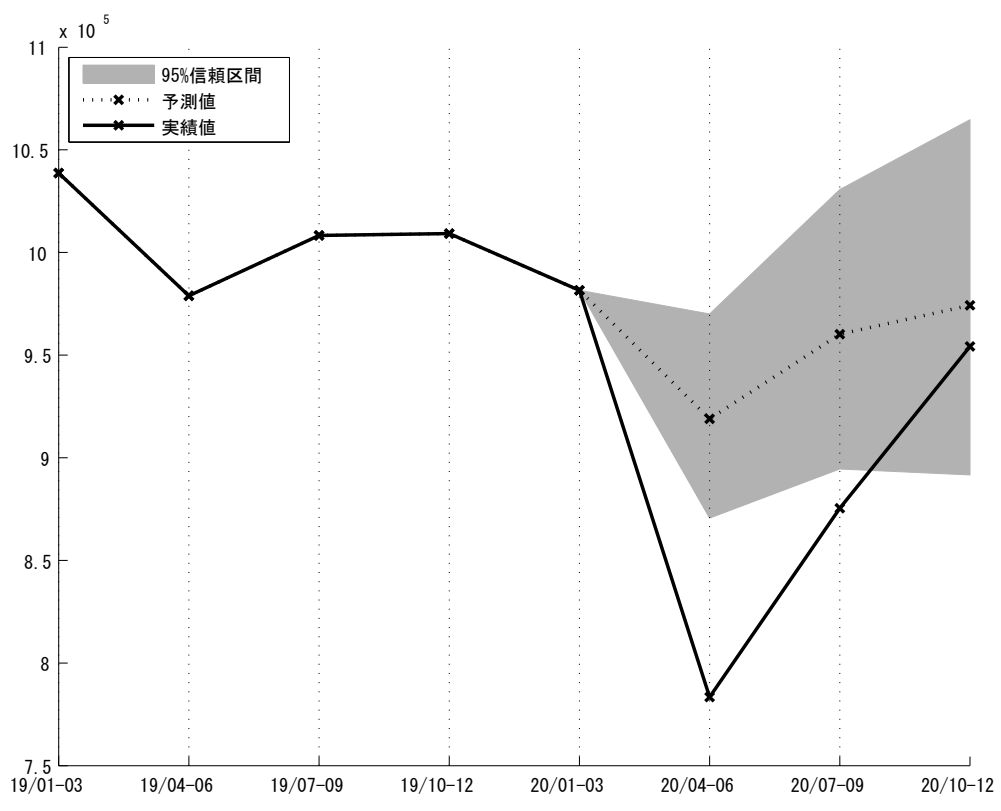


図 2-1: 売上高 (製造業)

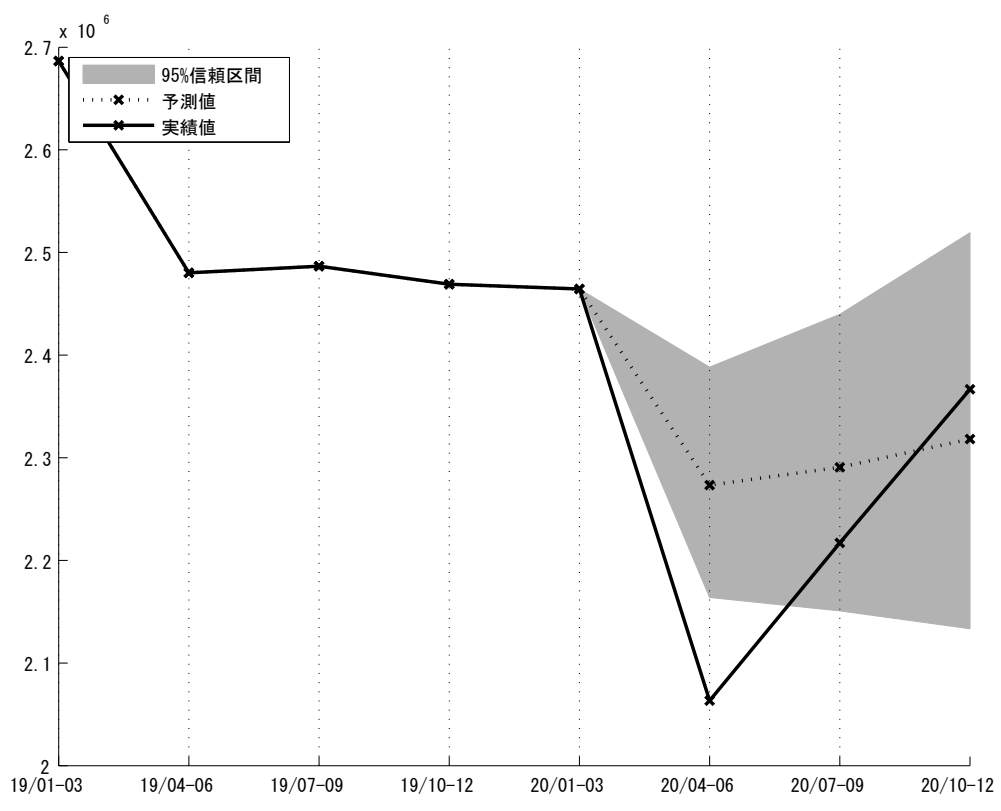


図 2-2: 売上高 (非製造業)

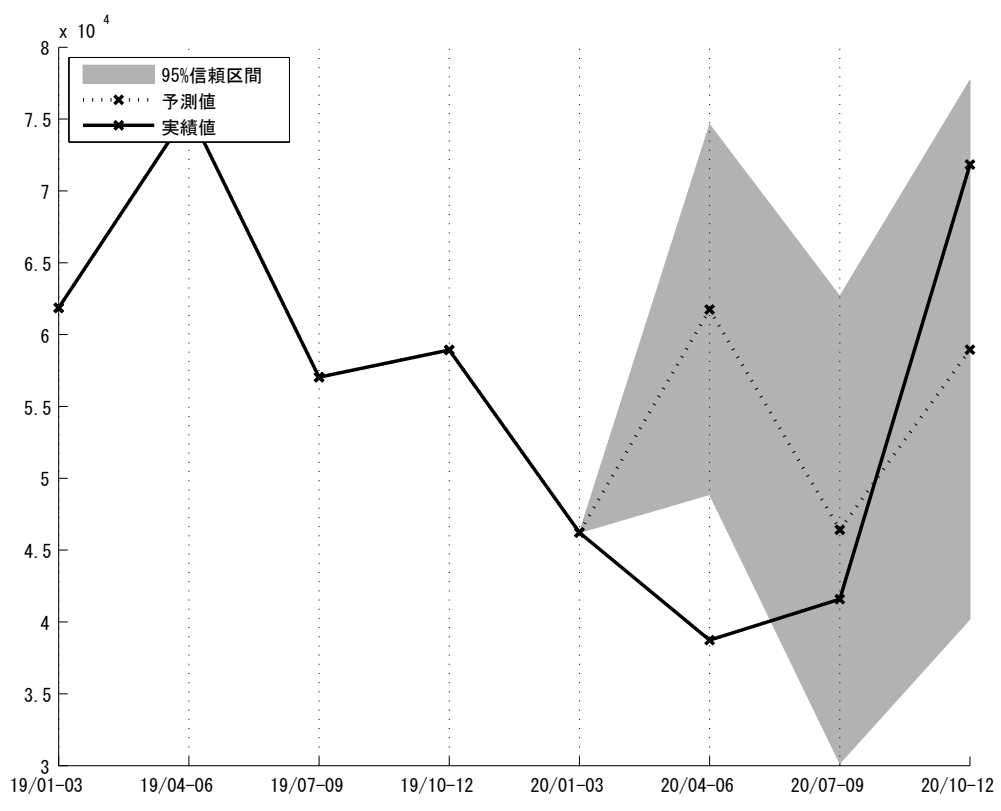


図 2-3: 経常利益 (製造業)

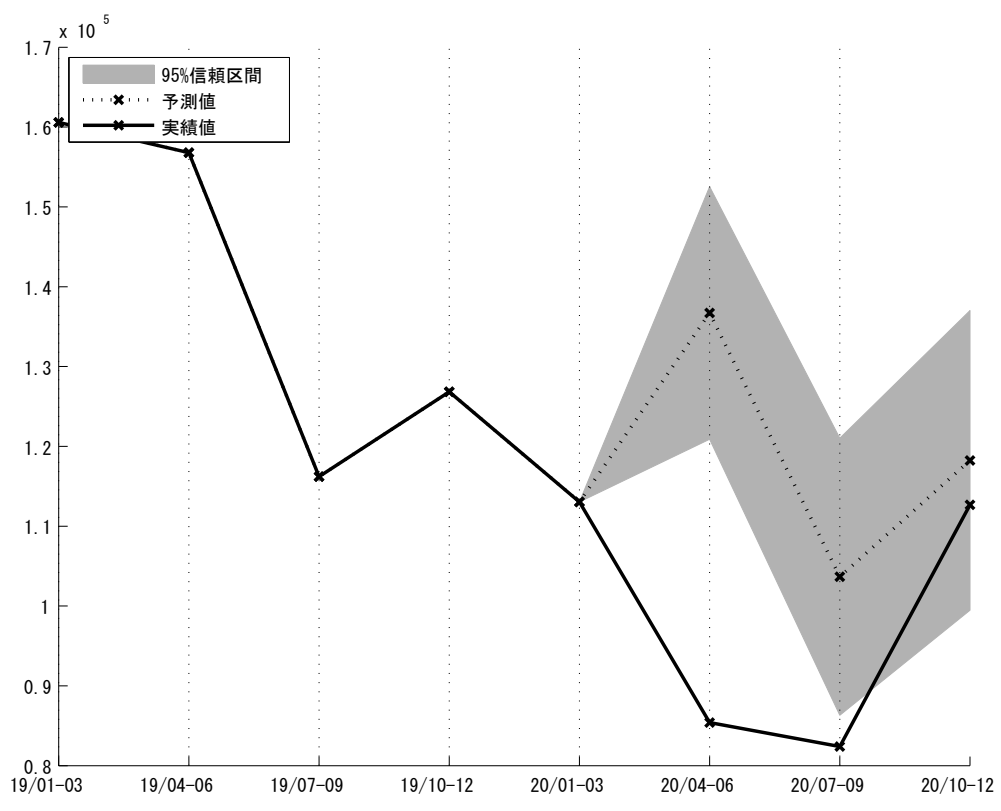


図 2-4: 経常利益 (非製造業)

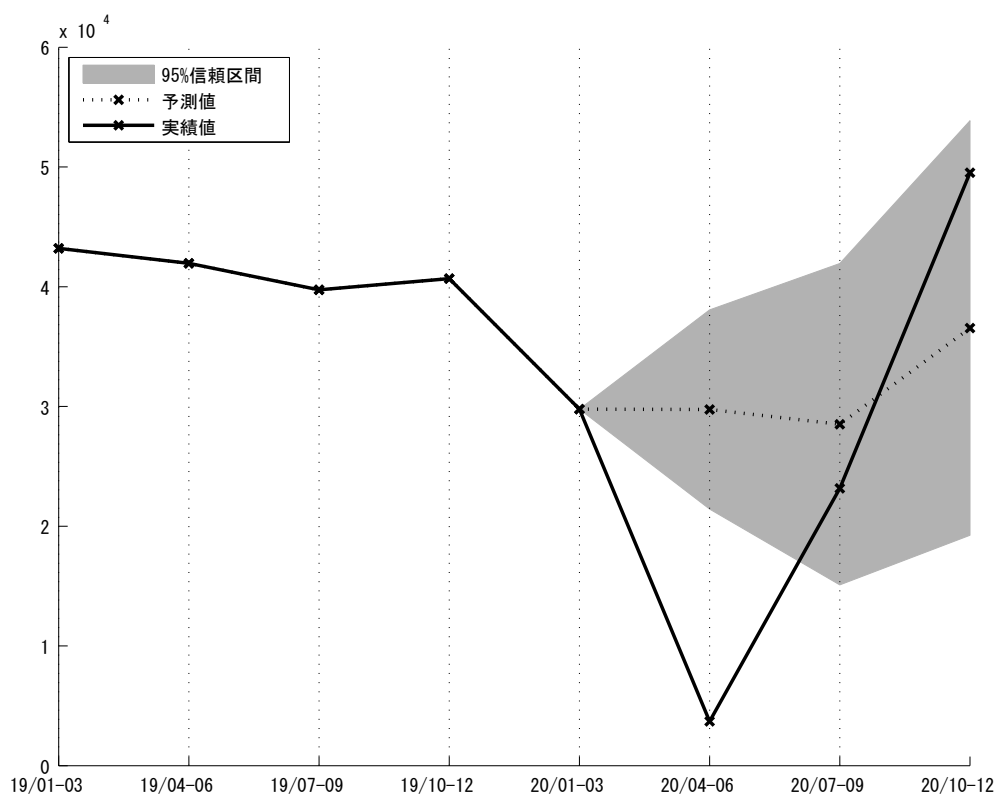


図 2-5: 営業利益 (製造業)

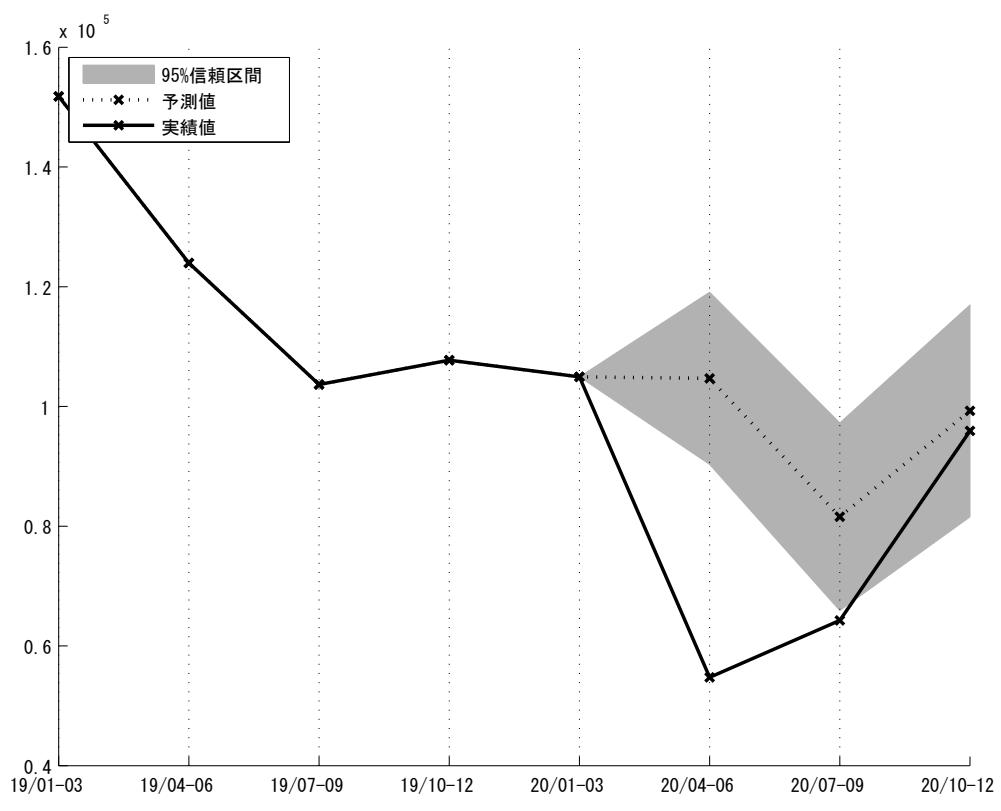


図 2-6: 営業利益 (非製造業)

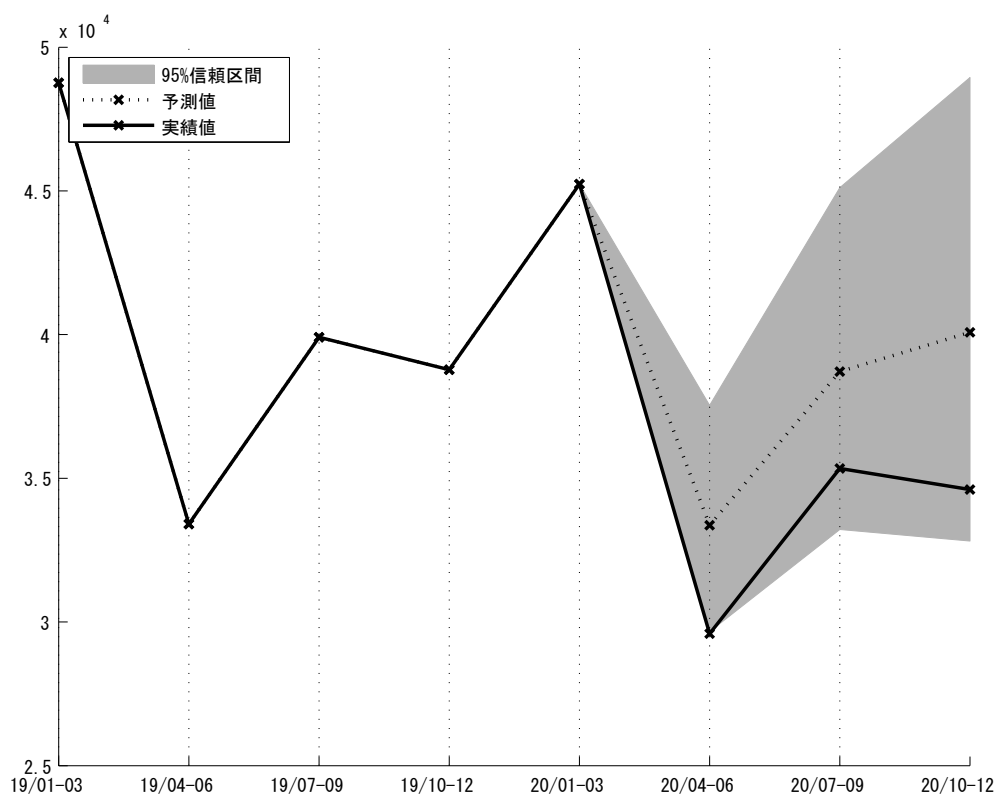


図 2-7: 設備投資・ソフト除く (製造業)

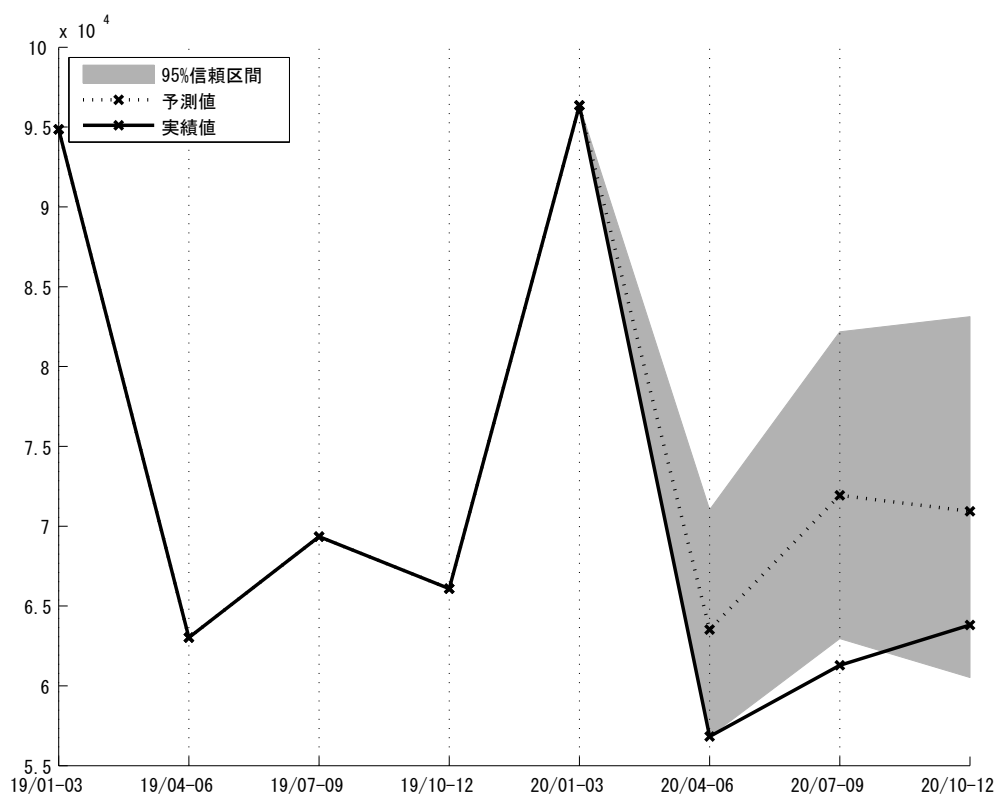


図 2-8: 設備投資・ソフト除く (非製造業)

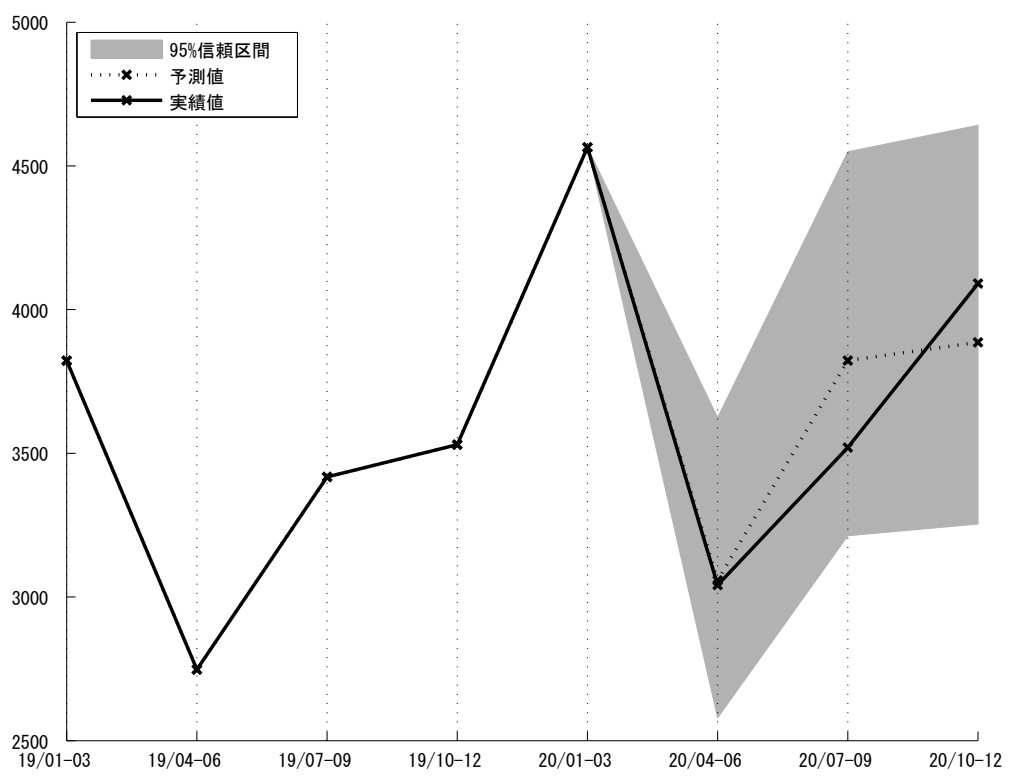


図 2-9: 設備投資・ソフト (製造業)

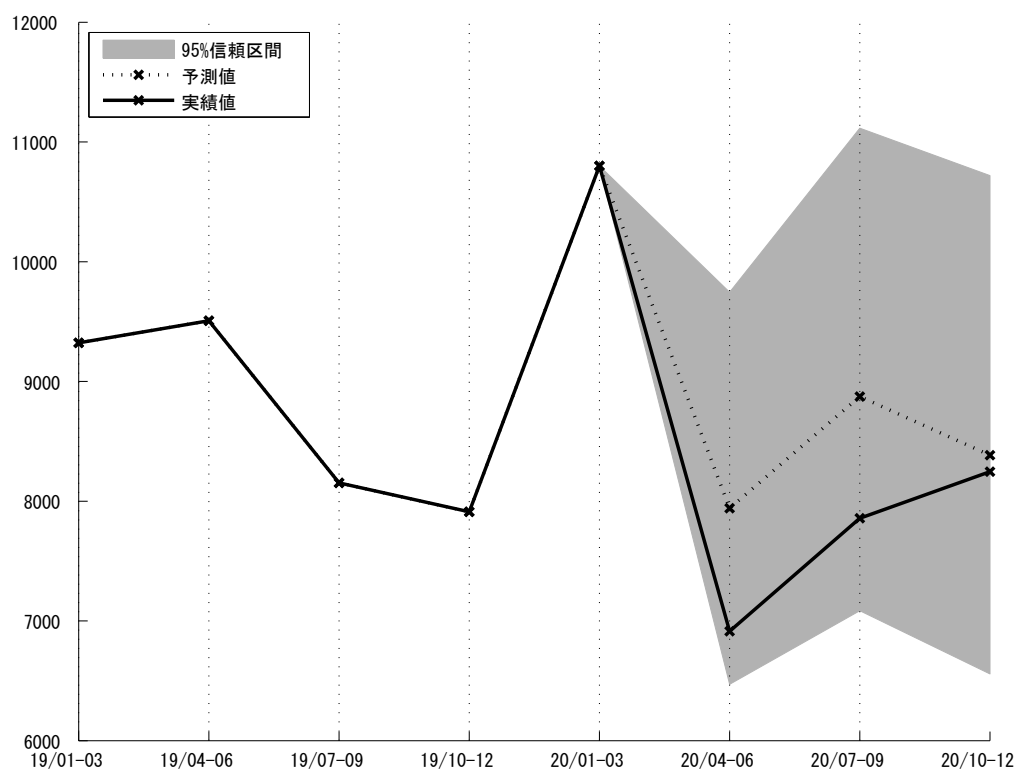


図 2-10: 設備投資・ソフト (非製造業)

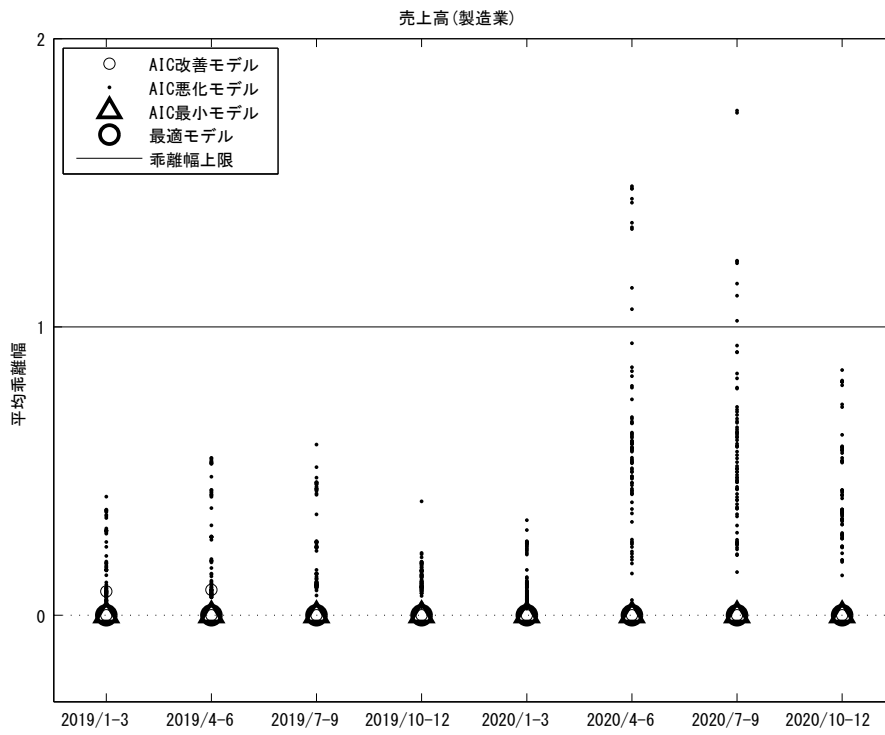


図 3-1: 売上高(製造業)

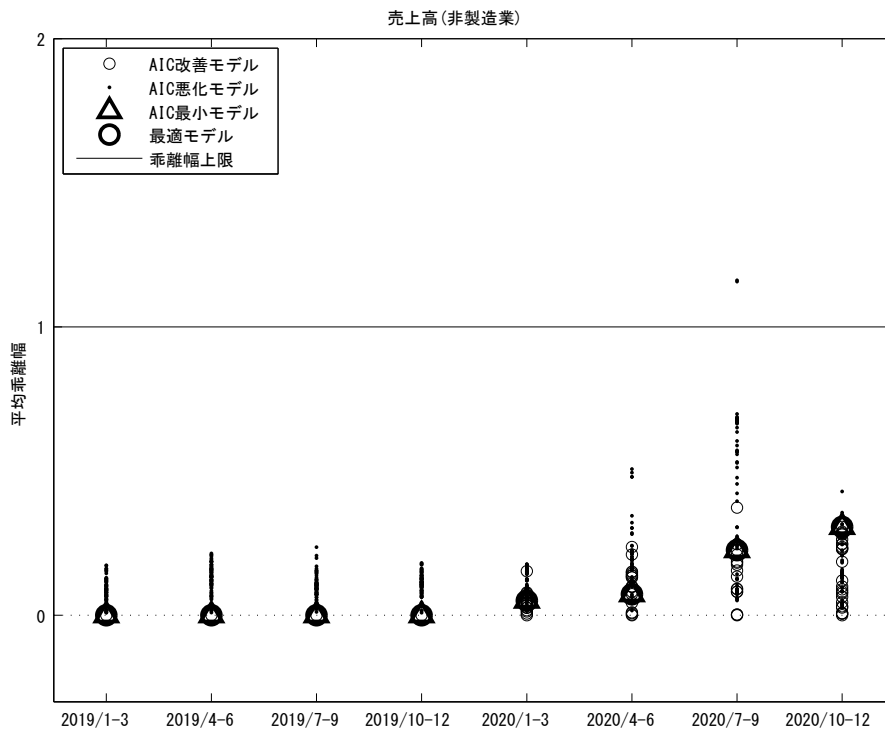


図 3-2: 売上高 (非製造業)

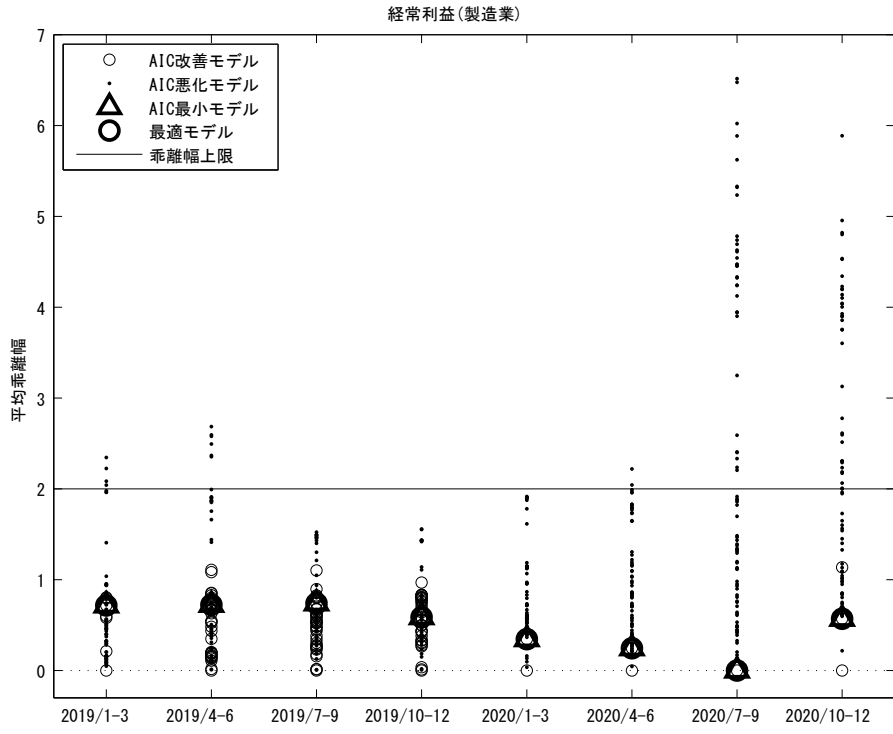


図 3-3: 経常利益 (製造業)

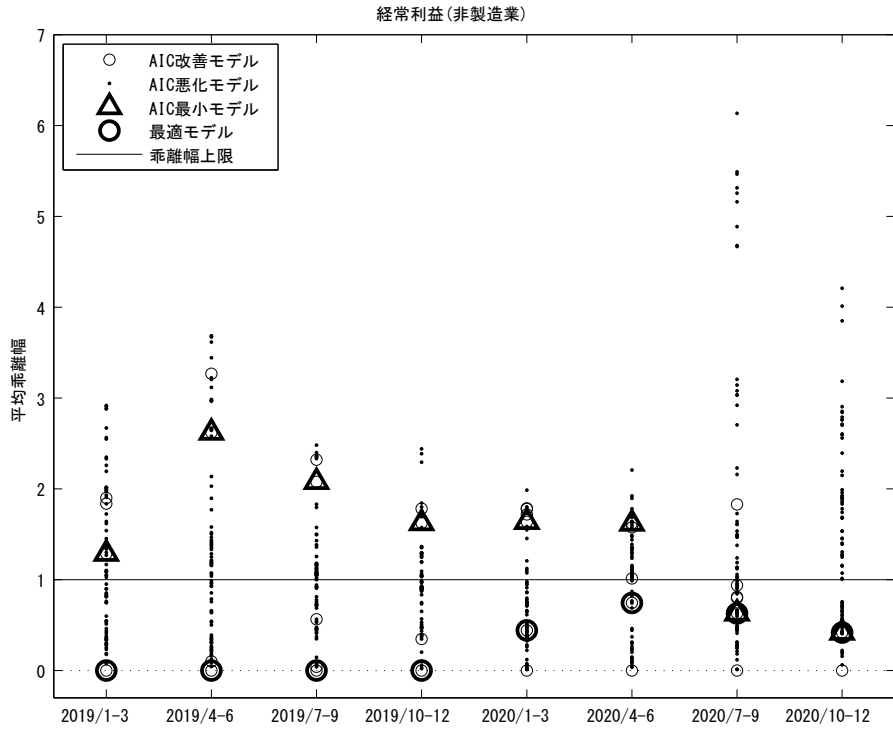


図 3-4: 経常利益 (非製造業)

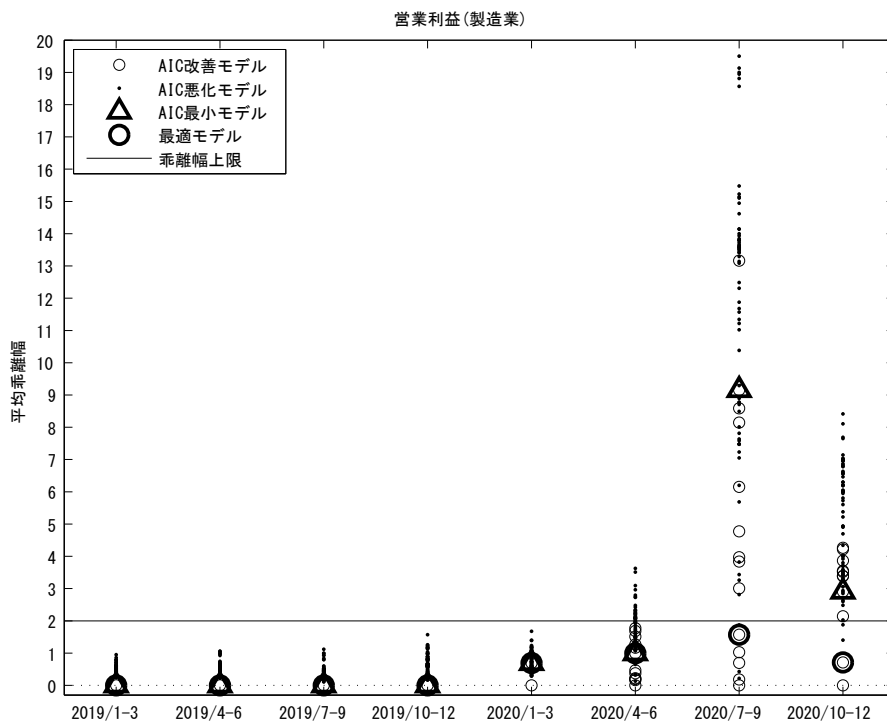


図 3-5: 営業利益 (製造業)

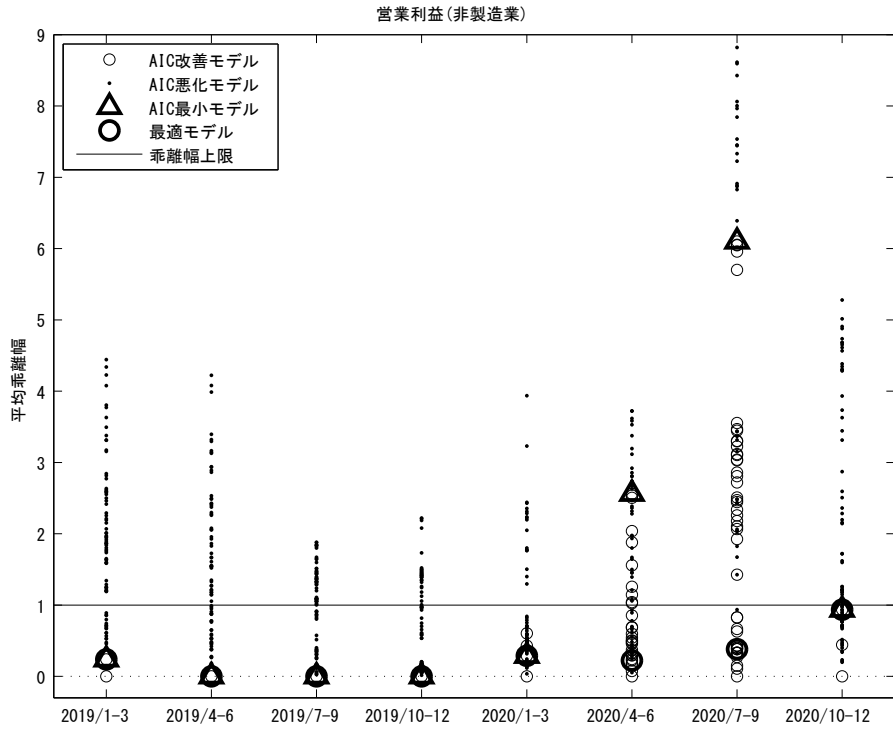


図 3-6: 営業利益 (非製造業)

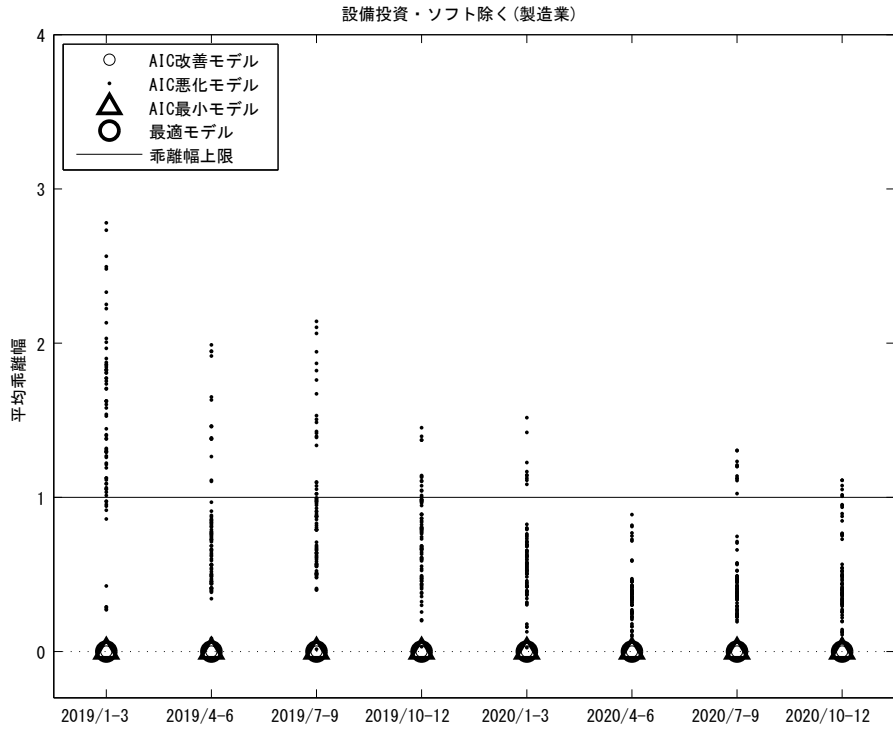


図 3-7: 設備投資・ソフト除く(製造業)

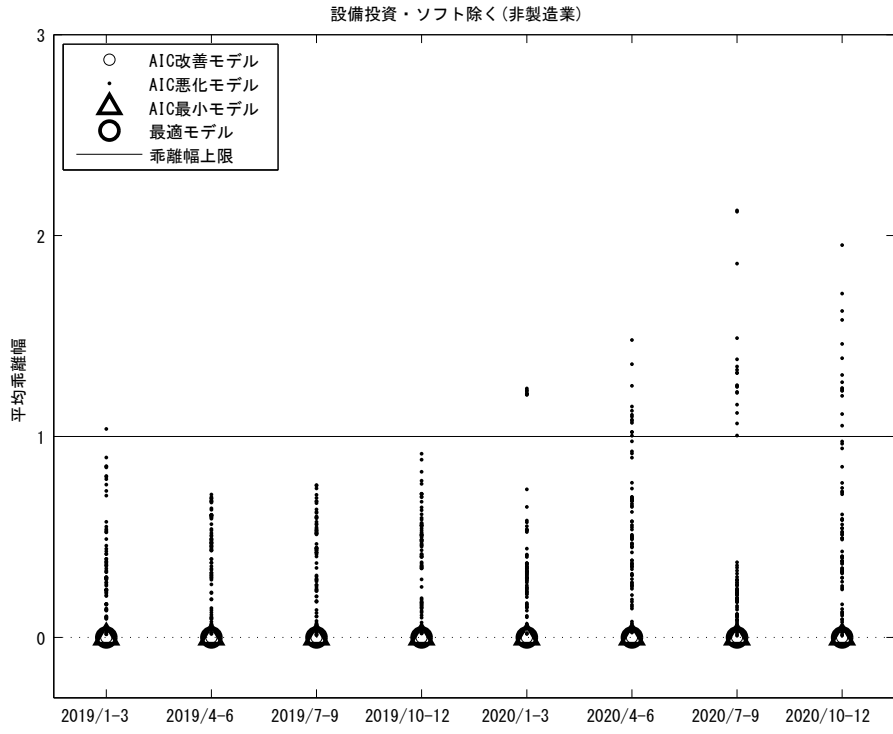


図 3-8: 設備投資・ソフト除く(非製造業)

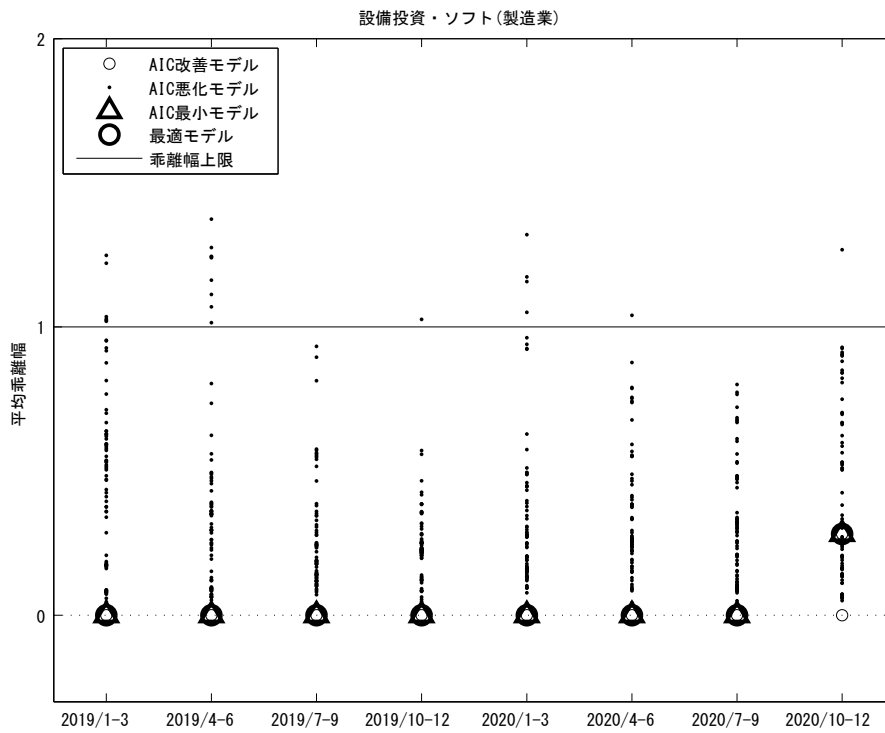


図 3-9: 設備投資・ソフト(製造業)

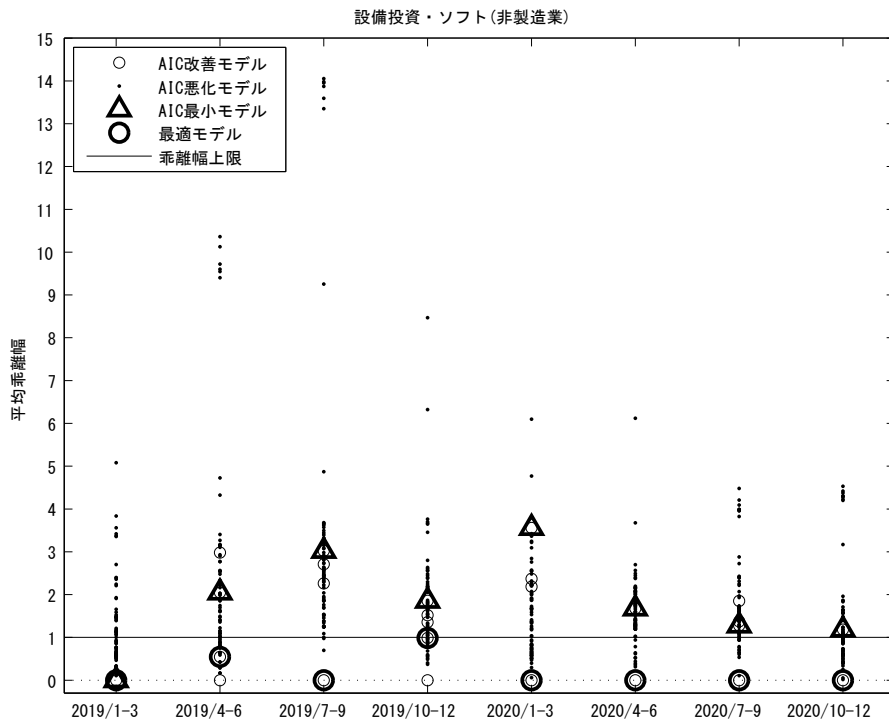


図 3-10: 設備投資・ソフト(非製造業)

A 新しいモデル選択方式について

A.1 季節調整値の安定性の指標とモデル替え

まず t 期までのデータが得られた場合の s 期 ($s \leq t$) における季節調整値を $A_{s|t}$ と表記し、この季節調整系列に基づく前期比成長率を

$$R_{s|t} = \frac{A_{s|t} - A_{s-1|t}}{A_{s-1|t}} \times 100 \quad (s \leq t) \quad (1)$$

と表すとする。一般に、 $R_{s|t}, R_{s|t+1}, R_{s|t+2}, \dots$ の変化の程度が小さいほど公表済み季節調整値の改訂幅が小さく、季節調整の安定性が高いと考えられる。

モデル替えが行われる場合には、直近の公表に使用したモデル（以下「現行モデル」と表記）と新たに選択されるモデルとの間で、さらに大きな乖離が発生する可能性がある。

現行モデルによる季節調整値および前期比成長率をそれぞれ $A_{s|t}^c$ および $R_{s|t}^c$ とする。さらに、 K 個のモデル候補があるときに、その中の一つを用いた季節調整値および前期比成長率をそれぞれ $A_{s|t}^{(k)}$ および $R_{s|t}^{(k)}$ ($k = 1, \dots, K$) とする。

いま t を直近として、 $\{R_{t|t}^c, R_{t-1|t}^c, R_{t-2|t}^c, \dots\}$ が直近の前期比成長率として公表された後に、 K 個の候補の中から $t+1$ 期以降の公表に用いるモデルの選択を行う状況を考える。ただし、選択の時点では $t+1$ 期における原数値はまだ利用可能ではないとする。

以上の設定の下で、モデル替えによる現行モデルからの改訂の程度を評価するために、指標 $SR_m^{(k)}$ (Standard Revision) を

$$SR_m^{(k)} = \frac{1}{m} \sum_{j=0}^{m-1} \left| R_{t-j|t}^c - R_{t-j|t}^{(k)} \right| \quad (1 \leq k \leq K) \quad (2)$$

と定義する¹。 $SR_m^{(k)}$ は、モデル替えによって過去 m 期間の前期比増加率が1期あたり平均何パーセントポイント改訂されるかを表している。 $SR_m^{(k)}$ を最小にするモデルは明らかに現行モデルであり、 $SR_m^{(k)}$ が小さいモデルほど過去の公表値からの改訂幅が小さいことを表す。

¹これは2つのデータ系列の間のある種の数学的距離を定義したものであるが、一般には様々な定義の仕方がある。例えば

$$SR_m^{(k)} = \left\{ \frac{1}{m} \sum_{j=0}^{m-1} \left| R_{t-j|t}^c - R_{t-j|t}^{(k)} \right|^w \right\}^{\frac{1}{w}} \quad (1 \leq w)$$

のような形式も考えられるが、ここでは直感的に理解しやすい $w = 1$ のケースを採用した。

A.2 安定性を考慮したモデル選択

前節で定義した指標を利用し、与えられた境界値 a ($0 \leq a$) に対して

$$\hat{k} = \arg \min_k AIC^{(k)} \text{ subject to } SR_m^{(k)} \leq a$$

として、モデル \hat{k} を最適モデルとする。ここで $AIC^{(k)}$ はモデル k の AIC を表すものとする。これを手順の形で書き下すと、

- (1) K 個の候補モデルの全てについて、 AIC と SR を計算する。
- (2) 与えられた境界値 a ($0 \leq a$) より SR が小さいモデルのみを候補として限定する。
- (3) 限定された候補の中から AIC を最小化するモデルを選択する。

となる。

この方法では、 $a = 0$ の場合には必ず現行モデルが選択され、 $a = \infty$ の場合には現在のモデル選択法と同一の結果が得られる。従って、 SR を利用してモデル候補を限定する方式は、従来のモデル選択法を特殊ケースとして包含する、より一般的な方法になっている²。

² t 期において、現行モデルによる成長率 $\{R_{t|t}^c, R_{t-1|t}^c, R_{t-2|t}^c, \dots\}$ が公表値として発表された後、 $t+1$ 期における原数値が内部的に利用できるようになった時点で、 $t+1$ 期以降で採用するモデルを選択するケースも考えられる。この場合は改定幅の指標として

$$\widetilde{SR}_m^{(k)} = \frac{1}{m} \sum_{j=0}^{m-1} \left| R_{t-j|t}^c - R_{t-j|t+1}^{(k)} \right|$$

を用いることができる。このような”連鎖方式”を利用すれば過去の公表値の改定幅をより確実にコントロールすることができる。ただし、現行の選択方式を含んだ方式にはならない。