

## 第2章 全国データのオプション設定

この章では、全国データのオプションを設定し、その結果を検証する。オプションについては、①JpDays と呼ばれる説明変数を導入すること、②RegARIMA モデルの回帰期間の長さを 10 年とすること、が結論付けられる。また、結果については、稼働日要因の除去と数値の安定性の両面で新オプションが良好なパフォーマンスを示すことが明らかになる。

### 1 作業の概要

#### (目標のブレイクダウン)

この章では、職業安定業務統計の全国データについて、オプションの設定を行う。X-12-ARIMA では、計算の細かい内容を「スペックファイル」と呼ばれるファイルで指定する。オプションとは、このスペックファイルで指示する内容のことである。

本研究の主要目標は稼働日要因による変動の除去であるが、前章の RegARIMA モデルの言葉を使えば、目標を次のようにブレイクダウンできる。すなわちスペックファイルで指示する説明変数、ARIMA パラメータ  $p, d, q, P, D, Q$ 、回帰期間、及び、その他の計算条件を適切に設定すること、である。

図表 2-1-1 ブレイクダウンされた作業目標

- |  |
|--|
| <ul style="list-style-type: none"><li>① 説明変数 <math>x_{it}</math> の適切な設定</li><li>② ARIMA パラメータ <math>p, d, q, P, D, Q</math> の適切な設定</li><li>③ 回帰期間の適切な設定</li><li>④ その他、計算条件の適切な設定</li></ul> |
|--|

#### (遡及修正の抑制にも配慮)

なお、これらのオプションの設定にあたって、稼働日要因の除去はもちろんであるが、最新データの追加にともなう遡及修正をできるだけ抑制することも、副次目標とする。

社会経済の構造変化により季節変動も年々変化しており、また、新しいデータが入手できることにより的確にその構造変化を捉えることができる。したがって、新しいデータを用いて季節調整を再度実施したときに遡及修正が発生すること自体は、悪いことでない。①しかし、季節調整の方法に問題があると、必ずしも構造変化を反映したものではない遡及修正が発生すること<sup>(注24)</sup>、及び、②大幅な遡及修正は統計利用者に不便を感じさせること、を考慮

(注24) 遡及修正の小ささ(季節調整値の安定性)自体が、モデルの適合性の良い判断指標になることが知られている (Feldpausch, Hood, Wills[17])

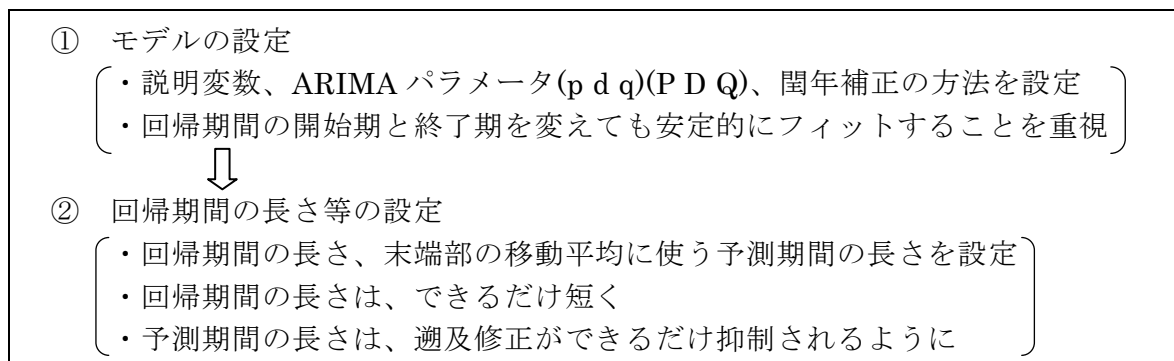
して、遡及修正の抑制を副次目標とするものである。

### (2つのパートに分けて作業)

最良のオプションを求めるためには、可能な選択肢の組み合わせをすべて検討するのが理想である。しかし、膨大な作業をとまなうので、これは事実上不可能である。本研究では、省力化のため、①モデルの設定、及び②回帰期間の長さ等の設定、の2つのパートに分けて作業を行うことにする。モデルの設定に当たっては、とくに運用上の安定性に配慮して、新規データの追加により回帰期間の開始期と終了期が変化しても安定的にフィットするモデルを選ぶように務めた。一方、回帰期間の長さについては、構造変化を迅速に反映できるように、できるだけ短く設定するようにした。

なお、X-12-ARIMA にはいくつかのバージョンがあるが、本研究で使ったのは、「X-12-ARIMA Version 0.2.10」(2002年7月26日リリース、2005年11月時点の最終バージョン)である。

図表 2-1-2 作業手順



## 2 モデルの設定

### (1) 基本方針

(説明変数 5 種類、ARIMA パラメータ 5 種類、閏年補正方法 3 種類、回帰期間 16 ケース)

この項では、説明変数、ARIMA パラメータ(p d q)(P D Q)、及び、閏年補正の方法を設定する。原則として 5 種類の説明変数、5 種類の ARIMA パラメータ、及び 3 種類の閏年補正方法をクロスさせた 75 種類の候補を立て、その中から比較的良いものを選出した。

選出は 2 段階に分けて行った。第 1 段階では、原則 75 種類の候補それぞれについて、16 ケースの回帰期間で計測を行い、その中で 8 割以上のケース (13 ケース以上) で最低の「適合基準」を満たしたものを第 1 次合格とする。これは、安定的にフィットするモデルのみを選ぶことを狙ったものである。ここで、次の 3 条件をすべて満たすことをもって適合基準と設定した。これは、X-12-ARIMA で標準で設定されている適合基準である。

図表 2-2-1 適合基準（第 1 次合格の要件）

- ① 直近 3 年間の推計誤差が平均 15%以下であること
- ② Ljung-Box の Q 統計量<sup>(注25)</sup>による検定で、攪乱項の自己相関が 5%水準で有意でないこと
- ③ 過剰階差でないこと（パラメータ d が大きすぎないこと）<sup>(注26)</sup>

**(AICC、外挿予測誤差、スペクトル、系列間の共通性)**

第 1 次合格した候補の中からは、基本的に AICC<sup>(注27)</sup> でみて優良なものを選出することとするが、外挿予測誤差<sup>(注28)</sup> やパワースペクトルも参照する。

また、新規求職申込件数、新規求人数、就職件数の 3 系列については、稼働日調整のための説明変数の候補（後出図表 2-2-2）から、できるだけ共通のひとつを選ぶことにする。これらの系列から新規求人倍率（新規求人数／新規求職申込件数）や就職率（就職件数／新規求職申込件数）といった加工系列が作成されるので、稼働日要因の除去方法が共通だと解釈や分析が楽になると考えるためである。

同様に、有効求職者数と有効求人数の 2 系列についても、これらから有効求人倍率（有効求人数／有効求職者数）が作成されることを考慮して、稼働日調整には、できるだけ同じ説明変数を採用するようにする。

**(大量の候補から選択するという考え方)**

なお、時系列分析の教科書である Box and Jenkins[2]には、モデル設定の標準的な方法として、次の手順が示されている。

- ① Identify 自己相関関数（ACF）や偏自己相関関数（PACF）からモデルの見当をつける。

(注25)  $m$  を正整数として、Ljung-Box の Q 統計量  $Q_m$  は、次の式で表される（Ljung and Box[3]）。

$$Q_m = n(n+2) \sum_{i=1}^m (n-i)^{-1} \bar{r}_i^2 \quad \bar{r}_i = \sum_{t=i+1}^n \bar{a}_t \bar{a}_{t-i} / \sum_{t=1}^n \bar{a}_t^2$$

$\bar{a}_t$  は、攪乱項の推計値。 $a_t$  に自己相関がないという帰無仮説のもとで、 $Q_m$  は、自由度  $m - (p + q + P + Q + k)$  の  $\chi^2$  分布で近似される。

本研究では、 $m = 24$  と設定している。

(注26) 「過剰階差でないこと」のチェックとは、図表 1-2-2 の記号で、 $\theta_1 + \dots + \theta_q \leq 0.9$  であること。

(注27) AICC(F-corrected AIC)は、AIC（赤池情報量基準）の変形で、自由度が小さいときの精度を改良したもの。観測数を  $N$ 、推計パラメータの個数を  $n_p$ 、推計式の対数尤度を  $L_N$  とするとき、AICC は、次のように定義される。

$$AICC = -2L_N + 2n_p (N / (N - (n_p + 1) / N))$$

AIC 及びそれから派生した AICC 等の指標は、モデルのあてはまり状況を手軽に推計できることから、広く普及している。AICC は、センサス局で重視されており、X-12-ARIMA に標準装備されている。

(注28) 外挿予測誤差（OSFE: Out-of-Sample Forecast Error）は、推計に使った最新データから 1 か月先ないし 12 か月先の予測値と実績値との差から測定される誤差（参考資料 5）。外挿予測誤差は、モデルのあてはまり状況を直接計測する指標だが、大量のデータが必要で計算に時間がかかる。本研究では、AICC の適用が難しいケースで外挿予測誤差を使った。

- ② Estimate 見当をつけたモデルについてパラメータを推計する。
- ③ Check 推計されたパラメータを使って、モデルの妥当性を検証する。

本研究の方法は、この①を省略する代わりに大量の候補を立てたものである。これは、モデルが複雑になるにしたがって①の作業はかなりの熟練技能が必要となり、一方、パソコンの性能が向上した現在、大量推計の方がかえって省力化につながると考えたからである。

## (2) 説明変数の候補

曜日構成及び祝日の調整のための説明変数として、tdnolpyear、td1nolpyear、JpDays、JpDays1、JpDays3 の 5 種類の候補を立てる（図表 2-2-2）。

図表 2-2-2 曜日構成及び祝日の調整のための説明変数の候補

変数名	意味
tdnolpyear	D1-D7, D2-D7, ..., D6-D7 という 6 個の変数。D1, D2, ..., D7 は、それぞれ、月曜、火曜、…、日曜の日数。
td1nolpyear	$(D1+D2+D3+D4+D5) - (5/2)(D6+D7)$
JpDays	日本の祝日等を考慮した 8 個の曜日変数の総称
JpDays1	日本の祝日等を考慮した稼働日数
JpDays3	JpDays1、1 か月前の JpDays1、2 か月前の JpDays1 の、3 変数の総称

(注) 1) tdnolpyear(trading day, no leap year)及び td1nolpyear は、X-12-ARIMA にあらかじめ準備されている変数。なお、X-12-ARIMA には、このほかイースター休日、サンクスギビング、クリスマスなどを調整する変数が準備されているが、日本の建国記念日、勤労感謝の日、天皇誕生日などを調整する変数は準備されていない。

2) JpDays は、具体的には次のとおり。

$$\begin{aligned} DJpMon &= (JpMon - Sun) - A1 \\ DJpTue &= (JpTue - Sun) - A2 \\ DJpWed &= (JpWed - Sun) - A3 \\ DJpThu &= (JpThu - Sun) - A4 \\ DJpFri &= (JpFri - Sun) - A5 \\ DJpSat &= (JpSat - Sun) - A6 \\ DJpHS &= (JpHS - Sun) - A7 \\ DJpHnotSS &= (JpHnotSS - Sun) - A8 \end{aligned}$$

JpMon から JpSat までは、月曜日から土曜日までのそれぞれの月間日数のうち、祝日等を除いた日数。Sun は、日曜日の月間日数。JpHS は、土曜日と重なる祝日等の日数。JpHnotSS は、土曜日と日曜日を除く祝日等の日数。土曜日については、週休 2 日制の普及を反映しやすくするため、祝日等と重なるものと重ならないものに分けた（ハローワークでは、週休 1 日半制→4 週 6 休制→完全週休 2 日制と推移してきた）。

A1 から A8 までは、季節調整値の全体の水準を調整するための定数であり、(JpMon - Sun)から (JpHnotSS - Sun)までのそれぞれの 2003 年 1 月～2030 年 12 月にわたる平均。2003 年は、祝日法（昭和 23(1948)年 7 月 20 日 法律第 178 号）の最近の改正の年。2030 年はその 28 年後。28 年とは、曜日構成が 1 巡する周期。A1、…、A8 は、RegARIMA モデルで階差をとっているとき (d+D>0 のとき)、係数パラメータの推計に影響しない。しかし、A1、…、A8 をこのように設定することにより、最近の季節調整値の年平均が原数値の年平均に概ね一致ようになる。ただし、長期の過去については、季節調整値と原数値の年平均に若干の乖離が生ずる。

3) JpDays1 は、具体的には次のとおり。

$$JpDays1 = (JpMon + JpTue + JpWed + JpThu + JpFri) - 2.019486 (Sat + Sun + JpH)$$

係数 2.019486 は、2003 年 1 月～2030 年 12 月にわたる JpDays1 の平均が 0 になるように設定

tdnolpyear は、X-12-ARIMA にあらかじめ準備されている変数であって、月曜日から土曜日までの各日数を基に設定された 6 個の変数の総称である。

td1nolpyear は、ウィークデイとウィークエンドの相対日数を基に設定された 1 個の変数である。もし月曜日から金曜日までの効果に大きな差がなく、また、土曜日と日曜日との効果にも大きな差がなければ、tdnolpyear より効率的と考えられる。これも X-12-ARIMA にあらかじめ準備されている変数である。

JpDays は、本研究のために新たに作った変数であって、曜日構成に加えて日本の祝日等（図表 2-2-3）も組み込んだ 8 個の変数の総称である（注29）。

JpDays1 は、td1nolpyear と同様の考え方で、JpDays の 1 変数版として設定した。

JpDays3 は、JpDays1 に加えてその 1 か月前、2 か月前のラグを入れた 3 変数の総称である。有効求職者数と有効求人数がおおむね 2 か月前からのストックとしての性格も持っていることから、この変数も候補としたものである。この変数は、有効求職者数と有効求人数の 2 系列についてのみ使うことにする。

図表 2-2-3 祝日等の一覧

祝日法による国民の祝日 年末年始（12 月 29 日～1 月 3 日） 国民の祝日が日曜日と重なったときの振替休日（1973 年 4 月 12 日以降） 5 月 4 日（1986 年以降） 昭和天皇の大喪の礼、今上天皇の即位の礼、明仁親王（今上天皇）の結婚の儀、徳仁親王の結婚の儀
--

### (3) ARIMA パラメータの候補

ARIMA パラメータの候補としては、図表 2-2-4 の「原則候補」を原則として用いる。これらは、X-12-ARIMA にデフォルトで準備されているものである。

ただし、原則候補に第 1 次合格するものがなかった場合は、「追加候補」を用いる。この追加候補には、いわゆるミックスモデル（ $p$  と  $q$  のどちらも 1 以上のモデル）が含まれていない。これは、キャンセルーション（注30）と呼ばれる推計上の不都合が生じるのを防ぐ意味

（注29） 祝日等を組み込む方法としては、JpDays のような方法とは別に、「tdnolpyear+稼働日数」という方法もある。本研究で JpDays の方を選んだ理由は、①JpDays では月曜日～金曜日のどの曜日と祝日が重なるかを識別できること、②JpDays の方法でも「tdnolpyear+稼働日数」でも説明変数の個数は基本的に変わらないこと（「tdnolpyear+稼働日数」が合計 7 変数なのに対して JpDays が合計 8 変数となっているが、これは JpDays が祝日と重なる土曜日の特掲しているため）、による。職業安定業務統計については、ウィークデイ（月曜日～金曜日）かどうか、祝日かどうか、といった要因だけでなく、祝日以外のウィークデイでどの曜日が多いか、といったことにも強く影響を受ける（補論の「1 曜日別の求人求職動向」参照）ので、JpDays のような変数の設定が有効と思われる。

（注30） 第 3 章の 3 参照

があり、また、 $p$  又は  $q$  が大きいときはミックスモデルを使う意味が薄れる<sup>(注31)</sup> ことも考慮したものである。

図表 2-2-4 検討対象とする ARIMA パラメータ

(原則候補)

モデル名	ARIMA パラメータ(p d q)(P D Q)
M1	(0 1 1)(0 1 1)
M2	(0 1 2)(0 1 1)
M3	(2 1 0)(0 1 1)
M4	(0 2 2)(0 1 1)
M5	(2 1 2)(0 1 1)

(追加候補)

モデル名	ARIMA パラメータ(p d q)(P D Q)
ma3	(0 1 3)(0 1 1)
ma4	(0 1 4)(0 1 1)
ma5	(0 1 5)(0 1 1)
ma6	(0 1 6)(0 1 1)
ar3	(3 1 0)(0 1 1)
ar4	(4 1 0)(0 1 1)
ar5	(5 1 0)(0 1 1)
ar6	(6 1 0)(0 1 1)

(注) 原則候補は、X-12-ARIMA で基本的に使われている候補と同じ。追加候補は、原則候補で 1 次合格するものがなかったときに検討対象とする。

#### (4) 閏年補正方法の候補

閏年補正については、RegARIMA モデルに説明変数を追加する方法 (lpyear)、2 月のデータに月の日数に反比例する係数を機械的に乗じる方法 (rescale)、閏年補正を行わないもの (nolpyear) の 3 種類を候補とする。ただし、ストックとしての性格も持っている有効求職者数と有効求人数については、rescale を使わない。

(注31) 一般に、ミックスモデルは、 $p$  が大きくて  $q=0$  のモデル (AR モデル) と  $q$  が大きくて  $p=0$  のモデル (MA モデル) のいずれでも近似できる (Box and Jenkins[2])。

図表 2-2-5 検討対象とする閏年補正方法

候補名	補正方法
lpyear	RegARIMA モデルに次の説明変数を追加 閏年の 2 月:0.75、閏年以外の 2 月:-0.25、2 月以外:0
rescale	事前に原数値を次のように補正 閏年の 2 月: $Y'_t = (28.25/29)Y_t$ 、閏年以外の 2 月: $Y'_t = (28.25/28)Y_t$ 、 その他の月: $Y'_t = Y_t$
nolpyear	閏年補正を行わない

(注) rescale は、新規求職申込件数、新規求人数、就職件数についてのみ候補とする。

### (5) 回帰期間の種類

下の 16 種類の期間でモデルの適合状況をテストした<sup>(注32)</sup>。この 16 種類の 8 割以上 (13 ケース以上) で適合基準を満たしたモデルを第 1 次合格とする。

図表 2-2-6 適合基準のテスト期間

ケース番号	開始年月	終了年月	回帰期間の長さ (年)
1	1980.1	1999.12	20
2	1981.1	2000.12	20
3	1982.1	2001.12	20
4	1983.1	2002.12	20
5	1984.1	2003.12	20
6	1985.1	2004.12	20
7	1986.1	2004.12	19
8	1987.1	2004.12	18
9	1988.1	2004.12	17
10	1989.1	2004.12	16
11	1990.1	2004.12	15
12	1991.1	2004.12	14
13	1992.1	2004.12	13
14	1993.1	2004.12	12
15	1994.1	2004.12	11
16	1995.1	2004.12	10

(注32) 回帰期間は最終的に 10 年に設定されるが、10 年の期間で適合基準がテストされたのはケース番号 16 の 1 ケースに過ぎない。ただ、これは大きな問題を生じないと考えられる。というのは、適合基準を満たさなかったケースのほとんどが Ljung-Box の Q 統計量による検定で不合格となったものであるが、この検定は、回帰期間が長いほど厳しくなる傾向があるからである。第 1 次合格したモデルは、1 例を除いて、すべてケース番号 16 で適合基準を満たしていた。適合しなかった 1 例は、最終的に採用されなかった。

## (6) 適合基準の検討結果

適合基準の検討の結果、安定的にフィットするとして第1次合格したモデルを図表 2-2-7 に示す。いくつか補足する。

- ① 有効求職者数を除き、M4 の ARIMA モデル(0 2 2)(0 1 1)は全滅した。これは、すべて「過剰階差」のテストで不合格になったものである。M4 は、候補とした ARIMA モデルの中で唯一  $d=2$  のものである（残りはすべて  $d=1$ ）。
- ② M4 以外の ARIMA モデルで不合格となったものは、ほとんどが Ljung-Box の Q 統計量による検定で不合格となったものである。
- ③ 新規求人数と有効求人数については、一般にあてはまりが悪く、ma6、ar5 といった次数の高い ARIMA モデルでようやくフィットした<sup>(注33)</sup>。

図表 2-2-7 第1次合格したモデル

(新規求職申込件数)

TD_M1_lpyear	TD_M1_rescale	TD_M2_lpyear	TD_M2_rescale	TD_M2_nolpyear
TD_M3_lpyear	TD_M3_rescale	TD_M3_nolpyear	TD_M5_lpyear	TD_M5_rescale
TD_M5_nolpyear				
TD1_M3_lpyear	TD1_M3_rescale	TD1_M3_nolpyear	TD1_M5_lpyear	TD1_M5_rescale
TD1_M5_nolpyear				
JD_M3_lpyear	JD_M3_rescale	JD_M3_nolpyear	JD_M5_lpyear	JD_M5_rescale
JD1_M5_lpyear	JD1_M5_rescale			

(新規求人数)

TD_ma6_lpyear	TD_ma6_rescale	TD_ma6_nolpyear
TD1_ma6_rescale		
JD_ma6_lpyear	JD_ma6_rescale	JD_ma6_nolpyear
JD1_ma6_lpyear	JD1_ma6_rescale	JD1_ma6_nolpyear

(有効求職者数)

TD_M3_lpyear	TD_M4_lpyear	TD_M4_nolpyear	TD_M5_lpyear	TD_M5_nolpyear
TD1_M3_lpyear	TD1_M4_lpyear	TD1_M4_nolpyear	TD1_M5_lpyear	TD1_M5_nolpyear
JD_M3_lpyear	JD_M3_nolpyear	JD_M4_lpyear	JD_M4_nolpyear	JD_M5_lpyear
JD_M5_nolpyear				
JD1_M4_lpyear	JD1_M4_nolpyear	JD1_M5_lpyear	JD1_M5_nolpyear	
JD3_M3_lpyear	JD3_M3_nolpyear	JD3_M4_lpyear	JD3_M4_nolpyear	JD3_M5_lpyear
JD3_M5_nolpyear				

(注33) 新規求人数について ma1, ma2, ..., ma6 の AIC を比較すると、ma6 が最小だった。新規求人数のモデル (ma6) は、AIC 基準に照らしても問題ない。一方、有効求人数について da1, da2, ..., da6 の AIC を比較すると、 $da1 < da2 < \dots < da6$  だった。有効求人数について、本研究で選択したモデル (da5) は、一見 AIC 基準に反するように見える。しかし、ここで用いた AIC は、厳密なものでなく、攪乱項に自己相関がないことを前提に計算された推計値である。そこで、本研究では、AIC 基準よりも、その有効性の前提となる Ljung-Box の Q 統計量による自己相関の検定を優先した。なお、後出の脚注 46 も参照のこと。



(有効求人数)

TD_ar5_lpyear			
JD_ar5_lpyear	JD_ar5_nolpyear	JD_ar6_lpyear	JD_ar6_nolpyear
JD3_ar5_lpyear			

(就職件数)

TD_M1_lpyear	TD_M1_rescale	TD_M1_nolpyear	TD_M2_lpyear	TD_M2_rescale
TD_M2_nolpyear	TD_M3_lpyear	TD_M3_rescale	TD_M3_nolpyear	
TD1_M1_lpyear	TD1_M1_rescale	TD1_M1_nolpyear	TD1_M2_lpyear	TD1_M2_rescale
TD1_M2_nolpyear	TD1_M3_rescale	TD1_M5_lpyear	TD1_M5_nolpyear	
JD_M2_rescale	JD_M2_nolpyear	JD_M3_nolpyear		
JD1_M2_lpyear	JD1_M2_rescale	JD1_M2_nolpyear	JD1_M3_lpyear	JD1_M3_rescale
JD1_M3_nolpyear	JD1_M5_lpyear	JD1_M5_rescale	JD1_M5_nolpyear	

(注) 表のモデルは、(変数名)\_(ARIMAモデル名)\_(閏年補正方法)という標記で示した。なお、変数名の TD、TD1、JD、JD1、JD3 は、それぞれ tdnolpyear、td1nolpyear、JpDays、JpDays1、JpDay3 を表す。

## (7) モデルの選定

(AICC では JpDays、JpDays1、JpDays3 が優良、しかし JpDays1 と JpDays3 は稼働日要因除去不十分 ⇒ JpDays を選択)

第 1 次合格したモデルについて AICC を調べてみると、どの系列でも JpDays、JpDays1、JpDays3 が tdnolpyear や td1nolpyear より優良であることが分かる (図表 2-2-8)。

JpDays、JpDays1、JpDays3 の中では、JpDays1 や JpDays3 が優良なものもみられたが、JpDays1 や JpDays3 では曜日構成の要因が十分に除去されないことがスペクトル分析で分かった (図表 2-2-9)。そこで、すべての系列で共通に JpDays を説明変数として採用することにした。

ARIMA モデルについては、とくに有効求職者数で M4 の AICC が最も優良という結果となった。しかし、12 か月先の外挿予測誤差をみると、M4 のモデルは明らかにあてはまりが悪い (図表 2-2-10) (注34)。そこで、M4 は採用しないこととして、M4 以外から AICC やモデルの単純性 (注35) を考慮して選ぶこととした。

閏年補正については、AICC を基準にすることにやや疑問があること (注36) から、外挿予測誤差を重視することとした。新規求職申込件数と新規求人数については、12 か月先の外挿予測誤差に大きな差がなく、1 か月先の外挿予測誤差で lpyear と rescale が優良であることが

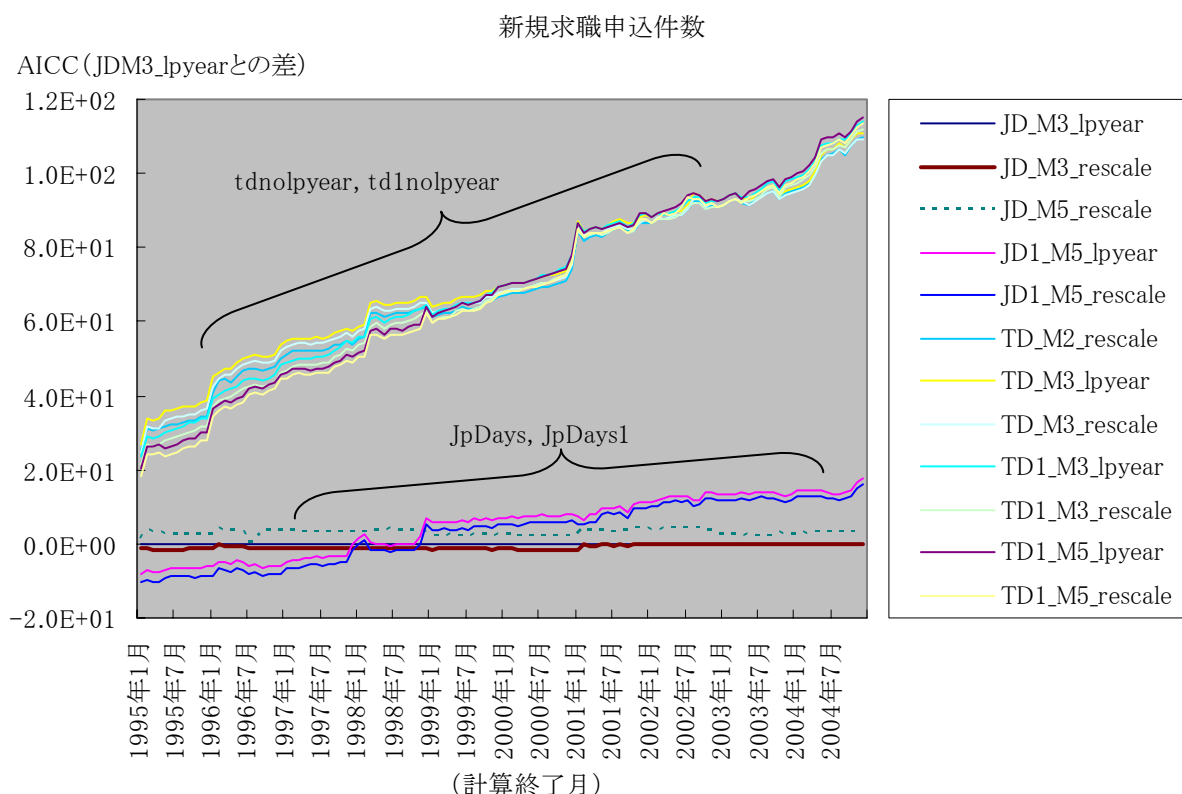
(注34) 本研究の過程で多くのシミュレーションを行った。その結果をみると、AICC の大小は、1 か月先の外挿予測誤差の大小とよく一致していた。しかし、12 か月先の外挿予測誤差と AICC は、あまり一致していないようにみえた。

(注35) AICC にそれほど差がないときは、p+q が小さい方を優先する。また、ミックスモデルである M5 は、できるだけ避けるようにする (第 3 章の 3 参照)。

(注36) rescale は原系列に操作を施しているため、lpyear や nolpyear と被説明変数が異なり、AICC で単純に比較することに疑問がある。

分かったので（図表 2-2-11）、説明変数が少ない **rescale** を採用した<sup>(注37)</sup>。有効求職者数と有効求人人数については、**lpyear** と **no\_lpyear** で差がなかったことから、説明変数が少ない **no\_lpyear** を採用した。就職件数については、説明変数と **ARIMA** モデルで絞り込んだところ **rescale** と **no\_lpyear** が残ったが、外挿予測誤差に大きな差がなかったことから、新規求職申込件数や新規求人数に合わせて<sup>(注38)</sup> **rescale** を採用した。

図表 2-2-8 AICC の比較  
 (“JD~” のモデルが “TD~” のモデルより優良)

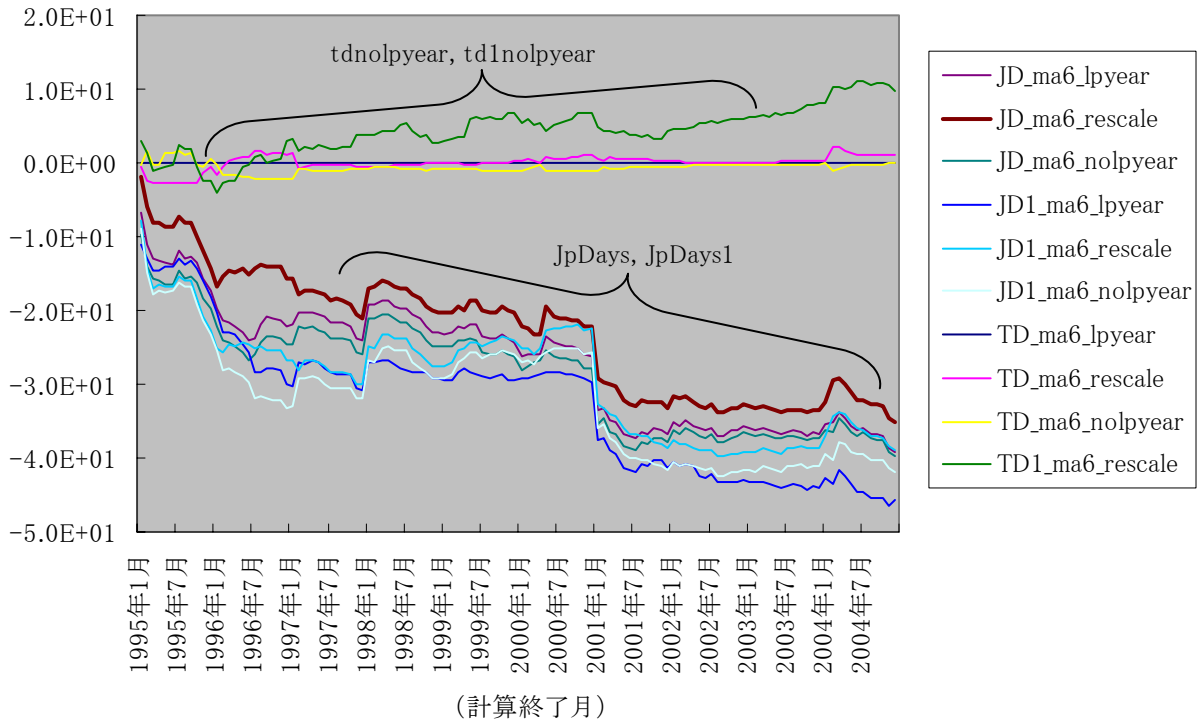


(注37) **lpyear** については、実績データから補正の幅を推計するものであるが、結果に影響を与える閏年 2 月は全期間の 48 分の 1 しかない。例えば、10 年間の計測では、2 ないし 3 個の閏年 2 月のデータが全体の補正幅を支配してしまうことになる。したがって、たとえ過去のデータで **lpyear** のあてはまりが良かったとしても、将来、たまたま閏年 2 月に異常値が発生したとき、**lpyear** の補正方法では季節調整を大きく歪めてしまう危険が大きい。

(注38) 就職件数は、新規求職申込件数や新規求人数と組み合わせて「就職率」、「充足率」といった指標に加工することがあるので、新規求職申込件数や新規求人数と説明変数をそろえた方が分かりやすい。

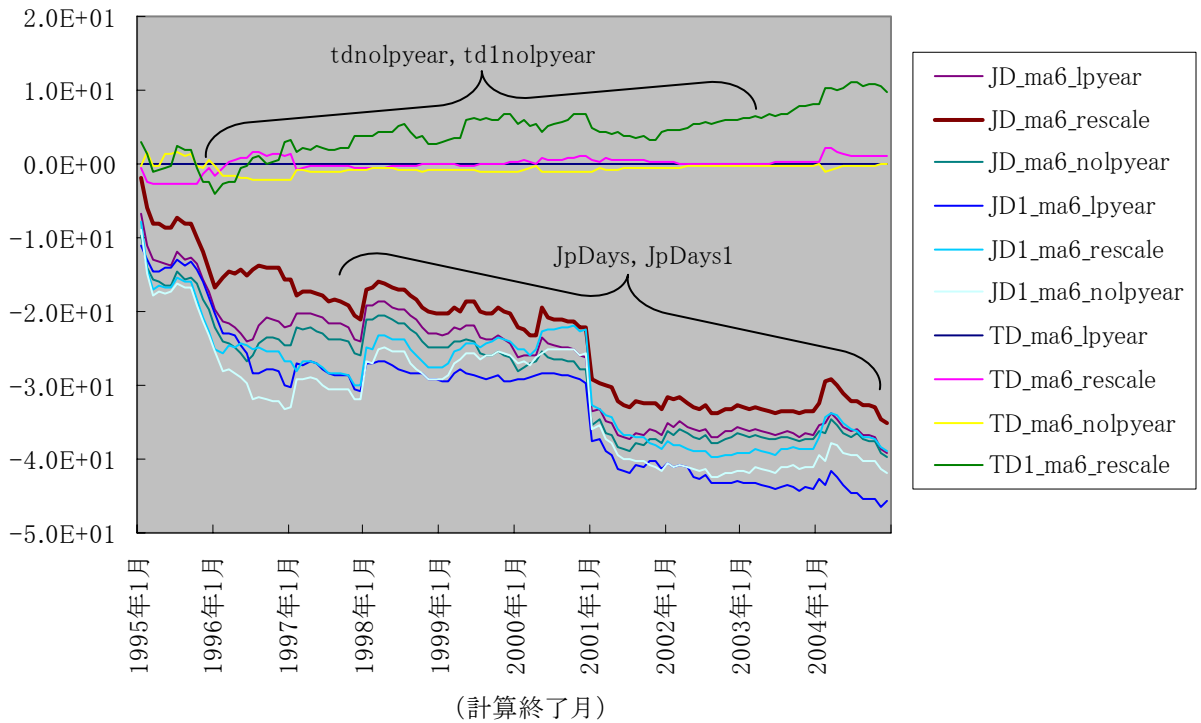
新規求人数

AICC (TD\_ma6\_lpyearとの差)



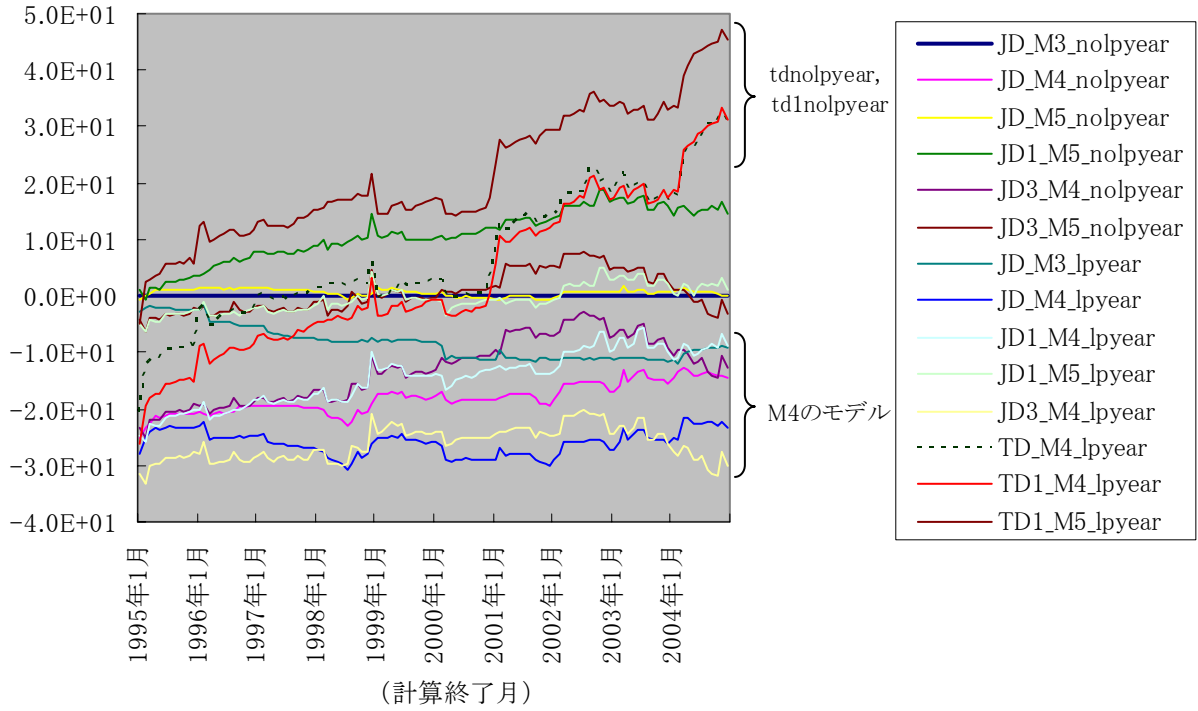
新規求人数

AICC (TD\_ma6\_lpyearとの差)



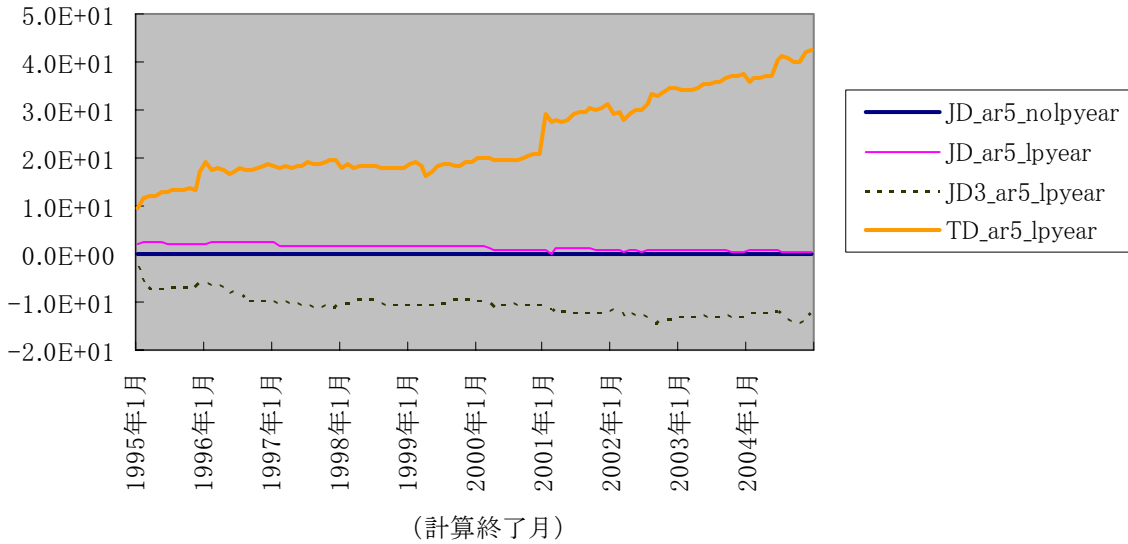
有効求職者数

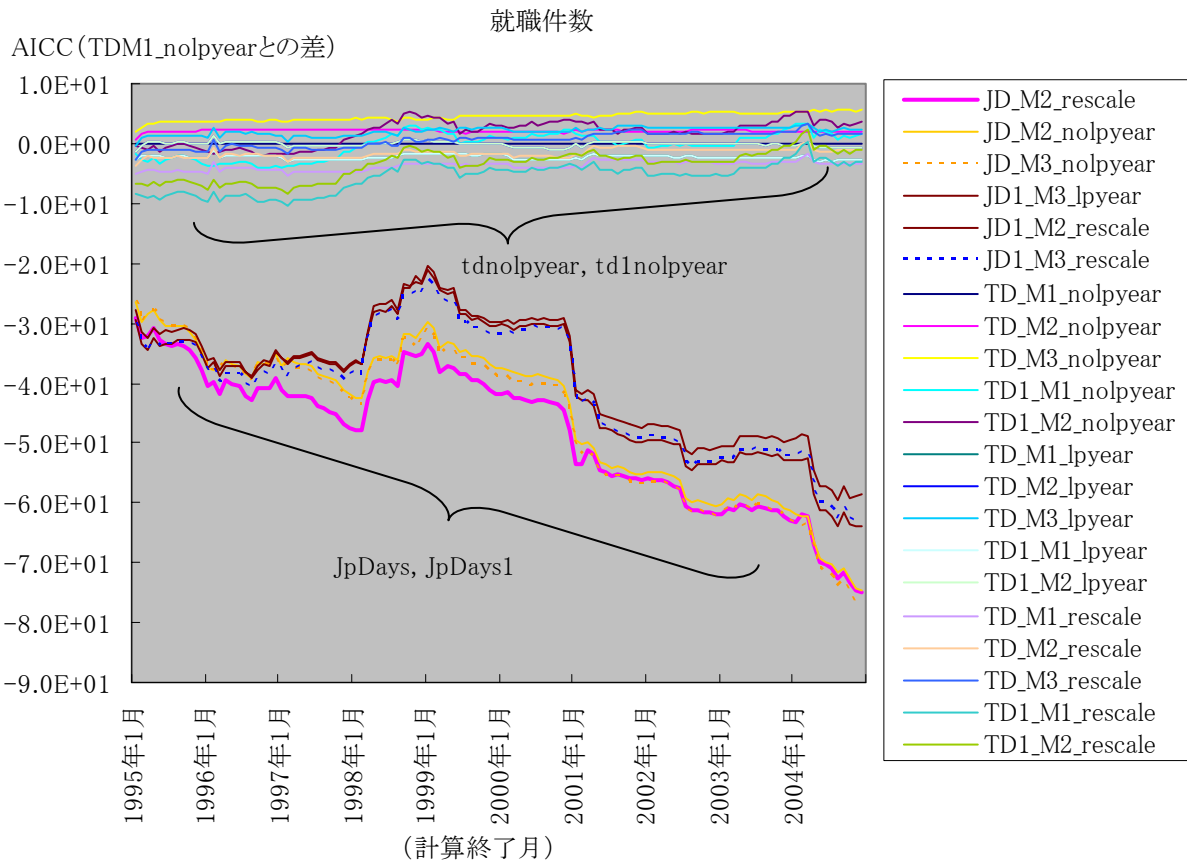
AICC (JD\_M3\_nolpyearとの差)



有効求人数

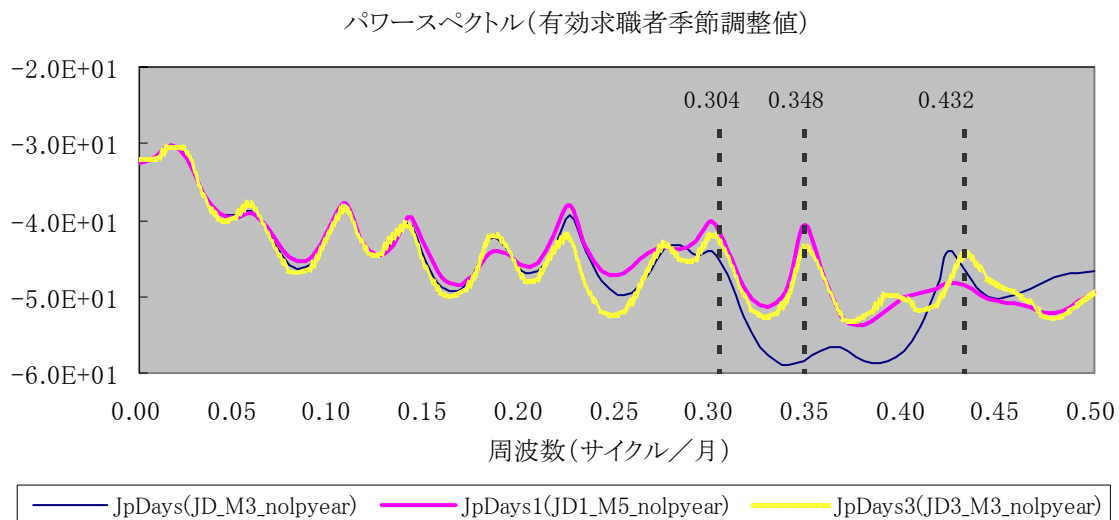
AICC (JD\_ar5\_nolpyearとの差)





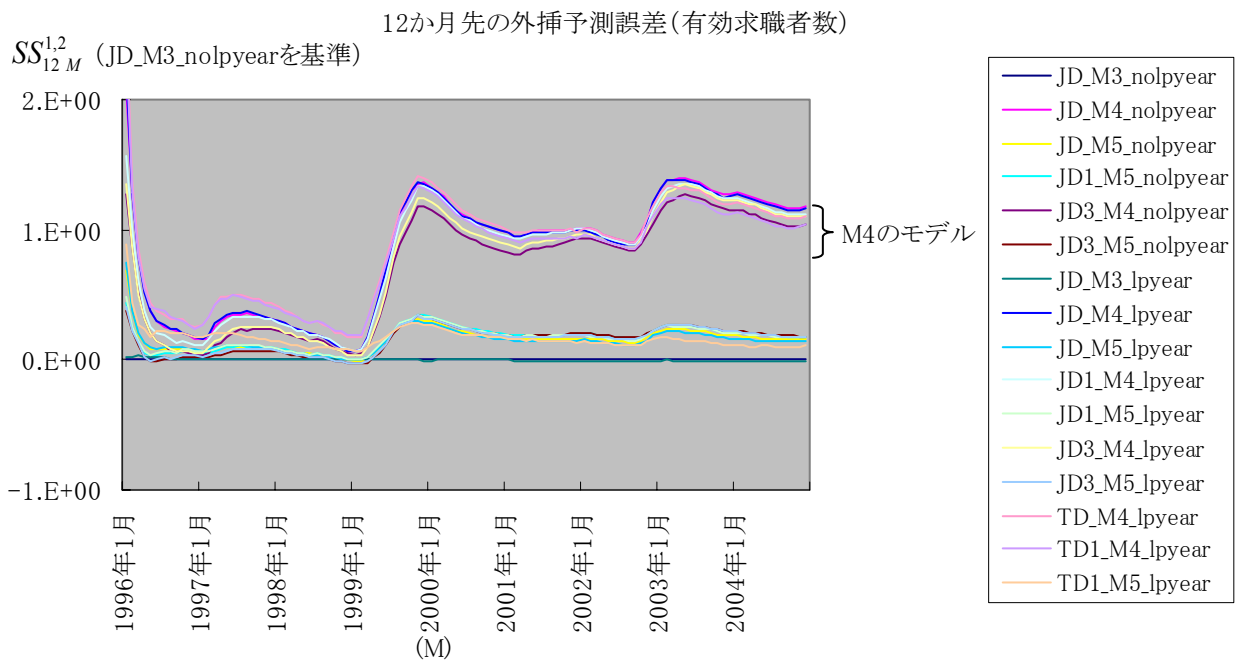
(注) 1) 1985年から図の計算終了月までの期間で推計した。  
 2) AICCは、下に位置するものほど優良とされる。

図表 2-2-9 JpDays1 と JpDays3 は曜日変動のスペクトルピークが残っている



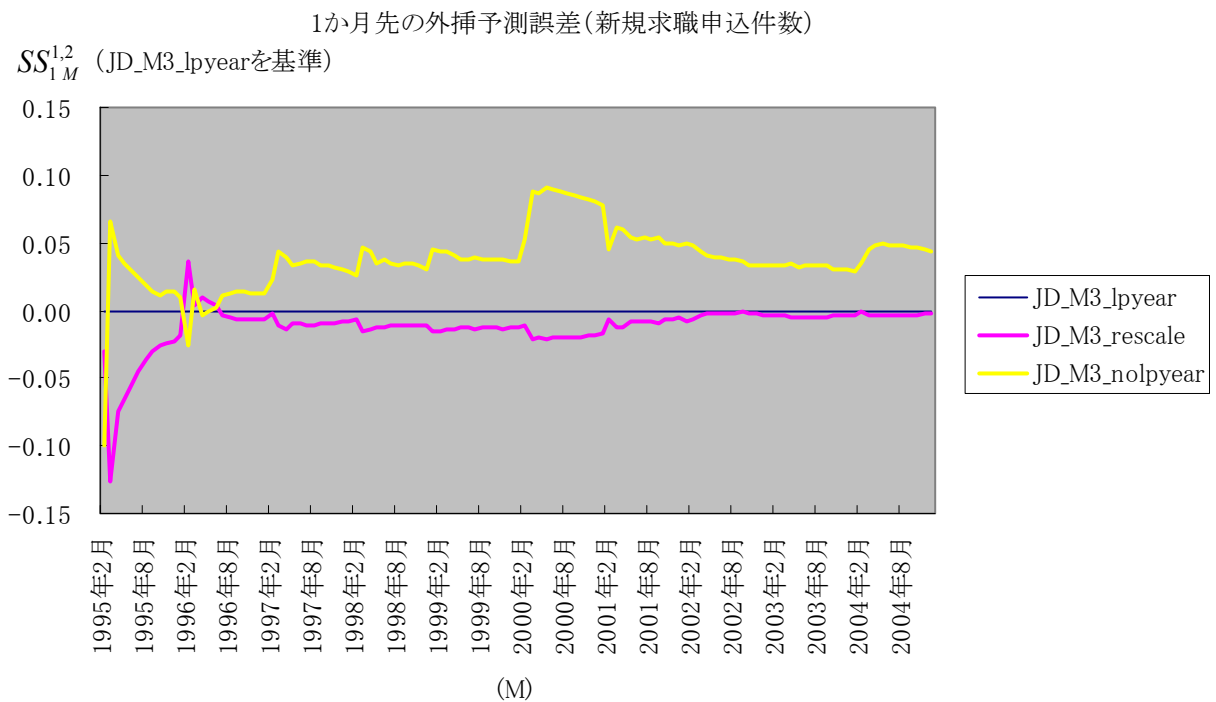
(注) 1) 季節調整値の対数の前期差についてパワースペクトルを計算し、さらにその対数をとった。計測期間は1997年1月から2004年12月。  
 2) 曜日構成の影響を受ける系列では、0.348、0.432、0.304などの周波数でスペクトルのピークが現れる。

図表 2-2-10 M4 のモデルは外挿予測誤差が大きい  
(有効求職者数、12 か月先の外挿予測誤差)



(注) 外挿予測誤差は、相対的に下に位置するもの、右下がりのものが優良 (参考資料 5 参照)。

図表 2-2-11 閏年補正では rescale と lpyear の外挿予測誤差が小さい



(注) 外挿予測誤差は、相対的に下に位置するもの、右下がりのものが優良 (参考資料 5 参照)。

### 3 回帰期間等の設定

#### (1) 回帰期間の長さ

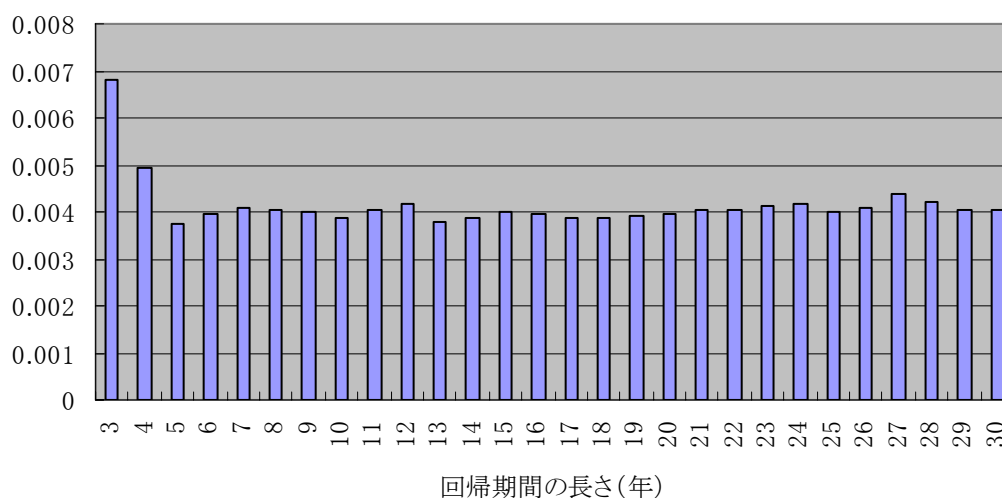
(回帰期間の長さが8年以上あれば推計上問題ない⇒10年回帰とする)

RegARIMAモデルの係数をどの期間で計測するかは、季節調整値の特性に直接影響する要素である。この回帰期間をあまりに長くとりすぎると、最近の構造変化を季節調整に迅速に反映できなくなる。例えば、日本では週休2日制が徐々に普及してきたが、これが普及する以前まで回帰期間に含めると、土曜日の効果は、休日よりはウィークデイに近い性格が強まってしまう。したがって、最近の状況を的確に捉えるためには、回帰期間を短くするのがよい。しかし、一方で、あまりに回帰期間を短くしすぎると、推計結果が不安定になってしまうという悪影響が生じる。

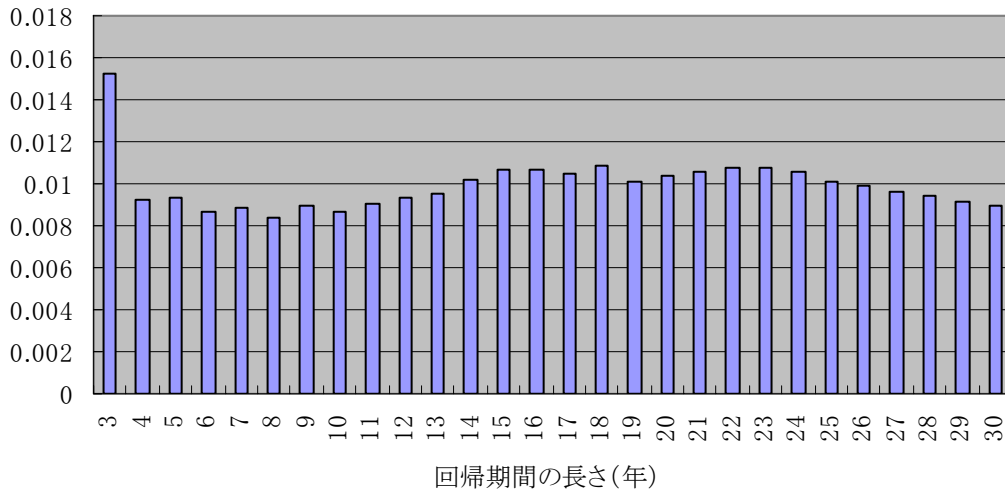
本研究では、職業安定業務統計を使ってシミュレーションを行い、どの程度回帰期間を短くすると外挿予測誤差が拡大するのかを調べた。RegARIMAモデルは前項で設定されたものを用いた。その結果、新規求職申込件数、新規求人数、有効求職者数、有効求人数では、誤差が拡大するのは回帰期間の長さが5年以下の部分であり、就職件数でも7年以下で誤差が拡大するに過ぎないことが分かった(図表2-3-1)。したがって、回帰期間の長さが8年以上あれば推計上特段の問題を生じないと判断される。そこで、本研究の設定が仮に実用化された場合の実務上の利便性を考え、区切りのいいところで、回帰期間の長さを10年にとすることとする。

図表 2-3-1 回帰期間の長さによる外挿予測誤差の比較  
(回帰期間の長さが8年程度以上あれば精度に問題ない)

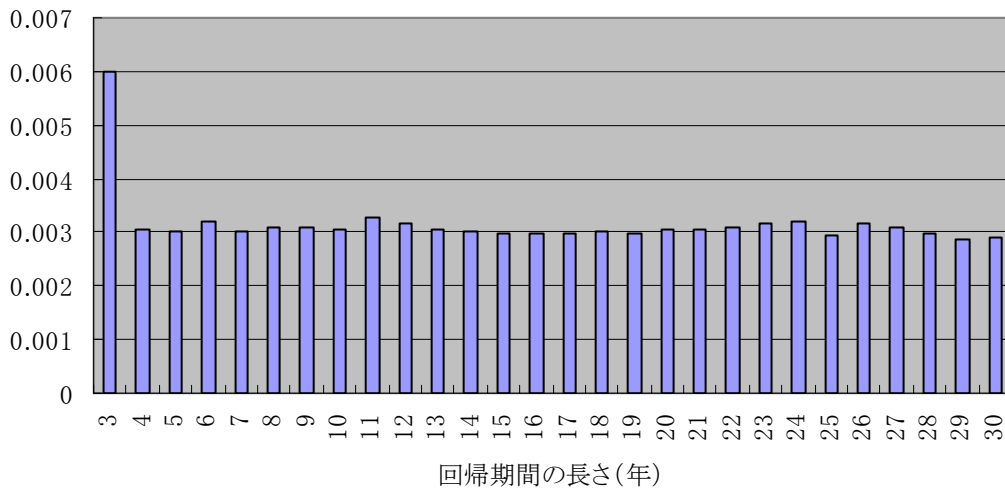
新規求職申込件数の平均平方誤差率(1-12か月予測平均)



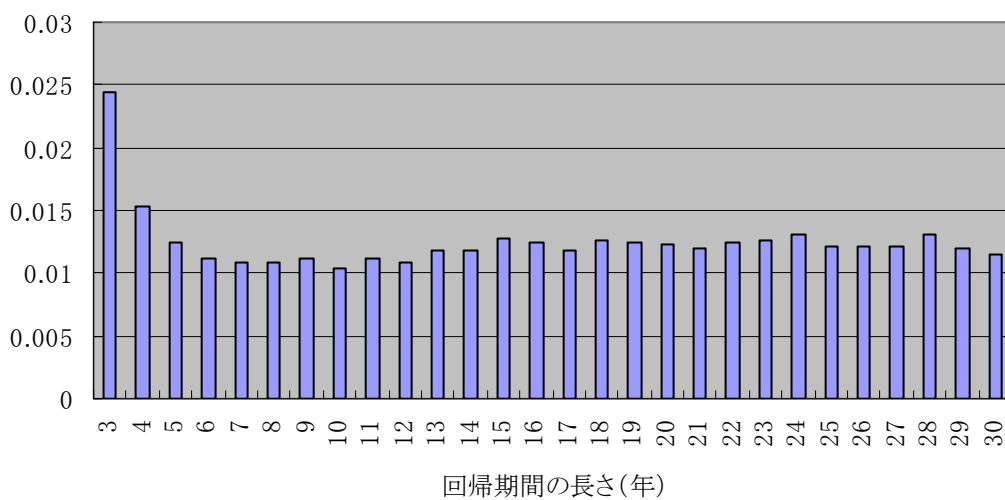
新規求人数の平均平方誤差率(1-12か月予測平均)



有効求職者数の平均平方誤差率(1-12か月予測平均)

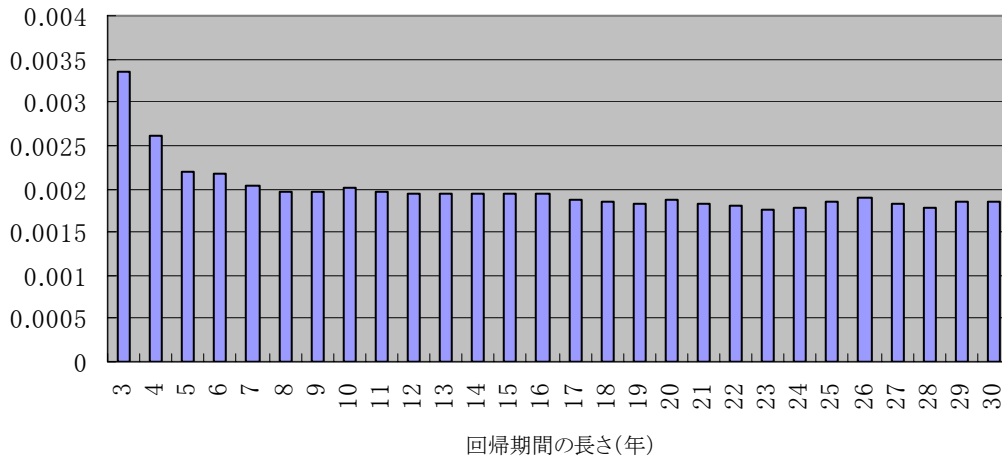


有効求人数の平均平方誤差率(1-12か月予測平均)





就職件数の平均平方誤差率(1-12か月予測平均)



(注) 3年から30年までのそれぞれの長さの回帰期間について、データの終了月を1996年1月から2003年12月まで1か月ずつずらした96のケースで計測を行った。さらに、それぞれのケースについて、1か月先の予測、2か月先の予測、・・・、12か月先の予測を行い、下のように算定される外挿予測誤差率の平方の全ケース(データ終了月96ケース×予測スパン12ケース=合計1,152ケース)に渡る平均をとった。  
 外挿予測誤差率 = (予測値 - 実績値) / 実績値

## (2) 予測期間の長さ

RegARIMAモデルの長所のひとつに、データの末端部分に予測値を使うことにより移動平均の性能を向上できることがある。しかし、誤差が大きいモデルでは、予測値を使うことによりかえって性能が低下することもある。予測値を使うかどうか、また、使うとき予測期間の長さをどう設定したらよいか、といったことは実際のデータに即して判断しなければならない。

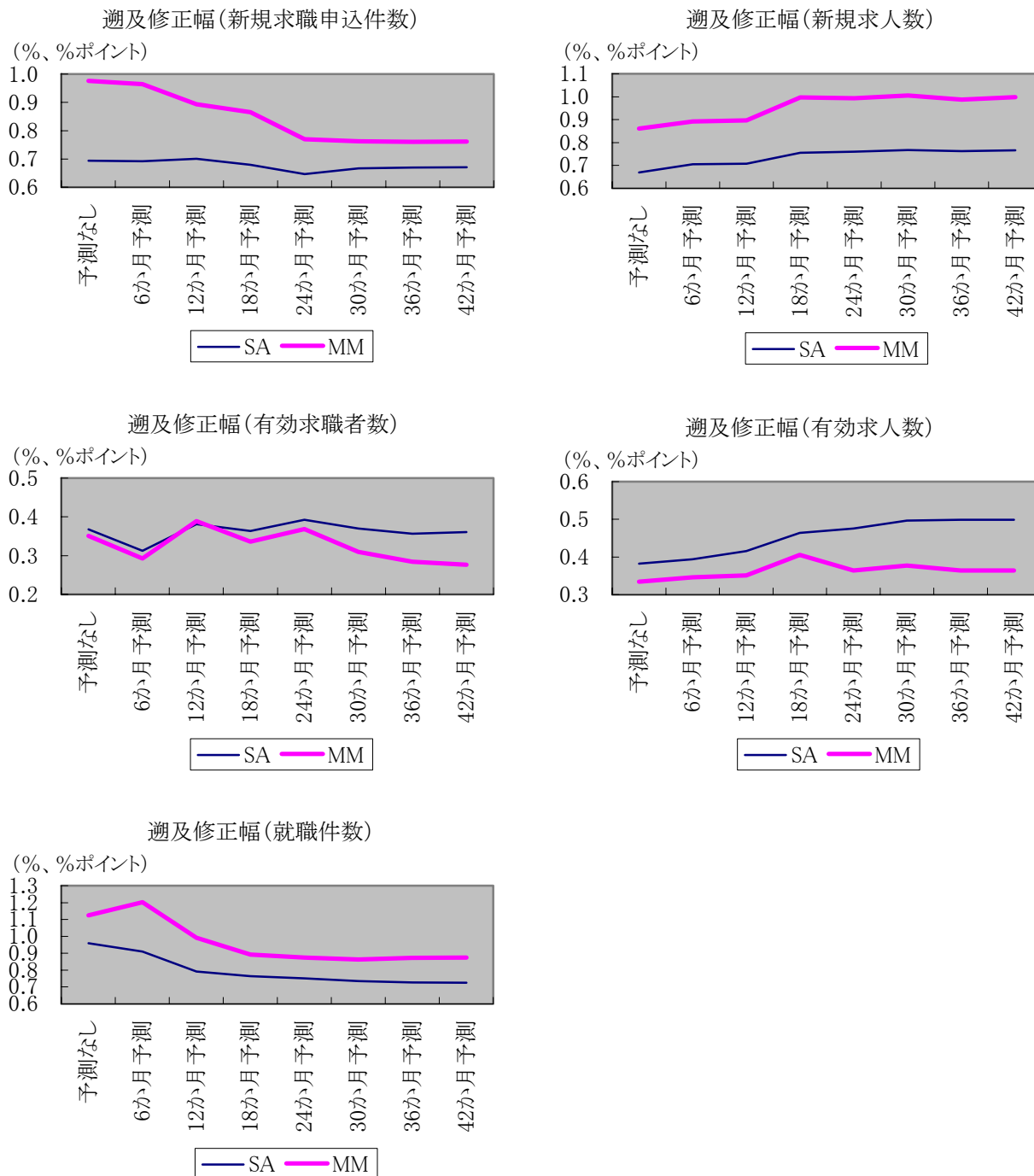
本研究では、季節調整値の安定性をよりどころに予測期間の長さを設定することとした。すなわち、データを1年分追加したときの過去1年間の遡及修正幅をみることにした。遡及修正幅としては、季節調整値の水準自体の修正幅(SA)と対前月比の修正幅(MM)をシミュレーションで推計した。ただ、判定には、実際の場面で使われることが多い前月比の修正幅(MM)の方を重視した。

推計結果から、新規求職申込件数、有効求職者数及び就職件数については長期の予測が優良であり、一方新規求人数と有効求人数については予測値を使うことにより遡及修正幅がかえって拡大することが分かった。そこで、前者の3系列については42か月予測<sup>(注39)</sup>を採用

(注39) 新オプションでは最長3×5項移動平均(3か月移動平均値をさらに5か月移動平均する)が使われる(参考資料9)。この場合、末端部分では42か月分のデータが不足する。予測期間をこれ以上延ばしても結果にあまり影響しないことから、最長予測期間を42か月に留めた。42か月分不足というのは、次の理由による。3×5項移動平均は、該当月の3年前、2年前、1年前、当月、1年後、2年後、3年後にそれぞれ0.067、0.133、0.200、0.200、0.200、0.133、0.067のウェイトを付けた加重平均と同じである。これを実行するためには、該当月の36か月後(3年後)までのデータが必要となる。3×5項移動平均の前に中心化12か月移動平均(第1章1(3)の手順1)が行われるため、さらに前後6か月延長したデータが必要になる。したがって、最新月では、36+6=42か月のデータが不足する。厳密にはヘンダーソン移動平均等のためにもっと不足するが、大きな影響はない。

し、後者の2系列については「予測なし」を採用することとした。

図表 2-3-2 データの1年追加にともなう遡及修正幅



(注) データが1年分追加された場合の過去1年間の季節調整値の修正幅をグラフにした。それぞれの長さについて、データ終了月を1994年12月から2003年12月まで1年ずつずらした10のケース、合計120か月(10年×12か月)について計測した。

SA 修正増減率の絶対値の平均 (%)

MM 対前月増減率の修正増減幅の絶対値の平均 (%ポイント)

例えば、2003年12月までのデータで計算した2003年1月から2003年12月までの季節調整値と、2004年12月までのデータで計算した2003年1月から2003年12月までの季節調整値とを比較して、2003年の12か月分の修正幅とした。

## 4 結果

### (1) 選択されたオプション

以上の検討作業の結果、次のオプションが選択された。

図表 2-4-1 選択されたオプション

	新規求職申込 件数	新規求人数	有効求職者数	有効求人数	就職件数
稼働日調整	JpDays	JpDays	JpDays	JpDays	JpDays
レベルシフト	ls1975.4	ls1975.4	ls1975.4	ls1975.4	ls1975.4
閏年補正	Rescale	Rescale	行わない	行わない	Rescale
ARIMA モデル	(2 1 0)(0 1 1)	(0 1 6)(0 1 1)	(2 1 0)(0 1 1)	(5 1 0)(0 1 1)	(0 1 2)(0 1 1)
回帰期間の長さ	10 年	10 年	10 年	10 年	10 年
計算期間の長さ	全期間	全期間	全期間	全期間	全期間
予測期間の長さ	42 か月	0 か月	42 か月	0 か月	42 か月

(注) 1) JpDays とは、各月の祝日数及び曜日別日数から作成される 8 個の変数の総称。

2) ls1975.4 は、次の値をとる変数。1975 年 4 月は、雇用保険法が施行された月。

1975 年 3 月以前 -1

1975 年 4 月以降 0

3) Rescale とは、データに事前に次の変換を施すこと。

閏年の 2 月  $Y_t = (28.25/29) Y_t$

閏年以外の 2 月  $Y_t = (28.25/28) Y_t$

その他の月  $Y_t = Y$

4) 回帰期間とは、稼働日調整及び ARIMA モデルの係数の推計に使用されるデータの期間。

5) 計算期間とは、季節調整値の算定に使用されるデータの期間。これは、季節調整値の出力期間と同じ。

6) 予測期間とは、最新データの付近で移動平均を実行するときに用いられる予測値の計算期間。

7) X-12-ARIMA で用いられる X-11 機能については、標準オプションを採用する。

いくつか補足する。

#### (レベルシフト)

5 系列すべてにおいて、1975 年 4 月を境にする断層が有意だった。そこで、この断層を表現するレベルシフト変数 ls1975.4 を RegARIMA モデルに加えた。1975 年 4 月は、雇用保険法が施行された月なので、これが求人・求職等の動向に影響を与えたものとみられる。

沖縄が日本に復帰したことに伴い、厳密には 1972 年 4 月からのデータにも断層が存在する。しかし、断層の存在を検定したところ、新規求職申込件数、新規求人数、有効求職者数、就職件数においては有意でなく、また、有効求人数では有意だが影響が小さいことが判明した。このため、有効求人数を含めすべての系列で、沖縄復帰に関して特段のレベルシフト変数を導入しないこととした。

なお、X-12-ARIMA には、断層の存在を自動的に検出して説明変数に追加する機能がある。しかし、季節調整値が予期しない振る舞いを示すことを避けるため、この自動検出機能

を使わないこととした。

### (計算期間)

回帰期間が RegARIMA モデルを推計する期間なのに対して、計算期間は、移動平均による季節調整を実行する期間である。

回帰期間に対応する直近 10 年の季節調整値を得るためには、末端部の移動平均を考慮して、14 年程度の長さの計算期間を指定すれば十分である。ただ、①計算期間を過去に延ばしても直近の季節調整値にほとんど影響しないこと、②計算期間の長さを特定の年数に固定するにはスペックファイル（オプションを指定するファイル）の該当部分を毎年変更する煩雑さがあること、を考慮して、省力化のため計算期間を全期間とした。

職業安定業務統計は、原則として 1963 年 1 月から利用できるのもので、「全期間」の指定は、原則として 1963 年 1 月から直近データまでの季節調整を実行することを意味する。

### (X-11 機能のオプション)

現行公表値では、X-11 機能で次の標準でないオプションを使っている（参考資料 9 の 2）。

- ① 移動平均項数 3×1
- ② 特異項管理限界 1.6、2.6

移動平均項数については、季節調整値の安定性を考慮して（後出（2）参照）、今回は標準オプションを採用することとした。また、特異項管理限界については、標準オプション（1.5、2.5）との差異がわずかであること、さらに、今回、稼働日調整を取り入れたことにより不規則要素(I)の性格が従来のもので大きく変わっており、従来の特異項管理限界は参考にならないことから、煩雑さを避けるために今回は標準オプションを採用した。

## (2) 現行公表値との比較

新しいオプションによる季節調整値を現行の公表季節調整値と比較してみる（図表 2-4-2～図表 2-4-8）。

### (稼働日要因の変動が除去された)

最近 3 年を拡大した図（各ページ上段）をみると、新規求職申込件数、新規求人数、及び就職件数で稼働日要因の変動が顕著に改善している。現行公表値では新規求職申込件数における稼働日要因の変動が新規求人より大きいことから、新規求人倍率では稼働日の多い月に倍率が高くなるという逆サイクルの動きを示している。この新規求人倍率の動きも、新オプションでは解消している。

有効求職者数と有効求人数については、現行公表値がもともと滑らかだったことから、新規の系列ほど顕著でないが、これも改善がみられる。ただ、有効求人倍率については、現行

公表値と新オプションであまり差がみられない。

長期時系列でみると、とくに近年、新オプションの方がグラフが滑らかで傾向を読みとりやすくなっているのが分かる（各ページ中段）。

以上のことは、パワースペクトル<sup>(注40)</sup>によっても確認できる（各ページ下段）。多くの系列で、曜日構成による変動を示す0.348、0.432、0.304の周波数のピークが、新オプションでは消えている。また、パワースペクトルの全体的な水準が、新オプションでは現行公表値より下に位置している。これは、新オプションの方が全体的に滑らかな動きになっていることを示している<sup>(注41)</sup>。

### （長期の過去には乱れも）

ただ、就職件数について顕著なように、長期の過去については、新オプションの方が数値が乱れている（図表 2-4-8 中段）。回帰期間を最近 10 年に限ったため、長期の過去については当時の現状にそぐわない「稼働日調整」が行われてしまったことが、これの一因として考えられる。

このような不都合を解消するため、新オプションの実際の運用に当たっては、後出の「6 年以前固定方式」を採用することが考えられる。

### （遡及修正は縮小）

職業安定業務統計の現行の公表方式では、各月の発表時にまず予測季節要素によって計算された季節調整値が発表される。そして、翌年 1 月分の発表時に、12 月までのデータから計算された季節調整値で第 1 回目の修正がなされる（図表 2-4-9）。その後毎年、1 月分の発表時に同様の修正がなされていく。

通常、最も修正幅が大きいのが第 1 回目の修正である。そこで、新オプションの第 1 回目の修正幅を現行公表値と比較してみる（図表 2-4-10）。SA が水準自体の修正幅で、MM が前月比の修正幅である。SA、MM のいずれでみても、7 系列すべてで新オプションの修正幅の方が小さいことが分かる。すなわち、新オプションの方が結果の安定性が高い。

なお、予測季節要素を使わない同時調整（Concurrent Adjustment：毎月季節調整プログラムを動かして再計算する）の方法で同様の修正幅をみると、すべての系列でさらに縮小している。遡及修正幅を縮小する目的からは、オプションの見直しだけでなく、公表方法の見直しも有効なのである。

---

(注40) 図表 2-4-2～図表 2-4-8 では、季節調整値を対数変換して前月差をとった系列のパワースペクトルを対数変換して示してある。計測期間は、1997 年 1 月から 2004 年 12 月までの 8 年間。

(注41) 一般に、 $x_t$  を離散定常系列として  $\sigma^2$  をその分散、 $p(\omega)$  をそのパワースペクトルとすると、

$$\sigma^2 = 4\pi \int_0^{0.5} p(\omega) d\omega$$

の関係がある (Priestley[5])。すなわち、分散が小さい系列ほどパワースペクトルが平均的に小さい。

### (移動平均項数変更の効果)

新オプションで修正幅が縮小したことについては、移動平均項数を変更したことの効果も大きい。

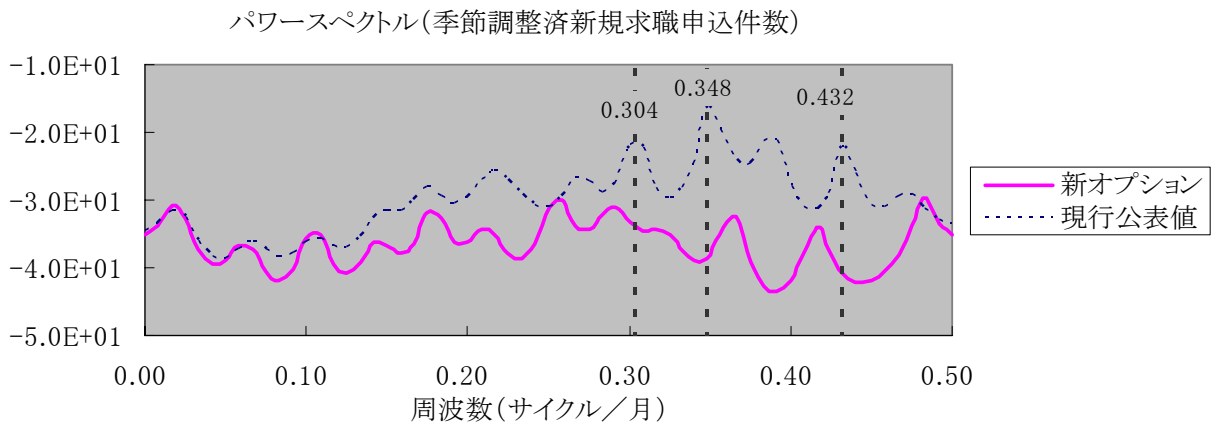
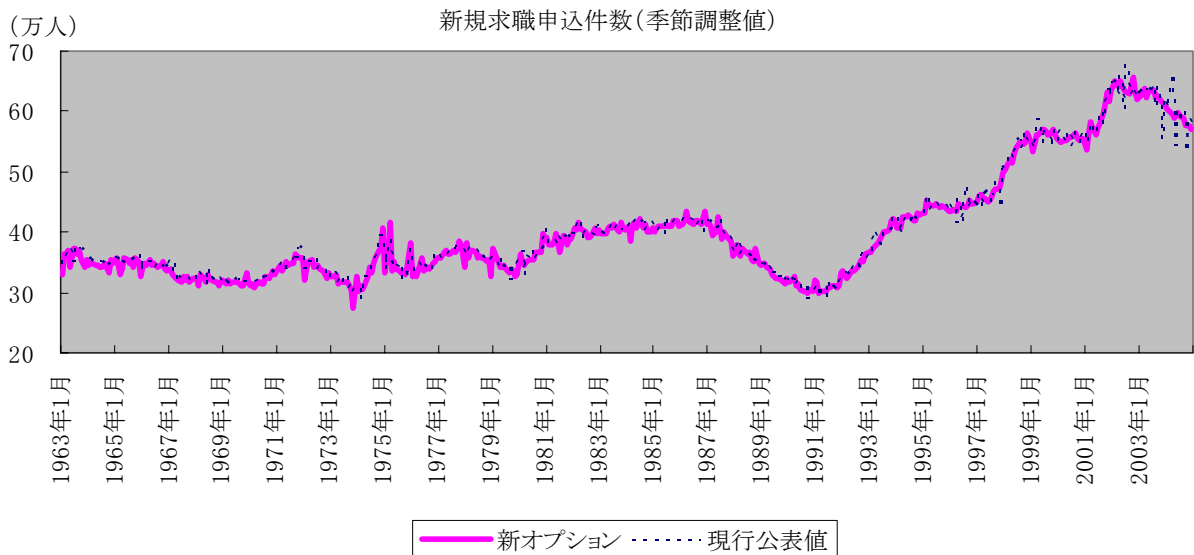
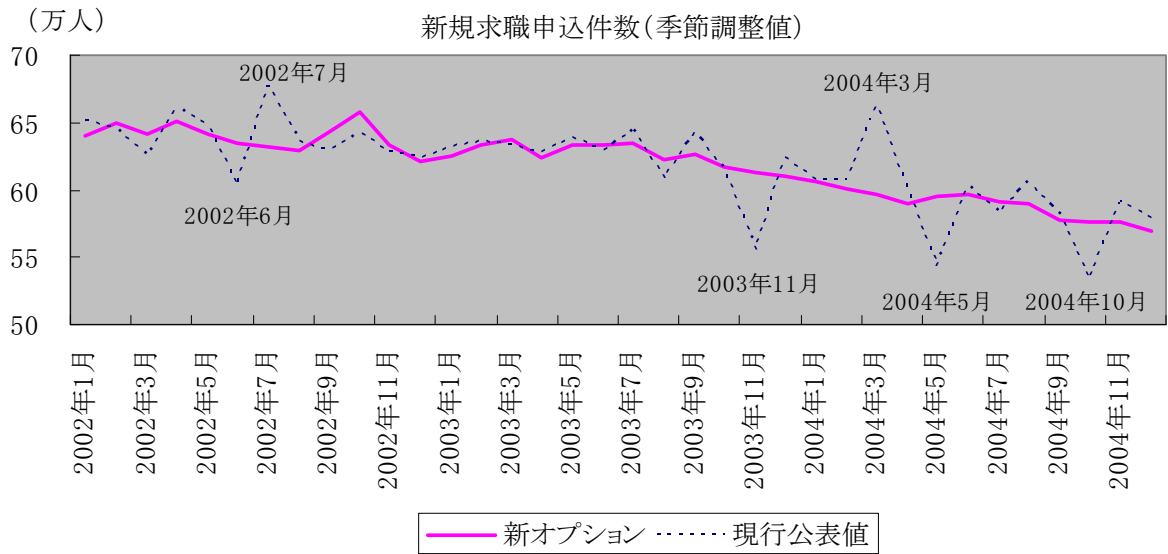
現行公表値では、季節要素×不規則要素 (S×I) から季節要素 (S) を抽出する過程 (第1章1(3)の手順3) で3項移動平均という極めて短い移動平均を用いている。一方、新オプションでは、X-12-ARIMAの標準オプションである3×3項ないしは3×5項の移動平均(3項移動平均したものをさらに3項ないし5項移動平均する)を用いている(参考資料9の2)。別途、現行の季節調整方法で移動平均項数を上記のように見直しただけのもの(RegARIMAモデルを使わないもの)を調べてみたところ、それでもかなりの遡及修正幅縮小が確認された。

現行の短い移動平均は、雇用保険法の施行等制度の変更にもなう特異な動きを季節要素の変動として捉える傾向が強く、結果的に滑らかな季節調整値が得られるが、一方で安定性が損なわれる可能性がある。新オプションでは、移動平均項数が長くなったことで、多少ごつごつした季節調整値になる効果<sup>(注42)</sup>があるが、安定性は向上する。なお、もし滑らかなデータが必要なら、移動平均項数を短くするのではなく、X-12-ARIMAで出力される趨勢循環要素(C)を使うのが良い。

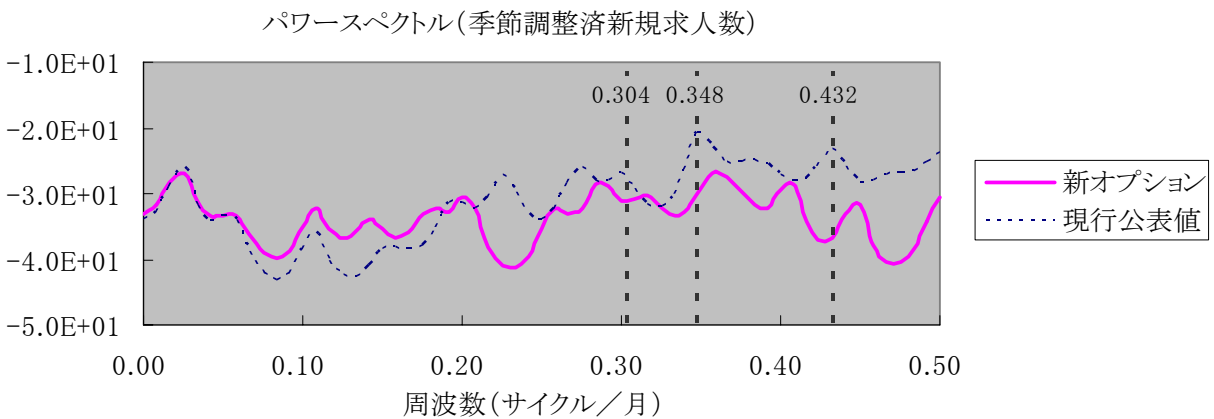
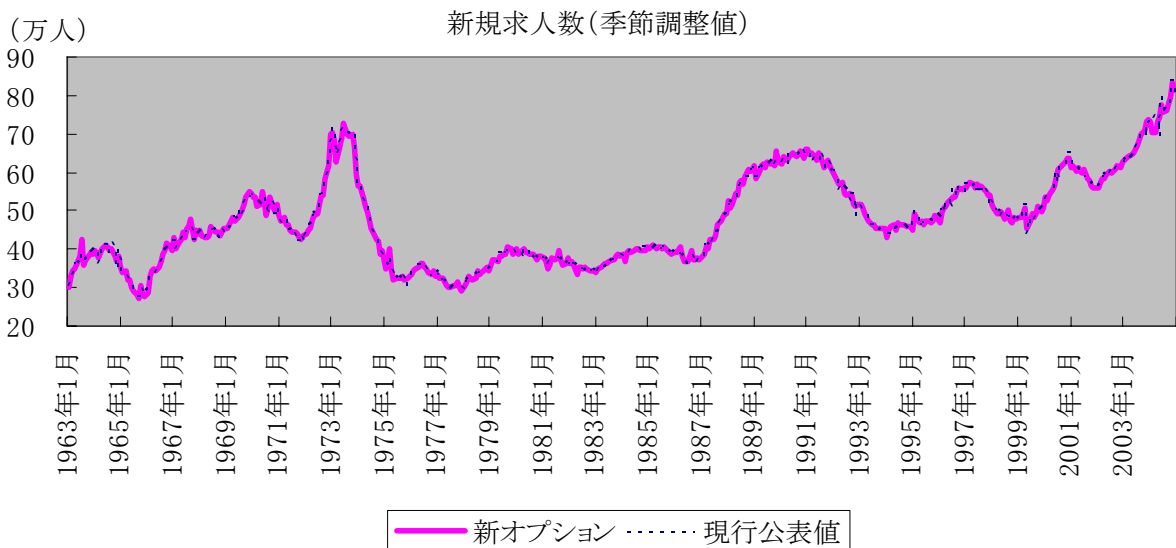
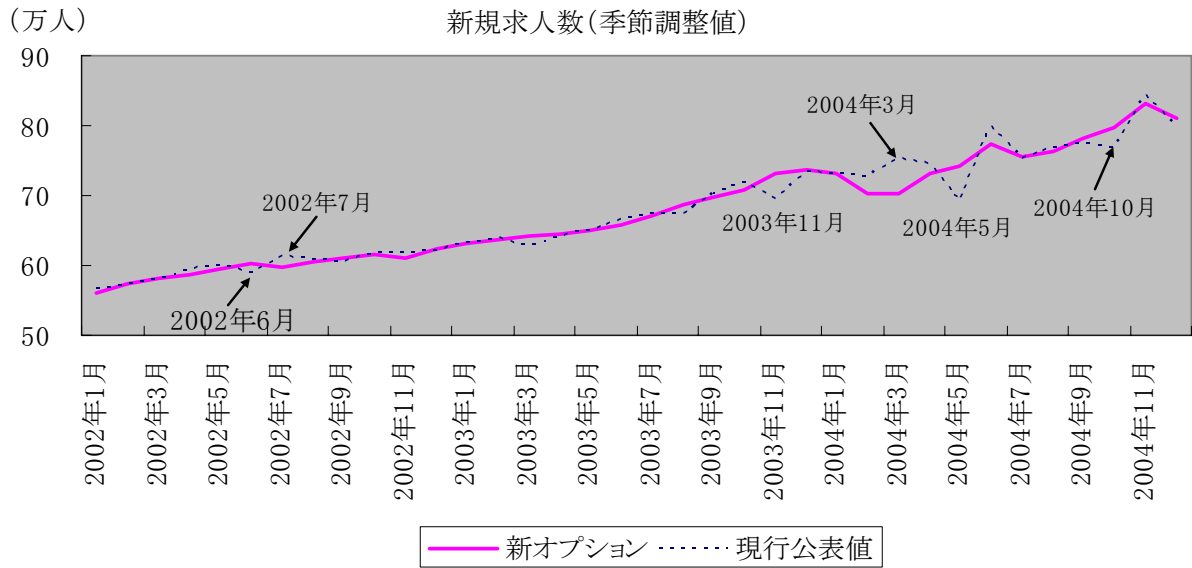
---

(注42) ただし、新オプションでは、稼働日要因を除去しているので、この効果により全体としてはむしろ現行公表値より滑らかになっている。すなわち、①移動平均項数を長くした効果(滑らかさ低下)、②稼働日要因を除去した効果(滑らかさ向上)のうち、実際には②の効果が勝っている。

図表 2-4-2 新オプションによる新規求職申込件数

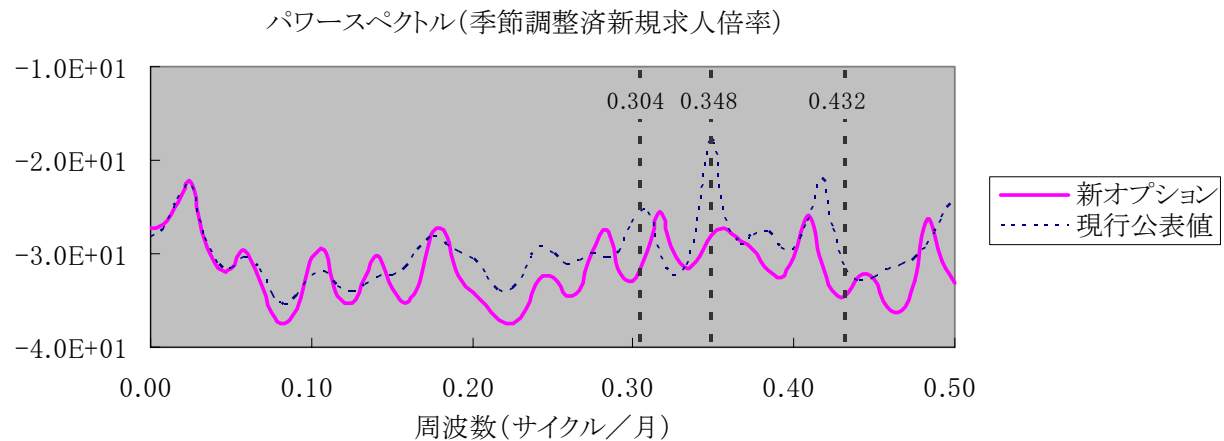
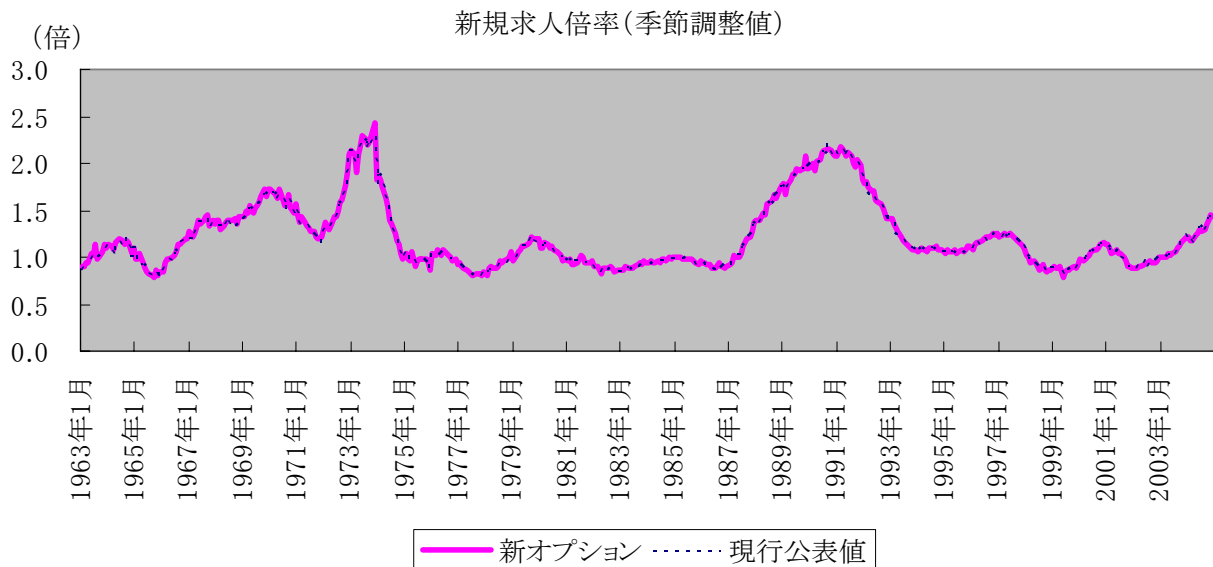
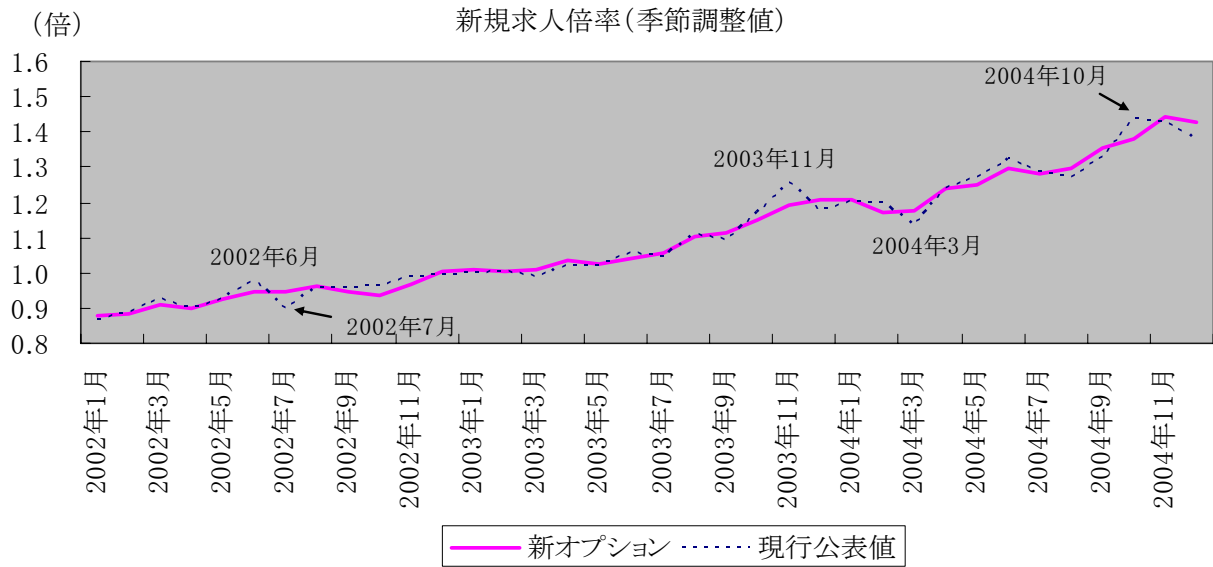


図表 2-4-3 新オプションによる新規求人数

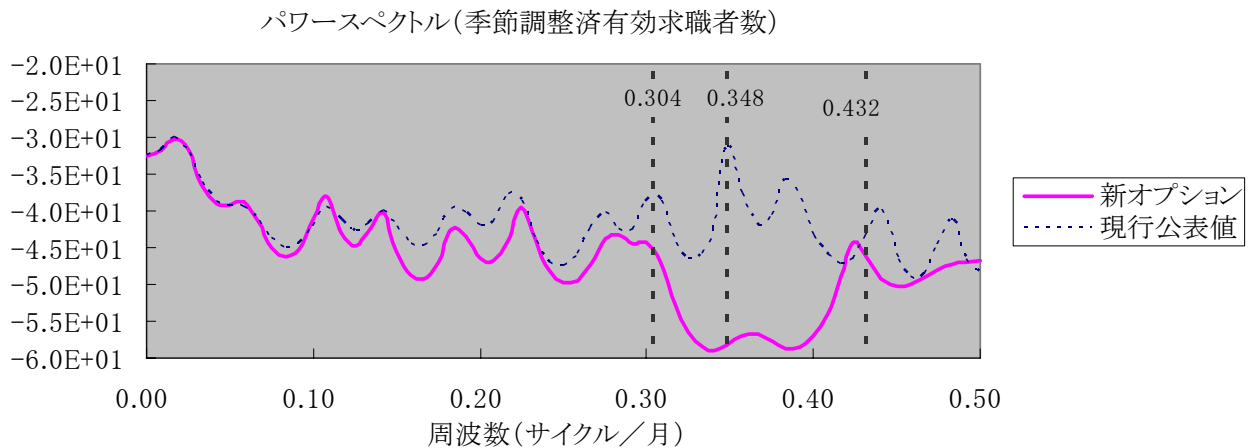
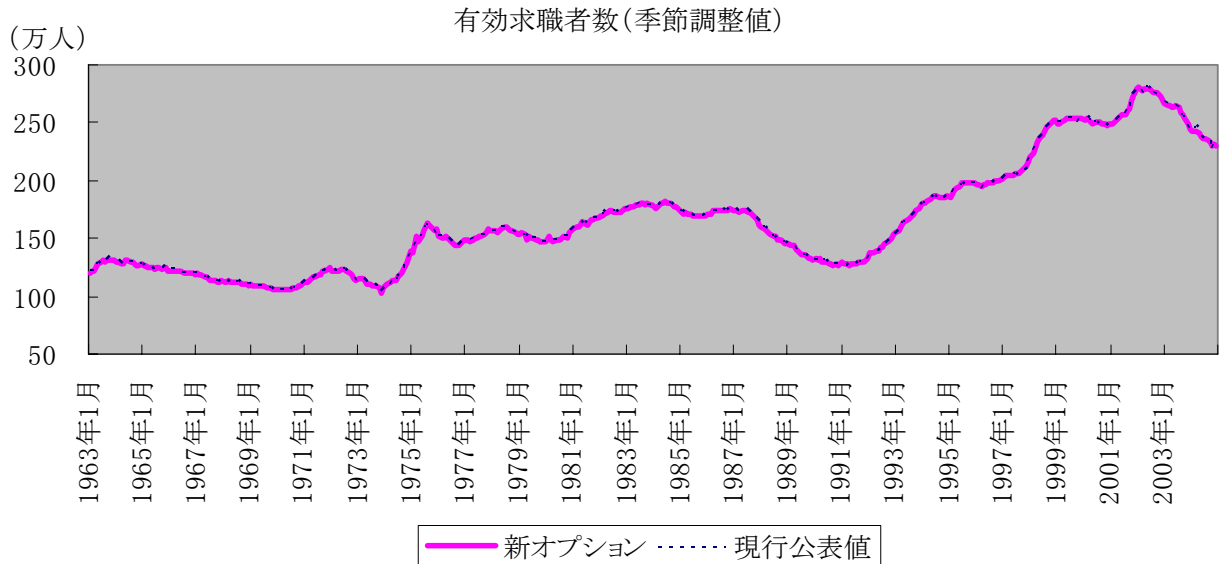
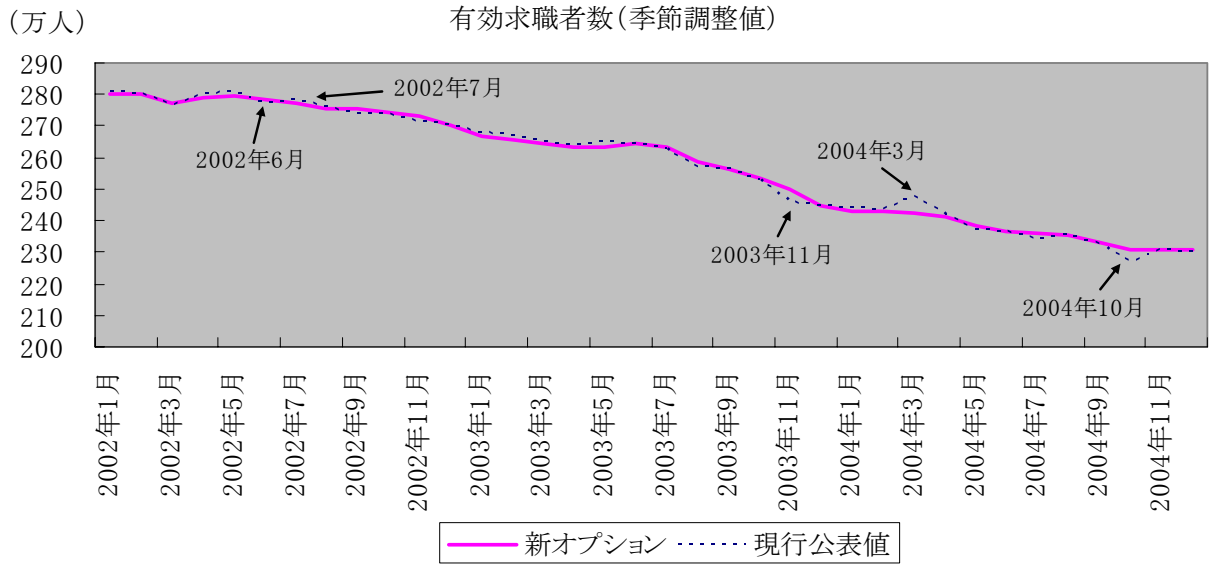




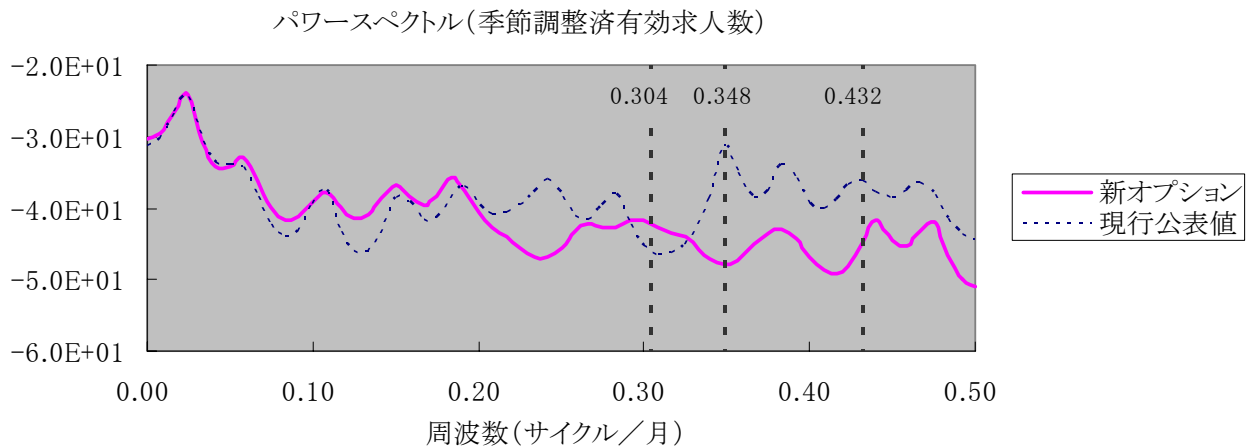
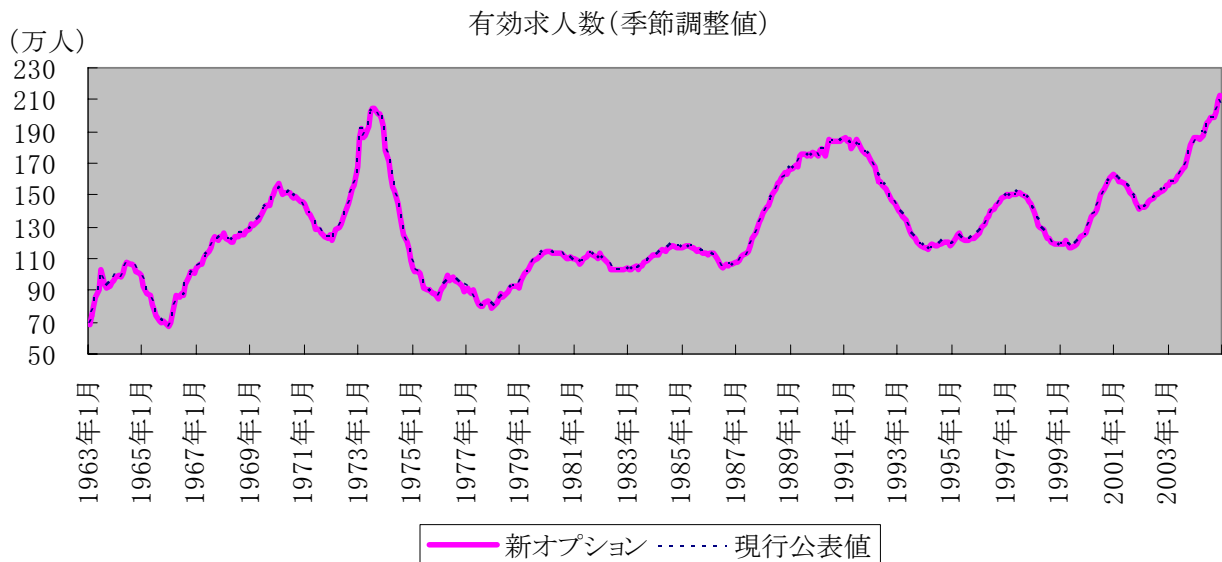
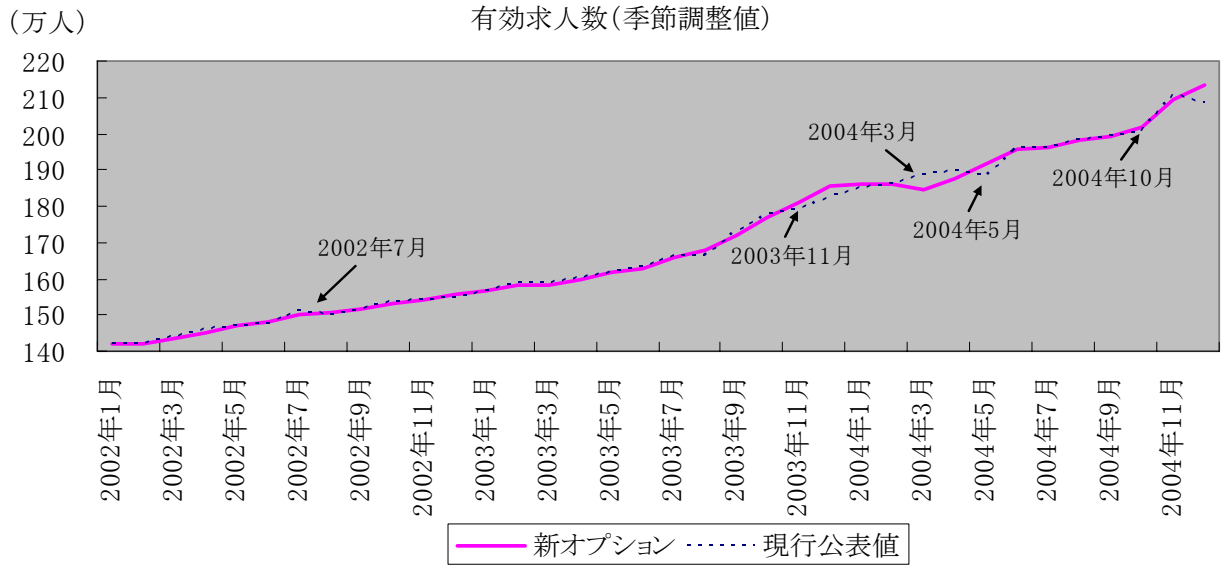
図表 2-4-4 新オプションによる新規求人倍率



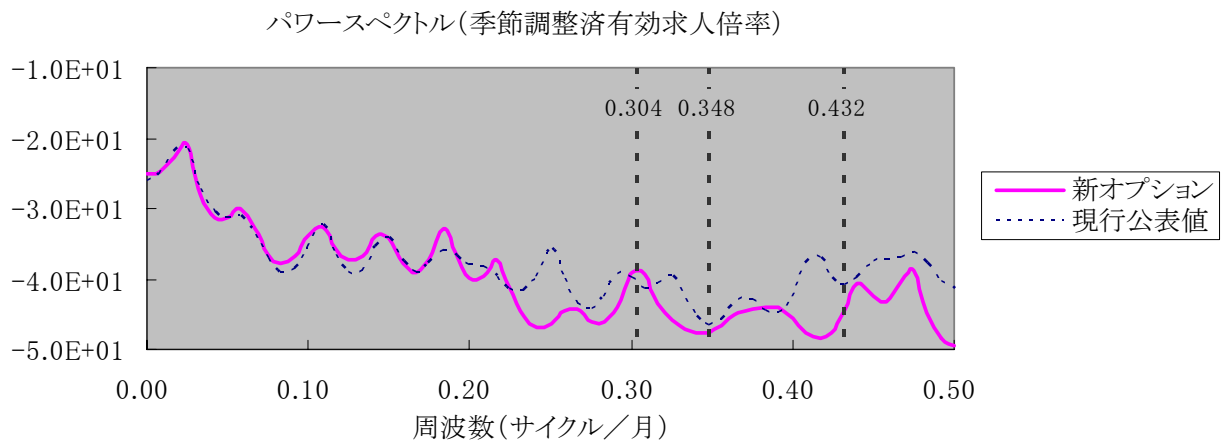
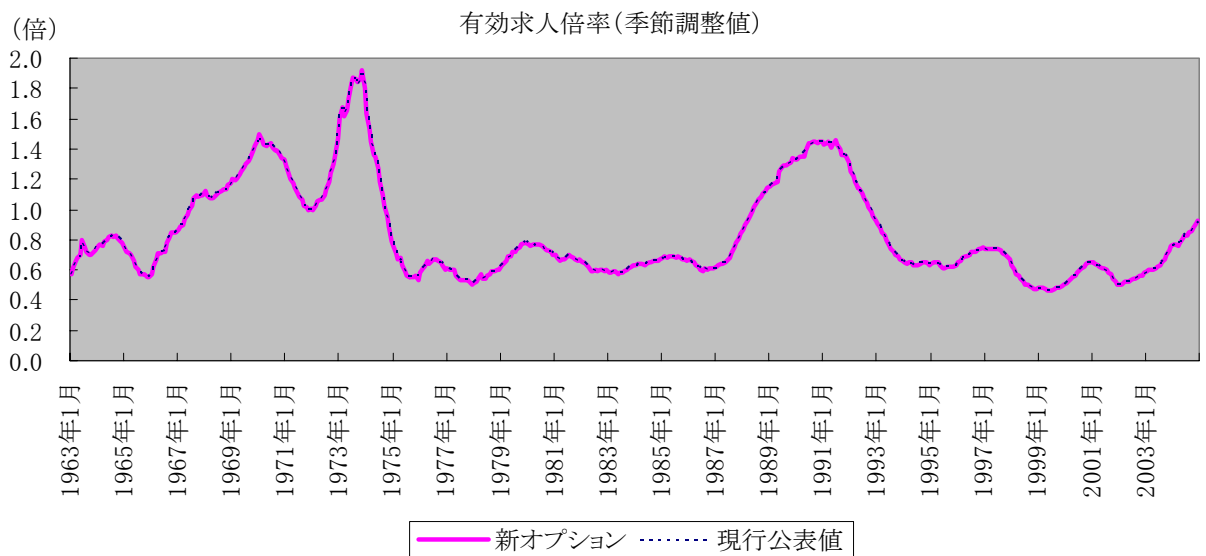
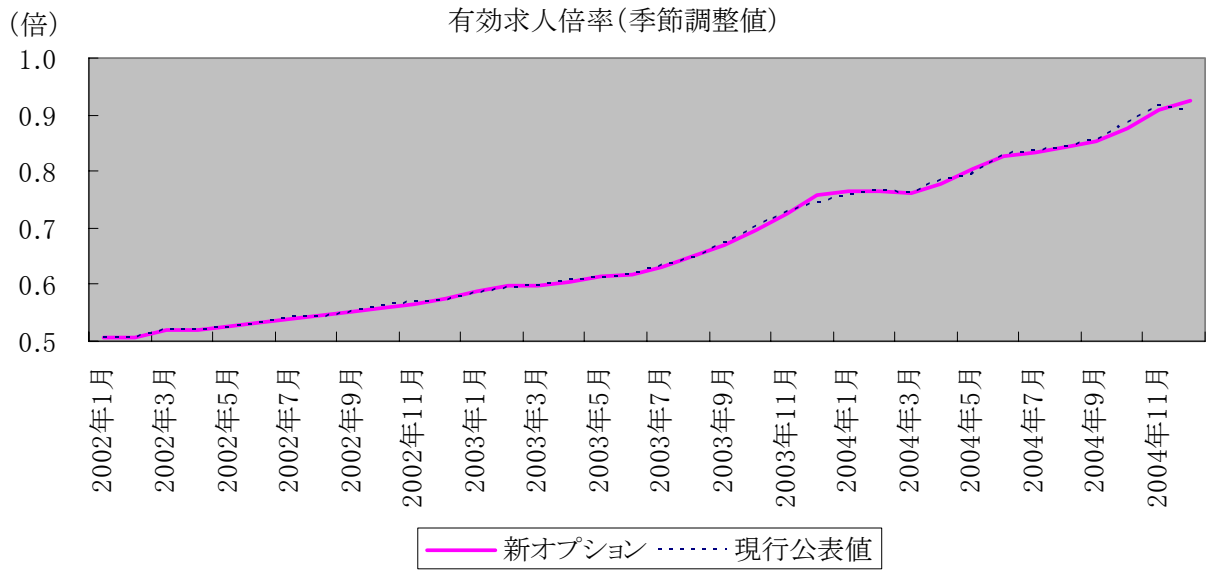
図表 2-4-5 新オプションによる有効求職者数



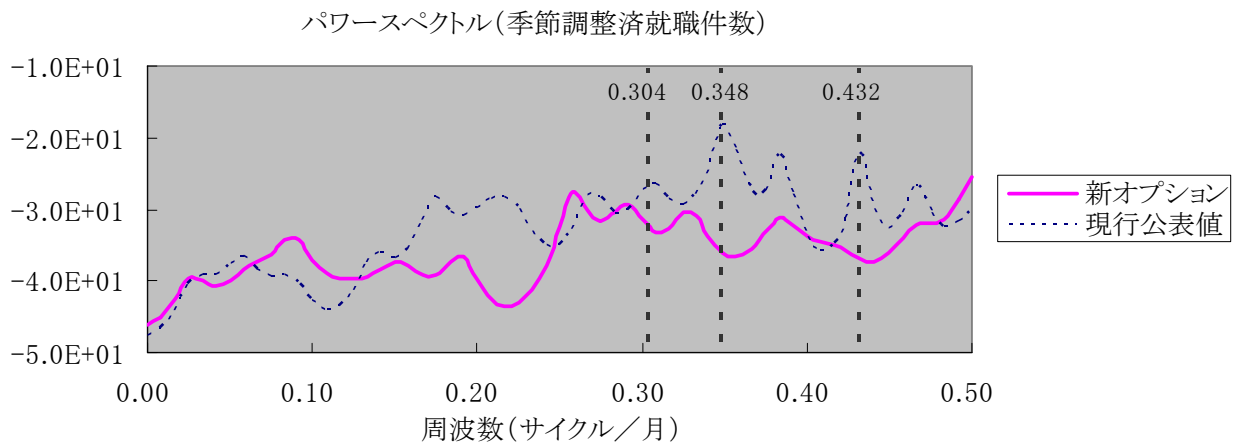
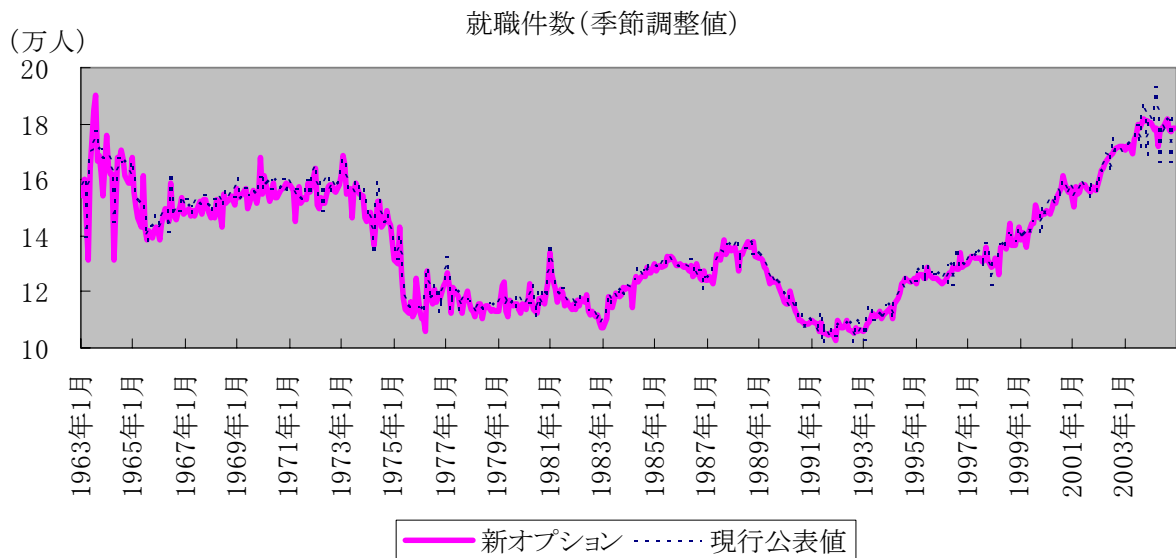
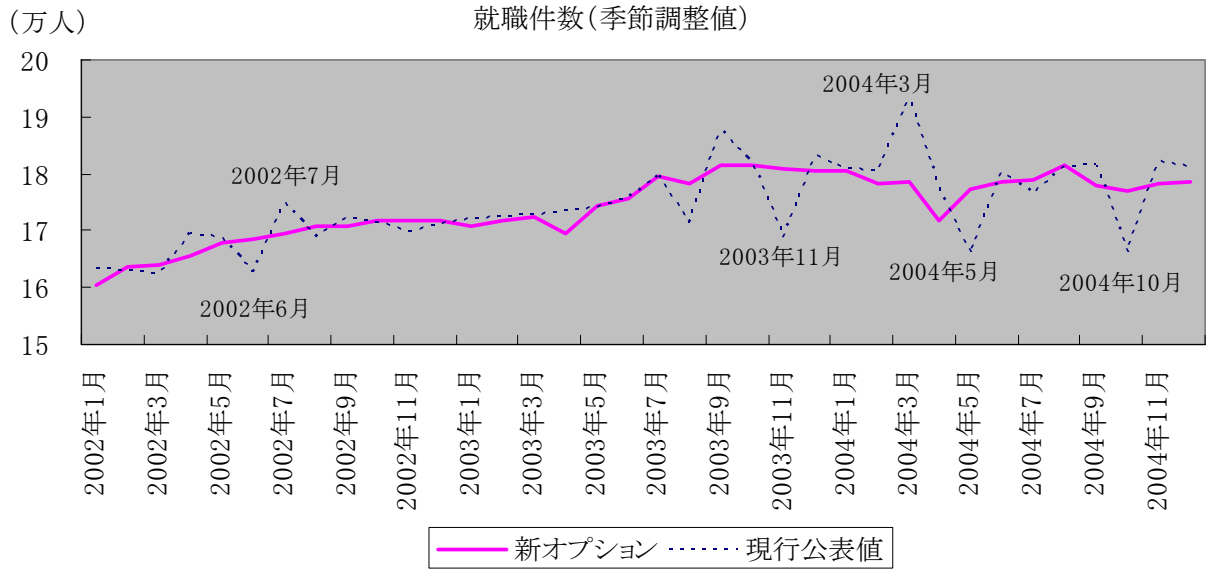
図表 2-4-6 新オプションによる有効求人人数



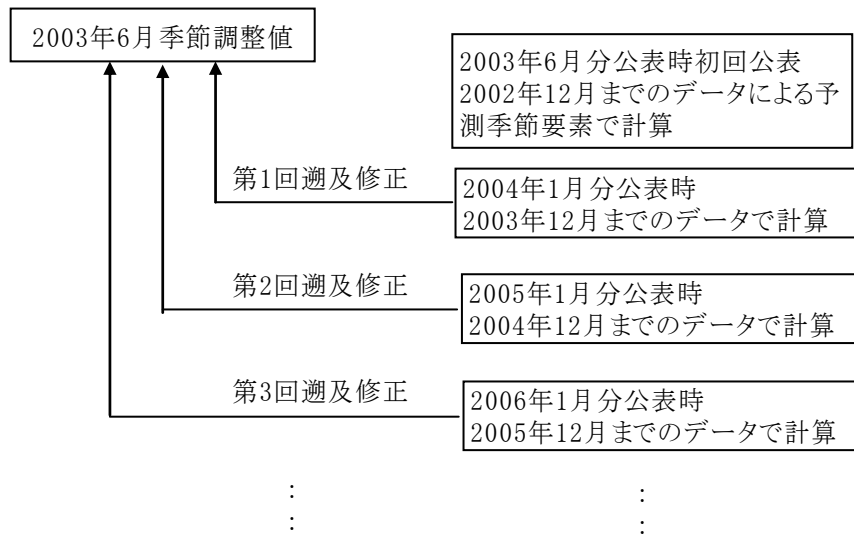
図表 2-4-7 新オプションによる有効求人倍率



図表 2-4-8 新オプションによる就職件数

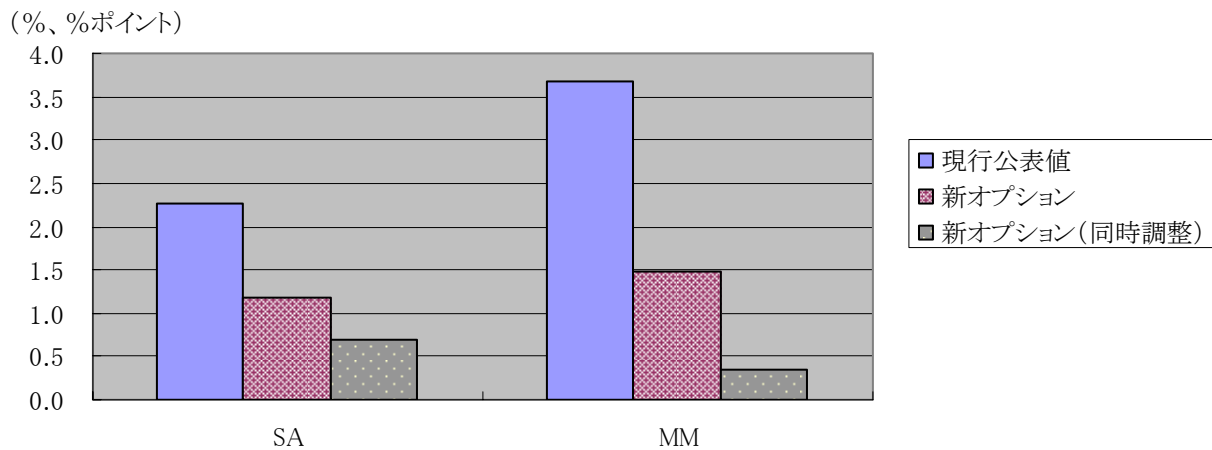


図表 2-4-9 遡及修正の繰り返し（2003年6月分の例）

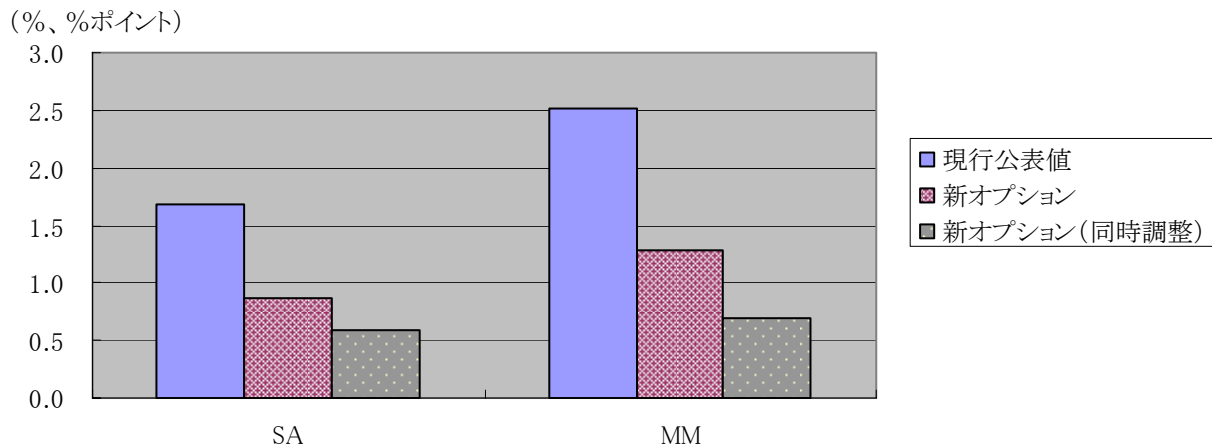


図表 2-4-10 新オプションによる第1回目の遡及修正幅

遡及改正幅(新規求職申込件数)

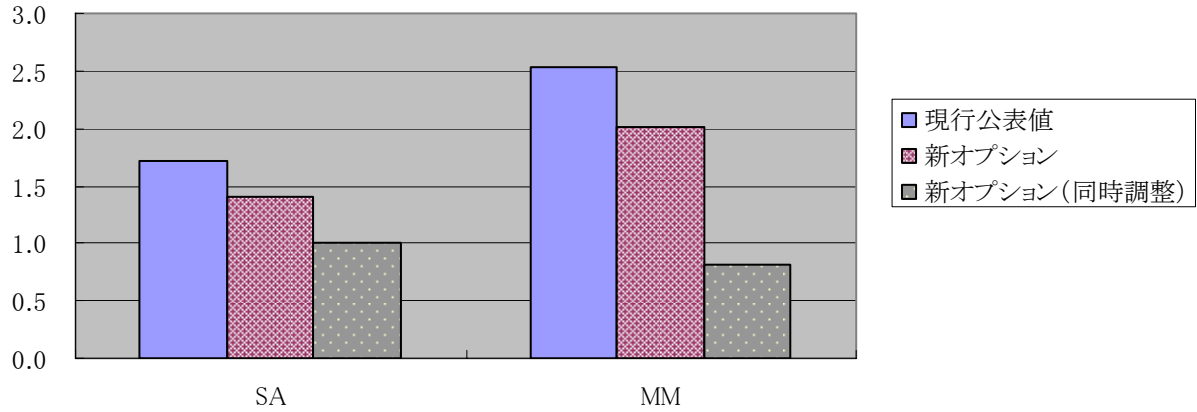


遡及改正幅(新規求人数)



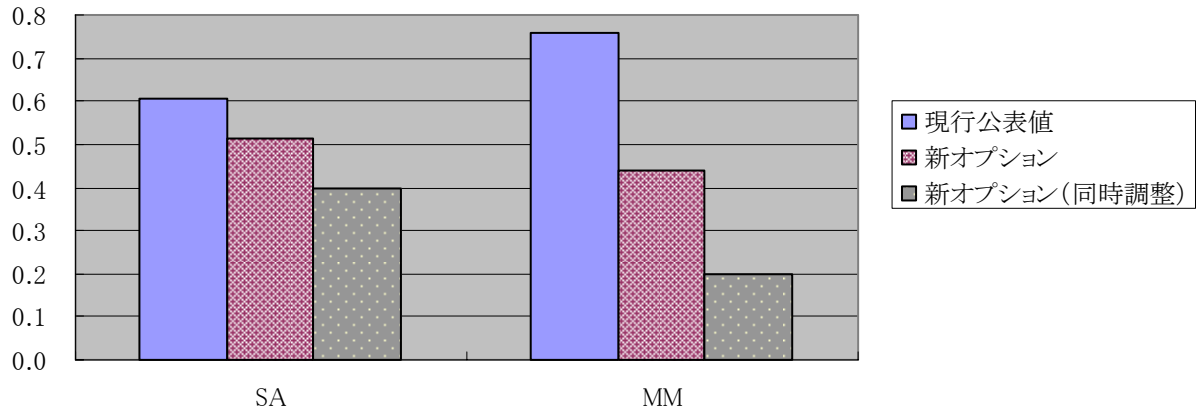
遡及改正幅(新規求人倍率)

(%、%ポイント)



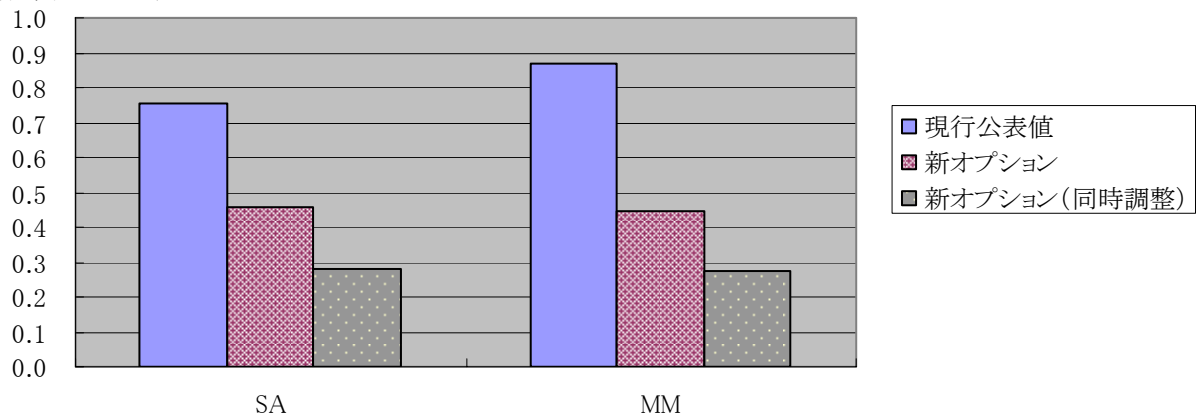
遡及改正幅(有効求職者数)

(%、%ポイント)

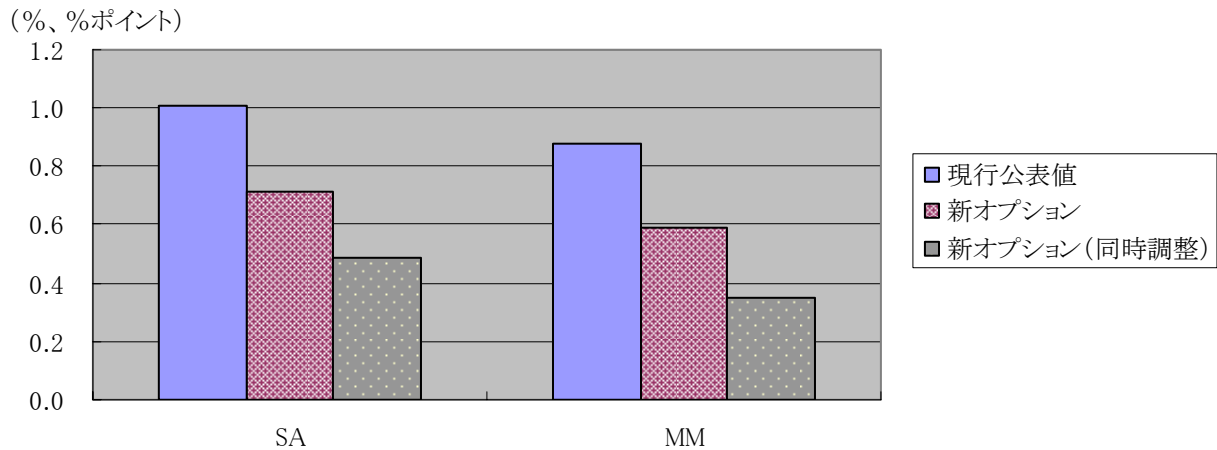


遡及改正幅(有効求人人数)

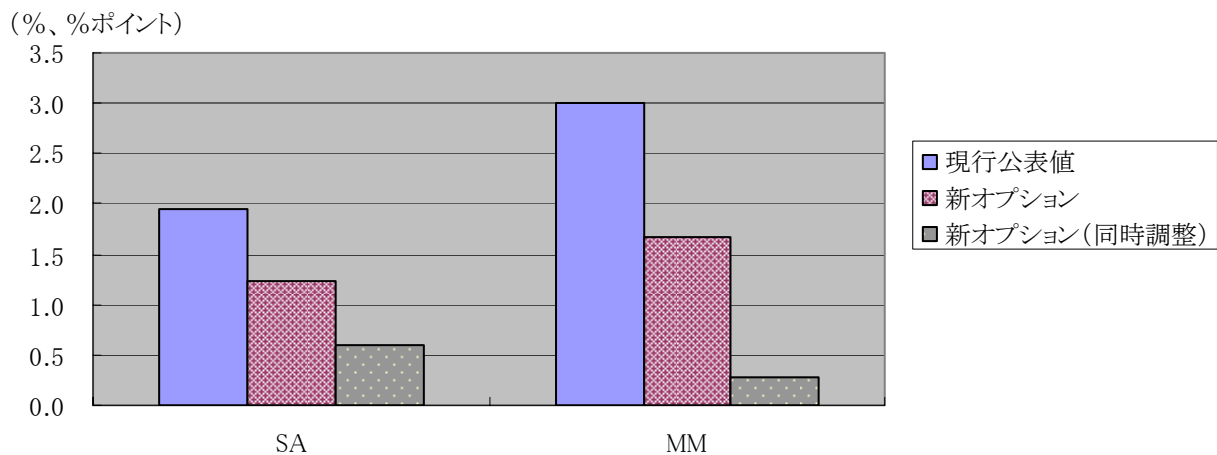
(%、%ポイント)



遡及改正幅(有効求人倍率)



遡及改正幅(就職件数)



(注)「現行公表値」と「新オプション」は、予測季節要素を使う現行の公表方式によるものである。1996年1月分発表時の前年12か月間(1995年1月分～12月分)の修正から、2005年1月分発表時の前年12か月間(2004年1月分～12月分)の修正まで、合計120か月(10年×12か月)について計測した。

「新オプション(同時調整)」は、同時調整(Concurrent Adjustment:毎月本当の季節調整を実行する)方式によるものである。上と条件を合わせるため、1995年12月分発表時の当年12か月間(1995年1月分～12月分)の修正から、2004年12月分発表時の当年12か月間(2004年1月分～12月分)の修正まで、合計120か月(10年×12か月)について計測した。

SA 当初の値から修正値への増減率の絶対値の平均(%)

MM 当初の前月比から修正前月比への差の絶対値の平均(%ポイント)