

JSPS Grants-in-Aid for Creative Scientific Research
Understanding Inflation Dynamics of the Japanese Economy
Working Paper Series No.37

大規模POSデータを用いた 価格変化の因子分析

外木 暁幸

May 8, 2009

Research Center for Price Dynamics
Institute of Economic Research, Hitotsubashi University
Naka 2-1, Kunitachi-city, Tokyo 186-8603, JAPAN
Tel/Fax: +81-42-580-9138
E-mail: sousei-sec@ier.hit-u.ac.jp
<http://www.ier.hit-u.ac.jp/~ifd/>

大規模POSデータを用いた 価格変化の因子分析*

外木 暁幸[†]

2009年5月6日

概要

本稿では大規模な小売スキャナー・データを用いて価格変動の因子分析を行った。因子モデルによって価格変化を共通因子と独自因子に分解し、価格変動が店舗特有の要因、あるいはマクロレベルの要因にどの程度影響を受けていたかを検討した。分析の結果、1990年代末から2000年代半ばにかけては個別の価格系列の独自性が上昇する傾向が見られた。一方、2000年代後半には共通因子の価格分散への寄与率が高まる傾向が顕著となった。特売の値下げ率についても同様な傾向が観察された。

JEL Classification : E31, C31, C32, C33

Keywords : 価格変化, フィリップス曲線, 因子分析, 分散分解

*本論文の作成にあたり、阿部修人先生（一橋大学）から多くの指導を受けた。また、渡辺努先生（一橋大学）、塩路悦朗先生（一橋大学）、祝迫得夫先生（財務総合政策研究所）から有益なコメントを頂いた。さらに、第3回若手経済学者のためのコンファレンスにおける報告に際して討論者の敦賀貴之先生（関西大学）から、一橋大学物価研究センターの研究会では白塚重典先生（日本銀行）、水野貴之先生（一橋大学）からも有益なコメントを頂いた。記して感謝を表したい。ただし、本文中に含まれる誤りは全て著者の責任であることは言うまでもない。また、日本経済新聞デジタルメディア社には貴重なデータセットを使用する機会を頂いたことに深く感謝する。なお、この論文は日本学術振興会の科学研究費学術創成プロジェクト「日本経済の物価変動ダイナミクスの解明」の成果の一つである。

[†]一橋大学大学院経済学研究科博士課程

1 はじめに

日本のフィリップス曲線 (Phillips Curve) は 1990 年代以降, それ以前の時代に比較してフラット化したとされている。この間, 日本の価格のマイクロデータの実証研究では価格改訂確率が上昇したことが報告されており, 価格改定確率の上昇がフィリップス曲線の傾きの上昇となって現れるニュー・ケインジアン・フィリップス曲線 (New Keynesian Phillips Curve : NKPC) の含意とは矛盾する結果となっている。もう 1 つのフィリップス曲線の理論としては Lucas (1972) の不完全情報モデルがある。この理論ではフィリップ曲線の傾きを決めるのは価格改定確率ではなく価格変動の分散に対する個別ショックの分散の比率であり, この個別ショックの分散寄与率が高いほどフィリップス曲線の傾きはフラットになる。日本におけるフィリップス曲線のフラット化はニュー・ケインジアンのメカニズムではなく, ルーカス型の不完全情報のメカニズムで起こっているのではないかという問題意識がこの研究の動機である。

ニュー・ケインジアン・モデルにおいては価格仮定頻度は NKPC の傾きを決定する重要なパラメーターである。最近では, Bils and Klenow (2004), Kehoe and Midrigan (2007), Nakamura and Steinsson (2007) など小売価格データに基づいて推計された価格改定確率を用いてニュー・ケインジアン・モデルにおける金融政策の実質効果を検証する多くの研究がなされてきた。しかし, いずれも月次や週次の小売データからの推計であり, 真の価格改定頻度を推計しているか疑問がある。日本の小売物価調査を用いて価格改定確率を推定した最近の論文としては Saita and Higo (2007) と Abe and Tonogi (2008) がある。Saita and Higo (2007) では月次の小売物価統計を用いて価格改定確率が推計され, 1990 年代から 2000 年代前半にかけて価格改定確率の上昇が報告された。Abe and Tonogi (2008) では日次

のスクャナー・データ（日経 POS データ）を用いて価格改定確率を推計しているが、既存研究のそれより著しく高く推計されている。POS データより推計した日次価格改定確率の推定値の推移を図 1 の（A）に示しておく。1980 年代から 1990 年代半ばにかけては 15-17 % 程であった価格改定確率は 1990 年代後半以降急激に上昇し、2000 年代半ば以降は約 30 % で横ばいとなっている。店舗プールデータの週次の最頻値価格を定価と定義して推計した日次の定価改定確率でも、水準は 2 ~ 3 % と低くなるものの、通時的な変動はほぼ同様の動きを示した（図 1（B））。図 2 には同じ時期の日経 POS データから計算したインフレ率と、総務省の消費者物価指数（CPI）の詳細品目価格うち、POS データがカバーする加工食品と家事用品に分類される品目の価格指数のみを抜き出して加重平均した物価のインフレ率を示した。1990 年代後半以降、いずれの POS 物価も CPI もデフレーションの傾向を示すようになり、2000 年代後半になって徐々にインフレーションに転じ、2008 年に入るとインフレ率は急激に上昇している。

この研究の目的は因子モデルによって日経 POS データの個別商品の名目価格を共通因子 (common factor) と独自因子 (idiosyncratic factor) に分解し、1990 年代末から最近までの共通因子の分散寄与率の推移を見ることで、デフレーションが進行し、フィリップス曲線がフラット化した時期に、価格変動のに占める独自因子の分散寄与率の上昇（つまり、共通因子の分散寄与率の低下）があったかどうかを検証することにある。

本稿では次のように議論を進める。第 2 節ではフィリップス曲線のフラット化とその背景にある二つの理論について議論する。第 3 節では日経 POS データの解説およびデータ加工の過程を議論する。第 4 節では価格変動を分散分解する目的で 2 つの実証分析が行われる。1 つは価格変動の固定効果モデルによる時間ダミーへの回帰分析である。時間ダミーとして抽出されたマクロ・ショックの決定係数が

1999 年代末から 2000 年代半ばにかけて低下し、その後上昇していることが明らかになる。もう一つは近似的因子モデルによる分散分析である。複数の直交した共通因子を抽出して共通因子の分散寄与率を推計した結果、1999 年代末から 2000 年代半ばにかけて低下し、その後上昇していることが示される。第 5 節では本稿の分析の結論を述べる。

2 フィリップス曲線の理論

1990 年代から 2000 年代中盤にかけて、価格改定確率が上昇したことを図 1 に示したが、この間、日本のフィリップス曲線はフラット化が進行した。木村・黒住・原（2008）では消費者物価指数、及び国内需要デフレーターを GDP ギャップにローリング回帰してフィリップス曲線の傾きの変化を検証しているが、同様の結論を得ている¹。

今日、フィリップス曲線についての代表的な理論であるニュー・ケインジアン・モデルにおいては、独占的競争を行う企業の価格設定に Calvo 型の価格改定を想定している。この理論では価格改訂確率の上昇はフィリップス曲線の傾きを上昇させることになる。 π_t を t 期のインフレ率、 x_t を GDP ギャップ、 β を時間割引要因、 ρ を価格改訂確率とすると、ニュー・ケインジアン・フィリップス・カーブ（NKPC）は次のように書ける。

$$\pi_t = A(\rho)x_t + \beta\pi_{t+1}, \quad \left(A(\rho) > 0, \frac{\partial A(\rho)}{\partial \rho} > 0 \right)$$

Rotemberg 型の価格改定を想定したとしても、トレンドのインフレ率が一定の

¹ただし、NKPC の構造推計を行ったわけではなく、誘導型のフィリップス曲線のフラット化について議論している。

下で価格改定コストの低下は価格改訂確率の上昇となって表れるため価格改訂確率の上昇とフィリップス曲線の傾き低下が起こっている日本の状況とは整合的ではない。

一方、ルーカス型の不完全情報モデルにおいては、フィリップス曲線の傾きはマクロの物価変動の分散と個別の商品の相対価格変動の分散の比率によって決定される。このとき、個別の商品の相対価格変動の分散が価格の全変動に占める割合が大きくなれば、一般物価の上昇を個別の財の需要増と錯覚し難くなるため、ルーカス型総供給曲線の傾きは大きくなる。つまり、フィリップス曲線の傾きは小さくなる。一般物価の分散を $Var(p)$ 、個別の商品価格の独自要因の分散を $Var(\varepsilon)$ とすると、ルーカス総供給関数は次のように書ける。

$$x_t = B(v) [\pi_t - E(\pi_t)], \left(v = \frac{Var(\varepsilon)}{Var(p) + Var(\varepsilon)}, B(v) > 0, \frac{\partial B(v)}{\partial v} > 0 \right)$$

もし、1990年代末から2000年代半ばのデフレーション期の日本で個別の財の価格変動に占める独自ショック（個別の需要ショック）の割合が共通ショック（一般物価）に比して上昇していることが観察されるなら、フィリップス曲線のフラット化の理論的解説としてはルーカス型の不完全情報モデルの方が適切である可能性が高くなる。本稿ではこの点に注目して、個別価格の変動を因子モデルにより共通因子と独自因子に分解し、その分散分解を試みるという実証分析を行う。

3 データ

本稿では価格変動についての実証分析に日経デジタルメディア社が提供する日経 POS データを用いる。日経 POS データは 1988 年から現在までの全国 373 店舗のスーパー、GMS といった小売店で販売された加工食品、日用品の JAN コード、販売額、販売数量を日次で記録した大規模データセットである。記録された商品点数は約 150 万、販売記録の観測総数は約 40 億点にのぼる。店舗には新規出店と退出があるため、この研究では 1998 年 1 月 1 日から 2008 年 4 月 30 日までの 3772 日間、常に存在した 84 店舗のデータのみを分析の対象とする。また、価格データの欠損を少なくするため販売日数が 3000 日を超え、なおかつ 10 店舗以上で販売されていた 657 商品に対象を絞る。84 店舗で販売された 657 商品の店舗・商品の結合 ID の数は 16383 のに上る (表 1)。この 16384 の結合 ID を振られた商品の販売額で、当該 84 店舗の 11 年 4 ヶ月間の売上の 9.18 % を占めている。また、その平均販売日数は 3370 日と全販売日数の 89.4 %、販売レンジ (販売初日から販売最終日までの日数) は 3714 日と 98.8 % とかなり高いカバー率となっている。

3.1 データ処理

3.1.1 データの月次化

日経 POS データは日次で商品ごとの販売金額と販売数量を記録したデータ・セットであるが、本稿では日次データを月次データに加工して分析を進める。価格を月次化するには、商品の実売価格を表すものとして販売数量加重平均価格、また商品の定価を表すものとして月間最頻値価格を、そして両者の価格比を表すバーゲン価格率の三つの価格データ・セットを作成する。

先ず、月次の販売数量加重平均価格の計算について説明する。日付 td において、店舗 s で販売された商品 i の販売金額を $S_{td}^{s,i}$ とし、その販売数量を $Q_{td}^{s,i}$ とすると、

日付 td に店舗 s で販売された商品 i の価格 $P_{td}^{s,i}$ は以下の通り .

$$P_{s,i,td} = \frac{S_{td}^{s,i}}{Q_{td}^{s,i}}.$$

日付 td , がある月 tm に属すると仮定すると , 月 tm において , 店舗 s で販売された商品 i の販売数量加重平均価格 $P_{s,i,tm}^{(a)}$ は以下の通り .

$$P_{s,i,tm}^{(a)} = \sum_{td \in tm} \left[\left(\frac{Q_{td}^{s,i}}{\sum_{td \in tm} Q_{td}^{s,i}} \right) P_{s,i,td} \right].$$

欠損値がある場合 , その月の属する年の平均価格 (月次販売数量加重平均価格の平均値) で埋めている . それでも欠損値が残る場合 , 3 年間の平均価格 (月次販売数量加重平均価格の平均値) で埋めている . 最後に , 前年同月比の伸び率 $p_{s,i,tm}^{(a)}$ に変換して実証分析に用いる .

$$p_{s,i,tm}^{(a)} = \left[\frac{P_{s,i,tm}^{(a)} - P_{s,i,tm-12}^{(a)}}{P_{s,i,tm-12}^{(a)}} \right] \times 100.$$

次に , 月次の最頻値価格の定義について説明する . 月 tm において , 店舗 s で販売された商品 i の最頻値価格 $P_{s,i,tm}^{(m)}$ は以下の通り .

$$P_{s,i,tm}^{(m)} = \text{mode}_{td \in tm}(P_{s,i,td})$$

ただし , $\text{mode}_{td \in tm}$ は月 tm に属する日付 td のうち , 最頻値の価格 (最頻値が二つ以上ある場合はそのうちの最大値) を返すオペレータである .

欠損値がある場合 , その月の属する年の最頻値価格 (月次最頻値価格の年間の最頻値) で埋めている . それでも欠損値が残る場合 , 3 年間の最頻値価格 (月次最頻値価格の 3 年間の最頻値) で埋めている . 最後に , 前年同月比のデータに変換

して実証分析に用いる。

$$p_{s,i,tm}^{(m)} = \left[\frac{P_{s,i,tm}^{(m)} - P_{s,i,tm-12}^{(m)}}{P_{s,i,tm-12}^{(m)}} \right] \times 100.$$

次に、月次のバーゲン価格率の定義について説明する。最頻値価格 $P_{s,i,tm}^{(m)}$ を定価とし、販売数量加重平均価格 $P_{s,i,tm}^{(a)}$ が実売価格であるとする、両者の差はバーゲンによる値下げ幅に相当する²。それを定価で割ったものを、バーゲン価格率と定義する。つまり、月次のバーゲン価格率 $p_{s,i,tm}^{(b)}$ は以下の通り。

$$p_{s,i,tm}^{(b)} = \left[\frac{P_{s,i,tm}^{(m)} - P_{s,i,tm}^{(a)}}{P_{s,i,tm}^{(m)}} \right] \times 100.$$

注記しなくてはならないのは、日経 POS データは 2003 年 11 月、12 月の記録が欠損していることである。実証分析はこの部分のデータは取り除いた上で行われている。価格の前年同月比伸び率を示す $p_{s,i,tm}^{(a)}$ のデータ、及び $p_{s,i,tm}^{(m)}$ のデータについては 2003 年 11 月、12 月に加えて、2004 年 11 月、12 月のデータを除いた上で実証分析を行う。

3.1.2 季節調整

$p_{s,i,tm}^{(a)}$ のデータ、及び $p_{s,i,tm}^{(m)}$ のデータについては前年同月比をとることで乗法的な意味での季節性を取り除けているはずであるが、 $p_{s,i,tm}^{(b)}$ のデータについては季節性が残っている。これについてはデータを 12 か月の月ダミーに回帰し、その残差を実証分析に用いることで季節性の影響を回避する。

²実際には、月次最頻値価格を月次販売数量加重平均価格が上回っているということもありうる。しかし、このような事例はごく少数である。

4 実証分析

4.1 パネルデータ分析

先ず、価格のパネルデータを時間ダミーに回帰して、マクロ・ショックの価格変動への影響度合いについておおまかな傾向をみる。個別価格（店舗と商品の結合 ID (s_i) を振った個別価格）を店舗・商品の結合 ID ダミー (D_{s_i}) と、月ダミー (D_{tm}) に回帰する固定効果モデルで、月ダミーの動きと、決定係数の時間を通じた変化を検討する。回帰モデルは次の通り。

$$\begin{aligned} p_{s_i,tm}^{(a)} &= \sum_{s_i} D_{s_i} \alpha_{s_i}^{(a)} + D_{tm} \beta_{tm}^{(a)} + \varepsilon_{s_i,tm}^{(a)}, \\ p_{s_i,tm}^{(m)} &= \sum_{s_i} D_{s_i} \alpha_{s_i}^{(m)} + D_{tm} \beta_{tm}^{(m)} + \varepsilon_{s_i,tm}^{(m)}, \\ p_{s_i,tm}^{(b)} &= \sum_{s_i} D_{s_i} \alpha_{s_i}^{(b)} + D_{tm} \beta_{tm}^{(b)} + \varepsilon_{s_i,tm}^{(b)}. \end{aligned}$$

図3は時間ダミーの係数（基準時の時間ダミー係数をゼロとし、全て定数項を足してある）を、販売数量加重平均価格、最頻値価格、バーゲン価格率のそれぞれについてプロットしたものである。販売数量加重平均価格、最頻値価格についてはほぼ近い動きをしている。1999年代末から2000年代半ばにかけて緩やかにインフレ率の低下が進行し、係数もほぼ0から-1の間を動いている。ところが、2000年代半ば以降はインフレ率が上昇に転じ、2008年には急上昇している。図2のPOS物価及び加工食品・家事用品のCPIのインフレ率の1990年代末以降の推移と比較すると、よく似た動きをしていることが見て取れよう。次に、バーゲン価格率（実売価格が低下からどれだけ低いかを示す）の時間ダミーをみると、1990年代末には低かった係数が、2000年代半ばにかけて少しずつ上昇していることがわか

る。これは定価と実売価格に次第に開きが出てきたことを意味している。しかるに、2007年から2008年の価格上昇時には係数が低下しており、この時期、定価と実売価格の差が全体的に縮小したことが示されている。

次に、3年毎に期間を区切って上記の固定効果モデルを推計し、決定係数の推移をみることで時間ダミーの影響が変化したかを検証しよう。決定係数の変遷をまとめたものが表2及び図4である。表2をみると販売数量加重平均価格と最頻値価格は1999年から2001年のサンプルで高かった決定係数が2000年代半ばにかけていったん低下している。その後、2004年から2006年のサンプルにおいて急上昇し、2006年から2008年のサンプルでは桁違いに大きくなっている。これはタイム・ダミーで捕捉されたマクロ・ショックの個別価格に対する影響の度合いが1990年代末から2000年代半ばにかけて低下し、2000年代後半にかけては上昇していることを示している。図4の(A)(B)のグラフではそれぞれ、販売数量加重平均価格と最頻値価格の固定効果モデルの推計における決定係数の推移を示した³。図4の(C)バーゲン価格率についてはマクロ・ショックの影響度合いの低下が2005年から2007年のサンプルまで継続したが、2006年から2008年のサンプルでは反転上昇している。

3つの価格の固定効果モデルによるマクロ・ショックの影響度合いの変化についての分析を総括すると、マクロ・ショックの個別価格に対する影響の度合いが1990年代末から2000年代半ばにかけて低下し、2000年代後半にかけては上昇していると総括できよう。ただし、ここでマクロ・ショックとして取り上げたのは時間ダミー1つのみであり、決定係数は極めて低い。マクロ・ショックとして2つ以上の要因が存在するとも考えられることから、以下ではこのようなマクロ要因を効率

³グラフでは2006年から2008年の決定係数は除いて表示している。この期間の決定係数は他の期間よりけた違いに大きいので、グラフに含めると2007年までの期間の決定係数の動きがみづらなるためである。

的に推計する手法である因子モデルを用いてマクロ・ショックと個別ショックの分散分解を試みる。

4.2 因子分析

前節では、価格を時間ダミーに回帰する固定効果モデルを用いてマクロ・ショックの影響の度合いの変化を調べたが、共通因子が時間ダミー1つのみで、データに対する説明力も極めて弱かった。固定効果モデルに時間ダミーのほかに様々なマクロ変数を加えて推計するという選択肢も存在するが、マクロ変数の選択は適切か、効率的にマクロ・ショックを抽出できているかといった問題をはらんでいる。ここでは共通因子同士に直交性を仮定することで効率的に共通因子を推定する手法である因子モデルを用いて価格変動の分散分解を試みる。

4.2.1 因子モデル

商品 i 、時点 t の価格を p_t^i とするとき、因子モデル (factor model) ではこれを q 個の共通因子 (common factor) $f_{t1}, f_{t2}, f_{t3}, \dots, f_{tq}$ と商品 i に固有な因子 (idiosyncratic factor) ε_t^i の線形結合で表わすことを仮定する。

$$p_t^i = \lambda_1^i f_{t1} + \lambda_2^i f_{t2} + \lambda_3^i f_{t3} + \dots + \lambda_q^i f_{tq} + \varepsilon_t^i, \\ (i = 1, 2, \dots, N; t = 1, 2, \dots, T).$$

本稿では p_t^i を先に定義した3種類の価格 ($p_{s,i,tm}^{(a)}, p_{s,i,tm}^{(m)}, p_{s,i,tm}^{(b)}$) として実証分析を行う。3種類の価格は全ての系列で平均ゼロ、分散1に基準化する。 λ_j^i は j 番目の

共通因子が i 番目の商品の価格に影響する程度を示したもので因子負荷量 (factor loading) という。 f_{tj} は j 番目の共通因子の時点 t での共通因子得点 (common factor score) という。 ε_t^i は i 番目の商品の価格の時点 t での独自因子得点 (idiosyncratic factor score) という。

$p_t^i (i = 1, 2, \dots, N)$ を要素とする $n \times 1$ ベクトルを \mathbf{p}_t とし、 $f_{tj} (j = 1, 2, \dots, q)$ を要素とする $q \times 1$ ベクトルを \mathbf{f}_t とし、 $\varepsilon_t^i (i = 1, 2, \dots, N)$ を要素とする $n \times 1$ ベクトルを $\boldsymbol{\varepsilon}_t$ とする。また、 $\lambda_j^i (i = 1, 2, \dots, n; j = 1, 2, \dots, q)$ を要素とする $N \times q$ 行列を Λ とする。このとき因子モデルは次のようにベクトル表現に書き換えることができる。

$$\mathbf{p}_t = \Lambda \mathbf{f}_t + \boldsymbol{\varepsilon}_t.$$

モデルには次のような 4 つの仮定を置く。一つ目の仮定は共通因子と独自因子は無相関というものである。

$$\text{Cov}(\mathbf{f}_t, \boldsymbol{\varepsilon}_t) = E(\mathbf{f}_t \boldsymbol{\varepsilon}_t') = \mathbf{0}.$$

二つ目の仮定は異なる二つの価格に対応する独自因子は無相関というものである。

$$\text{Var}(\boldsymbol{\varepsilon}_t) = E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') \equiv \Psi,$$

$$\Psi = \begin{bmatrix} \psi_1 & 0 & \cdots & 0 \\ 0 & \psi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \psi_n \end{bmatrix}.$$

三つ目の仮定は共通因子及び独自因子の平均はゼロというものである。

$$E(\mathbf{f}_t) = \mathbf{0}, E(\boldsymbol{\varepsilon}_t) = \mathbf{0}.$$

四つ目の仮定は共通因子の分散は全て1で異なる二つの共通因子は無相関というものである。

$$E(\mathbf{f}_t \mathbf{f}_t') = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix} \equiv I_q.$$

以上のような仮定を持つ因子モデルに基づいて、価格の分散 $Var(\mathbf{p}_t) \equiv \Sigma$ は次のように分解できる。

$$\begin{aligned} \Sigma &= E(\mathbf{p}_t \mathbf{p}_t') = E[(\Lambda \mathbf{f}_t + \boldsymbol{\varepsilon}_t)(\Lambda \mathbf{f}_t + \boldsymbol{\varepsilon}_t)'] \\ &= E[\Lambda \mathbf{f}_t \mathbf{f}_t' \Lambda' + \Lambda \mathbf{f}_t \boldsymbol{\varepsilon}_t' + \boldsymbol{\varepsilon}_t \mathbf{f}_t' \Lambda' + \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] \\ &= \Lambda \Lambda' + \Psi. \end{aligned}$$

つまり、価格 p_t^i の分散は次のように分解されることになる。

$$\begin{aligned} Var(p_t^i) &= Var(\lambda_1^i f_{t1} + \lambda_2^i f_{t2} + \lambda_3^i f_{t3} + \cdots + \lambda_q^i f_{tq}) + Var(\varepsilon_t^i) \\ &= (\lambda_1^i)^2 + (\lambda_2^i)^2 + \cdots + (\lambda_q^i)^2 + \psi_i. \end{aligned}$$

全価格, $p_t^1, p_t^2, \dots, p_t^N$ において共通因子により説明される部分の寄与は, 以下の通り.

$$\text{trace}(\Sigma - \Psi) = \text{trace}(\Lambda\Lambda').$$

全価格における共通因子により説明される部分の寄与率 (CR) は, 次のように計算される.

$$CR = \frac{\text{trace}(\Lambda\Lambda')}{\text{trace}(\Sigma)}.$$

本稿では因子得点及び因子負荷量については通時で推計し, 分散分解は3年の期間 ($period$) ごとに行う. 従って, 次の方法で共通因子の分散寄与率を計算する.

$$CR_{period} = 1 - \frac{\sum_{i=1}^N \text{Var}_{t \in period}^i(\varepsilon_t^i)}{\sum_{i=1}^N \text{Var}_{t \in peirod}^i(p_t^i)}.$$

ここで $\text{Var}_{t \in period}^i$ は, 商品 i について, 期間 $period$ に属する価格 (あるいは独自因子) について分散を計算するオペレータである.

ここで用いるサンプルは $N > T$ で古典的な因子モデルの推定法は利用できない. 従って近年推定理論が発展した近似的因子モデルを用いる. 推定方法については Bai (2003) の手法を用いた⁴.

4.2.2 因子数の決定

マクロの物価を多数の商品の価格から抽出したいと考えるとき, 抽出すべき因子数は1個で十分とも考えられるが, 小売系列が同じ店舗が同じ価格戦略をとる,

⁴近似的因子モデルの推定法については Appendix で解説する.

あるいは地域ごとに異なった共通要因がある，原材料の構成が近い等といった要因も共通因子に含めなければ，店舗独自の要因を抽出できない可能性がある．1 因子のケースは時間ダミーを説明変数とした固定効果モデルの推計で扱ったので，ここでは複数因子のケースを検討する．また，因子モデルの手法は複数の共通因子を効率的に抽出するのに適している．

共通因子数 q の決定については，標本相関行列の 1 より大きい固有値の数を共通因子の数とするという基準を用いる．この基準は分散を 1 に揃えた非説明変数よりも大きな分散を持つ因子を共通因子として採用し，それよりも小さな分散しか持たない因子を不採用とするという明確な基準を持つ．

4.2.3 結果

先ず，データから因子得点を推定する．図 5 は，販売数量加重平均価格，再頻値価格，及びバーゲン価格率のそれぞれについて近似的因子モデルをもちいて推定した因子得点のうち分散の大きいもの上位 5 つを取り上げてグラフにしたものである．図 5 の (A) (B) の第 1 因子をみると，図 3 の販売数量加重平均価格，再頻値価格の時間ダミー係数のグラフに極めて似通っていることが分かる．図 5 の (C) の第 1 因子についても図 3 のバーゲン価格率の時間ダミー係数のグラフと酷似しており，前節のパネルデータ分析でみた時間ダミーによる共通因子が因子モデルの第 1 因子をほぼとらえていたことが確認できる．第 2 因子以下をみると図 5 の (A) (B) では景気循環の動きに相当するような因子もみられる（第 4 因子）．第 2 因子以下の因子は，原材料価格（穀物，原油等）の動きや消費財の在庫率の状況，卸売物価といった様々な要因を反映していると考えられるが，ここではそれぞれの因子がどのようなマクロ変数に対応するのかを調べることは主題ではない．価格変動をマクロ・ショックと個別ショックに分解して分散寄与率の推移をみ

ることが目的である。

次に、共通因子として採用する因子の数を決定する。図 6 は販売数量加重平均価格、再頻値価格、及びバーゲン価格率のそれぞれについて、相関行列の固有値を降順に並べたものである。販売数量加重平均価格の固有値は 37 個目までが 1 以上、再頻値価格は 33 個目まで、バーゲン価格率は 35 個目までが 1 以上となっている。ここでは先にあげた因子数決定の基準に基づき、表 3 のように因子数 q を決定する。図 7 は販売数量加重平均価格、再頻値価格、及びバーゲン価格率のそれぞれの SBIC (シュワルツ・ベイズ情報量基準) を参考として計算したものだ。販売数量加重平均価格と再頻値価格では因子数 37 の時に SBIC が最小になっており、上記の固有値基準での因子数と近い結果となった。バーゲン価格率は因子数 3 で SBIC が最小になっており固有値基準の因子数とかけ離れた結果となった。

最後に、共通因子の分散寄与率を期間を分けて推定する。図 8 はその結果を示したものである。頑健性をみるために、採用した因子数より 1 及び 2 多く因子数を採用した場合と 1 及び 2 少なく因子数を採用した場合の分散寄与率も掲載した。いずれの因子数でも、共通因子の分散寄与率の動きは、基準となる因子数による分散寄与率の動きととほぼ同じ動きとなっている。図 8 の (A) (B) をみると 2000 年代を通じて低下してきた共通因子の分散寄与率が 2005 年から 2007 年の期間を底にして 2006 年から 2008 年の期間では上昇していることがみてとれる。ここでも 2000 年代半ば頃まではマクロ・ショックの影響度合いが低下し、最近になって上昇しているという結果が見出された。図 8 の (C) のバーゲン価格率のグラフではさらにそれが鮮明であり、2000 年代半ばを底にして共通因子の分散寄与率は V 字型となっている。デフレが進行していた 2000 年代半ばまでの日本では価格変動に占める個別ショックの影響力が増大し、その後のインフレ局面ではマクロ・ショックの影響が増大したということができよう。

5 結語

この研究の目的は因子モデルによって個別商品の名目価格にを共通因子 (common factor) と独自因子 (idiosyncratic factor) に分解し, 1990 年代末から直近までの共通因子の寄与率の推移を見ることで, この時期の価格変動に占めるマクロ・ショック, 個別ショックそれぞれの影響度合いの変化をみることにあった. 我々は先ず, 価格データを時間ダミーに回帰する固定効果モデルを推計し, その決定係数が 2000 年代半ばにかけて低下, その後は上昇するという結果を得た. 更に近似的因子モデルを用いて価格変動の共通因子を抽出することを試み, その第 1 因子が時間ダミー係数とほぼ近いものであることを確認した. 適切な因子数を採択した因子モデルによる分散分解では共通因子の因子寄与率が 2000 年代半ばにかけて低下し, その後上昇するという結果が再び得られた. デフレーションが進行していた 2000 年代半ばまでの日本では価格変動に占める個別ショックの影響力が増大し, その後のインフレーション局面ではマクロ・ショックの影響が増大したといえることができる.

2000 年代前半において物価変動が非常に小さく, 緩やかなデフレーションが進行する一方で, 価格改定確率が上昇していったことの背景にはマクロの物価変動ではなく, 個別店舗の独自の価格設定行動があることが示唆されよう. 2000 年代後半には共通因子の分散寄与率は上昇しており, 日本経済がデフレーションから脱却する過程においては, 個別の商品, 店舗の価格設定行動の影響力は低下してもマクロ・ショックの影響力が強まったことが示された. この事実を 1990 年代末から 2000 年代前半におけるフィリップス曲線のフラット化と 2000 年代半ば以降の再スティープ化と考え合わせれば, フィリップス曲線の理論的解説としてはルーカス型の不完全情報モデルの方が適切である可能性が高いといえよう.

Appendix : 近似的因子モデルの推定

$N > T$ の時, 価格系列の標本分散共分散行列 $S = \left(\sum_{t=1}^T \mathbf{p}_t \mathbf{p}'_t \right) / T$ がフルランクとならない. この場合, Bai and Serena(2002), Bai (2003) で展開された近似的因子モデルの推計手法を用いる. k 因子モデルを次のように表現する.

$$P = F_k \Lambda'_k + E,$$

ここで,

$$\begin{aligned} P_{(T \times N)} &= \begin{bmatrix} \mathbf{p}'_1 \\ \mathbf{p}'_2 \\ \vdots \\ \mathbf{p}'_T \end{bmatrix}, \quad F_k_{(T \times k)} = \begin{bmatrix} \mathbf{f}'_1 \\ \mathbf{f}'_2 \\ \vdots \\ \mathbf{f}'_T \end{bmatrix}, \quad E_{(T \times N)} = \begin{bmatrix} \boldsymbol{\varepsilon}'_1 \\ \boldsymbol{\varepsilon}'_2 \\ \vdots \\ \boldsymbol{\varepsilon}'_T \end{bmatrix}, \\ \Lambda_k_{(N \times k)} &= \begin{bmatrix} \lambda_1^1 & \lambda_2^1 & \cdots & \lambda_k^1 \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_k^2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^N & \lambda_2^N & \cdots & \lambda_k^N \end{bmatrix}. \end{aligned}$$

k 因子モデルの推定は次の最適化問題を解くことで行われる.

$$\begin{aligned} V(k) &= \min_{\Lambda^k, F^k} \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \left(p_t^i - \sum_{j=1}^k \lambda_j^i f_{tj} \right) \quad (\text{A.1}) \\ &\text{subject to } (F'_k F_k) / T = I_k, \quad \Lambda'_k \Lambda_k \text{ は対角行列} \end{aligned}$$

である．このとき， $\Lambda'_k = (F'_k F_k)^{-1} F'_k P = F'_k P / T$ とすると，(A.1) は次の主成分問題に等しい．

$$\max_{F_k} \text{trace}(F_k (PP') F'_k)$$

この最大値は，行列 PP' の固有ベクトルのうち固有値の大きい k 個の固有ベクトルに \sqrt{T} を掛けたものを F_k とすることで得られる．この方法で推定された共通因子を \tilde{F}_k とすると，因子負荷量のベクトルの推定値 $\tilde{\Lambda}'_k = (\tilde{F}'_k \tilde{F}_k)^{-1} \tilde{F}'_k P = \tilde{F}'_k P / T$ として推計できる．

参考文献

- Abe, N., and A. Tonogi (2008): Micro and Macro Price Dynamics over Twenty Years in Japan – A Large Scale Study Using Daily Scanner Data . *RCPD Working Paper Series*, No.18.
- Bai, J. (2003):Inferential Theory for Factor Models of Large Dimensions. *Econometrica*, 71, 135-171.
- Bai, J., and Serena Ng (2002): Determine the Number of Factors in Approximate Factor Models (with Serena Ng), *Econometrica* 70, 191-221.
- Bai, J., and Serena Ng (2002): *Large Dimensional Factor Analysis*, now Publishersn Inc.
- Bils, M., and P. J. Klenow (2004): Some Evidence on the Importance of Sticky Prices, *Journal of Political Economy*, 112(5), 947-985.
- Kehoe, P., and V. Midrigan (2007): Sales, Clustering of Price Changes, and the Real Effects of Monetary Policy, *Working Paper, University of Minnesota*.
- Lucas, Robert (1972) : Expectations and the Neutrality of Money, *Journal of Economic Theory*, 4(2): 103-124
- Nakamura, E., and J. Steinsson (2007): Five Facts about Prices: A Reevaluation of Menu Cost Model, *Working Paper, Harvard University*.
- Saita, Y., and M. Higo (2006): Price Setting in Japan: Evidence from CPI Micro Data, *Bank of Japan Working Paper Series*, No. 07-E-20.

- Shintani, M (2008): A Dynamic Factor Approach to Nonlinear Stability Analysis, *Journal of Economic Dynamics and Control*, 32(9), 2788-2808.
- Shiratsuka, S., (1999): *Measurement Errors in the Japanese: Consumer Price Index, Monetary and Economic Studies*.
- Sonoda, K (2006): An Empirical Analysis of Price Stickiness and Price Revision Behavior in Japan Using Micro CPI Data, *Bank of Japan Working Paper Series*, No. 06-E-8.
- Stock, J. H. and M. W. Watson (1998): Diffusion Indexes. *NBER Working Paper*, No. 6702.
- 阿部修人, 外木暁幸, 渡辺努 (2008): 「企業出荷価格の粘着性-アンケートとPOS データに基づく分析-」, *経済研究*, 59(4), 305-316.
- 木村武, 黒住卓司, 原尚子 (2008): 「日本のフィリップス曲線に何が起こったか - 企業の価格設定行動の変化と名目硬直性の高まり - 」, *Bank of Japan Working Paper Series*, No. 08-J -1.

表 1

Data Sampling from Nkkei POS

	Total	Items sold for over 3000 days	
			Items sold at over 10 stores
number of Items	703908	4223	657
number of stores	84	84	84
number of item-store ID	7863919	26358	16383
number of makers	15238	1079	182
number of the 6 classes	1681	720	289
number of the 3 classes	215	172	124
average range of sales	583	3,714	3,726
average day of sales	157.8	3,346	3,370
sum of sales(100 million Yen)	13,759	2,015	1,264
share of sales	100	14.64	9.18

Note: Total means sum of 84 stores for period from Jan 1, 1998 to Apr 31, 2008.

表 2

R-square of Fixed Effect Regressions to Time Dummies

	1999-2001	2000-2002	2001-2003	2002-2004	2003-2005	2004-2006	2005-2007	2006-2008
Weighted Average Price								
within	0.0020	0.0010	0.0011	0.0014	0.0014	0.0029	0.0031	0.0204
between	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
overall	0.0019	0.0010	0.0011	0.0013	0.0013	0.0027	0.0029	0.0187
Mode Price								
within	0.0012	0.0007	0.0007	0.0009	0.0008	0.0023	0.0027	0.0248
between	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
overall	0.0012	0.0007	0.0006	0.0009	0.0008	0.0021	0.0025	0.0227
Bargain Sale Ratio								
within	0.0034	0.0028	0.0026	0.0026	0.0031	0.0017	0.0016	0.0027
between	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
overall	0.0030	0.0025	0.0023	0.0024	0.0028	0.0015	0.0014	0.0023

The weighted average price variable is weighted average price of daily prices in a month by sales quantity weights. The mode price variable is mode price of prices in a month. The bargain sale ratio variable is difference between log of the mode price variable and log of the weighted average price variable. The weighted average price variable and the mode price variable are year on year change rate variables. This R-squares are calculated by fixed effect regression that each variable is regressed to store-item dummies and time dummies.

表 3

Determining Number of Factors

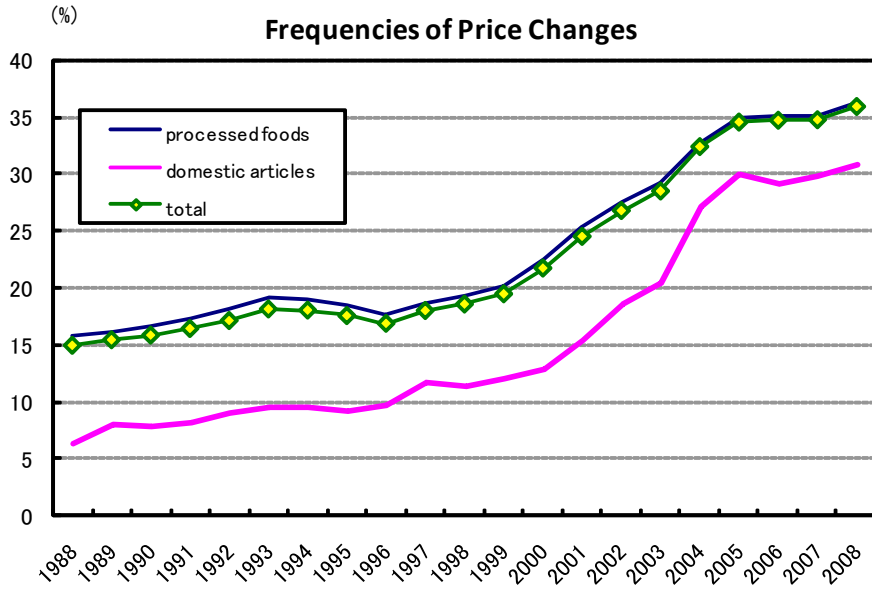
Weighted Average Price	37
Mode Price	33
Bargain Sale Ratio	35

We adapt the factor score as common factors by the criterion that correspondig eigenvalue is 1 or over.

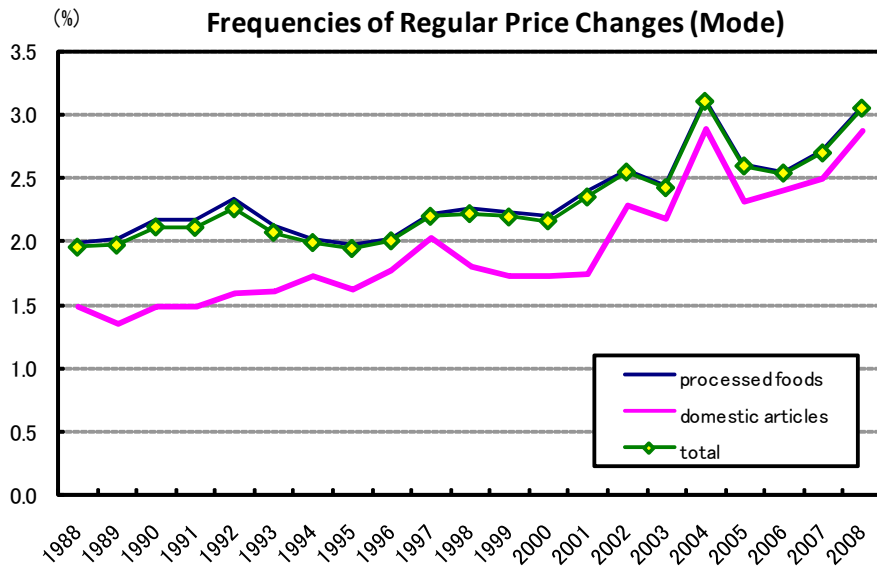
☒ 1

Price Change Probabilities

(A)

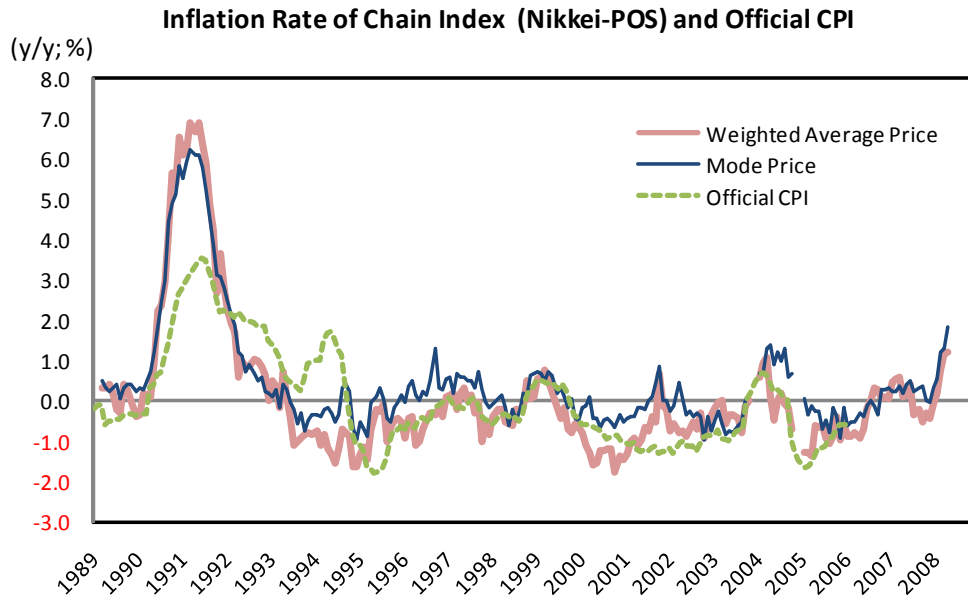


(B)



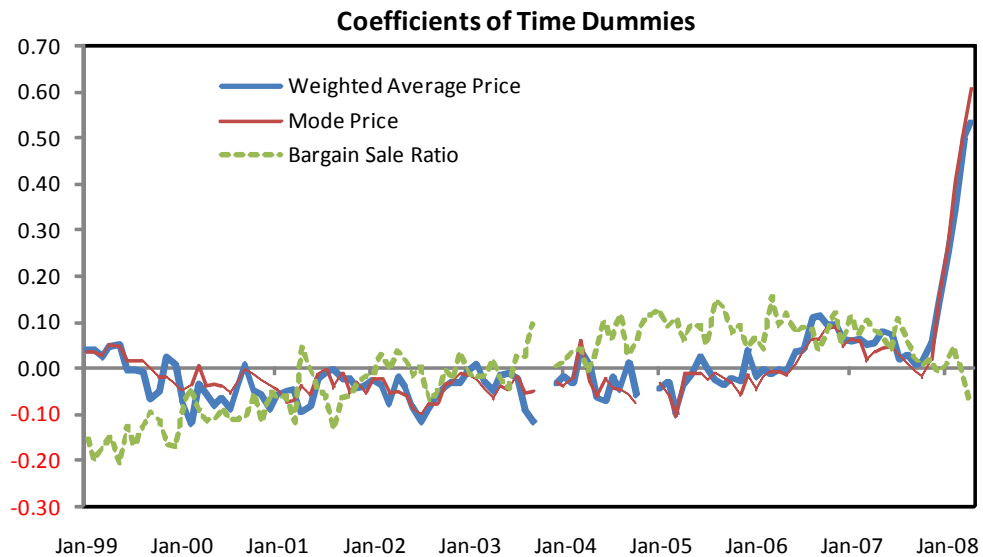
The frequencies of price changes are calculated from the samples which are selected item's raw prices by criterion that the sales days are over 1000 days. The frequencies of regular price changes are calculated from the samples which are selected item's mode prices by criterion that the sales days are over 1000 days. A mode prices are monthly and pooled stores mode by an item.

☒ 2



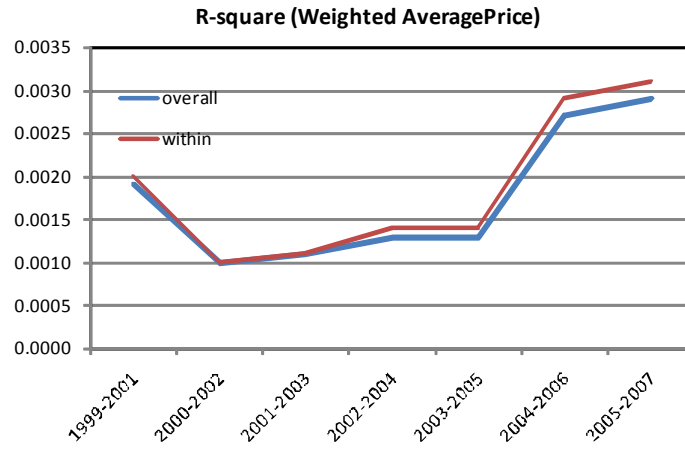
The inflation rates of weighted average price variable is aggregation of each y/y growth rate of item level monthly weighted average prices. A series of item level monthly weighted average prices is weighted average prices of daily raw prices by sales quantities of all stores that sell the item. The inflation rates of mode price variable is aggregation of each y/y growth rate of the item level monthly mode prices. An item level monthly mode price is a mode price of daily raw prices in a month and in all store that sell the item. The Official CPI variable is calculated from the official CPI data by items that belong to categories of processed food and household goods.

☒ 3

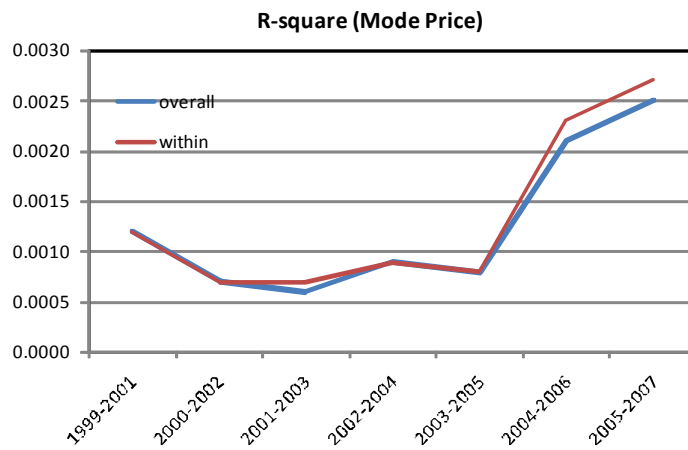


☒ 4

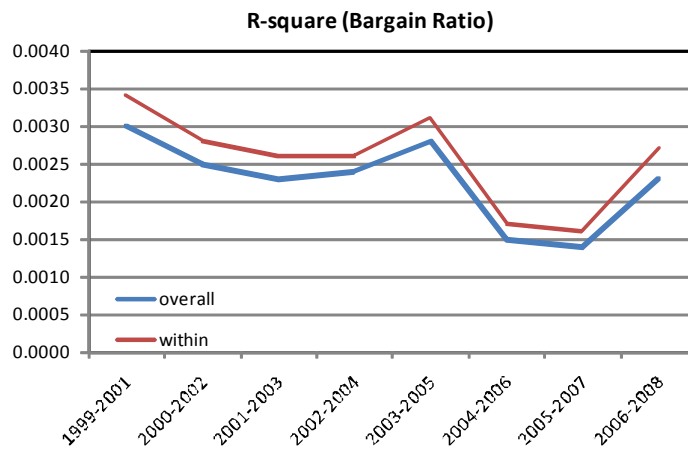
R-squares on the Fixed Effect Regressions to Time Dummies
(A)



(B)

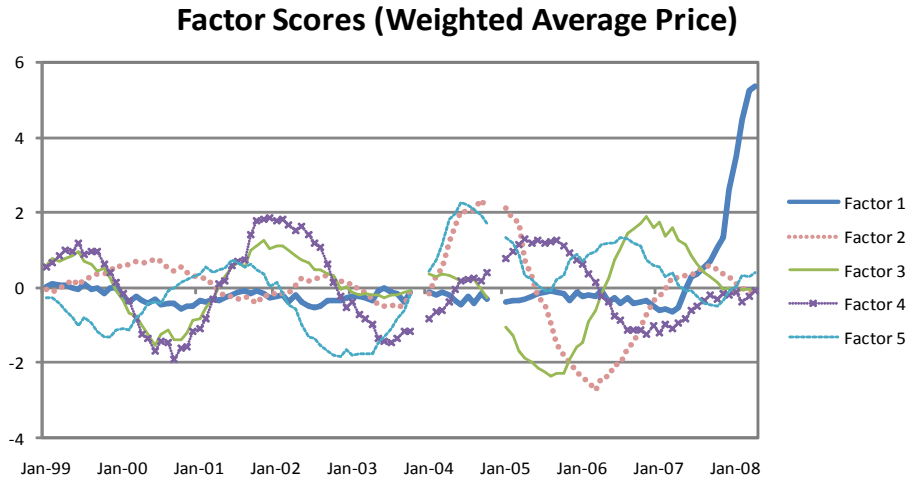


(C)

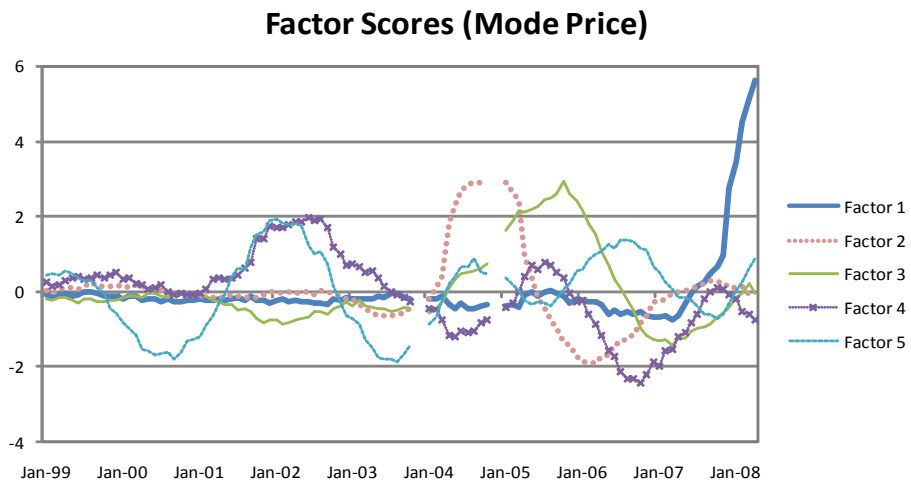


☒ 5

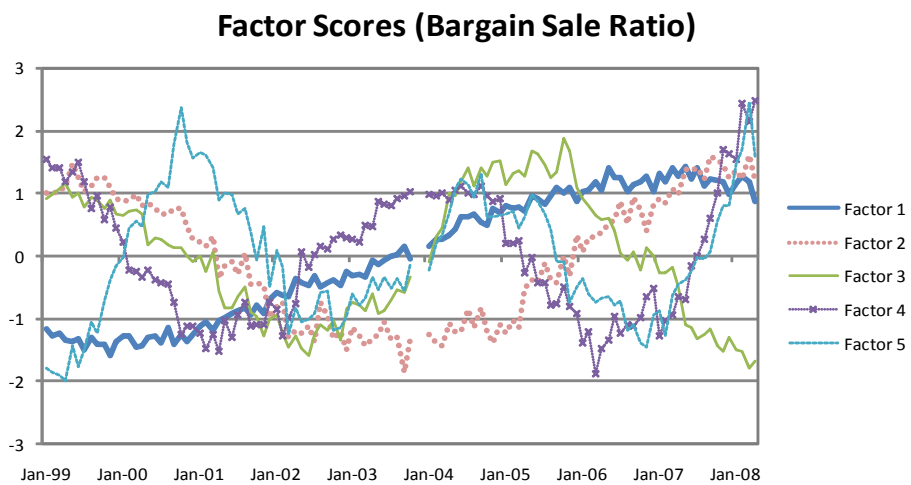
Estimated Factor Scores
(A)



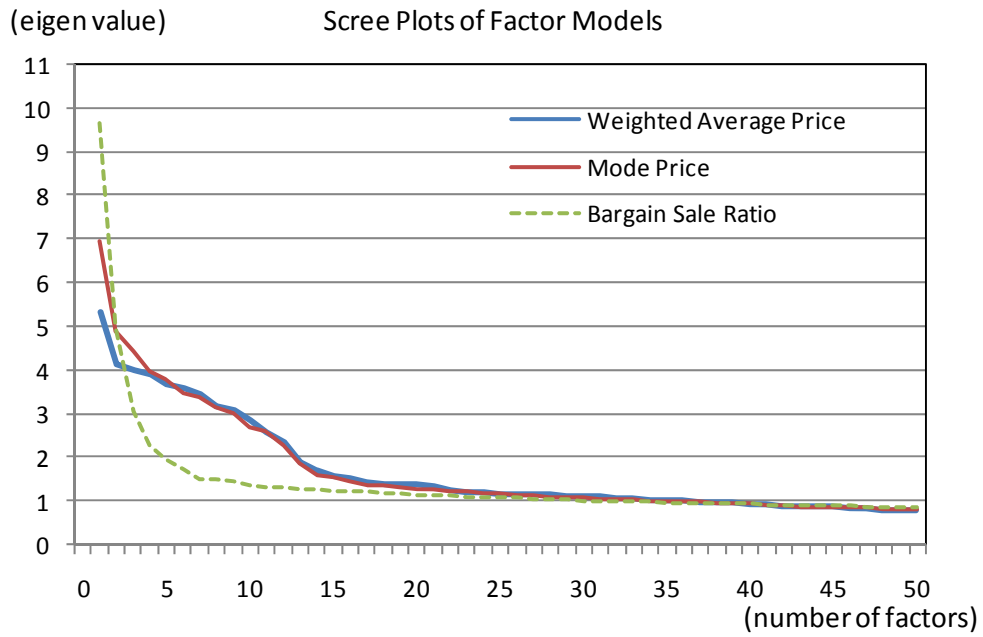
(B)



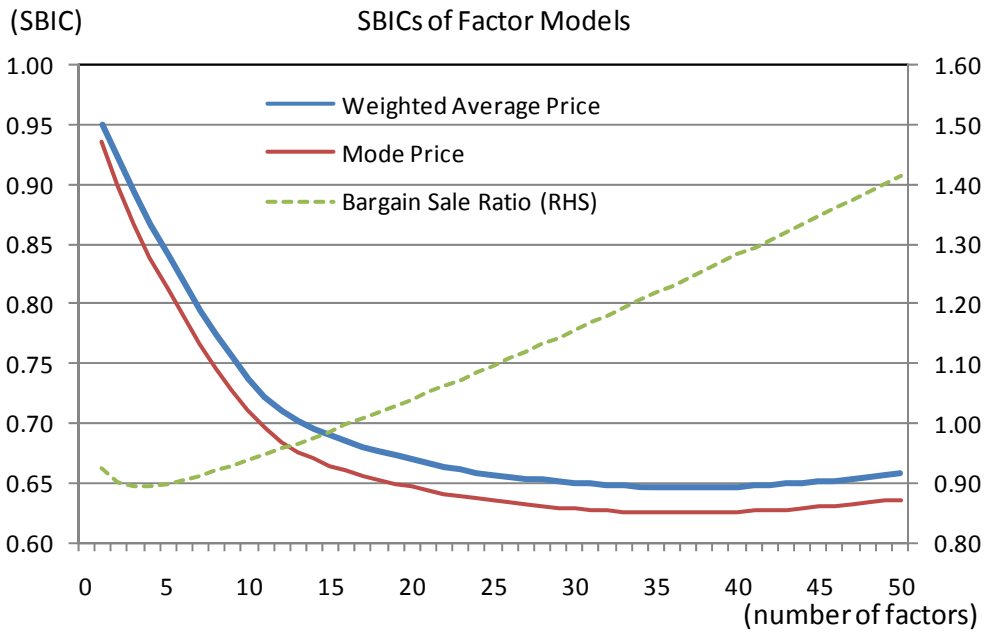
(C)



☒ 6

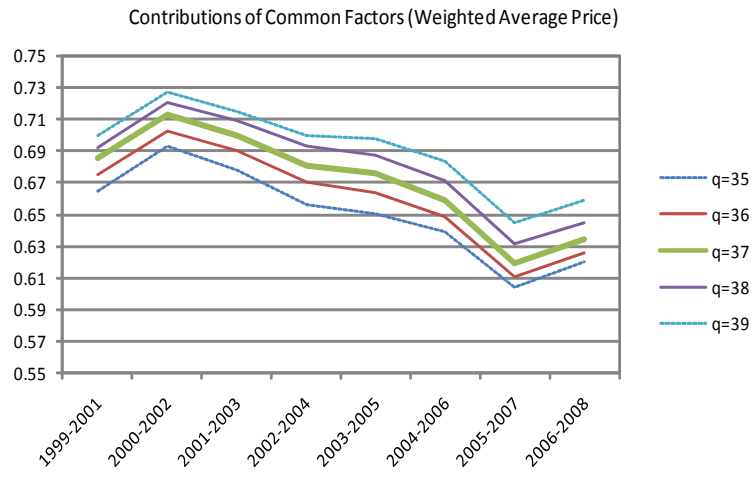


☒ 7

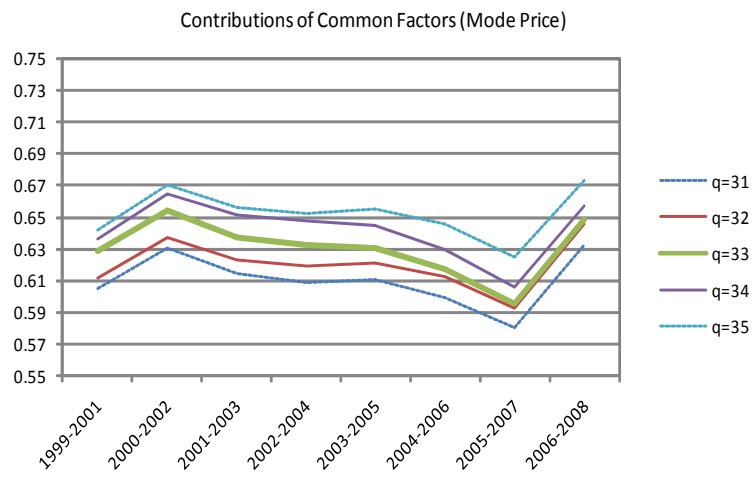


Results of Variance Decompositions

(A)



(B)



(C)

