

2-3 女子賃金率・男性所得に関する仮説 H0 検定の結果と結果利用の留意点

以下、本小節および(出生・結婚・雇用に関する)次小節で報告する結果は、定常性の仮定を(ほぼ)満たしていると考えられるものである。(11) なお、沖縄県は有配偶率の推定でかなり特異な振る舞いをしたので、すべての推定で除いている。

表 2-3-1 は大都市を含む都県について、AR(1)係数推定値を整理したものである。

表 2-3-1 都県別 AR(1)係数推定値のまとめ

(1階の階差をとった女子平均賃金率)

	係数推定値	95%信頼区間 (係数推定値)		95%信頼区間 (H0受容区間)		H0 を 受 容
		下限	上限	下限	上限	
宮城	-0.258	-0.332	0.046	-0.510	-0.132	*
東京	-0.391	-0.519	-0.320	-0.432	-0.234	*
神奈川	-0.357	-0.458	-0.274	-0.438	-0.254	*
福岡	-0.377	-0.607	-0.296	-0.548	-0.237	*

(1階の階差をとった男子平均年収)

	係数推定値	95%信頼区間 (係数推定値)		95%信頼区間 (H0受容区間)		H0 を 受 容
		下限	上限	下限	上限	
宮城	-0.161	-0.297	0.119	-0.474	-0.059	*
東京	-0.174	-0.177	0.007	-0.327	-0.143	*
神奈川	-0.170	-0.227	0.060	-0.387	-0.100	*
福岡	-0.347	-0.661	-0.294	-0.467	-0.100	*

表 2-3-1 の係数推定値列は各地域ごとに最小二乗推定を行った点推定量 $\{\hat{\beta}_r\}_{r=1}^R$ である。95%信頼区間(係数推定値)列は Bootstrap によって構築された、 $\{\hat{\beta}_r\}_{r=1}^R$ の信頼区間である。すなわち、*B 回の Bootstrap による点推定量のセットを $\{\hat{\beta}_r^{*b}\}_{b=1}^{*B}$ とし、あるセットの 95% 信頼区間の上限(下限)を UB(LB) であらわすと、 $[\hat{\beta}_r - UB\{\hat{\beta}_r^{*b}\}_{b=1}^{*B}, \hat{\beta}_r + LB\{\hat{\beta}_r^{*b}\}_{b=1}^{*B}]$ が 95% 信頼区間(係数推定値)である。最後の 95% 信頼区間(H0 受容区間)は、H0 のもとでの、 β_0 の ($\hat{\beta}_r$ に対する)95% 信頼区間 $[\beta_0 - UB\{\hat{\beta}_r^{*b}\}_{b=1}^{*B}, \beta_0 + LB\{\hat{\beta}_r^{*b}\}_{b=1}^{*B}]$ であり、H0 が正しければ $\hat{\beta}_r$ がこの信頼区間の外に出るのは 5% のチャンスしかないことになる(したがって $\hat{\beta}_r$ が最後の信頼区間に落ちていれば、変量の都道府県固有の動的展開にはマクロ時系列変動と差がないという帰無仮説を受容する)。

表 2-3-1 から、いくつかのことがすぐにわかる。まず、点推定値が 95% 信頼区間(係数推定値)の外にあるということはなく、また推定値の絶対値は 1 より小さく定常性の仮定を満たしている。のちに述べるように、このモデル・テストをパスしない地域も存在するが、ここで取り上げる都県は、経済変数と 20 歳代・30 歳代の年齢別推定(第 3 節)において、このモデル・テストをパスしている都道府県のうち代表的なものである。そして、表最右列 H0 テストに関する結果を見ると、これらモデル推定が整合的に行われた都県では、

女子賃金・男性所得に関しては、 H_0 を棄却することができない、したがって、都道府県に固有の動的展開はあるとはいえないことがわかる。

その他の都道府県の検定の結果は、同一の形式で図示している。水平な太実線はパラメータ制約を課したマクロ推定値 $\hat{\beta}_0$ を国勢調査で用いられる都道府県番号に対するプロットである。また、参考のため $\hat{\beta}_0$ の 95%信頼区間の上限、下限を、それぞれ太点水平線と太破点水平線で報告している。点マーカークの細実線が地域ごとに最小二乗推定を行った点推定量 $\{\hat{\beta}_r\}_{r=1}^R$ であり、95%信頼区間(H_0 受容)の上限と下限が、それぞれ細破線、細点線によって表されている。したがって、細い破点線と点線の間、実線上のポイントが落ちていない都道府県では第2-1節の仮説 H_0 を棄却し、当該都道府県に固有の動態があることになる。

図 2-3 都道府県別 AR(1)係数推定値: 1階の階差をとった女子平均賃金率

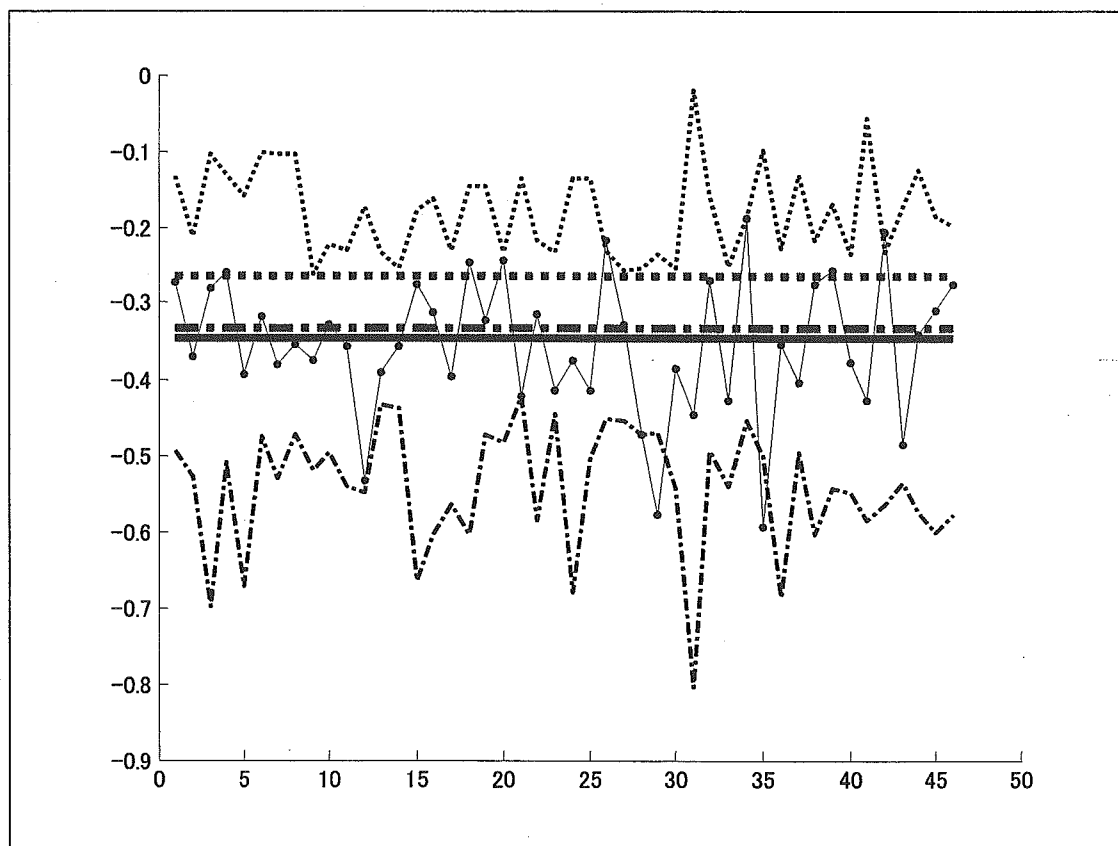
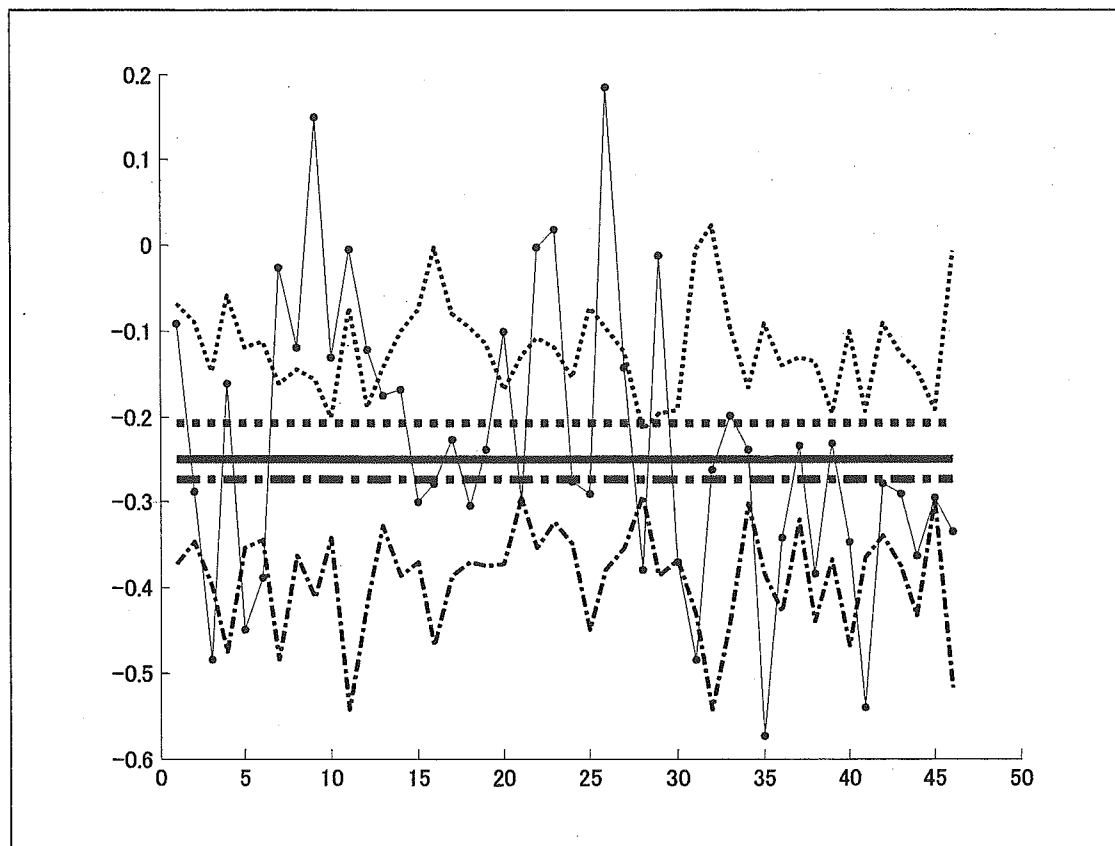


図 2-4 都道府県別 AR(1)係数推定値: 1 階の階差をとった男子平均年収



女子賃金率(図 2-3)についてはほぼすべての都道府県で、男性所得(図 2-4)については約 7 割の都道府県で、 H_0 を棄却できない、したがって、これら変数の動的な展開に都道府県差があるとはいえ、誤差を抜きにすれば、ほぼすべての都道府県が日本全国の時系列変動トレンドの動向にしたがっているといえる。これら経済変数の都道府県別の動的な展開は第 2-3 節で見る出生・結婚といった変数と極めて異なっており、時系列変動のパターンの分析が目的ならば、全国集計されたデータによって非現実でない近似値を得ることができることを示唆する。

ただし、ここでの分析は極めて限られたデータに基づくものであり、十分な注意が払われる必要がある。特にモデル特定化の問題について以下にいくつかの留意点をあげる。

第一に、いくつかの変数で、点推定値が(H_0 を制約として課さない Bootstrap による)95%信頼区間の下限の下に出してしまうケースがあった。これは、モデルの定式化が正しければありえないことなので、仮定のどれかが間違っている証拠となる。たとえば、誤差項が負の系列相関を持っている場合、推定量は下方にバイアスを持つ。(12)

第二に、我々の想定しているデータ生成過程では $\beta_{ij} = 0 \forall i \neq j$ と $\sigma_{ij} = 0 \forall i \neq j$ が仮定されているが、満たされる保証はない。

この制約が失敗する例として、たとえば、地域相関(Spatial correlation)が観測値にある場合が考えられる。前期の変数値に正の地域間(Spatial)系列相関があり、他の地域前期のデータが、当該地域今期の変数値に影響を及ぼすなら、他の地域前期データは除外されて

しまった変数であり、パラメータ推定量に上方バイアスを及ぼす。⁽¹³⁾ 地域間相関のパターンが逆ならば、推定量には下方バイアスを及ぼす。たとえば、東京都に居住していたが、結婚し、出産を考えたときに、近隣の県に移動するというようなことがあれば、東京都の出生率が低くなる代わりに、近隣の県の出生率が上がり(前期の変数間に負の相関があり)、近隣の県の今期の出生を行う女性の一部が前期に東京都に居住していたので、前期の東京都の変数値は今期の近隣の県の変数値に影響するというようなことがありうる。

これらの問題はデータ期間が長く取れば対策を立てることができる。たとえば、時系列相関に対しては、過去3期かそれ以上のラグをとった説明変数を操作変数として用いることが考えられる。また、地域間(Spatial)系列相関に対しては、より一般的な47都道府県のVARを考えことができる。 $\beta_{ij} = 0 \forall i \neq j$ の制約をおかず、47都道府県すべての相互関係を自由にできるほど長期のデータを取ることは困難であるが、たとえば近隣の都道府県との相互関係や、東京から周辺地域への外正的な影響の存在といったパターンを仮定して、検定することが可能である。十分なデータ期間があれば、最適化された目的関数のとる値を基礎としたさまざまなテストを容易に行うことができる。そのためにも、都道府県別の長期時系列データの整備は今後早急に取り組まれるべき重要な課題である。

2-3 出生率・有配偶・女子雇用労働率に関する仮説 H0 検定の結果

表2-3-2は、合計出生率、合計有配偶率、女子雇用労働力率に関する推計結果を表2-3-1と同じ都県について整理したものである。表2-3-1とは一転して、ほぼすべてのケースで仮説 H0 を棄却し、係数推定値はマクロ時系列変動から乖離していることが一見してわかる。前小節の最後に指摘したモデル特定化の問題は、これらの変数に対してより深刻である可能性があるが、経済変数と人口変数について、同じモデルの推定においてこれほど異なった結果をもたらすことは特筆に価するファインディングである。

その他の都道府県についての、合計出生率、合計有配偶率、女子雇用労働力率の検定の結果は次ページに図示した。

女子雇用労働力率(図2-5)では棄却は4割程度だが、合計出生率(図2-6)では、賃金率と好対照をなし、ほぼすべての地域で棄却、合計有配偶率(図2-7)でも約9割の地域で棄却という結果となった。したがって、出生・結婚といった変数の動的な展開は、都道府県格差があるだけでなく、時系列変動パターンそのものに地域差があることが示唆される。⁽¹⁴⁾

表 2-3-2 都道府県別 AR(1)係数推定値のまとめ

(1階の階差をとった女子雇用就業率)

	係数推定値	95%信頼区間 (係数推定値)		95%信頼区間 (H0受容区間)		H0 を 受 容
		下限	上限	下限	上限	
宮城	0.334	0.450	0.984	-0.261	0.273	
東京	0.722	1.014	1.894	-0.471	0.409	
神奈川	-0.904	-1.889	-1.721	-0.123	0.045	
福岡	-0.048	-0.302	0.236	-0.247	0.291	*

(1階の階差をとった出生率)

	係数推定値	95%信頼区間 (係数推定値)		95%信頼区間 (H0受容区間)		H0 を 受 容
		下限	上限	下限	上限	
宮城	-0.291	-0.251	-0.137	-0.443	-0.330	
東京	-0.022	0.245	0.425	-0.485	-0.305	
神奈川	-0.155	0.022	0.152	-0.443	-0.313	
福岡	-0.467	-0.614	-0.458	-0.455	-0.299	

(1階の階差をとった女子有配偶率)

	係数推定値	95%信頼区間 (係数推定値)		95%信頼区間 (H0受容区間)		H0 を 受 容
		下限	上限	下限	上限	
宮城	0.533	0.557	0.566	0.499	0.507	
東京	0.595	0.669	0.701	0.487	0.519	
神奈川	0.443	0.360	0.396	0.482	0.518	
福岡	0.568	0.613	0.654	0.485	0.525	

図 2-5 都道府県別 AR(1)係数推定値: 1階の階差をとった女子雇用就業率

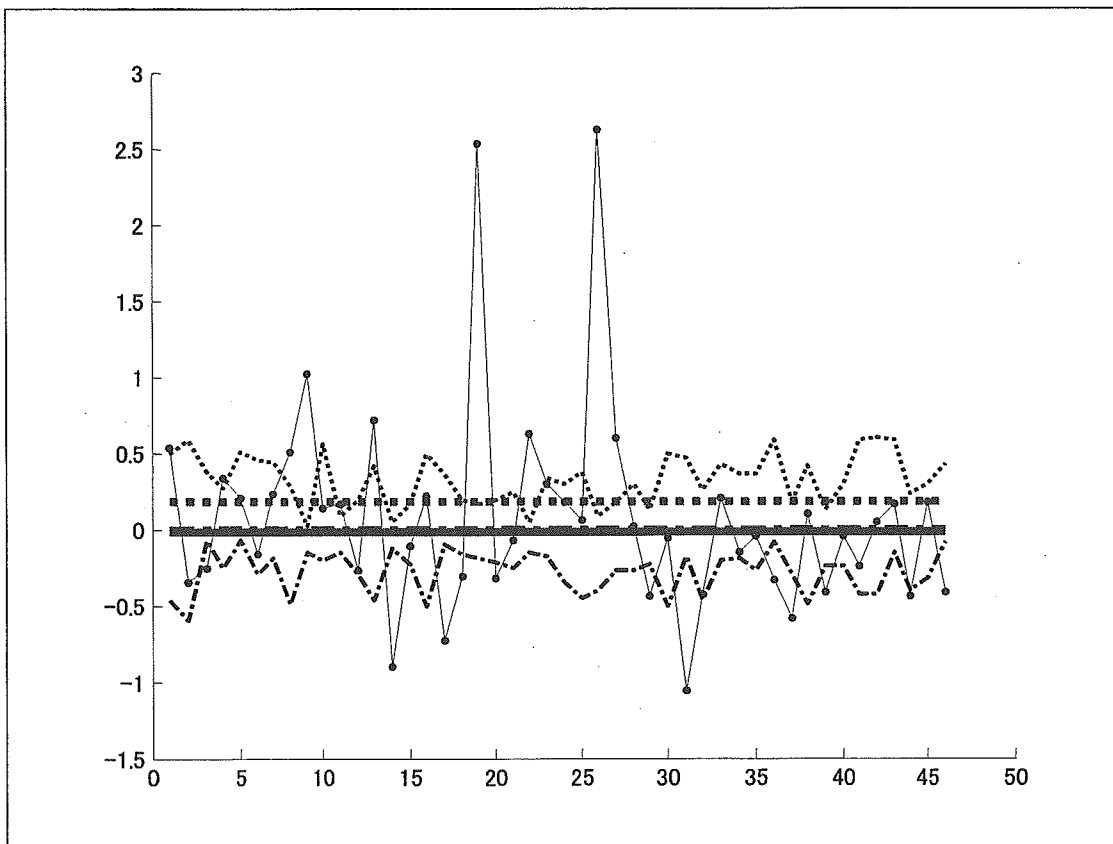


図 2-6 都道府県別 AR(1)係数推定値: 1 階の階差をとった出生率

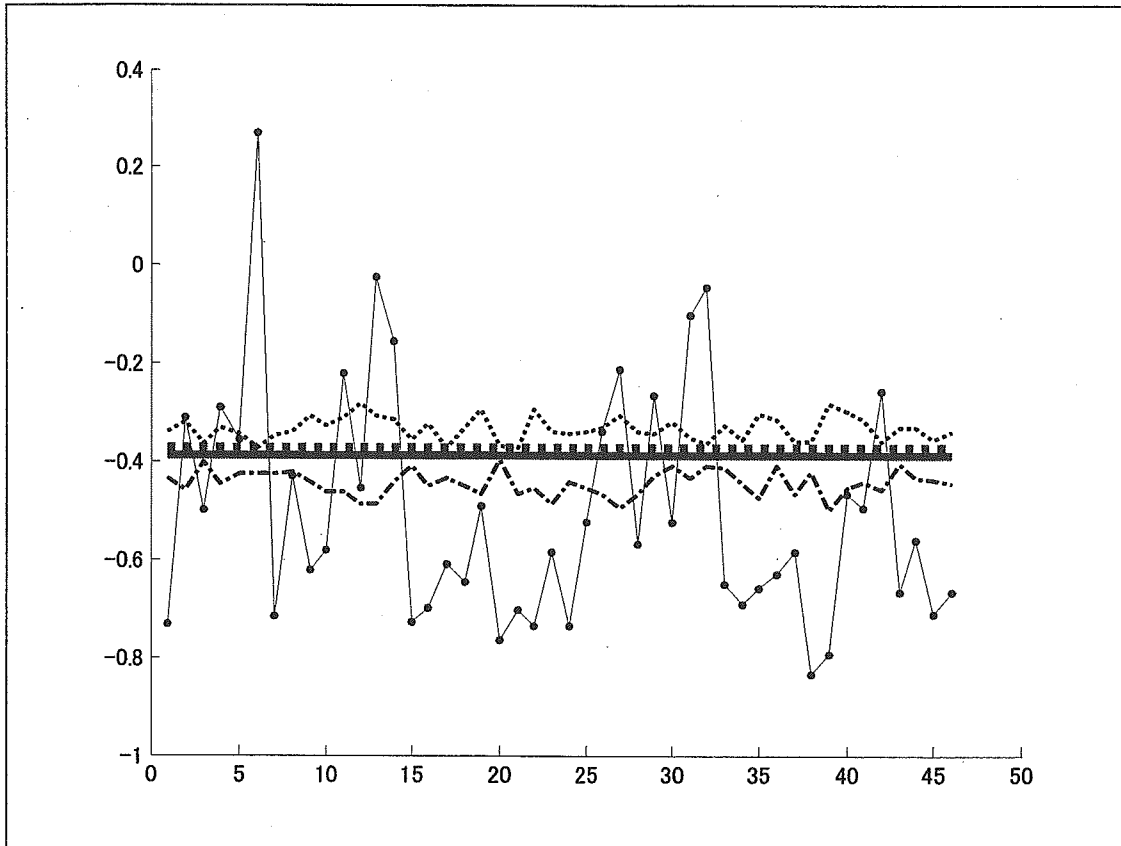
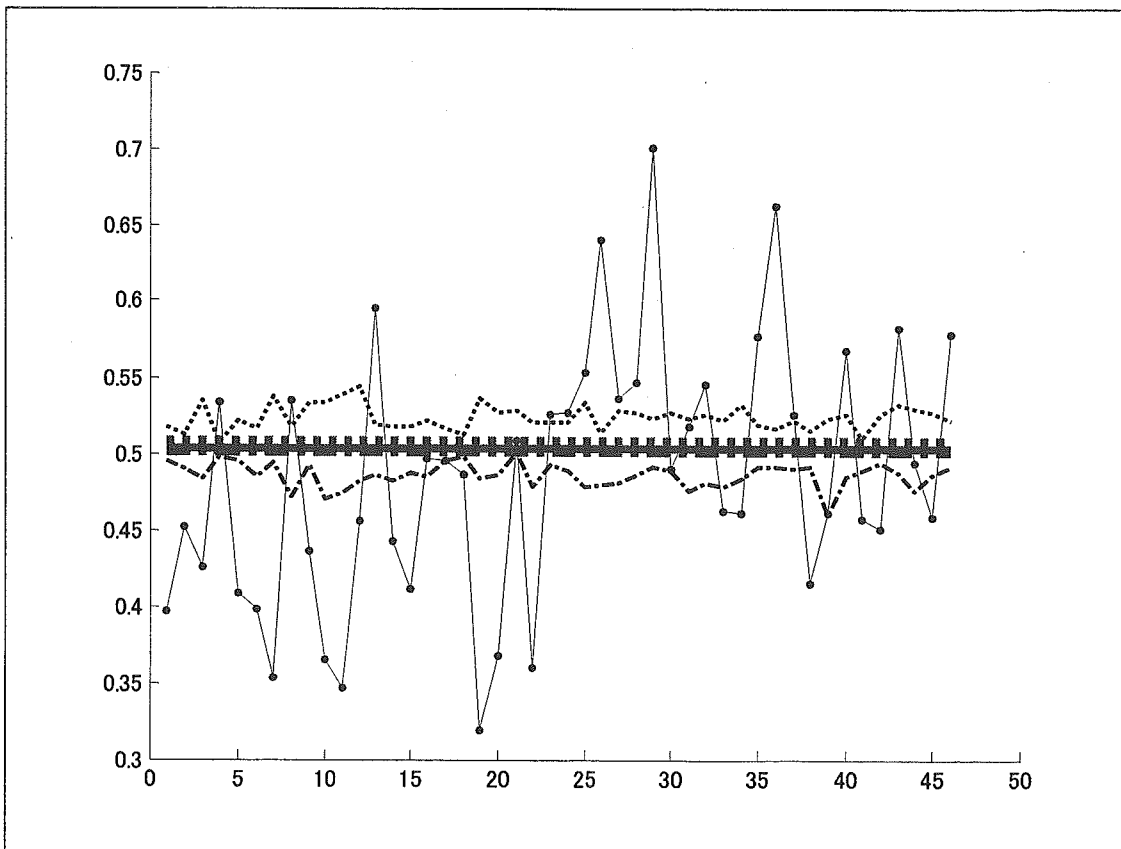


図 2-7 都道府県別 AR(1)係数推定値: 1 階の階差をとった女子有配偶率



3 結婚・出生・雇用の地域格差の出生コーホート変動

3-1 「価値変動」の地域パターンと結婚・出生・雇用における動的展開の地域格差

なぜ、経済変数と出生・結婚といった人口動態に関わる変数との間にこれほど大きな差が現れるのであろうか。

一つの、必要ではないが十分な説明は、出生コーホートに依存した(各コーホートが年齢を重ねることでは変化しない)「価値変動」(taste heterogeneity)の存在である。前 2-3 節での変量は、各時点で年齢層の合計であり、特定のコーホートの行動の結果は複数年にわたり重複している。そのため、 t 年と $t-1$ 年の二つの時期に出生・結婚行動を行うグループに、重複した出生コーホートの(推定式から除外された)「価値変動」により、推定式の誤差項には系列相関と内生性が生まれる。人口動態に特に強く関わる「価値変動」のパターンに都道府県差があるために、図 2-3~7 のような明確な差が現れたと考えることができる。

この点をより詳しく見るために、出生コーホート c の地域 r 、時点 t の観察される変量 $y_{c,r,t}^*$ が、観察されない「価値変動」 $\alpha_{c,r}$ ⁽¹⁵⁾の何らかの関数 $\varphi_{c,r,t}$ であると考える。

$$y_{c,r,t}^* = \varphi_{c,r,t}(\alpha_{c,r})$$

ここで $\alpha_{c,r}$ が $\alpha_{c-1,r}$ 独立であると想定しよう。⁽¹⁶⁾

合計変量(出生率・結婚率)に対する、われわれの AR(1)モデルは

$$y_{r,t} = y_{r,t-1}\beta_r + \varepsilon_{r,t} \Rightarrow \sum_{c \in C_t} y_{c,r,t} = \sum_{c \in C_{t-1}} y_{c,r,t-1}\beta_r + \varepsilon_{r,t}$$

であり、ここで C_t は t 時に出生・結婚行動を行うコーホートの出生年のセットである。これまで $\{\varepsilon_{r,t}\}_{t=1}^T$ は地域に固有のショックプロセスであると考えてきたが、「価値変動」との関係を考えて、これは当該期に結婚・出生行動を行うコーホートの「価値観」⁽¹⁷⁾も含んでいると見るべきである。そこで、 $\varepsilon_{r,t} = \varepsilon_{r,t}(\alpha_{c,r} | c \in C_t)$ と書くことにする。

$$\text{したがって、観測された変量によって推定可能なモデル} \quad \sum_{c \in C_t} y_{c,r,t}^* = \sum_{c \in C_{t-1}} y_{c,r,t-1}^* \beta_r + \varepsilon_{r,t}$$

は次のようになる。

$$\sum_{c \in C_t} \varphi_{c,r,t}(\alpha_{c,r}) = \sum_{c \in C_{t-1}} \varphi_{c,r,t-1}(\alpha_{c,r})\beta_r + \varepsilon_{r,t}(\alpha_{c,r} | c \in C_t)$$

$C_t \cap C_{t-1} : \text{non-empty}$ なので、

$$E\left(\sum_{c \in C_{t-1}} \varphi_{c,r,t-1}(\alpha_{c,r}) \cdot \varepsilon_{r,t}(\alpha_{c,r} | c \in C_t)\right) = \text{Cov}(\varphi_{c,r,t-1}(\alpha_{c,r}), \varepsilon_{r,t}(\alpha_{c,r} | c \in C_t) | c \in [C_{t-1} \cup C_t] \setminus [C_{t-1} \cap C_t]) \\ + \text{Cov}(\varphi_{c,r,t-1}(\alpha_{c,r}), \varepsilon_{r,t}(\alpha_{c,r} | c \in C_t) | c \in [C_{t-1} \cap C_t]) \neq 0$$

右辺の最初の項はゼロであるが、第二項がゼロではないので、通常の観測誤差を含むモデ

ルの場合と同様、最小二乗推定量は不偏性・一致性を満たさない。加えて、あるコーホート c について、 $Var(\varepsilon_{r,t}(\alpha_{c,r})) \neq 0$ for each r であり、 $C_t \cap C_{t-1} : non-empty$ なので、 $E(\varepsilon_{r,t}\varepsilon_{r,t-1}) \neq 0$ for each r 、すなわち推計式の誤差項は系列相関を持つ。なお、推定量のバイアスは

$$\hat{\beta}_r - \beta_r = \frac{Cov_t(\varphi_{c,r,t-1}(\alpha_{c,r}), \varepsilon_{r,t}(\alpha_{c,r} | c \in C_t) | c \in [C_{t-1} \cap C_t])}{Var_t\left(\sum_{c \in C_{t-1}} \varphi_{c,r,t-1}(\alpha_{c,r})\right)}$$

である。したがって、推定量のバイアスは「価値変動」 $\sum_{c \in [C_{t-1} \cap C_t]} \alpha_{c,r}$ の分散共分散構造に依

存していることがわかる。この分散共分散構造は、地域 r の条件付分散共分散であることに注意すると、推定量のバイアスは、このような「価値変動」の分散共分散構造の地域間差のパターンの関数であることがわかる。真の「価値変動」がコーホートの間でどのように変化してきているのかが明確ではないため、バイアスがどのような方向なのかを考えることは容易ではないが、少なくとも、たとえば晩婚化・晩産化といった変化が始まった時期が地域間で異なっていれば(晩婚・晩産の始まった出生コーホートが異なっていれば)、推定量に地域間の差が現れる。(18)

より具体的には、各出生年・時点毎の観測される変数の関係を参照いただきたい。(表 2-4)

表 2-4 各出生年・年次毎の観測される変数 $y_{c,r,t}^* = \varphi_{c,r,t}(\alpha_{c,r})$

	20~24 歳	25~29 歳	30~34 歳	...	$\alpha_{c,r} c \in [C_{t-1} \cap C_t]$
1970 年	<u>$\varphi_{-50,r,70}(\alpha_{-50,r})$</u>	<u>$\varphi_{-45,r,70}(\alpha_{-45,r})$</u>	$\varphi_{-40,r,70}(\alpha_{-40,r})$...	<u>$\alpha_{-45,r}$</u> , <u>$\alpha_{-40,r}$</u> , ...
1975 年	$\varphi_{-55,r,75}(\alpha_{-55,r})$	<u>$\varphi_{-50,r,75}(\alpha_{-50,r})$</u>	<u>$\varphi_{-45,r,75}(\alpha_{-45,r})$</u>	...	<u>$\alpha_{-50,r}$</u> , <u>$\alpha_{-45,r}$</u> , ...
1980 年	$\varphi_{-60,r,80}(\alpha_{-60,r})$	$\varphi_{-55,r,80}(\alpha_{-55,r})$	<u>$\varphi_{-50,r,80}(\alpha_{-50,r})$</u>	...	$\alpha_{-55,r}$, <u>$\alpha_{-50,r}$</u> , ...
...

(注) c に対応する下添字 t は出生年の上限を表す。たとえば、 -50 は 45~50 年生まれコーホートに対応している。

表 2-4 では 1945~50 年生まれコーホートについて下線を引いてあり、「価値変動」の当該コーホートに特有のものは、表に現れる 20~34 歳のもっともライフイベントが起こりやすい時期で、1970、75、80 年に重複してあらわれる。「価値変動」は年齢を重ねても変化しないが、生まれ年(と地域)に依存するもの(生まれ年(と地域)に対して確率変数)と仮定されており、複数の出生コーホートを集計した合計変数の推定式の誤差項には系列相関が

生まれ、また重複によって誤差項と説明変数に内生性があり、推定量もバイアスを持つ。

3-2 年齢階層別出生率・有配偶・女子雇用労働率に関する仮説 H0 検定の結果

前3-1節の最後で指摘されたコーホートに固有の「価値変動」(taste heterogeneity or preference shock)は年齢階層別の変量の動態を見ることで対処することができる可能性がある。なぜならば、年齢 a 歳に対する $(t, t-1)$ 時点の変量の差は、 $(t-a, t-a-1)$ 出生コーホートの変量の差であり、年齢別に出生・結婚といった変量の動的な展開を検討するということは、変量の出生コーホートが遅くなるほどどのような変化が起こっているのかという、出生コーホート間の比較になるからである。

具体的には、第2-3節と同様に推定可能な観察された変量は $y_{c,r,t}^* = \varphi_{c,r,t}(\alpha_{c,r})$ によってあらわされるとする。推計可能な年齢階層別の AR(1)モデルは $y_{a,r,t}^* = y_{a,r,t-1}^* \beta_{a,r} + \varepsilon_{a,r,t}$ であるが、これは $y_{c,r,t}^* = y_{c-1,r,t-1}^* \beta_{a,r} + \varepsilon_{c,r,t}$ と同値である。したがって、推定誤差は、

$$\hat{\beta}_{a,r} - \beta_{a,r} = \frac{\text{Cov}_t(\varphi_{c-1,r,t-1}(\alpha_{c-1,r}), \varepsilon_{c,r,t}(\alpha_{c,r}))}{\text{Var}_t\left(\sum_{c \in C_{r-1}} \varphi_{c-1,r,t-1}(\alpha_{c-1,r})\right)} = 0 \text{ if } \alpha_{c,r} \text{ indep. of } \alpha_{c-1,r}$$

となる。分母の共分散は、前小節で仮定された $\alpha_{c,r}$ と $\alpha_{c-1,r}$ の独立性によってゼロであり、⁽¹⁹⁾推定誤差はなくなることがわかる。

推定の結果は、20~24歳代について先の表2-3と同一の都県のものを表2-5に、またすべての都道府県について付図2-1-1~付図2-1-4(出生率)、付図2-2-1、付図2-2-2(有配偶率)を最後にまとめた。⁽²⁰⁾

表2-5を見ると、表2-4の結果とは打って変わって、点推定値は95%信頼区間(係数推定値)の中にあり、またデータも定常とみなせ、すなわちモデル・テストをパスしていて、かつすべてのケースで H0 を棄却できない(動態に都道府県差は認められない)という結果となった。この結果は各コーホートに対してランダムであるような「価値変動」の存在を強く示唆するものと考えることができる。

表 2-5 年齢別都道府県別 AR(1)係数推定値のまとめ：出生率と女子有配偶率(20～24 歳)

(1階の階差をとった出生率:20-24歳)

	係数推定値	95%信頼区間 (係数推定値)		95%信頼区間 (H0受容区間)		H0を受容
		下限	上限	下限	上限	
宮城	-0.066	-0.130	0.336	-0.459	0.007	*
東京	-0.301	-0.624	-0.148	-0.483	-0.006	*
神奈川	-0.367	-0.753	-0.254	-0.479	0.020	*
福岡	-0.332	-0.626	-0.234	-0.421	-0.029	*

(1階の階差をとった女子有配偶率:20-24歳)

	係数推定値	95%信頼区間 (係数推定値)		95%信頼区間 (H0受容区間)		H0を受容
		下限	上限	下限	上限	
宮城	-0.318	-0.567	-0.172	-0.467	-0.072	*
東京	-0.386	-0.725	-0.310	-0.490	-0.075	*
神奈川	-0.403	-0.753	-0.330	-0.484	-0.061	*
福岡	-0.397	-0.682	-0.363	-0.425	-0.105	*

このような結果は表 2-5 に掲げた 4 都県のみに見られるものではない。付図 2-1-1～付図 2-1-4(出生率)では、図 2-6(合計特殊出生率)と比べ、極めて明確なコントラストがあり、年齢別推計では、ほとんどの都道府県で H0 を棄却できない(動態に都道府県差は認められない)ことがわかる。ただし、25～29 歳については、棄却・受容が半々程度である。この年齢層は、第二次出生転換期以降で、もっとも大きな変化を経験した層であり、「価値変動」が系列相関をもって、推計に影響を与えた可能性がある。(21)

有配偶率についても、20～24 歳層については 9 割以上、25～29 歳以上については、7～8 割の都道府県で H0 を棄却できないという結果となった。

女性雇用労働力については、結果は報告していないが、20～24 歳層では約 3 分の 2 の都道府県で受容、25～29 歳で 7 割程度の受容、30～34 歳では 8 割弱、35～39 歳層で 8 割強の都道府県での受容となった。低年齢層の方が H0 を棄却するチャンスは大きいですが、年齢合計女子雇用労働力率は 6 割程度の受容であったことと比べると、総じて動的な展開に都道府県差は認められない傾向が確認された。

第 2-3 節～第 3-2 節の経済変数と人口動態に関わる変量の振る舞いの明瞭な差と、時点集計量とコーホート変動に関するこれらの結果は、出生・結婚といった人口動態に関わる変数に対しては、出生コーホートに依存した(各コーホートが年齢を重ねることで変化しない)「価値変動」(taste heterogeneity, or iid preference shock)が強く影響し、都道府県格差があるだけでなく、「価値変動」の地域間差によって時系列変動パターンそのものに地域差があることを示唆する。本稿の分析から「価値変動」がどのように展開したのかについて結論を下すことはできないが、各出生コーホートに対して独立なショックであるような「価値変動」があって、そのような「価値変動」の(実現値の)地域パターンが異なるために、出生・結婚変数には水準と時系列変動パターンに都道府県差があることが示唆される。

この「価値変動」は出生コーホート×地域に固有(地域別出生コーホートに固有)のものである。(22) したがって、地域に固有の固定効果を、何らかの方法で操作するというアプローチでは、分析の対象である変量に関わる行動を当該時点に行う出生コーホートの組み合わせが変化していると、適切に除去することができない。合計特殊出生率はその簡便性からよく分析の対象とされるが、その危うさを示唆するものである。(23)

4 今後の課題

本稿では、第二次人口転換期以後の日本の女性の結婚・出生行動の地域差の動態が検討された。第2節では、結婚・出生といった変数がAR(1)モデルによって定式化され、その動的な展開の都道府県差が検証された。そこでは出生・結婚といった人口動態に関わる変量と経済変数の間の明確な差の存在が指摘され、人口動態に関わる変量に対しては、出生コーホートに依存した(各コーホートが年齢を重ねることでは変化しない)「価値変動」(taste heterogeneity, or iid preference shock)が強く影響する可能性が指摘され、さらに第3節では「価値変動」の地域パターンが異なるために、都道府県格差があるだけでなく、時系列変動パターンそのものに地域差がある可能性が指摘された。

本稿の分析手法には、少なくとも二つの問題が残されている。

第一に、結婚・出生などの変量に関し、データの制約からその他の変数を含まない単独のAR(1)モデルが想定されていた。しかしながら、もし真のモデルが

$$y_{r,t} = y_{r,t-1}\beta_r + z'_{r,t}\delta + \varepsilon_{r,t}$$

で、 $E(y_{r,t-1}z_{r,t}) = 0$ でないようなL次元ベクトル $z_{r,t}$ があるなら、 $\hat{\beta}_r$ はこの除外された変数によりバイアスを持つ。たとえば、女子雇用労働力率には地域格差があることはよく知られており、今期の女子雇用労働力率の低さは前期に結婚・出生を行った女性が多いためであれば、 $\hat{\beta}_r < \beta_r$ であるようなバイアスが発生する。前節の分析では、提示された仮説 $\hat{\beta}_r \rightarrow \beta_0 \forall r$ も本来は正しいが、 $Cov(y_{r,t-1}, z_{r,t}) \neq 0$ によって $\hat{\beta}_r$ に地域差があらわれていたという可能性を排除できない。

第二に、地域別に各出生コーホートに対して独立であるような「価値変動」の存在が示唆されたが、この「価値変動」は出生コーホート×地域に固有(地域別出生コーホートに固有)のものである。そのため、集計されたデータを利用した多くの研究で出生・結婚行動の動的な構造は適切に識別されておらず、そのような観測されない変数の影響を除去して、結婚・出生行動の動態の地域格差が観測された経済変数によって、どのように、どの程度説明できるのかについて明確な答えは出ていないことを示唆する。すなわち、第二次出生転換期以後、各都道府県で特異なパターンで低下してきた出生率に対し、各時点での地域

差を経済変数がどの程度説明するのか、また経済変数など観測される変数の地域差がどの程度低下のパターンを説明するのかについて本稿の分析が答えることはできない。

本稿の分析は、1970年から2000年の国勢調査年の都道府県別データを用いた分析であり、時系列変動の検証としてはデータの制約が非常に強い。特に、集計データのみを利用して個人レベルで定式化されたモデルのパラメータの推定を行うためには集計の単位内での変量の散らばりが集計の単位に依存しないといった仮定が満たされる必要があるが、集計データのみでこの仮定を検証することはできない。長期時系列変動の検証に、個人レベルのデータが提供する情報を合わせて利用することで、第二次人口転換期以後、各都道府県で特異なパターンで低下してきた結婚・出生率に対して、観測された経済変数は各時点での地域格差をどの程度説明するのか、また観測される変数の地域差がどの程度低下のパターンを説明するのかを検討することは来年度以降の課題である。

注釈

- (1) E-mail: keitas@ab.mbn.or.jp
- (2) 標準偏差には若干の低下傾向があるようにも見えるが、これは多分に沖縄県の「特異な」高出生の変化のためで、むしろ75年の格差が過剰であったと見ることができる。しかし1995年、2000年の相対絶対誤差はむしろ拡大しており、地域格差の縮小があるわけではない。なお、記述統計量は、47都道府県に同じウェイト1をあてており、平均は日本全国合計特殊出生率と一致しない。これは、(移動パターンを所与とし)各都道府県に住む女性がその時点の出生年齢パターンで出生を続けた場合の出生数という解釈の元で、各都道府県間の出生の地域格差を評価するためである。
- (3) 地域差と時系列変動によって説明されない誤差が、地域・時間に依存しているなら、このような結論は下せないことには十分な注意が必要である。
- (4) 経済学的な視点からは例えば日本労働研究雑誌2005年6月号(No.539)が「地域雇用」に関する特集を行っている。
- (5) 米谷(1995)、廣島・三田(1995)、高橋(1997)などがある。
- (6) 高山他(2000)などがある。
- (7) 北村・宮崎(2005)などがある。
- (8) 本稿の仮説H0の検定において、真の分布が正規分布であることは本質的でないが、1次2次のモーメントに関する制約により検定を行っていることから採用した表記上の簡略化である。
- (9) TFR、(年齢によらない)経済変数の動態を検証するためには本来、単年度のデータを利用すべきであるが、次節のコホート変動との比較のため5年毎のデータを利用している。
- (10) Bootstrap(推定と実証的誤差分布からのシミュレーション)の手順は紙面の都合上省略する。手順の概要は筆者より入手可能である。
- (11) データの定常性をチェックするために、まず、観測される系列自体で都道府県別にAR(1)推定を行い、都道府県別推定値に基づくインパルスレスポンス関数の信頼区間をBootstrapによって構築した。このインパルスレスポンスの信頼区間が発散するならば、定義からデータは定常でないので、 $\sqrt{T}(\hat{\beta}_r - \beta_0) \xrightarrow{d} N(0, \sigma_r^2)$ for $\forall r$ という仮説が意味を持たない。そこで、1階の階差をとって改めて定常性のチェックを行った。すべての系列がトレンドを持っているように見えることから推測されるように、1階の階差をとると都道府県別に見てもほぼすべての系列のインパルスレスポンスは

収束した。

- (12) 誤差項が $\varepsilon_{r,t} = \rho v_{r,t-1} + v_{r,t}$ で、 $\rho < 0$ のとき、 $y_{r,t-1} = y_{r,t-2}\beta_r + \rho v_{r,t-2} + v_{r,t-1}$ なので、 $Cov(y_{r,t-1}, v_{r,t-1}) > 0$ であるが、 $\rho < 0$ から

$Cov(y_{r,t-1}, \varepsilon_{r,t}) = Cov(y_{r,t-1}, \rho v_{r,t-1}) < 0$ である。したがって、推定量は下方にバイアスを持つ。 $\hat{\beta}_r = \beta_r + \frac{Cov(y_{r,t-1}, \varepsilon_{r,t})}{Var(y_{r,t-1})} < \beta_r$

- (13) 例えば、R 地域の变量が VAR(1) にしたがっているとする。ここで、地域相関(Spatial correlation) $Cov(y_{r,t-1}, y_{r',t-1}) \neq 0, r \neq r'$ があって、当該の population parameter が $\beta_{r,r'} \neq 0$ であるならば、本節のモデルのように先験的な制約 $\beta_{r,r'} = 0$ を課すことは当

該のパラメータ推定量 $\hat{\beta}_{rr}$ に $\left(\sum_{t=1}^T y_{r,t-1}^2 \right)^{-1} \sum_{t=1}^T y_{r,t-1} y_{r',t-1}$ だけバイアスをもたらす。R

本の方程式一本ずつの推定はこのようなバイアスをすべて含んだ形のものと考えべきである。

特に、出生・有配偶率には負の時系列傾向があるため、 $Cov(y_{r,t-1}, y_{r,t}) > 0$ であり、

$\hat{\beta}_r > 0$ である。負の時系列傾向は多くの地域で見られるが、そのような低下傾向のある地域の間では Spatial correlation $Cov(y_{r,t-1}, y_{r',t-1}) > 0$ があらわれる。そのため、 $\hat{\beta}_r$ には上方にバイアスがかかる。

- (14) 繰り返しとなるが、モデルの定式化が正しいという仮定の下での仮説の検定である。

- (15) 「価値変動」 $\alpha_{c,r}$ 自体がどのようなものと考えればよいかは明確ではないが、ここでは Non-degenerate な確率変数であると想定する。

- (16) これは、Preference shock のプロセスがについて明確モデルが存在しないため、一次接近のための単純化の仮定である。

- (17) このコーホートの「価値観」は taste heterogeneity というより、cohort specific, time invariant preference shock といった方がわかりやすいかもしれない。

- (18) 晩婚・晩産の始まったタイミングが地域間で異なっていることは十分だが必要ではないし、タイミングが同じであっても、その変化のパターンが異なっていれば(α の分散の和は異なるので)推定量に地域差は現れる。なお、 $\alpha_{c,r}$ は $\alpha_{c-1,r}$ と独立であり、

$\sum_{c \in [C_{t-1} \cap C_t]} \alpha_{c,r}$ の分散共分散は、各 $\alpha_{c,r}$ の分散の和となるが、想定から $Var(\alpha_{c,r}) \neq 0$ であ

るので、バイアスは残る。

- (19) 「価値変動」がコーホートごとに「価値変動」の母集団から無作為に抽出されたという想定は非現実的な可能性があり、誤差項にどの程度の系列相関があるのかは検証されることが望ましい。

- (20) ここでは紙面の都合上出生率と有配偶率の結果のみをあげた。また、30歳以上の年齢層の有配偶率は、1階の階差をとってもデータを定常とみなすことはできないため、検討の対象としない。

- (21) ここでの誤差分散の推定は系列相関がないものとして行っているが、系列相関があるなら推定される分散は大きくなる。したがって、系列相関によるバイアスを修正し、誤差分散の一致推定量を用いれば、結果は大きく変わる可能性がある。

- (22) たとえば、小椋・ディークル(1992)では α_{ar} (地域・特定年齢階層に固有の固定効果)が固定効果として処理されているが、本節の含意は固定効果として処理されるべきは $\alpha_{c,r}$ (地域・出生コーホートに固有の固定効果)である。また、同モデルで出生・結婚関数は同時方程式体系として捉えられているが、ランクコンディションから識別されるのは結婚関数であり出生関数は識別されない点、さらに出生・結婚関数の誤差項に相

- 関があるなら、係数推定量に同時方程式バイアスをもたらすが、この不適切な「価値変動」の除去により、この誤差項の相関の問題が残る可能性は高い。
- (23) 本稿で利用するデータでは、分析の対象とすることができないが、人口の地域間移動によって、地域を構成する出生コーホートがミックスしている(変化している)というような地域相関(Spatial correlation)がある場合、「地域に固有の効果」の問題は大きくなる。

参考文献

1. 伊達雄高・清水谷論「日本の出生率低下の要因分析: 実証的研究のサーベイと政策的合意の検討」『経済分析』176号、内閣府経済社会総合研究所、2005年6月
2. 廣嶋清志・三田房子「近年における都道府県別出生率格差の分析」『人口問題研究』No.50(4)、1995年1月
3. 北村信行・宮崎毅「結婚経験率と出生力の地域格差: 実証的サーベイ」Hi-stat Discussion Paper Series、2005年10月
4. 国土庁計画・調整局『地域の視点から少子化を考えるー結婚と出生の地域分析ー』大蔵省印刷局、1998年
5. 小椋正立・ロバートディーケル「1970年代以降の出生率低下とその原因」『日本経済研究』No.22、1992年3月
6. 高橋眞一「出生力の地域的分析」濱英彦・山口喜一編著『地域人口分析の基礎』古今書院、1997年2月
7. 高山他「結婚・育児の経済コストと出生率・少子化の敬意学的要因に関する一考察」『人口問題研究』No.56(4)、2000年12月
8. 米谷信行「わが国出生率低下の要因分析ー都道府県別データに基づくクロスセクション分析」『フィナンシャルレビュー』1995年2月

付表 2-1 最近の日本における結婚・出生の都道府県格差の研究

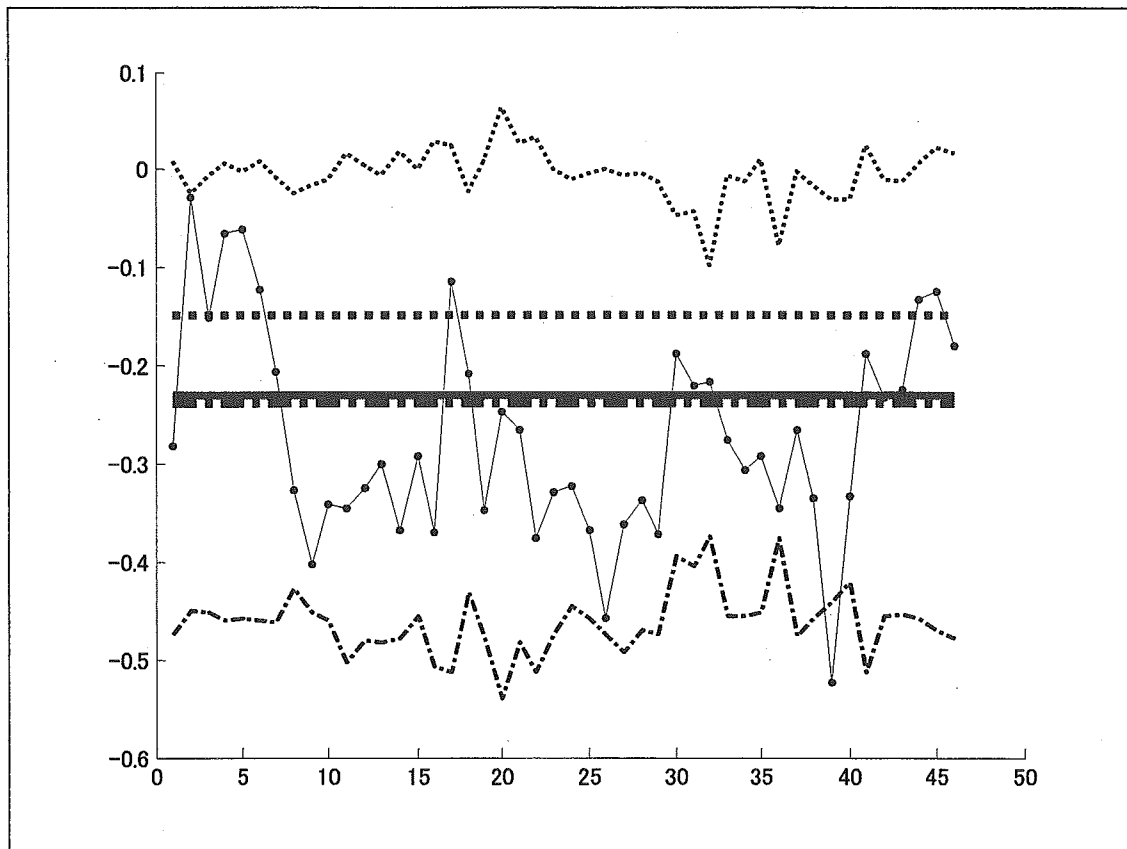
Article	関心のある変数	要因	Model	sign	Remark
小原・ティークル(1992)	都道府県年齢別出生率(1970→85歳)	女性賃金 男女賃金格差 高学歴割合(女性) 高学歴割合(男性) 地価 家賃	Fixed effect	- - / + - + - -	25歳、30歳 20歳 / 30歳 20歳、30歳 20歳 25歳 25歳、30歳
米谷(1995)	都道府県TFR(1970、80、92年)	住居費(消費支出に占める住居費比率) 教育費(消費支出に占める教育費) 女性の賃金 高学歴化 第三次産業 0-6歳人口に対する保育所定員比率 沖縄ダミー	OLS(year by year)	- - - - - - + +	70年を除く 70年を除く
廣島・三田(1995)	都道府県TFR(1985→90) 既住出生率 子供を産み終わった年齢35-39歳 における同居子供数 同居児数 既婚同居児数 既婚率 既婚出生率 就業率(既婚) 非就業者既婚率 都道府県同居児数の全国との格差 都道府県同居児数の全国との格差 都道府県同居児数の全国との格差 全国との既婚率格差 全国との就業率格差 同居児数	有配偶率初婚率 有配偶出生率 都道府県TFR 既婚TFR 就業率 就業率 既婚率 既婚就業率 既婚就業率 就業率(未婚) 就業率既婚率 既婚率 就業率、就業出生率 非就業者の既婚出生率 既婚率 就業率 同居児数の変化係数(標準偏差÷平均)	Simple cross tabulation	+ - + + + + + + - - + - + + + - -	東京のTFRの高さは夫婦出生率が低いにもかかわらず合計初婚率が比較的高いことによる。東京の低出生率は未婚率の高さそのものではなく初婚発生率の低さによる。 r=0.807(1980) TFRと同居字数はほぼ同じ地域傾向を持っている r=0.878(1985) r=0.901(1990) r=0.674(1985) r=0.843(1990) 大都市地域、就業者で低い 大都市地域、就業者で低い 大都市地域、就業者で低い+沖縄などいくつかの地方県 就業者 非就業者 就業者 非就業者 就業出生率が格差の85%を説明 就業出生率の寄与は就業既婚率と就業既婚出生に2分される 就業既婚率が出生率格差全体の約半分を説明 就業既婚出生率が76%を説明し、非就業出生率が35%を説明する 大都市圏でVariationが小さい

高橋(1997)	都道府県TFR 市区町村別出生(間接標準化出生率) 市区町村別出生(間接標準化出生率)	time time: 70→90 有配偶出生率	地域分布の変化 地域分布の変化	東京、大阪大都市周辺で大きく低下 大都市から離れた農業色のつよい県(沖縄島根鳥取岩手山形福島福井滋賀佐賀鹿児島)で低下しなかった。都区部の中心区部の出生がもつとも低く、都区部の周辺にある区部、その周辺市部→周辺に向かうほど出生力が高い(1970) 周辺部ほど出生力の低下が著しく、差は縮小、大部分の市区町村で低い水準。東京は同じパターン(1970→90)で変化なし、特に周辺部で結婚年齢の上昇→出生減。島根では結婚年齢の上昇があるが、有配偶出生は上昇。
高山他(2000)	都道府県TFR(1985→94)	25-9歳男性賃金 25-9歳男性賃金 ² 25-9歳の女性賃金 教育娯楽支出 教育費の物価指数 幼稚園定員 保育園定員 住居費(子育てコスト) 児童手当支給(公的) 子育てコストの軽減 児童福祉費支出(公的) 子育てコストの軽減 離婚率 平均初婚年齢(晩婚化) 離婚率 妊産婦保健指導数 (公的)子育てコストの軽減 社会保障収入 世代間移転収入 税保険料負担	weightedLS + - - - + - + - - + + + - - + +	

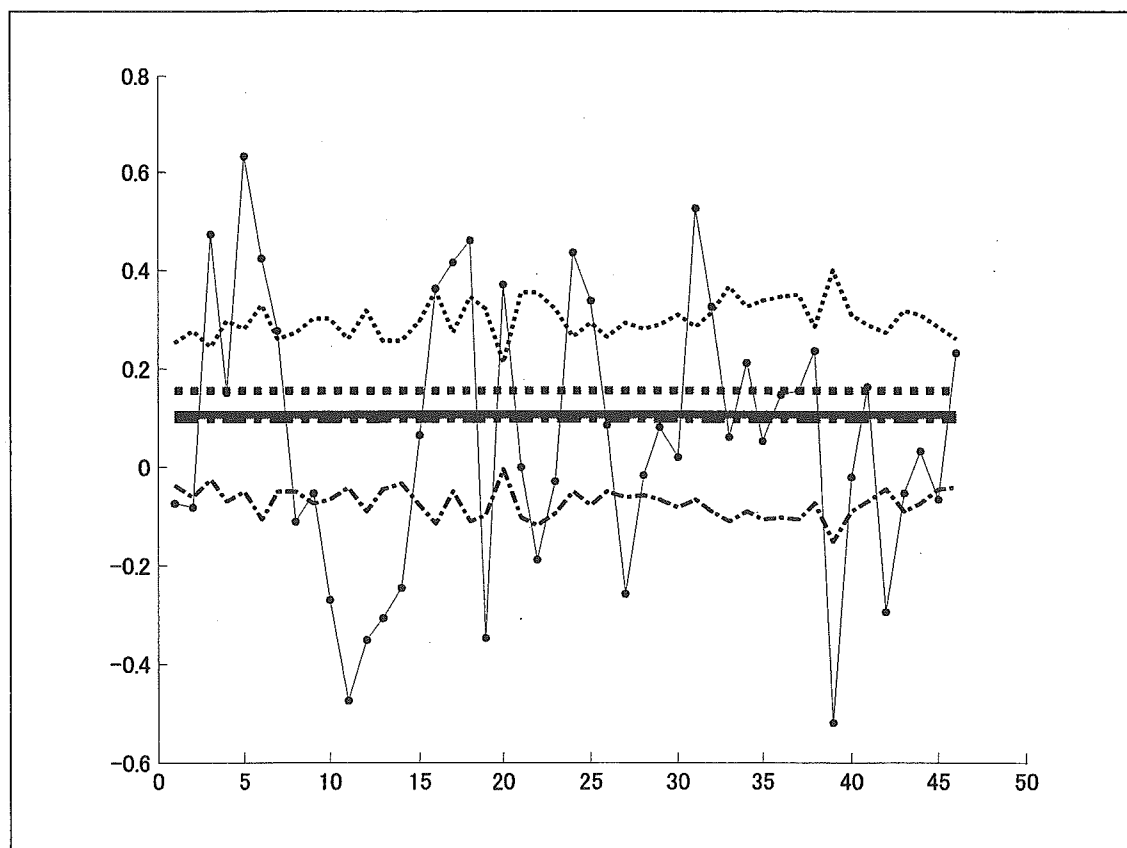
北村・宮崎(2005)	市区町村別出生TFR(2002年) 年齢階層別出生率 都道府県固定効果(都道府県ダミーの	人口密度 男性就業率 30歳女性の就業率 既婚女性の就業率 未婚女性の就業率 女性の30-34歳結婚経験率 結婚経験率 人口密度 男性就業率(20-59歳) 既婚女性就業率 人口増加率 大卒者比率 15歳未満人口比率 65歳以上人口比率 男性の初婚年齢 核家族比率 待機児童比率	Cross-tab (内生)2SLS Cross-tab	- + 無相関 + 無相関 + "/" "/" + - - + + - -	県間格差+県内で人口密度の高い都市部で出生は低くなっている 国土庁(1998)で市区町村データをを用いて同様の結果を確認 東京都 都道府県格差あり 男性/女性 女性(前年齢)+男性40-1歳、男性25-9歳 女性-30歳を除く 女性、男性20-4、35-歳を除く 女性-30歳を除く 東京を除く 鹿児島宮崎を除く
-------------	--	---	--	---	--

*Scope of literature limited to the study analyzing published Japanese data

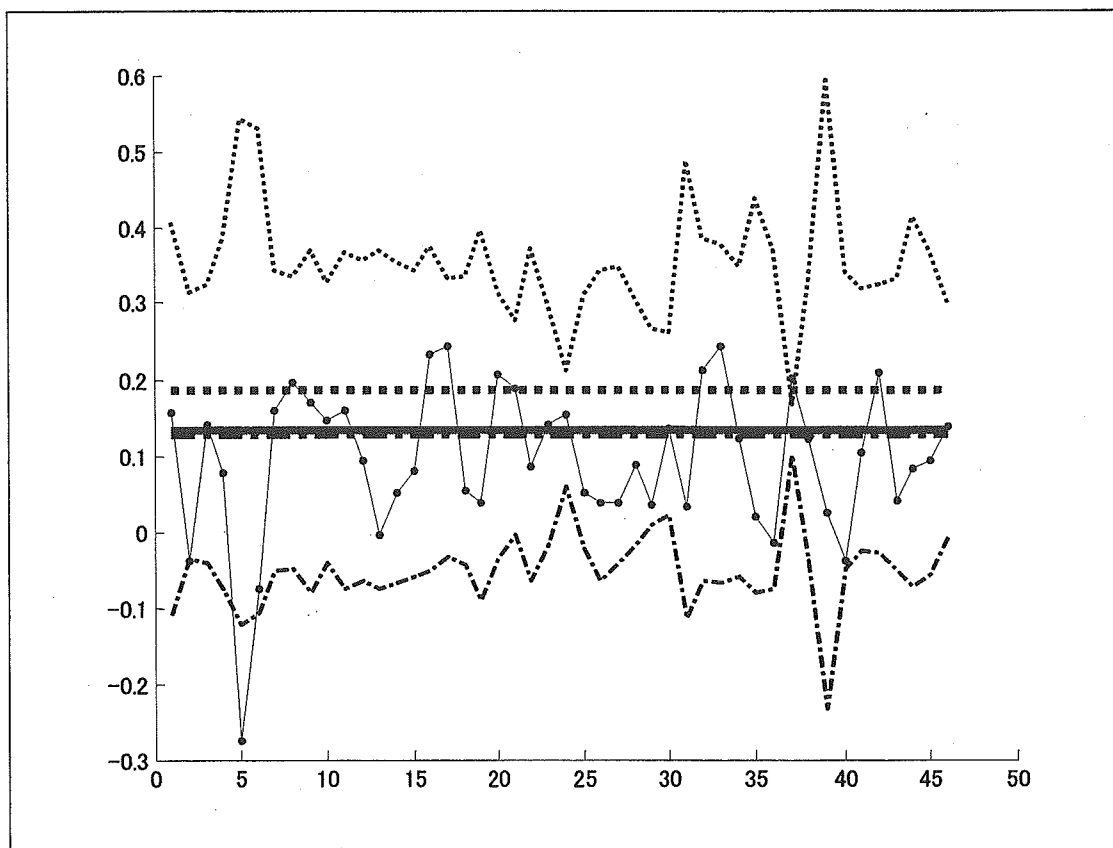
付図 2-1-1 都道府県別 AR(1)係数推定値: 1 階の階差をとった出生率(20-24 歳)



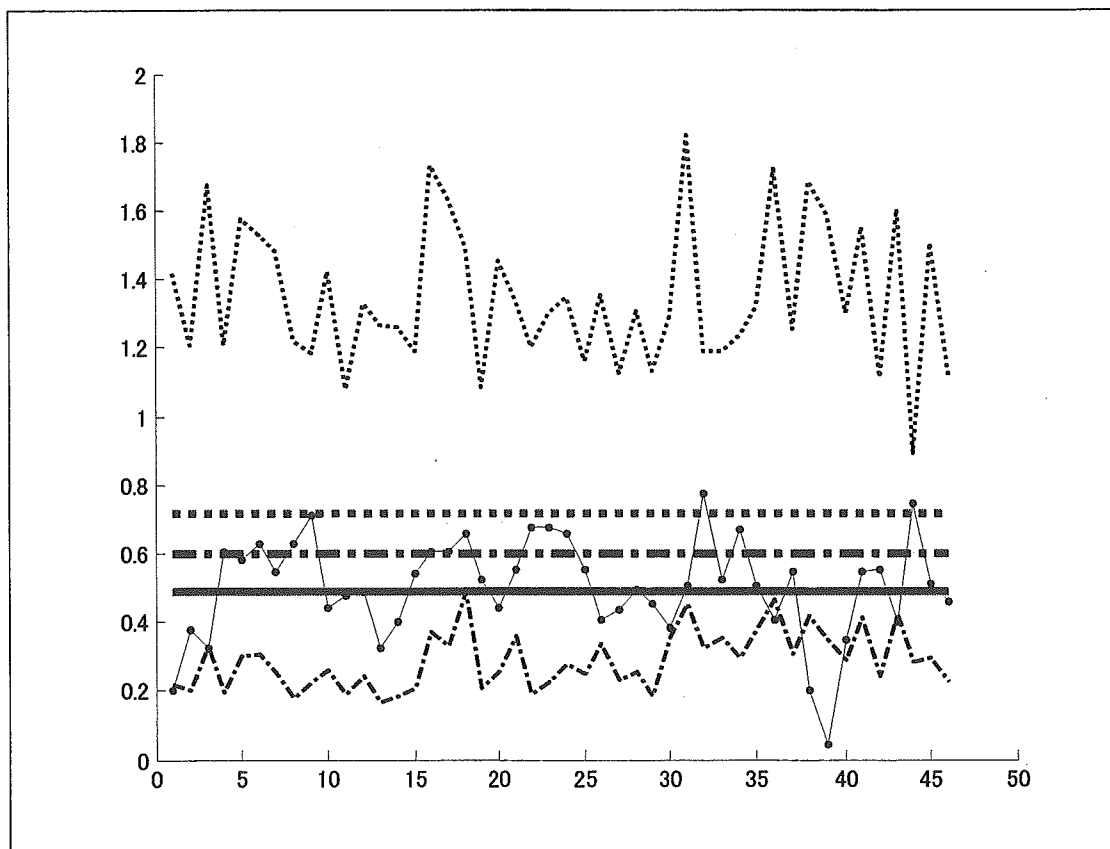
付図 2-1-2 都道府県別 AR(1)係数推定値: 1 階の階差をとった出生率(25-29 歳)



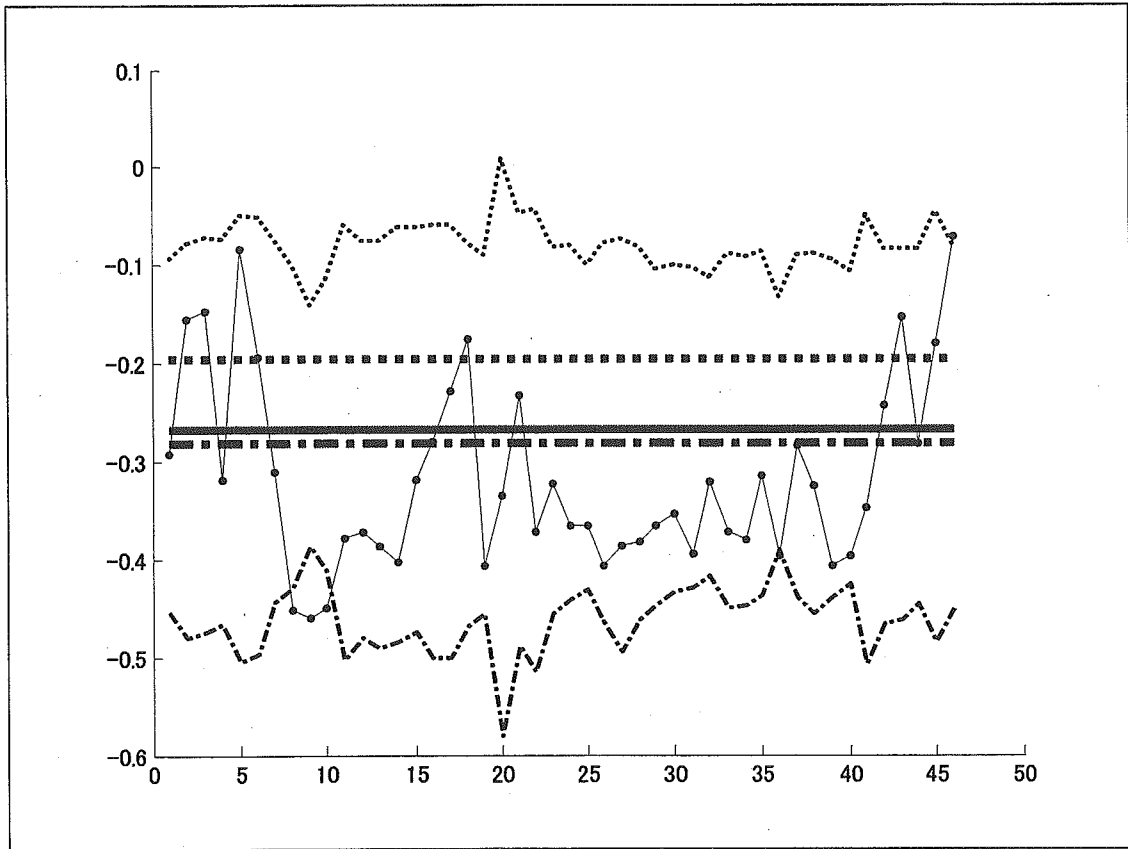
付図 2-1-3 都道府県別 AR(1)係数推定値: 1 階の階差をとった出生率(30-34 歳)



付図 2-1-4 都道府県別 AR(1)係数推定値: 1 階の階差をとった出生率(35-39 歳)



付図 2-2-1 都道府県別 AR(1)係数推定値: 1階の階差をとった女子有配偶率(20-24 歳)



付図 2-2-2 都道府県別 AR(1)係数推定値: 1階の階差をとった女子有配偶率(25-29 歳)

