経済世界のアロメトリー
“見せかけの相関はなぜ生じるか？”

日下部眞一

広島大学大学院総合科学研究科

On the spurious correlations among the ratio data induced by “the Law of Power Function”

Shinichi KUSAKABE

Graduate School of Integrated Arts and Sciences, Hiroshima University

Abstract

Most of the statistical data have power functional relationships with the size of the population, and therefore these statistical data fall in the powerfunctional relationships each other. These relations were analyzed through the derivations based on the simple model, and the law underlying these relations was called as “the Law of Power Function”. These power functional relationships necessarily cause the spurious correlations in the regression analyses. These analyses cast doubt the basic methodology of the econometrical analyses.

１．はじめに

アロメトリー（Allometry）とは，生物の発生成長にともなう形や大きさの変形を支配する原理を意味することばで，訳語としては，「相対成長」ということばがあてられている。生物個体の発生成長過程を測定するとそれらの測定値は，一般的に次のようなベキ関数で表される。

$$y = \alpha x^{\beta} \qquad (1)$$

例えば，x は体重で y は体の部位（脳の重量，体長など）の測定値などである。両側対数をとれば，例えば需要関数などの経済学で広く用いられるモデルで，β（弾力性）を傾きとする１次直線として表現される。生物界では，このベキ関数は生理学から生態学まで，さまざまな現象を表したり，種間にわたるさまざまな多様な測定値を表現，考察するのに用いられ，経験則としてアロメトリーが知られてきた（清水三雄 1959，McMahon and Bonner 1983）。

ここでは，経済世界の様々な現象がベキ関数によって表されるという，いわば「経済世界のアロメトリー」が重要であることを述べる。そして計量経済学的分析に用いられているような，per capita や率（％）で表した多くの統計指標値間で回帰分析する時に，ベキ関数関係が必然的に生じ，その結果，多くの場合に“みせかけの相関”が生じることを示す。あまりにも基本的な問題であるため問われていないのか，多くの計量経済学テキ

スト（たとえば Woolridge（2010）など）では注意がはらわれていない。しかし，計量経済学の基礎に関わる課題であるし，政治経済における政策提言などにも影響を与えるきわめて重要な課題である。このような分析対象となる統計指標値に内在する問題と見かけの相関についてすでに問題提起したが（日下部 2002a，2002b，2002c，2004），初等代数学的に演繹できることが明らかになったのでここに報告する。

2．ベキ関数の法則について

経済世界の様々な現象を表す多くの統計指標は，ベキ関数で表現される次のような性格を示し，４つの法則としてまとめられる。

「法則　１」　社会現象を表す統計指標の多くは人口のベキ関数で表される。

「法則　２」　任意の二つのベキ関数は，独立変数を同じくすれば，たがいにまたベキ関数として表される。

「法則　３」　ベキ関数として表されるような統計指標が，per capita（人口あたり）や率（%）で表されると，次数が１次小さくなった（$\beta-1$）乗のベキ乗関数として表され，弾性値 β の大小により必然的に正や負の相関が生じる。

「法則　４」　per capita や率（%）で表した任意の二種類の統計指標は，たがいにベキ関数として表現される。したがって，per capita や率（%），で表したこれら二種類の統計指標の間には，必然的に正や負の相関が生じる。

3．検　　証

分析した都道府県データはすべて2005年付けの値で，『民力2007』（2007），『地域経済総覧2009』（2009），『データで見る県勢』（2007）からとった。任意にとった18種類の統計指標値を“対数線形モデル”で人口に１次回帰した β〈傾き〉と R^2〈決定係数〉の値を表１に示している。18例のうち「地方交付税」だけが無相関であるが，ほかの多くは高度に有意な相関を示している。決定係数の値が0.9前後で，“対数線形モデル”によく適合している。つまり，多くの統計指標は人口のベキ関数として表現されるのが理解できる。

表１　各種統計指標を人口に対して１次回帰した傾き（β）と決定係数（R^2）の値（分析はすべて2005年付けの資料による）

統計指標	傾きβ	決定係数 R^2
犯罪認知件数	1.2798	0.9422
生活保護者数	1.2007	0.7131
地方税収	1.1670	0.9472
県民所得	1.1166	0.9796
賃金所得	1.1084	0.9829
県民総生産（GDP）	1.0996	0.9654
失業者数	1.0200	0.9394
NPO 数	1.0099	0.9219
就労者数	0.9947	0.9959
死亡数	0.8588	0.9756
教育費	0.8302	0.9850
歳出総額	0.7315	0.8856
一般公共事業費	0.7256	0.8786
行政投資	0.6870	0.8806
基準財政需要額	0.6790	0.9431
老人福祉施設数	0.6736	0.9309
国庫金支出金	0.5391	0.7236
地方交付税	−0.1035	0.0156

「法則　２」は，次のように代数的に導かれる。

二つの統計指標値を人口（x）のベキ関数としてあらわすと，

$$y_1 = \alpha_1 * x^{\beta_1} \qquad (2)$$
$$y_2 = \alpha_2 * x^{\beta_2} \qquad (3)$$

これらの式から x を消去すると，

$$y_1 = c * y_2^{\beta_1/\beta_2}，\text{ただし，} c = \alpha_1/(\alpha_2^{\beta_1/\beta_2}) \qquad (4)$$

となる。したがって，二つの統計指標値 y_1，y_2 はたがいにベキ関数として表される。

上述した表１の18の統計指標の中から傾き（弾性値 β）が適度に異なった８つの指標を選び，“両側対数線形モデル”で表現して任意の二つの指標間で相関をとった。可能な組み合わせ28の中で，「生活保護者数」との組み合わせの４つの場合（決定係数が0.7弱）をのぞき，多くの組み合

表２　各種統計指標の対数値をたがいに１次回帰した傾き（β）と決定係数（R^2）の値

被説明変数

老人福祉施設	行政投資	一般公共事業費	死亡数2005	失業者数	県民所得	生活保護者数	犯罪認知件数	説明変数
0.5878 0.8879	0.6052 0.8558	0.6460 0.8722	0.7552 0.9449	0.8729 0.8617	1.0020 0.9880	1.0067 0.6278	1.1286 0.9178	GDP
	0.9814 0.8758	1.0307 0.8640	1.2204 0.9603	1.3953 0.8569	1.5246 0.8902	1.6891 0.6878	1.7295 0.8388	老人福祉施設
		1.0371 0.9620	1.1212 0.8913	1.3028 0.8215	1.4152 0.8434	1.5752 0.6578	1.5738 0.7638	行政投資
			1.0612 0.8928	1.2164 0.8007	1.3511 0.8595	1.4486 0.6220	1.5063 0.7823	一般公共事業費
				1.1536 0.9084	1.2635 0.9482	1.3928 0.7253	1.4332 0.8933	死亡数
					1.0046 0.8781	1.2326 0.8322	1.1866 0.8970	失業者数
						1.0030 0.6333	1.1331 0.9401	県民所得
							0.7594 0.6708	生活保護者数

上の値が傾き　β
下の値が決定係数　R^2

わせで表２に示したように決定係数は0.9前後となった。したがって，任意の二種類の統計指標値がたがいにベキ関数で表されることが理解される。表２の最上段には，参考のため GDP（県民総生産）との相関を示している。

（２），（３）式におけるベキ関数で表されるような統計指標の性質として「法則　３」が次のように導かれる。

$y = \alpha * x^{\beta}$ の両辺を x で割ると，$y/x = \alpha * x^{\beta-1}$　（５）

この式は，per capita で表した統計指標が（$\beta-1$）乗，すなわち弾性値 β が１より大きいか小さいか，つまりそれぞれの統計指標値を人口に対して両側対数回帰したときの傾きが１より大きいか，小さいかで単調増加か単調減少になる関数として表現されることを明らかにしている。したがって，per capita で表した統計指標は，多くの場合，人口に対して必然的に正や負の相関を生じることになる。

以上のベキ関数で表されるような二つの統計指標の性質として「法則　４」が次のように導かれる。これは，「法則　２」の特別な場合と考えて良い。

（２），（３）の両辺を x でわった２変量（$W = y_1/x$，$Z = y_2/x$）は，

$$y_1/x = \alpha_1 * x^{(\beta_1-1)} = W \quad (6)$$

$$y_2/x = \alpha_2 * x^{(\beta_2-1)} = Z \quad (7)$$

これら二つの式から W と Z は，次の式で表される。

$$W = c * Z^{\rho}\,;\ \rho = (\beta_1-1)/(\beta_2-1),$$
$$c = \mathrm{Log}\,\alpha_1 - \rho * \mathrm{Log}\,\alpha_2 \quad (8)$$

すなわち，両辺を x でわった２変量（$W = y_1/x$，$Z = y_2/x$）は，たがいにベキ関数として表される。Z のベキ乗を表している ρ は正から負の値をとり，単調増加関数か一定値関数か，単調減少関数の性質を示す。ρ の値は，弾性値 β が１より大きいかどうか，つまり，比較する二つの統計指標が人口や GDP にたいして規模の効果が１より大きいか，小さいかで正か負になる。一般に，β がともに１より大きいか，ともに１より小さいと ρ は正，したがって W は Z の単調増加関数になるので正の相関として表される。どちらかが１より大きく，他方が１より小さいと負，したがって単調減少関数になるので負の相関として表されることになる。つまり，per capita や率（％）で表された統計指標どうしの回帰分析をすることは，（８）

表 3 各種統計指標を per capita で表してたがいに相関をとった傾き（β）と決定係数（R^2）の値

被説明変数

説明変数	老人福祉施設	行政投資	一般公共事業費	死亡数2005	失業者数	県民所得	生活保護者数	犯罪認知件数
GDP	−0.5265 0.1803	−0.4644 0.1004	−0.3049 0.0464	−0.2258 0.1036	−0.3740 0.1277	0.6885 0.7699	−0.2829 0.0104	0.5999 0.1338
老人福祉施設		0.9046 0.5855	0.7947 0.4845	0.4814 0.7239	−0.1176 0.0194	−0.3845 0.3691	−0.3964 0.0314	−0.9395 0.5044
行政投資			0.8880 0.8457	0.3213 0.4507	−0.0353 0.0024	−0.2822 0.2779	−0.1584 0.0070	−0.7618 0.4635
一般公共事業費				0.3302 0.4438	−0.0967 0.0171	−0.2286 0.1700	−0.2267 0.0134	−0.6878 0.3523
死亡数					−0.1855 0.0155	−0.5982 0.2860	−0.1643 0.0017	−1.5407 0.4343
失業者数						−0.3196 0.1817	1.7867 0.4540	0.2037 0.0169
県民所得							−0.6476 0.0335	1.0761 0.2650
生活保護者数								0.1352 0.0523

上の値が傾き β
下の値が決定係数 R^2

式で示したベキ関数関係を表しただけなのである。導かれた相関関係は，per capita や率（%）で表された任意の統計指標間で因果関係の有無にかかわらず必然的に相関現象が惹起されることを意味している。

表 2 で示した 8 つの統計指標について per capita，つまり人口あたりの数値で表した指標数値間でたがいに 1 次回帰すると表 3 で表すような回帰係数と，決定係数が得られる。ただし比較しやすいように，すべての統計指標値の per capita の平均値がほぼ50になるように調節している。表 3 の最上段には，参考のため GDP（県民総生産）との相関を示している。可能な28の組み合わせのうち，有意な相関を示す決定係数が得られる場合（R^2がほぼ0.1より大きい場合）は17あり，「県民所得と失業者数」の相関以外は，すべてにおいて先に述べた（8）式が示す，ρ の値のちがいに対応した予測に一致した正や負の相関を示している。つまり，ρ が 1 より大きい失業者数，県民所得，生活保護者数，犯罪認知件数と，ρ が 1 より小さい老人福祉施設数，行政投資費，一般公共事業費，死亡数との組み合わせでは相関係数が負，ρ が 1 より大きい指標と 1 より大きい指標との間では相関係数が正，ρ が 1 より小さい指標と 1 より小さい指標との間では相関係数が正となっていることが表 3 を見ると明らかである。

4．見せかけの相関

経済世界では，ここで分析したような“必然的な相関図”を示して，さらなる“因果関係”を十分検討せず，政策論や政策提言に議論を運ぶ場合がよく見られる。例えば，表 2 のそれぞれ per capita で表した一般公共事業費と県民所得の負の相関図（図 1（1））は，実に負の相関を強く印象づけ，「いずれの時点においても，所得水準と公共投資依存度の間には負の相関があり，公共投資が所得水準の低い地域に手厚く配分されてきたことがうかがわれる」（中里 1999）と結論づけたくなる。また，公共事業費や行政投資と県民所得に見られる負の相関ついての同様な議論は，例えば河野（2001）や赤井（2001）などで見られるが，すべて「相関＝因果関係」と即断して，単なる“見せかけの相関”である可能性について論じていない。

表 3 で検討した，per capita や率（%）で表された統計指標値の任意の二つ（「統計指標 1 」，「統計指標 2 」）と「人口」，「GDP」を 4 種の観測変数として，最も単純な逐次モデルで共分散構造分析を行った。回帰分析によって有意な相関が検出されていた17の組み合わせのうち共分散構造分析において有意なパスが検出されたのは〈県民所得〉と〈失業者数〉との間だけであった（図 2

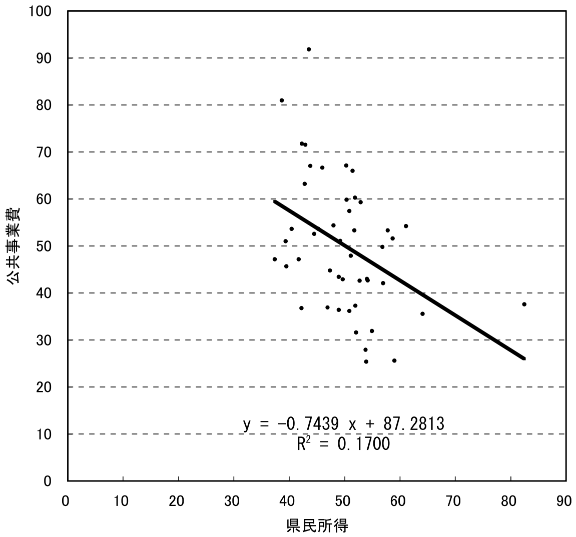

図１（１） per capita で表した公共事業費と県民所得との相関

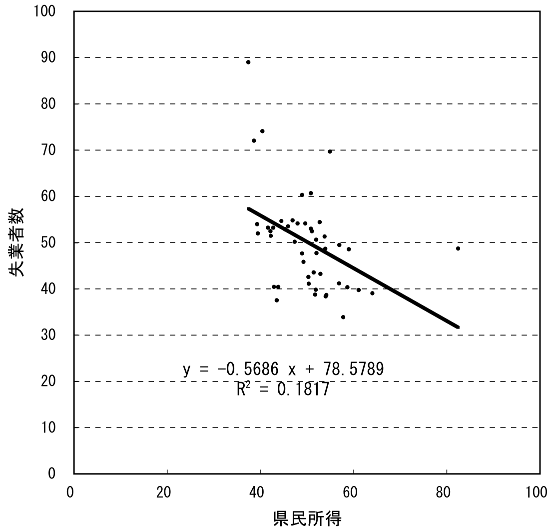

図１（２） per capita で表した失業者数と県民所得との相関

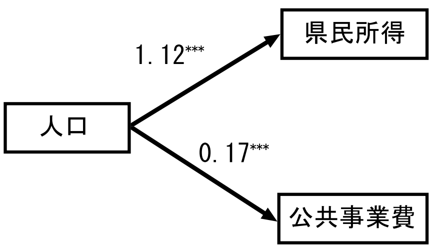

図２（１） 「人口」,「県民所得」,「公共事業費」のパス分析図
パス係数は，AMOS で計算した標準化係数値で，***は p＜0.01。
「県民総生産」は，有意なパスが検出されなかったので図には入れていない。
モデル適合度は0.94で，GFI＝0.99，AGFI＝0.99，NFI＝1.00，AIC＝10.01。

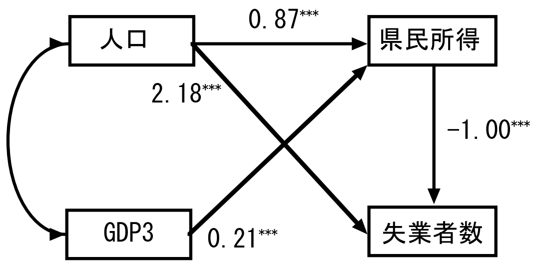

図２（２） 「人口」,「第３次産業の県民総生産」,「県民所得」,「失業者数」のパス分析図
パス係数は，AMOS で計算した標準化係数値で，***は p＜0.01。
「県民総生産」としては，「第３次産業の県民総生産」を用いた。
パス係数は，AMOS で計算した標準化係数値である。
モデル適合度は0.39で，GFI＝0.99，AGFI＝0.92，NFI＝0.99，AIC＝18.73。

（２））。つまり，〈県民所得〉と〈失業者数〉との間には有意な負の効果がはたらいていることが明らかになった。しかし，図１（１）で示した〈一般公共事業費〉と〈県民所得〉の間には有意なパスは検出できなかった（図２（１））。したがって，per capita の相関で有意な相関を示した17例のうち，「ベキ関数の法則」からはずれていた１例においてだけ有意な効果が検出されたことになる。per capita の統計指標値間で得られた有意な相関関係はほとんどが“見せかけの相関”であると判断される。「公共投資が所得水準の低い地域に手厚く配分されてきたことがうかがわれる」などと結論づけることはとうてい出来ない結果である。

５．考　　察

自然科学においては分析対象となる観測値の性質は十分検討されて研究解析がなされているが計量経済学や一般経済学において分析対象となっている統計指標についてはどうであろうか。本論で提起した per capita や率（％）で表された統計指標値間の“見せかけの相関”の問題については回帰分析の祖，Karl Pearson が生物器官のアロメトリーにおいて指摘していたようであるが（Kuh and Meyer 1955 による），経済学や計量経済学に

おいては不問にされてきたとしか言いようがない。唯一，Maddala（2001）が回帰分析における「分散の不均一性」の問題として論じている。しかし，「分散の不均一性」を解消する方法としてMaddala（2001）が提示したのは，１）対数変換すること，２）デフレーターで除した値で分析することである。後者の「デフレーターで除する」というのは，本論で問題点を指摘している“per capita や率（%）で表す”ということにほかならない。Maddala（2001）は Kuh and Meyer（1955）が指摘した“見せかけの相関 spurious correlation”についてふれてはいるが，本論で指摘したような経済統計指標が内在する性質（ベキ関数を示す性質）とその重要性については全く気づいていない。おそらく，経済研究者の多くは「任意の経済指標値は人口や GDP で除することによって無名化され，どんな分析にでも値する」と暗黙のうちに了承しているのではないだろうか。しかし，「人口や GDP で除すること」自体が暗黙のうちに「人口効果」や「GDP 効果」を織り込んでいることになるのである。

ここでは，都道府県の統計資料等からなるクロスセクションデータについて考察したが経済世界のアロメトリーは，時系列データでも多く見られる可能性がある。経済学で有名な「フィリップ曲線」は，インフレ率（%）と失業率（%）の座標軸で示され，多くの場合，負の相関が現れる。経済活動の時系列的成長に対して規模の効果が異なる二つの統計指標を率（%）で表して相関を取ると単調減少関数か単調増加関数になる可能性が強いことは，本論の解析で容易に推察される。スティグリッツ（2001）の経済学テキストや山本（1995）の計量経済学の教科書に示されているアメリカと日本におけるフィリップ曲線が減少関数としてシフトしていく過程は，単純に，率（%）で表した二つの経済指標間の相関が，異なる経済成長時期に対応してシフトした結果としても容易に理解されるのである。本論のper capita の比較分析の中で唯一例外であった〈県民所得〉と〈失業者数〉の間で見られる負の有意な相関（図１（２），図２（２））は，クロスセクションデータにおける「フィリップ曲線」を示していると考えることができる。

ここで示した「経済世界のアロメトリー」現象が表す課題は，“見せかけの相関”にとどまらず“誤った順位付け”にまで発展してしまうおそれがある（日下部 2002a，2002b，2002c)。クロスセクションデータ分析を行う研究者の目的は，相関関係を検出するだけでなく，その相関に基づいて調査対象域を順位づけて政策提言へ発展させようとするところにも見られる。この時，“見せかけの相関”にもとづいて順位付けられた評価は，人口や GDP の大きな地域は過小評価（または過大評価）され，逆に小さな地域は過大評価（または過小評価）されるような結果になる傾向にある。したがって，per capita や%などの経済指標を用いて分析すると，本来の調査研究者の意図とは異なって，単なる「人口効果（都会度，地方度)」を測定するだけの“誤った”統計インデックスになってしまうことになる。相関をもとにして順位づけることには大きな問題があるのである。

近年，国連開発計画や OECD，世界銀行，国内では内閣府などを中心として各国および国内で調査した各種データを用いてさまざまな総合化インデックスがつくられ，これらの値に立って各国各地域が評価され様々な社会課題が論議されている。しかし，これらのインデックスには本論で展開したような“見せかけの相関”と“誤った順位付け”が，per capita 表示や率（%）表示することによって無意識のうちに埋め込まれているということに経済世界の調査研究者はもっと注意をはらうべきであろう。

例えば，ここ数年，Putnum（1992，2000）らが主張しているソーシャル・キャピタル論が注目されてきて，国内では内閣府の主導のもと“ソーシャル・キャピタル・インデックス”なるものが作られ，国内地域が評価されるようにまでなってきている（例えば，稲葉2007の図10－１，10－２など)。しかし，そもそも Putnum（1992，2000）が立論している統計データ分析は因果関係分析を抜きにした，南北勾配をもつイタリアやアメリカ地域の単なる相関分析にしか過ぎない。本論で述べたように“単なる都会度，地方度”を見ているだけでしかない可能性が強いのである。調査研究

者の意図に反して，目的とする効果を測定しているのではなく，単なる“人口効果（都会度，地方度）”を見ている可能性が強い。これらの困難さを克服する統計量として著者（日下部　2002a，2002b，2002c）は回帰偏差値を提唱した。

本論では「ベキ関数の法則」として，４つにまとめたが，代数学的には「独立変数を共有する二つのベキ関数はたがいにベキ関数として表される」という単なる一つの規則の変形にすぎない。代数的には自明のことであるが，生物学や経済学が関わる自然社会現象においてはその持つ意味は大きい。per capita の統計指標間で検出される相関のうち，どれが“見せかけの相関（spurious correlation）”でないのか，判断することは難しい。しかし，ベキ関数として表現される多くの「経済世界のアロメトリー」現象は必然的に相関を生み出すのである。このような統計数値を基本的分析対象としている計量経済学や経済学一般は，まずもってこの課題を解決すべきであろう。これが解決されない限り計量経済学のよって立つ多くの線形分析は，その有効性に疑問が生じるのである。

参考文献

赤井伸郎（2001）「地方交付税削減が不可欠」『日本経済研究センター会報』pp. ４－７，11．15

朝日新聞社（2007）『民力2007』

稲葉陽二（2007）『ソーシャル・キャピタル』生産性出版

河野龍太郎（2001）「90年代の公共投資は景気対策ではなかった！」『金融ビジネス』JUN，76－80.

日下部眞一（2002a）「回帰偏差値の考案とその効用：地域間格差を相対評価する偏差値」広島大学総合科学部紀要Ⅳ　2002（理系編），109－126.

日下部眞一（2002b）「NPO の規模を規定する要因の解析と“回帰偏差値”による地域 NPO セクターの規模の相対評価」広島大学総合科学部紀要Ⅱ　2002，35－53.

日下部眞一（2002c）「NPO の規模をはかる回帰偏差値，“NPO 指数”の考案」The Nonprofit Review, Vol. 2, No. 2，2002，177－185.

日下部眞一（2004）「地域経済格差の正しい理解のために（３）行政投資と擬似相関」『統計』2004年２月号，pp. 74－82.

（財）矢野恒太記念会（2007）『データでみる県勢』

清水三雄（1959）『相対成長』協同医書出版

東洋経済新報社（2009）『地域経済総覧2009』

中里　透（1999）「公共投資と地域経済成長」『日本経済研究』39，pp. 97－115.

山本　拓（1995）『計量経済学』新世社.

Kuh, E. and J.R. Meyer（1955）Correlation and Regression Estimates When the Data Are Ratios, Econometrica, 1955, pp. 400–416.

McMahon, T.A. and J.T. Bonner（1983）*On size and Life*. Scientific American Books, Inc.

Maddala, G.S.（2001）*Introduction to Econometrics*. John Wiley & Sons.

Putnam, R.（1992）*Making Democracy Work*. Princeton Univ. Press.

Putnam, R.（2000）*Bowling Alone*. Simon & Schuster.

Stiglitz, J.E.（1997）Economics, 2nd Edition. W.W. Norton & Company, Inc. （『スティグリッツ　マクロ経済学　第２版』東洋経済新報社，2001）

Wooldridge, J.M.（2010）*Econometric Analysis of Cross Section and Panel Data, 2nd Edition*, MIT press.