サンプル・セレクション・バイアスを除去した リピート・セールス価格指数の計測

Repeat-sales price index estimates based on correcting sample selection bias

○ University of Toyama　Koji Karato　富山大学　唐渡　広志
Reitaku University　Chihiro Shimizu　麗澤大学　清水　千弘

This paper presents an improved estimation model with intertemporal covariance of error terms between selection function and repeat sales regression model to correct sample selection biases caused by repeat sales method, and expands an methodology in Gatzlaff and Haurin (1997). Empirical analysis of housing market in Setagaya-Wards Tokyo selling from 1986 to 2009 show that the estimated coefficient of weighted inverse Mills ratio has significantly negative. The result indicate that repeat sales price index is downward biased with sample selection.

キーワード：住宅価格指数，サンプル・セレクション，リピート・セールス法
keywords: house price index, sample selection, repeat-sales method

1. はじめに

これまで住宅価格指数を推定するために，いくつかの手法が検討されてきた. なかでも，Rosen (1974) によって理論的に整理されたヘドニック法や, Bailey, Muth and Nourse (1963), Case and Shiller (1987, 1989) によって精緻化されたリピート・セールス法（以下 RS 法とよぶ）は最も利用されている手法である．ヘドニック法は, 英国政府や Halifax 等の英国を代表するモーゲージバンクによって採用されている．米国では，シカゴ・マーカンタイル市場に上場されている住宅価格指数であるスタンダード・アンド・プアーズ・ケース・シラー価格指数に代表されるように, RS 法が利用されている．

ヘドニック法とは，ある不動産価格をその不動産のさまざまな属性（性能や機能，購入者や販売者の特徴）の価値に関する集合体（属性の束）とみなし，主に回帰分析を利用してそれぞれの属性価格を推定する手法である．プーリング・データを利用すれば，品質調整された価格指数を計測することができる．これに対して RS 法は，複数回売買された商品をサンプルとして選び，ペアになった同一商品の異時点間の対数差分価格を取引時点の時間ダミー変数に回帰させることで，価格指数を推定する．

ヘドニック法を利用した不動産市場の実証分析では，住宅の機能や特徴，あるいは住宅購入者の所得やタイプを説明変数として利用したものが多い（例えば，Sirmans *et al*. 2006 によるサーヴェイ）．しかしながら，本来は Rosen (1974) が提案しているように購入者だけでなく，販売者の経済行動も市場価格の決定に影響する．

住宅が売買される財として市場に登場するためには，販売者にとってのオファー価格が彼の留保価格を凌駕しているという条件を想定することは自然である．Gatzlaff and Ling (1994) および Gatzlaff and Haurin (1998) は，もしも住宅売買の意思決定が，オファー価格と留保価格の決定に影響するならば，実際に販売された住宅サンプルはランダム・サンプリングでない可能性があることを検証している．すなわち，実際に観察される取引価格は

オファー価格と留保価格を生ぜしめた確率過程に依存していると考え，Heckman (1979) による 2 段階推定法 (Heckit) を適用したセレクション・バイアスの除去を行っている．

RS 法もまた, そのデータ発生過程はヘドニック回帰モデルに依存しているため同様の問題が生じる．1 回目と 2 回目の販売時点において，販売者のオファー価格が留保価格を上回っている場合においてのみ，ペアになったデータ・セットとして観察されるため，セレクションされたサンプルを利用した分析にならざるを得ない.

Gatzlaff and Haurin (1997) におけるセレクション・バイアスの修正は Bailey *et al*. (1963) で提案された最もシンプルな RS 回帰モデルを基本として，Heckit を適用している．同論文では，セレクション関数と RS 回帰モデルとの間の共分散が通時的に一定であると仮定している．しかしながら，住宅の売り手が直面する経済状況は年々刻々と変化しているわけであるから，実際には誤差構造が変化している可能性がある.

本研究では，この点を改善してセレクション・バイアスを修正した RS 価格指数の推定方法を提案する．Gatzlaff and Haurin (1997) の最も簡単な拡張として，Case and Shiller (1987, 1989) のように誤差項におけるランダム・ウォークを仮定し，取引期間を長期化すればするほど，誤差分散が拡大するケースでのセレクションの除去を行う. この場合, Gatzlaff and Haurin (1997) のように RS 回帰モデルとセレクション関数との間の共分散が通時的に一定であったとしても，ランダム・ウォークによって生じる分散不均一を修正した重み付き RS 回帰モデルにおける逆ミルズ比も修正されなければならない.

以下では，次節においてセレクション・バイアスを制御した重み付き RS 回帰モデルを定義し，3 節において利用するデータの特徴について述べる．4 節において推定結果を示し，セレクション・バイアスの有無を検証する．5 節において結論を述べる.

2. モデル

2.1 リピートセールス価格指数

物件 i の取引時点 t 期における取引価格を $PRICE_{it}$ の対数値を $p_{it} = \log PRICE_{it}$，その物件の属性ベクトルを $\mathbf{x}_i$ とする．取引価格の対数値を属性ベクトルに回帰させたヘドニック・モデルを次のように書く.

$$p_{it} = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_{it}\,,\quad \varepsilon_{it} = \alpha + \delta_t + \upsilon_{it} \qquad (1)$$

$\boldsymbol{\beta}$ は属性のパラメタ・ベクトル（属性価格）である．ε_{it} は誤差項を示しており，モデル全体の定数 α，取引時点の時間効果 δ_t，撹乱項 υ_{it} で構成されているものとする.

Case and Shiller (1987, 1989) の RS 法は撹乱項において次のランダム・ウォークを仮定している.

$$\upsilon_{it} = \upsilon_{it-1} + u_{it}\,,\quad u_{it} \sim N\left(0, \sigma_u^2\right) \qquad (2)$$

ここで，u_{it} はホワイト・ノイズを仮定した系列相関のない互いに独立な確率変数である.

Bailey *et al*. (1963) および Case and Shiller (1987, 1989) は次の仮定を設けて，(1)を再定式化している．[仮定 1]. すべての属性は時間通じて不変である．[仮定 2]. すべての属性パラメタは時間通じて不変である. RS 法は同一物件の 1 回目と 2 回目の取引価格の対数差分を回帰モデルで検討する分析手法である．したがって，二つの仮定がある場合は 2 時点間の価格差は個々の物件の属性やその価格の変化ではなく，市場全体に共通して生じる変化（時間効果）によって説明することができる.

いま，1 回目の取引が行われた時期を t 期，2 回目の取引が行われた時期を $t+S$ 期としよう．2 時点間の物件 i の価格の対数差分は

$$y_i = p_{i,t+S_i} - p_{it} = \delta_{t+S_i} - \delta_t + e_i, (i=1,2,\cdots,n)$$

と示すことができる．ここで，誤差項は $e_i = \upsilon_{i,t+S_i} - \upsilon_{it}$ であり，(2) より $\upsilon_{i,t+S_i} = \upsilon_{it} + \sum_{j=0}^{S_i-1} u_{i,t+(S_i-j)}$ と変形できるので，その平均は $E(e_i)=0$，分散は $Var(e_i) = S_i\sigma_u^2$ となる．

ヘドニック回帰モデル (1) において時間効果を推定するには，取引時点に対応したダミー変数を利用する．したがって RS 回帰モデルにおいても時間効果を推定するために 2 時点間のダミー変数の差分を利用する．このことを考慮すると，クロス・セクション方向に積み上げた行列表示の RS 回帰モデルは次のように書ける．

$$\mathbf{y} = \mathbf{D}\boldsymbol{\delta} + \mathbf{e} \tag{3}$$

ここで，$\mathbf{y} = (y_1,\cdots,y_i,\cdots,y_n)'$，$\mathbf{D}$ は各取引時点に対応した 2 時点のダミー変数の差分の行列であり，1 回目の取引時点のとき –1，2 回目の取引時点のとき 1，それ以外のとき 0 となる．また，$E(\mathbf{ee}') = \sigma_u^2\overline{\mathbf{S}}$，ただし

$$\overline{\mathbf{S}} = \mathrm{diag}[S_i] = \begin{pmatrix} S_1 & & 0 \\ & \ddots & \\ 0 & & S_n \end{pmatrix}$$

であるから，誤差分散は取引時点の間隔に依存しており分散不均一である．Case and Shiller (1987, 1989) はこれを修正して，推定量の効率性を高めるための一般化最小 2 乗推定を提案している．すなわち重み行列

$$\mathbf{Q} = \mathrm{diag}[S_i^{-1/2}] = \begin{pmatrix} 1/\sqrt{S_1} & & 0 \\ & \ddots & \\ 0 & & 1/\sqrt{S_n} \end{pmatrix}$$

を利用して，(3) の両辺に $\mathbf{Q}$ を乗じた次の重み付き RS 回帰モデルを推定している．

$$\mathbf{y}^* = \mathbf{D}^*\boldsymbol{\delta} + \mathbf{e}^* \tag{4}$$

ここで，$\mathbf{y}^* = \mathbf{Qy}$, $\mathbf{D}^* = \mathbf{QD}$ および $\mathbf{e}^* = \mathbf{Qe}$ である．$\overline{\mathbf{S}}^{-1} = \mathbf{QQ}'$ であるから，$E(\mathbf{e}^*\mathbf{e}^{*\prime}) = \sigma_u^2\mathbf{Q}\overline{\mathbf{S}}\mathbf{Q}' = \sigma_u^2\mathbf{I}$ となり，最小 2 乗法が適用できる．

時間効果ベクトル $\boldsymbol{\delta}$ の一般化最小 2 乗推定量は $\hat{\boldsymbol{\delta}} = \boldsymbol{\delta} + (\mathbf{D}'\overline{\mathbf{S}}^{-1}\mathbf{D})\mathbf{D}'\overline{\mathbf{S}}^{-1}\mathbf{e}$ より，$E(\hat{\boldsymbol{\delta}}) = \boldsymbol{\delta}$ であるから，基準時点を s 期とする t 期の RS 価格指数は，(4)の推定量を利用して

$$I_{t/s}^{RS} = PRICE_t / PRICE_s = \exp(\hat{\delta}_t - \hat{\delta}_s)$$

から計測される．

しかしながら，取引時点のダミー変数は $\sum_{t=1}^{T} D_{it} = 0$ なる線型関係で結びつけられるので，多重共線性を考慮すると，すべての取引時点をダミー変数として利用できない．多くの文献にならって，取引の初期時点となるダミー変数 D_{i1} をモデルから落とし，$\delta_1 = 0$ を想定する．基準時点を $s=1$ とするとき，RS 価格指数系列は

$$I^{RS} = \{\exp(0), \exp(\hat{\delta}_2), \cdots, \exp(\hat{\delta}_T)\}$$

となる．

2.2 サンプル・セレクション

観測期間において物件は複数回取引されるものもあれば，1 度だけ取引されるものもある．したがって，RS 法には 2 回取引された物件だけを扱うことによって生じる，サンプリング・バイアスが存在する可能性がある．

Gatzlaff and Haurin (1997) は Bailey *et al.* (1963) による RS 法において，セレクション関数と RS 回帰モデルとの間の共分散が通時的に一定であると想定した上で，セレクション関数を導入することによって，セレクショ

ン・バイアスを取り除いた価格指数を計測している．しかしながら，住宅の売り手が直面する経済状況は年々刻々と変化しているわけであるから，実際には誤差構造が変化している可能性がある．本研究では，これに対応するために，前節で示された Case and Shiller (1987, 1989) による誤差項の分散不均一性を考慮し，セレクション・バイアスを修正した新しい RS 回帰モデルを提案する．

物件が市場で売買されるためには，売り手のオファー価格が留保価格を上回る必要があり，そしてそのときだけ取引価格が観察される．したがって，打ち切られたヘドニック価格の誤差分布の条件付き期待値が 0 でないことにより，ヘドニック価格にはセレクション・バイアスが生じる．

サンプル・セレクションがもたらすバイアスの修正方法は Heckman (1979) で示されている（Heckit とよばれる手法）．この方法は，実際には観察できない 2 回目の取引価格がない場合でも，時間効果の一致推定量をもたらす．いま，$t+S_i$ 期において住宅を売りに出すかどうかの選択を表す閾値を $Z^*_{i,t+S_i}$ とし，次の回帰式でその選択メカニズムが示されるものとしよう．

$$Z^*_{i,t+S_i} = \mathbf{w}'_{i,t+S_i}\boldsymbol{\gamma} + \nu_{i,t+S_i} \tag{5}$$

ここで，$\mathbf{w}_{i,t+S_i}$ は売り手の意思決定に影響すると考えられる住宅の属性やマクロ経済要因を表す変数ベクトル，$\boldsymbol{\gamma}$ は対応するパラメタ・ベクトルであり，$\nu_{i,t+S_i} \sim N(0,1)$ を仮定する．実際には，閾値 $Z^*_{i,t+S_i}$ は観察されないので，$Z^*_{i,t+S_i} > 0$ ならば $Z_{i,t+S_i} = 1$，それ以外の場合 0 となる 2 値変数を考え，これらの確率を

$$\Pr(Z_{i,t+S_i} = 1) = F(\mathbf{w}'_{i,t+S_i}\boldsymbol{\gamma})$$
$$\Pr(Z_{i,t+S_i} = 0) = 1 - F(\mathbf{w}'_{i,t+S_i}\boldsymbol{\gamma})$$

とする．ただし，$F(\mathbf{w}'_{i,t+S_i}\boldsymbol{\gamma})$ は標準正規分布の累積確率密度関数である．2 値変数はオファー価格が留保価格を上回るときだけ観察されるので，次のように定義することができる．

$$Z_{i,t+S_i} = \begin{cases} 1 & \text{if } \ p^O_{i,t+S_i} - p^R_{i,t+S_i} \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

いま，(3) における RS 回帰モデルの誤差 e_i と (5) におけるセレクション関数の誤差 $\nu_{i,t+S_i}$ が二変量正規分布にしたがい，その共分散が σ_{ev} であるものとしよう．観察される重み付き RS 回帰モデルは次の条件付き期待値を持つ．

$$\begin{aligned} E\left(y^*_i \,\middle|\, Z^*_{i,t+S_i} > 0\right) &= \mathbf{D}^{*\prime}_i\boldsymbol{\delta} + E\left(e^*_i \,\middle|\, \nu_{i,t+S_i} > -\mathbf{w}'_{it}\boldsymbol{\gamma}\right) \\ &= \mathbf{D}^{*\prime}_i\boldsymbol{\delta} + \sigma_{ev}\cdot S_i^{-1/2}\lambda_i \end{aligned}$$

ここで，$y^*_i = y_i \big/ \sqrt{S_i}$，$\mathbf{D}^*_i = \mathbf{D}_i \big/ \sqrt{S_i}$ および λ_i は (5) のプロビット推定から得られる逆ミルズ比である．セレクション・バイアスを制御した重み付き RS 回帰モデルは

$$y^*_i = \mathbf{D}^{*\prime}_i\boldsymbol{\delta} + \sigma_{ev}\cdot\frac{\lambda_i}{\sqrt{S_i}} + \eta_i \tag{6}$$

となる．

Heckit を利用したわれわれの推定プロシージャは次のようになる．

i. 2 回取引された住宅と 1 回だけ取引された住宅を合わせたサンプルについて，(5) のプロビット推定を行い，逆ミルズ比を計算する．このとき，2 回取引された住宅ならば $Z_{i,t+S_i} = 1$，それ以外は 0 となる 2 値変数を用いる．
ii. $1\big/\sqrt{S_i}$ で重み付けを行い，Case and Shiller 型の RS 回帰モデルのセレクション・バイアスを制御した (6) に最小 2 乗法を適用して時間効果 $\boldsymbol{\delta}$ および共分散 σ_{ev} 推定す

る．ここで，サンプルは 2 回取引された住宅だけを利用する．

なお，Heckit における 2 段階目の最小 2 乗法による推定値の標準誤差は，(6) の誤差項η_iの分散が

$$V(\eta_i | Z_i = 1) = \sigma_u^2 - \rho^2 \sigma_u^2 (d_i / S_i)$$

（ただし，$d_i = \lambda_i(\lambda_i + \mathbf{w}_i'\gamma)$）

となることから，別の分散不均一が生じてしまう．Greene (1981, 2008) より，次の手順で$\boldsymbol{\delta}$およびσ_{ev}の推定値の分散共分散行列の推定値を再計算する．はじめに，セレクション関数の推定結果より，$\hat{\lambda}_i$および$\hat{\boldsymbol{\gamma}}$を得るので，$\hat{d}_i = \hat{\lambda}_i(\hat{\lambda}_i + \mathbf{w}_i'\hat{\boldsymbol{\gamma}})$が定義できる．次に (6) の推定結果より，残差$\hat{\eta}_i$および修正された逆ミルズ比の係数推定値$\hat{\sigma}_{ev}^2$が得られるので，(2) の誤差分散が$\hat{\sigma}_u^2 = n^{-1}\left(\hat{\boldsymbol{\eta}}'\hat{\boldsymbol{\eta}} + \hat{\sigma}_{ev}^2 \sum \hat{d}_i\right)$より推定できる．また，誤差分散$\hat{\sigma}_u^2$と係数推定値$\hat{\sigma}_{ev}^2$より相関係数が$\hat{\rho}^2 = \hat{\sigma}_{ev}^2 / \hat{\sigma}_u^2$と計算できる．最後にこれらの結果を利用して，以下の分散共分散行列の推定値を計算する．

$$Var(\hat{\boldsymbol{\delta}}, \hat{\sigma}_{ev}) = \hat{\sigma}_u^2 (\mathbf{X}'\mathbf{X})^{-1} (\mathbf{X}'\boldsymbol{\Gamma}\mathbf{X} + \mathbf{R})(\mathbf{X}'\mathbf{X})^{-1}$$

ただし，$\mathbf{X} = (\mathbf{D}^*, \mathbf{Q}\hat{\boldsymbol{\lambda}})$，$\boldsymbol{\Gamma} = \mathrm{diag}\left[1 - \hat{\rho}^2 \hat{d}_i\right]$，$\mathbf{R} = \hat{\rho}^2 \mathbf{X}' \mathrm{diag}\left[\hat{d}_i\right] \mathbf{w} Var(\hat{\boldsymbol{\gamma}}) \mathbf{w}' \mathrm{diag}\left[\hat{d}_i\right] \mathbf{X}$である．

3. データ

本研究で利用するデータは，株式会社リクルートの情報誌である「週刊住宅情報」および「住宅情報タウンズ」に掲載された戸建住宅を対象として，東京都世田谷区において 1986 年 1 月から 2009 年 7 月までの期間に取引された物件である．情報が不完全なものや異常値と思われるものを除いた物件数は 30,916 である．このうち 1 度だけ取引された住宅は 28,131 件, 2 度取引された住宅は 2,785 件である．

2 度取引された住宅は RS 法のサンプルとして利用するが，本論文では時間効果を推定するための時間の単位を年次にするため，取引間隔が 1 年未満の物件を除いている．このことは，投機目的で短期的に売買が繰り返されるケースを排除して，できるだけ居住目的の価格指数に近い値を計測するための必要な措置である．

表 1. 住宅属性の記述統計

変数	単位	平均	標準偏差	最小	最大
Z		0.090	0.286	0	1
都心までの時間距離	[分]	12.8	5.3	1	25
占有面積	[m^2]	111.4	51.4	10.0	1475.4
土地面積	[m^2]	104.5	64.7	13.1	1648.1
最寄駅までの徒歩時間	[分]	10.0	4.8	1	36
建築後年数	[年]	5.6	8.4	0	69
取引年次	[年]	2000.6	5.3	1986	2009
竣工年次	[年]	1995.1	10.4	1929	2009

注．サンプル・サイズは 30,916 であり，このうち RS 法が適用される 2 度取引された住宅は 2,785 件である．その比率 (2 値変数 Z の平均) は 2785/30916＝0.090 である．

表 2. 取引価格の記述統計（単位: 万円）

		平均	標準偏差	最小	最大
取引が 1 回の物件		8,848	6,531	1,000	200,410
取引が 2 回の物件	1 回目	10,151	7,047	1,850	87,550
	2 回目	8,756	5,581	1,800	70,000

注．サンプル・サイズは取引が 1 回の物件が 28,131，取引が 2 回の物件は 2,785 件である．

表 1 は住宅属性の記述統計を示している．サンプル・サイズは 30,916 であり，このうち RS 法が適用される 2 度取引された住宅は 2,785 件である．その割合 (2 値変数 Z の平均) は 2785/30916＝0.090 である．『平成 18 年度国土交通白書』によると，既存住宅の流通量のシェアは米国で 77.6%, イギリスで 88.8%, フランスで 66.4%となっているのに対して，日本は 13.1%と極端に低い値になっている．2

度目の取引サンプルは確実に中古物件であるから，本研究で利用するデータは特に偏っているわけではない．

表 2 は取引価格の記述統計を示している．1 回だけ取引された住宅と 2 回取引された住宅では価格に大きな差異がある．また 2 回取引されたサンプルでは，1 回目に比べて 2 回目の価格は低くなり，ばらつきも小さくなっている．

表 3 はリピートセールス価格の取引年次別のサンプル・サイズ，価格，竣工年を示している．我々のデータベースは 80 年代後半の物件が若干少なく，再販売という性質上，2 回目は特に少ない．

表 3. RS 価格の年次別記述統計

	1 回目の取引			2 回目の取引		
取引年次	観測値の数	平均価格	平均竣工年	観測値の数	平均価格	平均竣工年
1986	27	6,436	1980	N.A.	N.A.	N.A.
1987	29	15,235	1980	4	17,220	1979
1988	14	15,734	1979	2	14,150	1982
1989	74	13,168	1980	12	10,351	1979
1990	61	16,739	1980	26	13,093	1979
1991	97	15,165	1981	27	15,816	1977
1992	61	16,124	1980	36	12,756	1981
1993	80	13,184	1985	46	11,971	1979
1994	84	9,132	1985	58	10,412	1983
1995	95	10,260	1985	77	7,932	1984
1996	90	9,931	1987	78	7,193	1984
1997	246	9,333	1991	85	9,601	1984
1998	189	9,645	1991	173	7,997	1990
1999	164	10,199	1991	187	8,406	1989
2000	164	9,735	1992	149	8,636	1989
2001	178	9,413	1995	200	7,688	1990
2002	207	8,094	1997	205	8,285	1994
2003	186	7,942	1999	227	8,247	1995
2004	168	9,926	2000	243	7,630	1996
2005	192	8,110	2001	206	9,227	1997
2006	120	10,223	2001	227	8,214	1999
2007	164	11,071	2004	160	10,110	1999
2008	95	8,547	2004	220	9,838	2000
2009	N.A.	N.A.	N.A.	137	7,615	2002

表 4. セレクション関数のプロビット推定

説明変数	推定値	t 値
定数項	−2.30	−15.91***
都心までの時間距離	0.00	1.68*
対数占有面積	0.13	2.83***
対数土地面積	−0.04	−1.18
最寄駅までの徒歩時間	0.00	−0.61
建築後年数	0.03	22.43***
完全失業率	0.09	6.58***
経済成長率	−0.02	−3.33***
観測値の数	30916	
対数尤度	−9057	
尤度比	605	[.000]

注．*** は 1%水準，** は 5%，* は 10% で有意であることを示している．尤度比は定数項を除くすべてのパラメタがゼロであるという帰無仮説についての検定統計量であり，[ ] 内はカイ 2 乗分布での p 値を示している．失業率および経済成長率は前年の値を利用している．

表 5. RS 回帰モデルの推定結果

	Eq. (6) SEL		Eq. (4) CS	
変数	推定値	t 値	推定値	t 値
D_1987	0.51	11.05 ***	0.49	10.55 ***
D_1988	0.52	8.74 ***	0.51	8.37***
D_1989	0.47	9.53 ***	0.45	8.83***
D_1990	0.54	10.81***	0.49	9.78***
D_1991	0.52	10.37***	0.47	9.22***
D_1992	0.35	6.90***	0.29	5.52***
D_1993	0.18	3.58***	0.11	2.11**
D_1994	0.08	1.57	0.00	−0.03
D_1995	−0.03	−0.64	−0.12	−2.38**
D_1996	−0.10	−2.02**	−0.20	−3.87***
D_1997	−0.08	−1.57***	−0.18	−3.48***
D_1998	−0.14	−2.61***	−0.26	−4.99***
D_1999	−0.18	−3.53***	−0.32	−6.20***
D_2000	−0.22	−4.24***	−0.36	−7.05***
D_2001	−0.27	−5.00***	−0.42	−8.08***
D_2002	−0.31	−5.85***	−0.48	−9.28***
D_2003	−0.33	−6.04***	−0.51	−9.86***
D_2004	−0.31	−5.57***	−0.51	−9.81***
D_2005	−0.28	−4.91***	−0.50	−9.52***
D_2006	−0.24	−4.13***	−0.49	−9.23***
D_2007	−0.18	−3.05***	−0.45	−8.47***
D_2008	−0.23	−3.82***	−0.54	−9.96***
D_2009	−0.33	−5.14***	−0.66	−12.08***
$\lambda/\sqrt{S}$	−0.02	−10.08***		
$(\hat{\rho})$	−0.1880			
$(\hat{\sigma}_u)$	0.1158		0.1353	

注．*** は 1%水準，** は 5%水準で有意であることを示している．(6) 式の推定値の標準誤差は Greene (1981) にしたがって再計算している．

4. 推定結果

表4は (5) 式のプロビット推定の結果を示している．占有面積が広く，建築後年数の古い物件は，売りに出す確率を有意に高めていることがわかる．また，失業率が高くなり，経済成長率が低くなる場合にも確率を高める要因になる．売りに出されるかどうかは，マクロ経済の景況に依存しており，景気の悪化は再販売の確率を高めると予想できる．

プロビット推定から逆ミルズ比λ_iが計測できる．λ_iの平均は1.75，標準偏差は0.20である. また, (6)の RS 回帰モデルにおいて, セレクション・バイアスを制御する$\lambda_i/\sqrt{S_i}$の平均は1.46，標準偏差は0.52になる．

表5にRS回帰モデル (6) および (4) の推定結果を示した．(6) はセレクション・バイアスを修正したRS回帰モデル (SEL) の, (4) はCase-Shiller 型の分散不均一のみを修正したRS回帰モデル (CS) の結果を示している．

セレクション・バイアスを制御した結果を見ると，重み付きの逆ミルズ比 $\lambda_i/\sqrt{S_i}$の係数推定値($\hat{\sigma}_{ev}$)は有意に負であり，相関係数も負になる．このことから，2回取引された物件固有の性質が価格の変化率に対して負の影響をもたらしていることが予想される．

セレクション・バイアスを制御していないCase-Shiller 型の推定結果を見ると, 時間効果はやや低い値で推定されている. このことは, 除外変数バイアスによる影響が考えられる．図1は時間効果の推定値より計算した1986年を基準年とした価格指数である．(4) の価格指数はやや下方にバイアスを持っていることがわかる．価格変化率は平均で−1.45%，最大で−2.94% だけ(6)に比べて過少になる．

仮に (6) の推定結果を正しいと考えてみよう．2回の取引が行われる物件だけをサンプルとして分析するとき，(4) の結果のように時間効果が負のバイアスを持つという予想にはどのような経済的意味が考えられるであろうか．Clapp and Giaccotto (1992) は，再販売される確率が高い住宅には, 粗悪な品質(いわゆる「レモン」）の住宅が混在していることを指摘している．本研究で利用したデータにおいても2回取引された物件にそのようなデータが混在している可能性がある。

また，Ooi (2006) や Pope (2008) で分析されているように，一度取引されたことのある中古住宅について，購入者が観察できない不確実性がある場合には，リスク回避的な購入者は期待されるよりも過少に住宅の質を評価する．他の事情が等しければ，住宅の確実性等価が低くなり，リスク中立的な購入者に比べて付け値が低くなる可能性がある．

住宅が再販売されるのは，上記のような理由があるからであり，本研究の結果はRSサンプルが母集団に比べ偏っている可能性があることを示唆している．

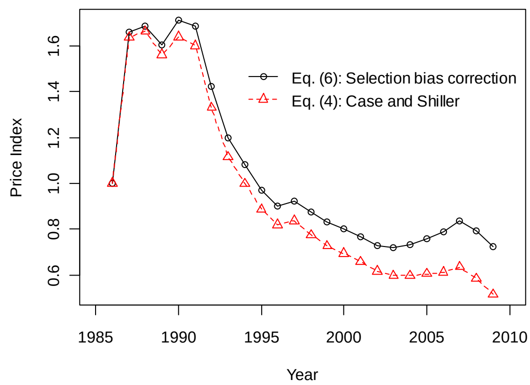

図 1. 価格指数（1986 年価格を 1 に基準化）

5. まとめ

住宅の売り手が直面する経済状況は，年々刻々と変化しているわけであるから，誤差構造が変化していると考えるのは自然な想定である．本研究は，Case-Shiller 型のリピート・セールス法がもたらすサンプル・セレクション・バイアスの修正において，セレクション関数とRS回帰モデルとの間の共分散を推定

するために逆ミルズ比の修正し，Gatzlaff and Haurin (1997) の改善を行った．

東京都世田谷区のデータによる分析結果より，セレクション・バイアスを制御した重み付きの逆ミルズ比の係数推定値は有意に負であり，相関係数も負になる．このことから，2回取引された物件の価格には、住宅市場の母集団に比べて負のバイアスが生じる可能性がある．

ただし，本研究では，セレクション関数と価格の誤差構造が，通時的に異なってしまう原因を Case and Shiller 型のランダム・ウォークに特定して分析を行った. RS 法のように方程式の階差をとる場合には，ランダム・ウォークは定常過程となり，分析が容易である．誤差構造の系列相関を想定する場合，もはや重み付きの回帰モデルを適用することはできない．この場合，尤度関数を特定化して部分最尤法を適用する必要がある．

また、データの制約から 1 度も取引されていない物件を含めることができなかった。実際には、そのような物件は非常に少ないと思われるが、セレクション関数を 1 回目の取引時点についても考慮して分析する必要がある。このような問題点は今後の課題である．

参考文献

[1] Bailey, M. J., R. F. Muth and H. O. Nourse (1963). “A Regression Model for Real Estate Price Index Construction,” *Journal of American Statistical Association* 58, 933−942.

[2] Case, K. E. and R. J. Shiller (1987). “Prices of Single Family Homes since 1970: New Indexes for Four Cities,” *New England Economic Review*, (Sept./Oct.) 45−56.

[3] Case, K. E. and R. J. Shiller (1989). “The Efficiency of the Market for Single−Family Homes,” *The American Economic review*, 79 (1), 125−137.

[4] Clapp, J. M. and C. Giaccotto (1992). “Estimating Price Trends for Residential Property: A Comparison of Repeat Sales and Assessed Value Methods,” *Journal of Real Estate Finance and Economics* 5, 357−374.

[5] Gatzlaff, D., and D. Ling. (1994). “Measuring Changes in Local House Prices: An Empirical Investigation of Alternative Methodologies,” *Journal of Urban Economics* 35 (2), 221-244.

[6] Gatzlaff, D. H. and D. R., Haurin (1997). “Sample Selection Bias and Repeat-Sales Index Estimates,” *Journal of Real Estate Finance and Economics* 14, 33−50.

[7] Gatzlaff, D. H. and D. R., Haurin (1998). “Sample Selection and Biases in Local House Value Indices,” *Journal of Urban Economics* 43, 199−222.

[8] Greene, W. H. (1981), “Sample Selection Bias as a Specification Error: Comment,” *Econometrica* 49, 795−798.

[9] Greene, W. H. (2008). *Econometric Analysis* 6th ed. (Chapter 24, 884-887), Prentice Hall.

[10] Heckman, J. (1979). “Sample Selectio n Bias as a Specification Error,” *Econometrica* 47, 153−161.

[11] Pope, J. C. (2008). “Do Seller Disclosures Affect Property Values? Buyer Information and the Hedonic Model,” *Land Economics* 84 (4), 551−572.

[12] Ooi, J.T.L., C.F. Sirmans and K.T., Geoffrey (2006). “Price Formation Under Small Numbers Competition: Evidence from Land Auctions in Singapore,” *Real Estate Economics* 34, 51−76.

[13] Rosen, S. (1974). “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition,” *Journal of Political Economy* 82, 34−55.

[14] Sirmans, G.S. and MacDonald, L. and Macpherson, D.A. and Zietz, E.N. (2006), “The value of housing characteristics: a meta analysis,” *Journal of Real Estate Finance and Economics* 33, 215-240.