

事例研究(ミクロ経済政策・問題分析 III)

- 規制産業と料金・価格制度 -

(第9回 - 手法(5) 応用データ解析/措置効果評価)

2011年 6月30日

戒能一成

0. 本講の目的

(手法面)

- 応用データ解析の手法のうち、措置効果評価 (Treatment Evaluation) の概要を理解する

(内容面)

- 計量経済学・統計学を実戦で応用する際の留意点を理解する (3)

1. 措置効果評価の基礎

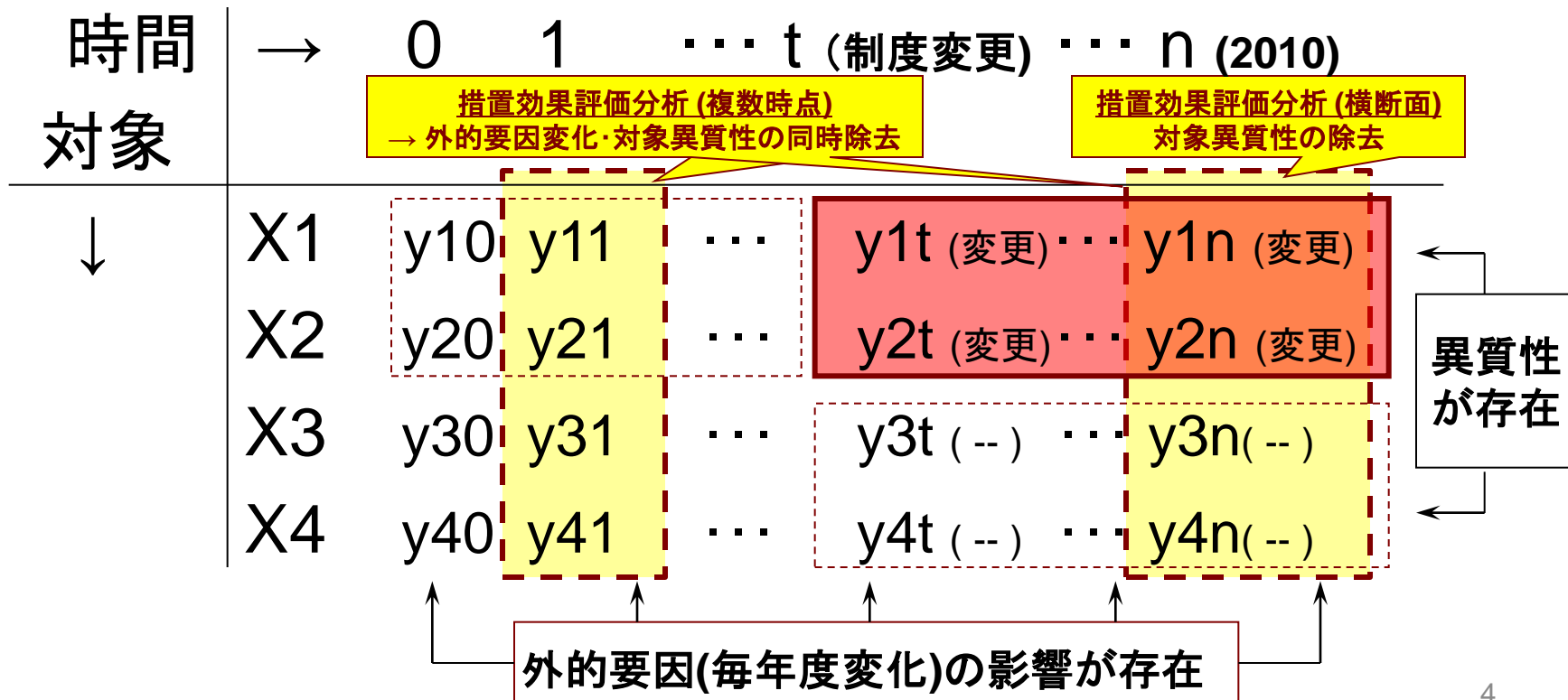
1-1. 措置効果評価(TE)の概念

- **措置効果評価 (Treatment Evaluation)** とは、対象の異質性を念頭に、措置(規制・補助・融資・教育訓練など)の新設・変更が、対象に及ぼす効果を推計するための一群の分析手法を指す
- 経済学分野では比較的新しい('90-)分野のため、用語・手法は様々で進歩・変化が著しい
- 措置効果評価では、通常は**対象が多数ある前提で横断面分析** 又は **パネルデータ分析**を行う(←対象が少数の場合 時系列分析 又は パネル)

1. 措置効果評価の基礎

1-2. 措置効果評価の要点

- 料金・価格制度の措置効果評価では、「異質性」(複数時点の場合 + 「外的要因」) の除去が必要



1. 措置効果評価の基礎

1-3. 措置効果評価の基本用語

D_i : 対象 i への**措置の有無** (= 1 or 0)

N_1, N_0 : 措置群($D_i=1$), 対照群($D_i=0$) の試料数

$y_i \equiv y_i(D_i) = y_i(1) \text{ or } y_i(0)$ (単に y_1 or y_0 も多用)

$y_i(1)$: 対象 i が措置群となった場合の結果指標

$y_i(0)$: 対象 i が対照群となった(= 措置されない)
場合の結果指標

$p(x)$: **措置率** Propensity score

条件 x に該当する全試料中、措置群となる
確率 (= $N_1 / (N_1 + N_0) | x$, 但し要推計の場合有)

1. 措置効果評価の基礎

1-4. 措置効果評価の前提条件

- 措置効果の独立性 “Unconfoundness”

→ 制度(変更)の影響が、措置群・対照群の間で独立と見なせること (影響の異質な対象だけ選分けて措置群としていないこと)

$$\Leftrightarrow (y_i(1), y_i(0) \perp D_i \mid x) \text{ for } \forall i$$

- 対照群の存在 “Overlap”

→ (条件 x を満たす) 試料中に、措置群・対照群が両方とも存在すること

$$\Leftrightarrow 0 < p(x) \text{ (措置率)} < 1$$

1. 措置効果評価の基礎

1-5. 平均措置効果・措置群平均措置効果

– **平均措置効果 (Average Treatment Effect: ATE)**

措置群と対照群の間の平均的な指標の差
(= 他の条件を一定とした措置の効果)

$$\rightarrow E(y(1) - y(0))$$

– **措置群平均措置効果 (ATE on treated: ATET)**

措置群において、措置が行われた前後の平均的な指標の差 (注意: 異時点比較のみ: 措置後の措置群には $y(0)$ が存在しない)

$$\rightarrow E(y(1) - y(0) \mid D = 1)$$

1. 措置効果評価の基礎

1-6. 平均措置効果の推計

- (条件 x を満たす) 平均措置効果 (ATE(x))の推計

$$\begin{aligned} \text{ATE}(x) &= E(y(1) - y(0) \mid x) \\ &= E(y(1) \mid x) - E(y(0) \mid x) \\ &= E(y(1) \mid x, D = 1) - E(y(0) \mid x, D = 0) \\ &\quad (\leftarrow \because (y_i(1), y_i(0) \perp D_i \mid x) \text{ for } \forall i) \end{aligned}$$

$$\begin{aligned} \therefore \text{ATE}^*(x) &= (1/N_1 * \Sigma_i (y_i(1)) - 1/N_0 * \Sigma_j (y_j(0)) \mid x) \\ &\quad (\text{措置群の平均値}) - (\text{対照群の平均値}) \end{aligned}$$

→ 本来は存在しない措置群の $y(0)$ を、措置効果の独立性条件から対照群の $y(0)$ で代用

1. 措置効果評価の基礎

1-7. 平均措置効果と結果指標

- (条件 x を満たす) 結果指標 y

$$\begin{aligned} y \mid x &= D * y(1) + (1 - D) * y(0) \mid x \\ &= D * (y(1) - y(0)) + y(0) \mid x \\ &= D * ([ATE^*(x)] + \varepsilon(1) - \varepsilon(0)) + y(0) \mid x \end{aligned}$$

$$\because y(1) = 1/N_1 * \sum_i (y_i(1)) + \varepsilon_i(1)$$

$$y(0) = 1/N_0 * \sum_j (y_j(0)) + \varepsilon_i(0)$$

→ y は 対照群の $y(0)$ に、措置群の平均措置効果と誤差($\varepsilon(1) - \varepsilon(0)$) が D の値(=1/0)に応じ **非線形に加算されたもの**と見なせる

1. 措置効果評価の基礎

1-8. 平均措置効果と手法選択 **[重要]**

- 措置群・対照群の選択が、観察可能な変数 z_i で決定される場合 ($D_i = g(z_i) = 1 \text{ or } 0$)
 - **ダミー変数(非線形)モデル**による分析
- 措置群・対照群の選択の変数 (z_i) が不明の場合
 - 措置前後の異時点の横断面データ利用可能
 - **DID分析** (Difference In Difference) 他
 - 措置ダミーを用いたパネルデータ分析
 - 措置後の横断面データのみ利用可能
 - 横断面分析 (Cross Section-)

2. 措置効果評価の手法(1)

2-1. ダミー変数(非線形)モデル(1) 二択モデル

- 離散値 D_i の選択が、ある観察可能な変数 z_i で決まる確率に従う場合、当該過程は **二択モデル** (Binary Outcome Model)が適用可

(離散値 D_i の選択)

$$D_i = \begin{cases} 1 \\ 0 \end{cases}$$

$$D_i = \Pr(D_i=1, z_i' \beta) + \varepsilon_i$$

$$\Pr(D_i=1, z_i' \beta) = \int_{-\infty}^{\infty} (z_i - z_0)' \beta / \sigma (2\pi\sigma^2)^{-1/2} * \exp(-1/2 * s^2 / \sigma^2) ds$$

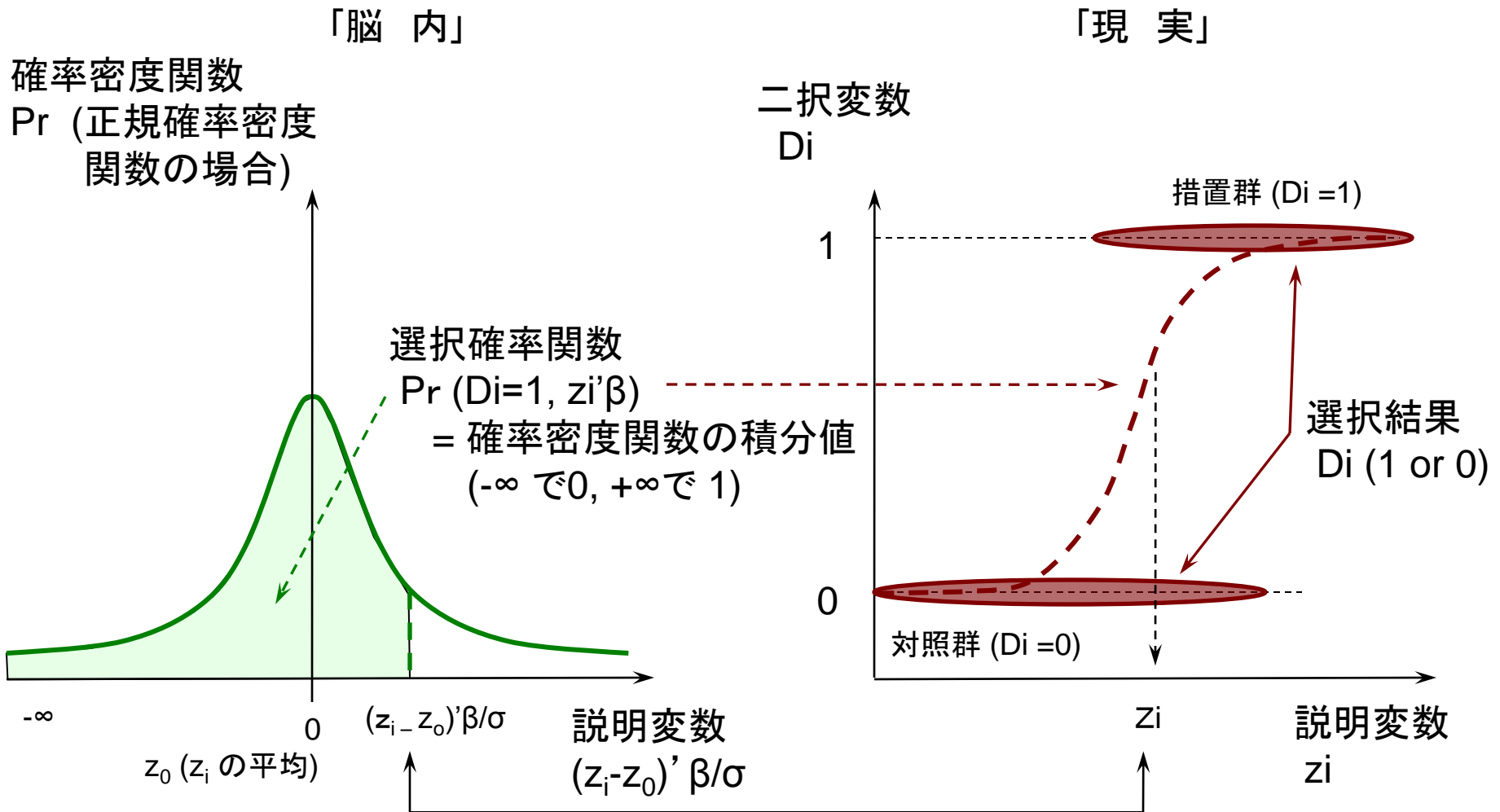
(Probit ; 正規確率密度関数 $\varphi((z_i - z_0)' \beta / \sigma)$ の積分)

$$\Pr(D_i=1, z_i' \beta) = 1 / (1 + \exp(-z_i' \beta))$$

(Logit; 対数確率関数 $\Lambda(z_i' \beta)$ (=「積分済」))

2. 措置効果評価の手法(1)

2-2. ダミー変数(非線形)モデル(2) 二択モデル概念



(例: D_i - 家計 i 太陽光発電装置購入の有無 z_i - 家計 i の所得)

2. 措置効果評価の手法(1)

2-3. ダミー変数(非線形)モデル(3) ダミー変数モデル

- 離散値 D_i の選択に応じ、 **$D_i = 1$ の場合のみ**
結果指標 y_i が z_i により決定される場合、ダミー
変数モデル(Dummy Dependent Model)が適用可

(第1段階: 離散値 D_i の選択)

$$D_i = \begin{cases} 1 & \text{if } D_i^* > 0; \\ 0 & \text{if } D_i^* \leq 0; \end{cases} \quad D_i^* = z_{i1}' \beta_1 + \varepsilon_{1i}$$

(通常 誤差 ε_{1i} は正規分布と仮定し Probit型で β_1 を推定)

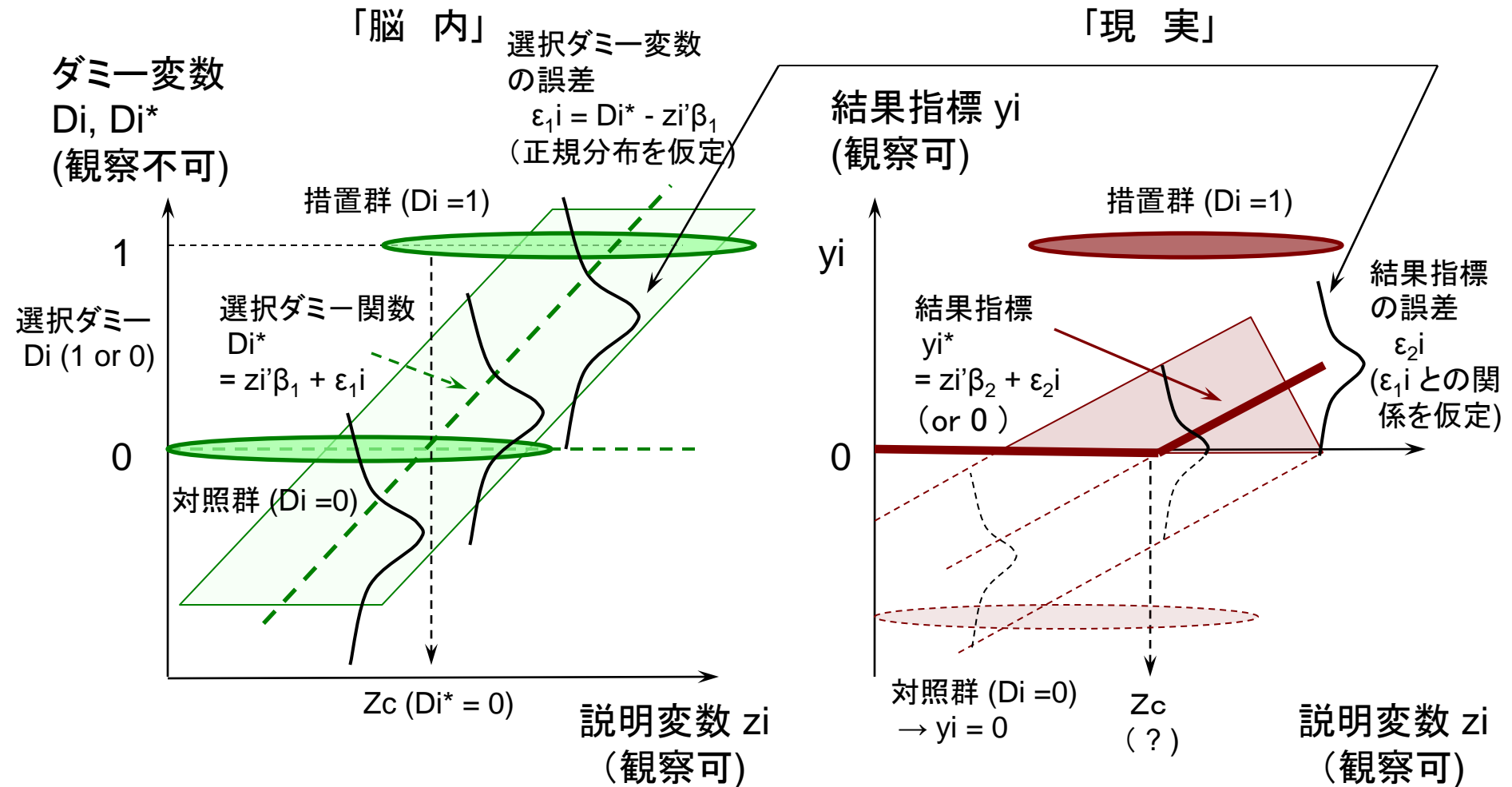
(第2段階: 結果指標 y_i の決定)

$$y_i = \begin{cases} y_i^* & \text{if } D_i^* > 0; \\ 0 & \text{if } D_i^* \leq 0; \end{cases} \quad y_i^* = z_{i2}' \beta_2 + \varepsilon_{2i}$$

(← 観察不能)

2. 措置効果評価の手法(1)

2-4. ダミー変数(非線形)モデル(4) 推計の概念



(例: y_i - 家計 i の太陽光発電量 D_i - 装置購入ダミー(観察不可) z_i - 家計 i の所得) ¹⁴

2. 措置効果評価の手法(1)

2-5. ダミー変数(非線形)モデル(5) モデルの種類

– Two Part モデル

第1段階を Probit型で推計し、第2段階で正の観察値のみ回帰推計
(= 第1段階・第2段階の確率や誤差の関係を仮定しないが、第1段階での
選択の有無 (= 第2段階が不存在か “0” が存在か) を識別する必要有)

– Tobit モデル (Type 2)

第1段階・第2段階の誤差が二元正規分布に従うと仮定し、第1段階の
Probit型推計の結果(回帰係数 β_1) を用い、第2段階を推計

– Heckman 2段モデル

第1段階・第2段階の誤差が線形関係と仮定し、第1段階の Probit型推計
の結果(回帰係数 β_1) を用い、第2段階で正の観察値のみを推計

$$E(y_i | D_i > 0) = z_i' \beta_2 + \sigma_{12} * \lambda(z_i' \beta_1) \quad \leftarrow \text{最尤値(ML)推計}$$

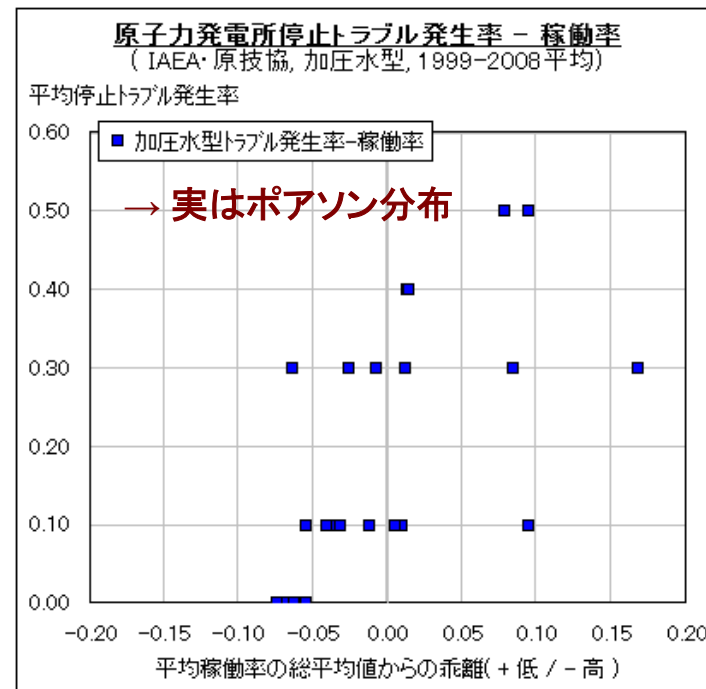
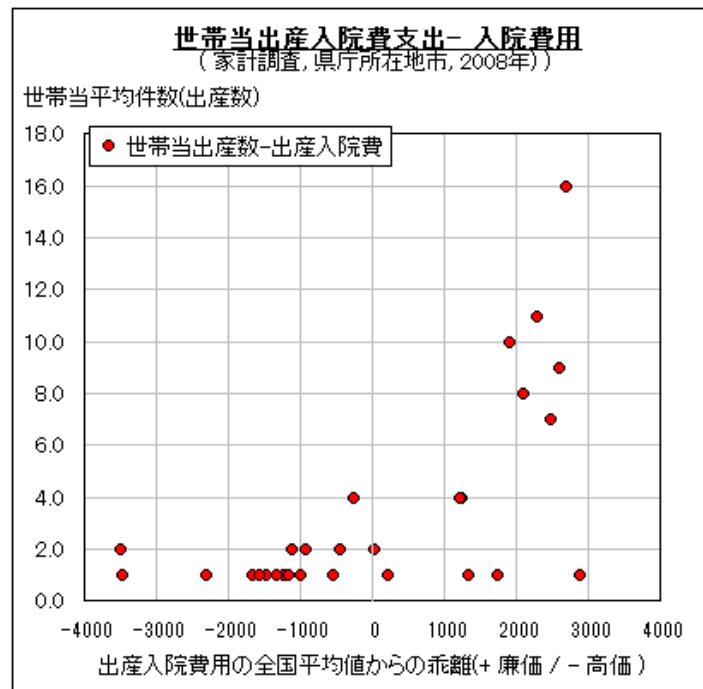
σ_{12} : 誤差間の線形回帰係数

$\lambda(z_i' \beta_1)$: 逆ミルズ比 正規分布確率密度関数 φ と確率の比
 $= \varphi(-z_i' \beta_1 / \sigma) / (1 - \int_{-\infty}^{-z_i' \beta_1 / \sigma} \varphi(s) ds)$

2. 措置効果評価の手法(1)

2-6. ダミー変数(非線形)モデル(6) 仮定と検定

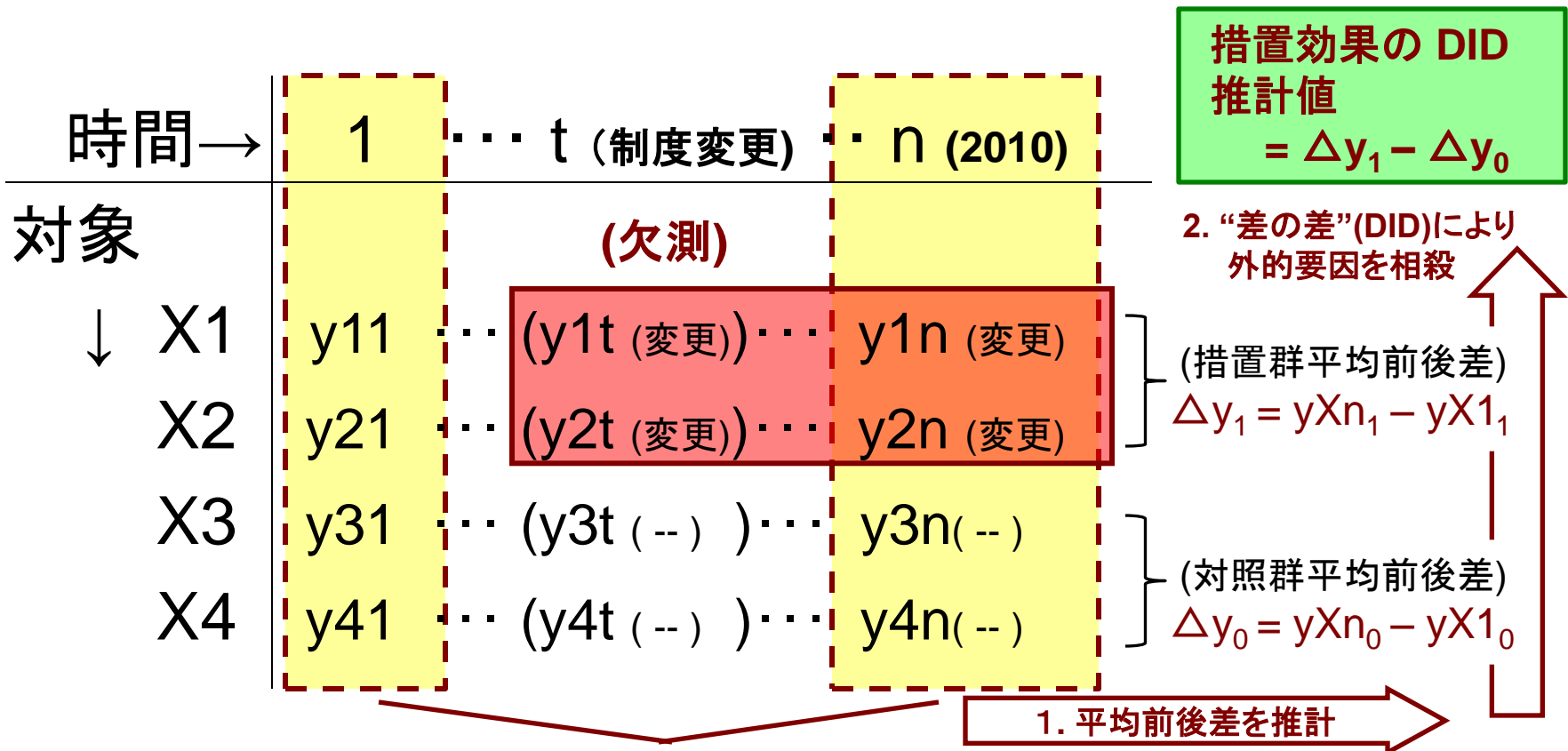
- ダミー変数モデルの多くは、少なくとも**第1段階の選択過程の誤差が正規分布に従う**と仮定
 - 誤差の正規性検定 (- linktestなど) が必須



3. 措置効果評価の手法(2)

3-1. DID (Difference In Difference) 分析(1)

- 措置の前後で 1つずつ、少なくとも 2時点の横断面データがあれば **DID分析** が可能



3. 措置効果評価の手法(2)

3-2. DID (Difference In Difference) 分析(2)

- DID推計では、時間とともに変化する **外的要因** δ (厳密には $\Delta\delta$) **が、措置群・対照群共通と仮定**
→ 措置群・対照群の外的要因に対する異質性が可能な限り排除されている必要有(後述)

$$y(i,t-h)_1 = \alpha + x_i' \beta + \delta(t-h) + \varepsilon(i,t-h)$$

$$-y(i,t+h)_1 = \alpha + x_i' \beta + \delta(t+h) + TE(i,t+h) + \varepsilon(i,t+h)$$

$$\Delta y(i)_1 = \Delta\delta + TE(i,t+h) + \Delta\varepsilon(i)$$

$$-\Delta y(j)_0 = \Delta\delta + \varepsilon(j)$$

$$\therefore E(\Delta y_1 - \Delta y_0) = E(TE(t+h) + \varepsilon_{DID}) = ATE$$

3. 措置効果評価の手法(2)

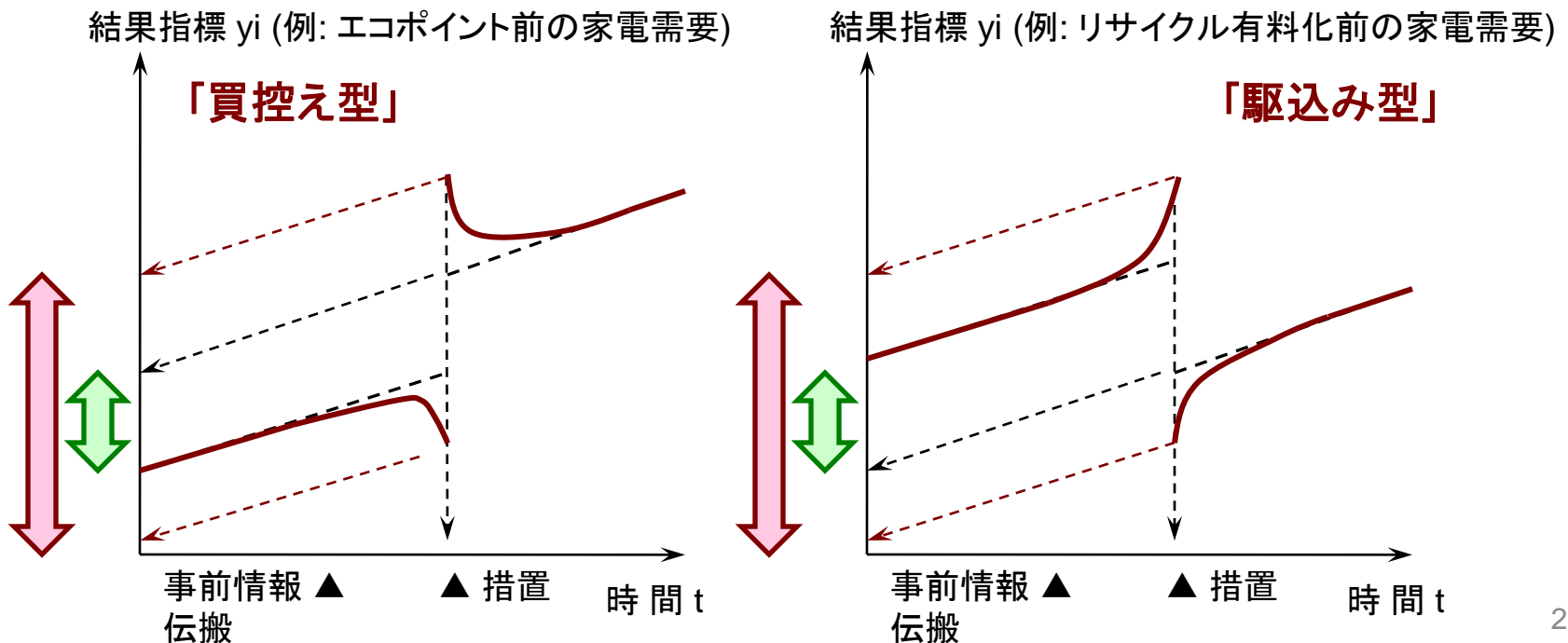
3-3. DID分析 と パネルデータ分析

- DID分析は、措置前後の 2期によるパネルデータ分析において **1階階差をとり、措置群にダミーを設定して分析した結果と同じ**である
 - 外的要因を階差の回帰分析で除くか、引算で除くか、という違い
- DID分析の結果は「**平均値の差**」の検定で判定
 - パネルデータ分析では措置群ダミー係数の有意性検定で判定
- 但し、多期のパネルデータが利用できる場合は、パネルデータ分析の利用が望ましい
 - 「**駆込み・買控え効果**」(Ashenfelter's Dip)

3. 措置効果評価の手法(2)

3-4. DID分析と「駆込み・買控え効果」(1)

- 観察する措置前後の2期で**措置効果が特異な時間変化**をしている場合に注意が必要
→ 特に「**駆込み・買控え効果**」は要注意



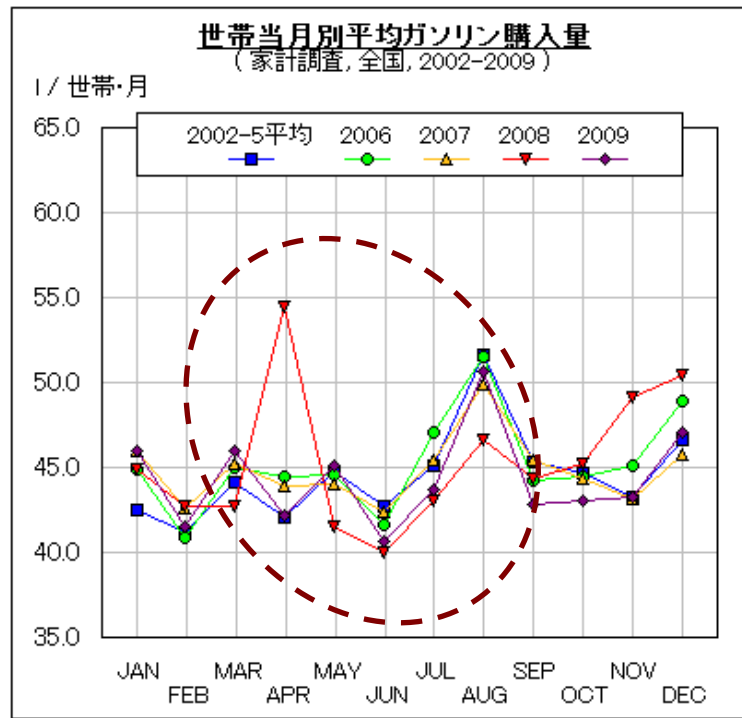
3. 措置効果評価の手法(2)

3-5. DID分析と「駆け込み・買控え効果」(2)

－「駆け込み・買控え効果」の実例

2008年4月: ガソリン税暫定税率一時廃止

(年度末での法案審議遅延による期限切・翌月衆院再可決)



- － 2008年4月の暫定税率廃止の結果、ガソリン価格は1リットル140円から約20円低下(当時)
- － その結果、家計において平年月の約20%相当の「駆け込み」需要が発生し(“ガソリン満タン化”)
5～8月頃迄その反動が残ったと推定される
- － 2008年秋からの増加は、高速道路料金社会実験(深夜割引時間帯拡大)の影響と推察される

4. 措置効果評価の手法(3)

4-1. マッチングによるデータ予備処理(1)

- DID分析 や パネルデータ分析 などにおいて、措置群・対照群間での**対象の異質性を軽減**する方法として、**マッチング**によるデータ予備処理が注目されている
- マッチングとは、措置群・対照群の横断面データを幾つかの変数(“**条件変数**”)を用いて分類し、当該分類された措置群・対照群が対応するよう(“Match”, = 異質性が少ない)データを選別した上で分析すること

4. 措置効果評価の手法(3)

4-2. マッチングの種類(1)

– 完全一致型 (Exact Matching)

措置群が少数で、多数の対照群がある場合、措置群と**条件変数が完全に一致する対象だけ**を対照群として抽出し試料化（逆も可、事例少）
（例：中小企業の場合：業種・売上・従業員数等を条件変数に設定）

– 最近接距離型 (Nearest Neighbor Matching)

措置群・対照群のデータから、**条件変数の「距離」**（例： $= ((Z_{a1} - Z_{a0})^2 + (Z_{b1} - Z_{b0})^2)^{0.5}$ ）が**最小**となる対を作り試料化（例：a = 売上高, b = 従業員数）

4. 措置効果評価の手法(3)

4-3. マッチングの種類(2)

– 措置率型 (**Propensity Score Matching : PSM**)

措置群・対照群のデータのうち、**措置率の刻みや範囲が一定となるように条件変数を区切って**
措置群・対照群を抽出して試料化 (事例多数)

例1: 層別・区間式措置率型マッチング Stratified/Interval type PSM

措置率 p が 0～100% の間で10%刻みとなるような条件変数の区間(例: 売上高1千万, 5千万, 2億…)を設定し、各区間内の措置群・対照群を試料化(但し措置率 0～10, 90～100%は廃棄)

例2: 半径式措置率型マッチング Radius type PSM

措置群内の試料別に、その試料から措置率 p が一定範囲内(例. $\pm 5\%$)となる条件変数の企業を当該試料の対照群に設定

4. 措置効果評価の手法(3)

4-4. 措置率型(PSM) と 平均措置効果(ATE)

- 措置率型(PSM)において**平均措置効果(ATE)**は措置率 $p(x)$ を用いて下記のように表現される

$$\begin{aligned} \text{ATE}(x) &= E((D - p(x)) * y / (p(x) * (1-p(x))) \\ &= E(y(1) | x, D = 1) - E(y(0) | x, D = 0) \end{aligned}$$

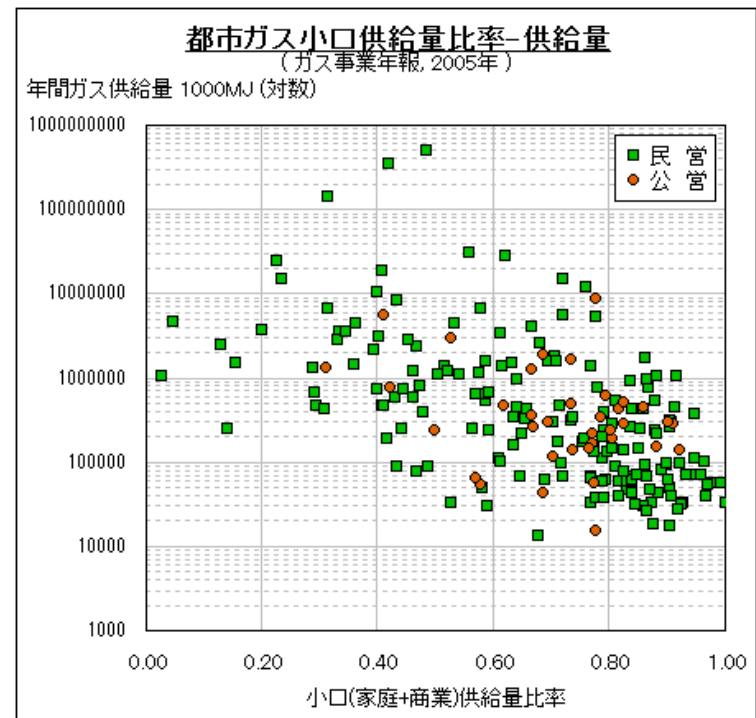
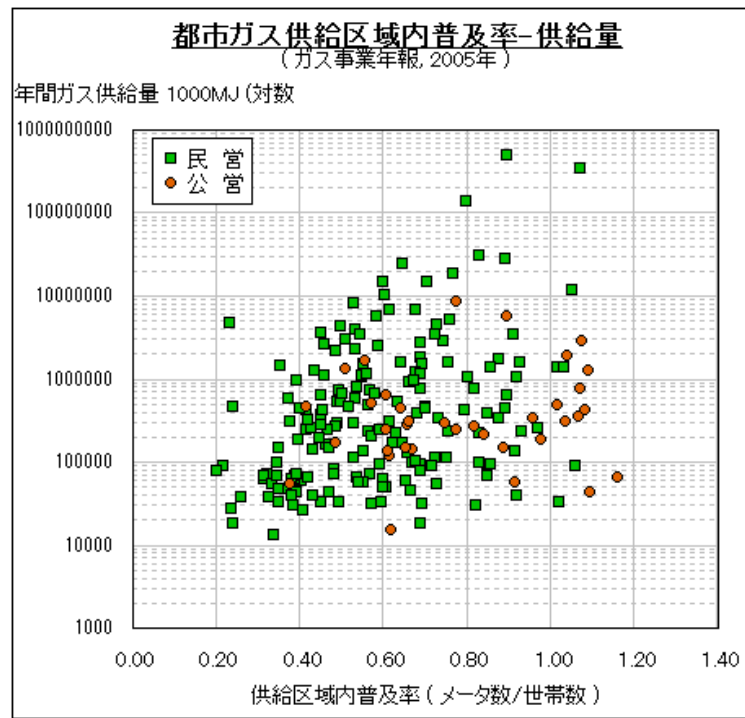
従って、試料数 N の場合の平均措置効果の推計値(ATE*)は、 $p^*(x_i)$ (← Logitモデルなどにより x_i から推計), y_i , D_i から下記のとおり推計できる

$$\text{ATE}^*(x) = N^{-1} * \sum_i ((D_i - p^*(x_i)) * y_i / (p^*(x_i) * (1 - p^*(x_i))))$$

5. 措置効果評価 - 実戦編 -

5-1. 横断面分析とマッチング(1)

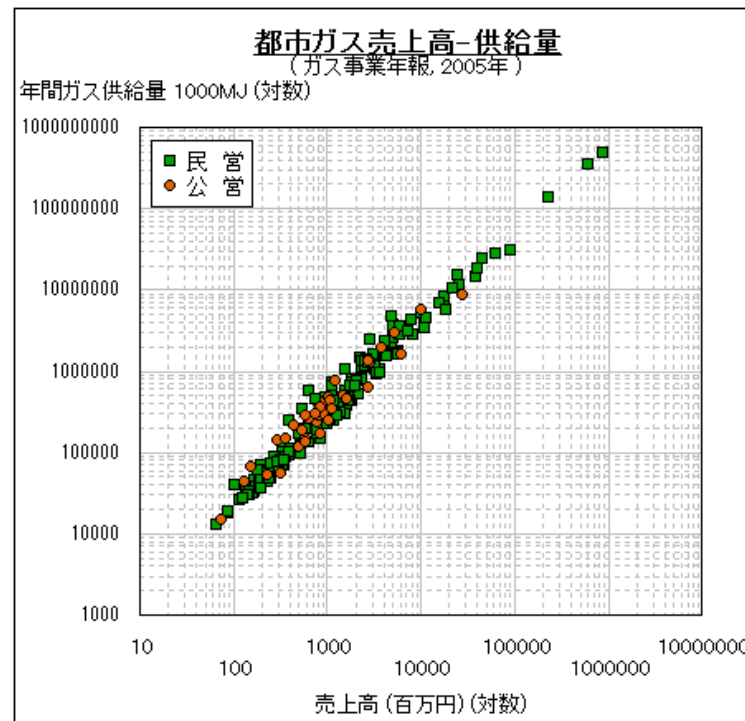
- 都市ガス事業は 2005年時点で民営178社・公営36社が存在するが、**生産性を民間・公営間で比較する場合、事業の「異質性」に注意が必要**



5. 措置効果評価 - 実戦編 -

5-2. 横断面分析とマッチング(2)

- マッチングの条件変数の設定に際し、分析する内容と過度に相関が高い変数を設定すると「**異質性**」の除去が分析指標に影響を与えてしまう



5. 措置効果評価 - 実戦編 -

5-3. 横断面分析とマッチング(3)

- 最近接距離マッチングの手順

1) 条件変数 z を選択

例: z_1 = 「供給区域普及率」 z_2 = 「小口供給量比率」

2) 各条件変数を規格化 (平均 0, 標準偏差 1)

$$\hat{z}_{i1,2} = (z_{i1,2} - \mu_{z1,2}) / \sigma_{z1,2}$$

($\mu_{z1,2} = z_{1,2}$ 各平均値 $\sigma_{z1,2} = z_{1,2}$ 各標準偏差)

3) 措置群 i の $\hat{z}_{i1,2}$ に対し最小の「距離」となる

$\hat{z}_{j1,2}$ を持つ j を対照群から抽出し試料化

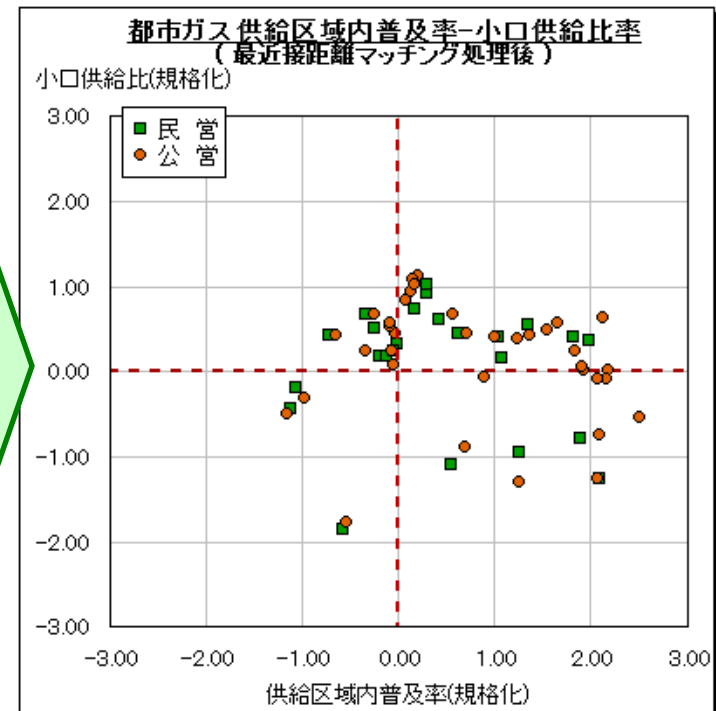
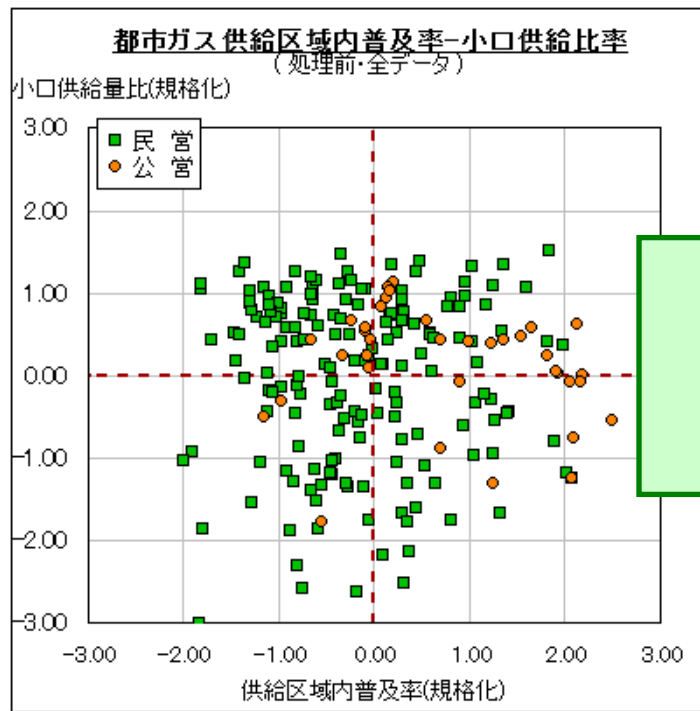
$$| \hat{z}_{ij} | = \min ((\hat{z}_{i1} - \hat{z}_{j1})^2 + (\hat{z}_{i2} - \hat{z}_{j2})^2)^{0.5}$$

注) 複数年の場合、複数年の加重平均「距離」が最小の j を選択

5. 措置効果評価 - 実戦編 -

5-4. 横断面分析とマッチング(4)

- **最近接距離マッチング**による処理
措置群と対応する対象だけを対照群から抽出、
「異質性」を軽減した試料を作成



5. 措置効果評価 - 実戦編 -

5-5. 横断面分析とマッチング(5)

- 最近接距離マッチングによる効果

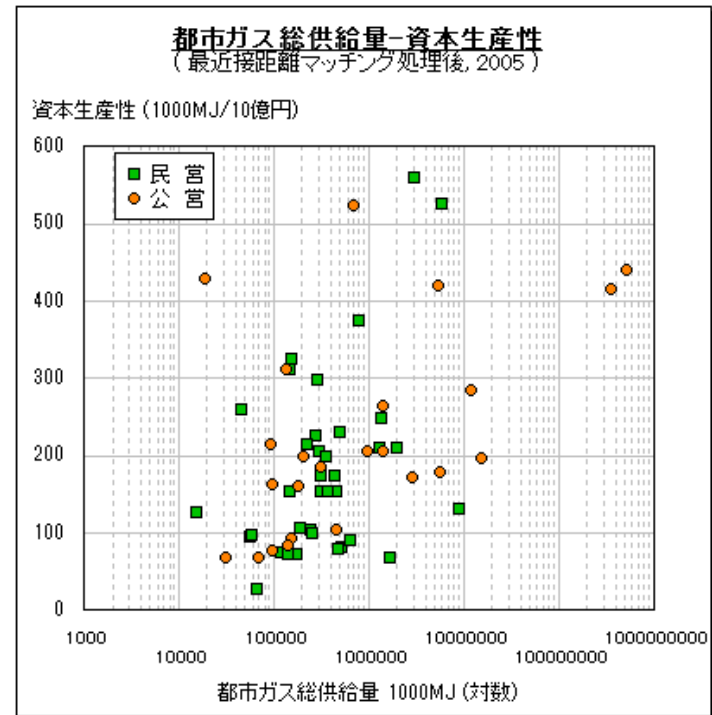
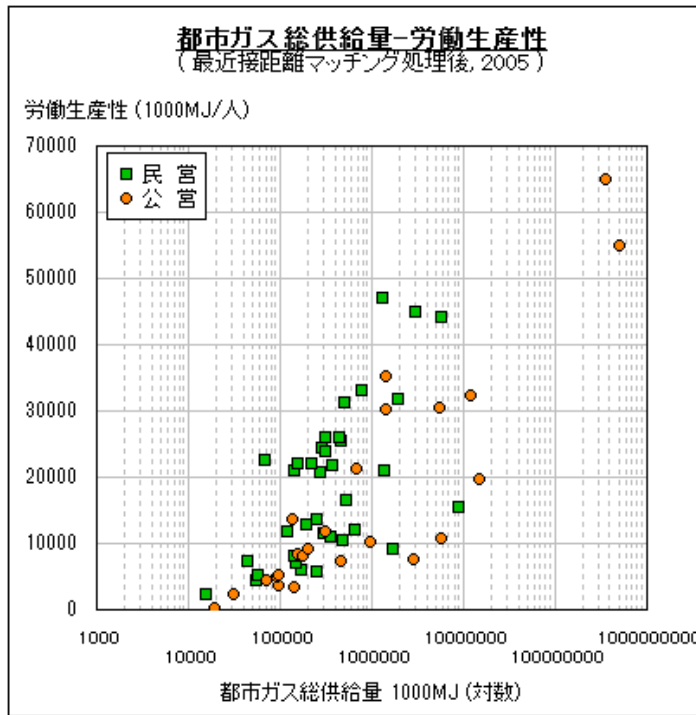
(2005年)	公営都市ガス事業			民営都市ガス事業		公営 - 民営	
	95%信頼上下限			未処理	処理済	未処理	処理済
供給区域普及率	0.782	1.123	0.441	-0.161	0.748	0.943	0.034
同標準偏差	1.044			0.914	1.008 (p値)	0.000	0.445
小口供給量比率	0.155	0.379	-0.069	-0.032	0.197	0.187	-0.042
同標準偏差	0.686			1.052	0.678 (p値)	0.092	0.396
平均資本生産性	186.70	226.11	147.29	251.81	232.16	-65.11	-45.46
同標準偏差	120.65			234.72	117.36 (p値)	0.008	0.500
平均労働生産性	19076	22878	15273	14492	18907	4584	168
同標準偏差	11640			15936	15546 (p値)	0.024	0.479

平均値の差の検定(Ho: 差なし)

5. 措置効果評価 - 実戦編 -

5-6. 横断面分析とマッチング(6)

- マッチングにより予備処理した試料を、さらに別の説明変数で回帰分析に掛けて分析すること多し (**混合推計**; Mixed Estimation)



5. 措置効果評価 - 実戦編 -

5-7. 横断面分析とマッチング(7) - 混合推計 -

- マッチング処理した試料を再度回帰分析

```
. reg lprpr lsplypr, robust
```

```
Linear regression
```

```
Number of obs = 36
F( 1, 34) = 11.47
Prob > F = 0.0018
R-squared = 0.3745
Root MSE = 9340.3
```

民営都市ガス事業

労働生産性を
総供給量(対数)
で回帰

lprpr	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
lsplypr	5335.243	1575.496	3.39	0.002	2133.451	8537.035
_cons	-48931.97	19227.97	-2.54	0.016	-88007.9	-9856.042

```
. reg lprgo lsplygo, robust
```

```
Linear regression
```

```
Number of obs = 36
F( 1, 34) = 128.84
Prob > F = 0.0000
R-squared = 0.7767
Root MSE = 7453.6
```

公営都市ガス事業

労働生産性を
総供給量(対数)
で回帰

lprgo	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
lsplygo	5502.705	484.7788	11.35	0.000	4517.516	6487.894
_cons	-56678.69	6034.581	-9.39	0.000	-68942.43	-44414.94

5. 措置効果評価 - 実戦編 -

5-8. 横断面分析とマッチング(8) - 結果の解釈 -

- 都市ガス事業の生産性を民間・公営間で比較する場合、供給区域内普及率・小口供給量比率など事業の「異質性」を管理せずに比較すると有意な差異が観察される
- 一方、これらの「異質性」を管理して比較すると、労働生産性・資本生産性ともに**生産性に有意な差異があるとは言えない**結果となる
(→ 生産性の差異は、民間・公営間での供給区域内普及率の差に起因する可能性大)