パネルデータ (2)

別所俊一郎

2006 年 7 月 5 日

Today's attraction

- 固定効果モデルとは
- 時間効果固定モデルとは
- 推定の条件，具体例

固定効果モデルとは

- **固定効果モデル（fixed effects model）**とは，主体によって異なるが，時間を通じて一定の値を取るような省略された変数（omitted variable）を制御する方法のひとつ
- n 個の異なる切片を持ち，各切片が各主体に対応する
- 省略変数の影響を吸収するためにダミー変数群が用いられる
- 時点の数は 3 以上でもよい
- (注) パネルデータの分析ではしばしば，固定効果モデルと並んで，変量効果モデル（random effect model）が紹介されるが，ここでは扱わない．経済データでは，固定効果モデルのほうが望ましい推定量を与えるケースが多いように（全てではない）思われる（Durbin-Wu-Hausman 検定）.

固定効果モデルとは

- 「固定効果」を Z_i とすると，回帰式は

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Z_i + u_{it}$$

- ここで，$\alpha_i = \beta_0 + \beta_2 Z_i$ とおくと，

$$Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it}$$

- $\{\alpha_i\}_{i=1}^n$ は推定されるべき切片.
- 回帰直線は主体ごとに異なるが，傾きは共通.
- 主体ごとのダミー変数と見てもよい

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \gamma_2 D2_i + \gamma_3 D3_i + \cdots + \gamma_n Dn_i + u_{it}$$

ここで，$\alpha_1 = \beta_0, \alpha_i = \beta_0 + \gamma_i (i \geq 2)$.

固定効果モデルの仮定

固定効果モデルはダミーを入れた重回帰モデルだから，β_1 の推定量が一致性をもつには以下の仮定が満たされなければならない．

1. $E[u_{it}|X_{i11}, X_{i12}, \cdots, X_{ikT}, \alpha_i] = 0$
2. $(X_{i11}, X_{i12}, \cdots, X_{ikT}, Y_{i1}, Y_{i2}, \cdots, Y_{iT})(i = 1, ..., n)$ が i.i.d.
3. (X_{ijt}, u_{it}) が有限でゼロより大きい 4 次モーメントを持つ
4. 完全な多重共線性はない
5. $\mathrm{cov}(u_{it}, u_{is}|X_{ijt}, X_{ijs}, \alpha_i) = 0$, for $t \neq s$.

（注）パネルデータでは，全ての観測値が i.i.d. と考えるのはふさわしくないので，各主体ごとの観測値の流列を一組と考えて，その組が i.i.d. と仮定している (2)．そのうえで，推定量の分散共分散行列の一致推定量を得るために，その組のなかでの誤差項の分布に制約を置いている (5).

固定効果モデルの推定と推測

ダミー変数を使った定式化であれば，OLS で推定が可能．ただし，説明変数の数が $k+n$ 個になってしまうので，より簡便な方法も用いられている．

- **“entity-demeaned” OLS アルゴリズム**では，まず主体ごとに変数の平均値を求め，各変数からその平均値を引いておく（demeaned）．次に，その加工された変数について OLS を施す．たとえば，

$$Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it}$$

から，各変数の主体ごとの平均値をとった

$$\overline{Y}_i = \beta_1 \overline{X}_i + \alpha_i + \overline{u}_i$$

を引くと

$$Y_{it} - \overline{Y}_i = \beta_1 (X_{it} - \overline{X}_i) + (u_{it} - \overline{u}_i)$$

新しい変数を定義して

$$\tilde{Y}_{it} = \beta_1 \tilde{X}_{it} + \tilde{u}_{it}$$

を OLS 推定すればダミーを用いた場合と同じ β_1 の推定値が得られる

- $T = 2$ のときの before-after の比較による推定との違いはない.
  - だから，$T = 2$ のときには，before-after の比較，ダミー変数を用いた特定化，“entity-demeaned” OLS の 3 つの方法が同一の推定値を与える.
- 適切な仮定のもとで，OLS の heteroskedastisity-robust（分散不均一に頑健な）標準偏差を検定・信頼区間の形成に用いることができる
- 主体ごとに異なり，時間を通じて不変な「観測できる」変数の係数は推定できない

時間固定効果とは

- **時間固定効果（time fixed effects）**とは，主体間では等しいが，時間とともに変化するような変数
- たとえば，全国的に導入された自動車の安全性の向上など.
- Omitted variables になりやすい
- 時間固定効果 S_t を考慮すると回帰式は

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Z_i + \beta_3 S_t + u_{it}$$

時間固定効果がある場合の推定

- 時間固定効果のみがあり，主体固定効果がない場合.

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_3 S_t + u_{it}$$

  – β_1 の推定が目的ならば，時点ごとのダミー変数（time effects）を T 個作って推定すればよい

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \delta_2 B2_t + \cdots + \delta_T BT_t + u_{it}$$

  $(\delta_2, \cdots, \delta_T)$ が推定される未知の係数.

- 両方向の固定効果がある場合（time and entity fixed effects / two-way fixed effects）

$$\begin{aligned} Y_{it} &= \beta_0 + \beta_1 X_{it} + \gamma_2 D2_i + \cdots + \gamma_n Dn_i \\ &\quad + \delta_2 B2_t + \cdots + \delta_T BT_t + u_{it} \end{aligned}$$

  両方向の固定効果モデルに対応するコマンドを持つソフトウェアもある.

酒税-交通死亡事故データへの応用

- 通常の固定効果モデル (8.15)
  - ビール税率と死亡数に統計的に有意な，大きな負の関係を検出
  - 値が大きすぎる．ビール税率の上昇傾向と，自動車の安全化傾向を拾っているだけ?
- 両方向の固定効果モデル (8.19)
  - 説明変数は 55 個！
  - ビール税率と死亡数に統計的に有意な，大きな負の関係を検出
  - 値が大きすぎる．ほかにも見落とされた変数が?
- 法律変数・経済変数の追加（table 8.1）
  - 飲酒を認める年齢や，違反時の罰則に差異がある
  - 経済状況も事故に影響する

酒税-交通死亡事故データへの応用

- 法律変数・経済変数の追加（table 8.1）
  - (1)〜(3) は本文中にもあったもの
  - (4) が base case で，(5)(6) は頑健性の確認
  - ビール税の係数は，法律・経済変数を追加すると小さく推定されるがそれでもまだ大きい
  - 飲酒許可年齢の違いの与える効果は小さい
  - 最初の罰則の効果も大きくない.
  - 経済変数の効果は大きい：好況時に交通量が増加するため?

TABLE 8.1 Regression Analysis of the Effect of Drunk Driving Laws on Traffic Deaths

Dependent Variable: Traffic Fatality Rate (Deaths Per 10,000).

Regressor	(1)	(2)	(3)	(4)	(5)	(6)
Beer tax	0.36** (0.05)	−0.66** (0.20)	−0.64* (0.25)	−0.45* (0.22)	−0.70** (0.25)	−0.46* (0.22)
Drinking age 18				0.028 (0.066)	−0.011 (0.064)	
Drinking age 19				−0.019 (0.040)	−0.078 (0.049)	
Drinking age 20				0.031 (0.046)	−0.102* (0.046)	
Drinking age						−0.002 (0.017)
Mandatory jail?				0.013 (0.032)	−0.026 (0.065)	
Mandatory community service?				0.033 (0.115)	0.147 (0.137)	
Mandatory jail or community service?						0.031 (0.076)
Average vehicle miles per driver				0.008 (0.008)	0.017 (0.010)	0.009 (0.008)
Unemployment rate				−0.063** (0.012)		−0.063** (0.012)
Real income per capita (logarithm)				1.81** (0.47)		1.79** (0.45)
State effects?	no	yes	yes	yes	yes	yes
Time effects?	no	no	yes	yes	yes	yes

These regressions were estimated using panel data for 48 U.S. states from 1982 to 1988 (336 observations total), described in Appendix 8.1. Standard errors are given in parentheses under the coefficients, and *p*-values are given in parentheses under the *F*-statistics. The individual coefficient is statistically significant at the *5% level or **1% significance level.

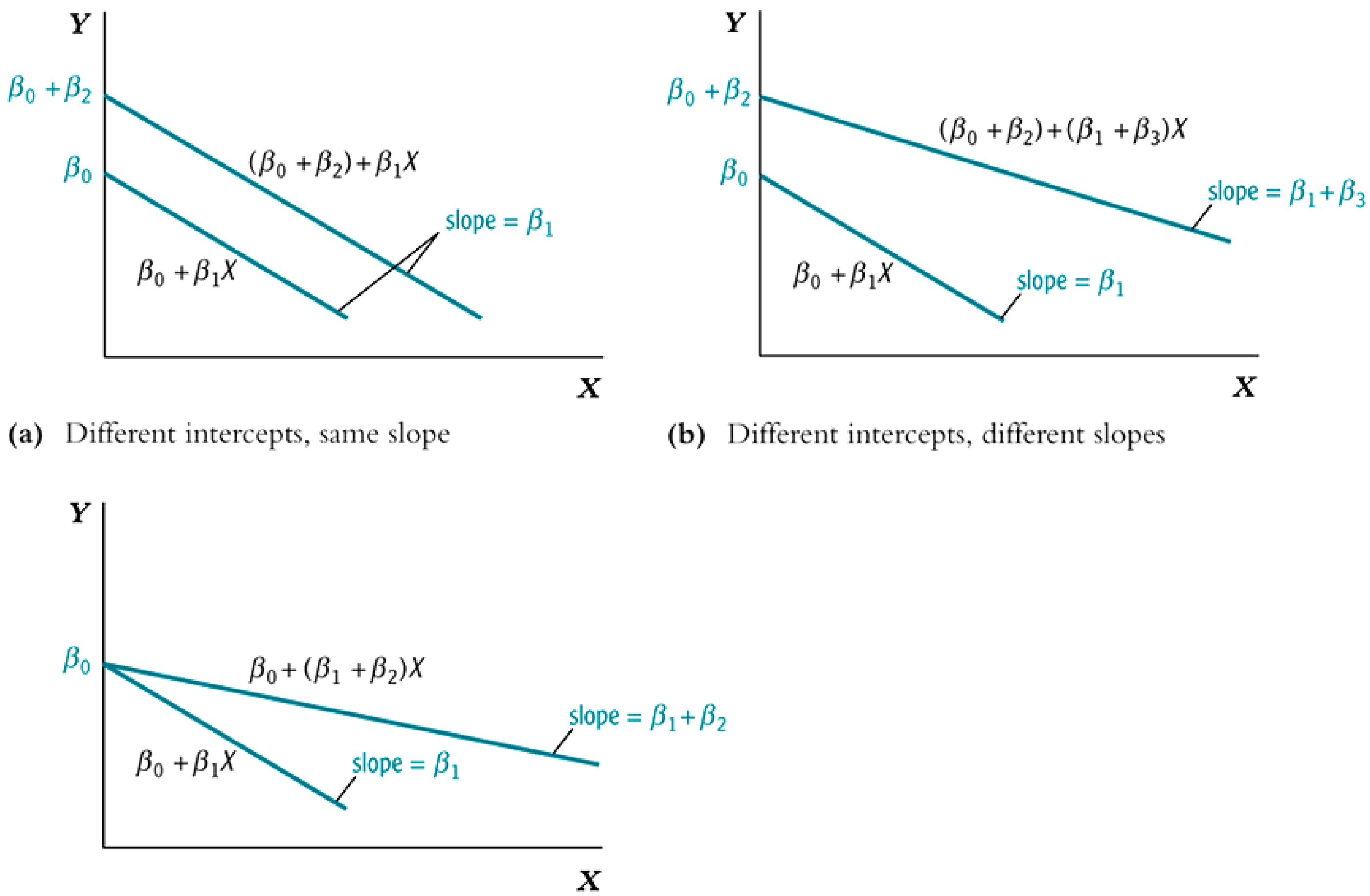
TABLE 8.1 Regression Analysis of the Effect of Drunk Driving Laws on Traffic Deaths

Dependent Variable: Traffic Fatality Rate (Deaths Per 10,000).

	(1)	(2)	(3)	(4)	(5)	(6)
***F*-statistics and *p*-values Testing Exclusion of Groups of Variables:**						
Time effects = 0			2.47 (0.024)	11.44 (<0.001)	2.28 (0.037)	11.59 (<0.001)
Drinking age coefficients = 0				0.48 (0.696)	2.09 (0.102)	
Jail, community service coefficients = 0				0.17 (0.845)	0.59 (0.557)	
Unemployment rate, income per capita = 0				38.29 (<0.001)		40.12 (<0.001)
$\bar{R}^2$	0.090	0.889	0.891	0.926	0.893	0.926

These regressions were estimated using panel data for 48 U.S. states from 1982 to 1988 (336 observations total), described in Appendix 8.1. Standard errors are given in parentheses under the coefficients, and p-values are given in parentheses under the F-statistics. The individual coefficient is statistically significant at the *5% level or **1% significance level.

FIGURE 6.8 **Regression Functions Using Binary and Continuous Variables**



(a) Different intercepts, same slope

(b) Different intercepts, different slopes

(c) Same intercept, different slopes

Interactions of binary variables and continuous variables can produce three different population regression functions: (a) $\beta_0 + \beta_1 X + \beta_2 D$ allows for different intercepts but has the same slope; (b) $\beta_0 + \beta_1 X + \beta_2 D + \beta_3 (X \times D)$ allows for different intercepts and different slopes; and (c) $\beta_0 + \beta_1 X + \beta_2 (X \times D)$ has the same intercept but allows for different slopes.