How to interpret the logistic regression with fixed effects

Klaus Pforr

5th ESRA Conference,
Ljubljana, Slovenia, July 15–19, 2013
Outlook

- Fixed-effects logit
  - Advantages
  - Disadvantages
- Interpretation
  - Standard technique
  - Alternative interpretations
  - Alternative model
- Conclusion
Fixed-effects logit (Chamberlain, 1980)

Individual intercepts instead of fixed constants for sample

\[ Pr(y_{it} = 1) = \frac{\exp(\alpha_i + x_{it}\beta)}{1 + \exp(\alpha_i + x_{it}\beta)} \]

Advantages

- Implicit control of unobserved heterogeneity
  - Forgotten or hard-to-measure variables
  - No restriction on correlation with indep. var’s
- Reduces problem of self-selection and omitted-variable bias
Fixed-effects logit

Disadvantages

• Panel data
• Only constant heterogeneity controlled
• Neglected heterogeneity weakened, but remains
• Interpretation severely limited
  • Part of index function unspecified
  • No predicted probabilities of outcome
  • No partial/discrete change effects

⇒ How do we interpret fixed-effects logit?
Interpretation alternatives

1. Odds ratio effects
   - OR-effect:
     \[
     \frac{\Pr(y_{it} = 1 | x_{it} + 1)}{\Pr(y_{it} = 0 | x_{it} + 1)} \div \frac{\Pr(y_{it} = 1 | x_{it})}{\Pr(y_{it} = 0 | x_{it})} = \exp(\beta)
     \]
   - Straightforward
     "All else equal with increase of \( x \) by 1 unit, odds of \( y = 1 \) vs. \( y = 0 \) increase by factor \( \exp(\beta) \)."
   - Odds non-intuitive
   - Polytomous DV: not necessarily same sign as change in prob.
Interpretation alternatives

2. Effect on cond. probability (Cameron & Trivedi, 2010)

- Probability to realize sequence of outcomes conditional on the number of occurrence of outcome within person
- Cond. prob. independent of $\alpha_i$

$$\Pr \left( y_i \mid x_i, \sum_{t=1}^{T_i} y_{it} \right) = \frac{\exp \left( \sum_{t=1}^{T_i} y_{it}x_{it}\beta \right)}{\sum_{d_i \in B_i} \exp \left( \sum_{t=1}^{T_i} d_{it}x_{it}\beta \right)}$$

😊 Predicted probabilities and average marginal/discrete changes possible
😊 Conditional probability non-intuitive
Interpretation alternatives

3. Effect on simplified cond. prob. (Cameron & Trivedi, 2009)

- Only $t = 2$ and $y_{i1} = 0$, $y_{i2} = 1$

$$
Pr (y_{i1} = 0, y_{i2} = 1 | x_i, y_{i1} + y_{i2} = 1) = \frac{\exp((x_{i2} - x_{i1}) \beta)}{1 + \exp((x_{i2} - x_{i1}) \beta)}
$$

😊 Cond. prob. of reduced case makes sense
😊 $T > 2$: Which time points to choose?
😊 Assumption on $\alpha_i$ introduced without basis in data

⇒ Iff $T = 2$, this is a reasonable option!
Interpretation alternatives

4. Effect on probabilities for prototypes (Schröder, 2010)

- Assume probability for outcome $y_{it}$ for prototypical unit with $x_{it}$ ⇒ Derive $\alpha_i$

😊 Now intuitive effect on prob's etc. possible

😊 Assumption on $\alpha_i$ based on aggreg. data

😊 Relevance of prototype depends on $\alpha_i | x_i$

$$E(\alpha_i | y_{it}, x_i) = \frac{E(y_{it} | \alpha_i, x_i)E(\alpha_i, x_i)}{E(y_{it})}$$

⇒ Relevance of estimated effects unknown ⇒ Only more intuitive interpretation of OR-effect
Alternative Model

Correlated random effects probit (Mundlak, 1978)

- Estimate random effects probit with across-time-means of covariates

 água

Stronger assumptions than full fixed-effects

\[ \alpha_i | x_i \sim \mathcal{N}(\gamma + \bar{x}_i \delta, \sigma_{\alpha_i}^2) \]

⇒ Simple correlation between \( \alpha_i \) and \( x_i \) allowed

 sikus

Effects on probabilities possible

 sikus

Average marginal effects possible
Conclusion

- Standard interpretation of fixed-effects logit limited to odds-ratio effects
- Other interpretation strategies within fixed-effects:
  - Conditional probability
  - Simplified conditional probability
  - Probability of prototype
- Correlated random effects probit
  - Stricter assumptions
  - Correlation between unobs. heterogeneity and covariates still allowed
  - Effect on probabilities possible

⇒ For $T > 2$, either accept odds-ratio effects or one step back with abandoning assumptions
Thank you
Back-up

Fixed-effects logit with person-dummies

• Linear fixed-effects models can be estimated with panel group indicators
• Non-linear fixed-effects models with group-dummies:
  • Person panel data (large $N$ and fixed $T$)
    ⇒ Estimates inconsistent for person-level heterogeneity, consistent for period dummies
  • Persons within countries (fixed “$N$” and large “$T$”)
    ⇒ Estimates consistent for country-level heterogeneity, inconsistent for person dummies

😊 Problem of omitted variables at one level remains
Back-up

Linear probability models with fixed-effects

- Linear probability models (OLS) can include fixed-effects
- Interpretation of effects on probabilities etc. possible
- Serial correlation across time can be allowed
- Neglected heterogeneity problem weakened
- Predicted probabilities unbounded

⇒ Works for marginal effects, not for predicted probabilities
References


Cameron, A. Colin, und Pravin K. Trivedi. 2010. Microeconometrics using Stata. College Station and TX: Stata Press.

