Partial effects in fixed effects models

J.M.C. Santos Silva

School of Economics, University of Surrey

Gordon C.R. Kemp

Department of Economics, University of Essex

22nd London Stata Users Group Meeting

8 September 2016

1. Introduction

- Models for panel data are attractive because they may make it possible to account for time-invariant unobserved individual characteristics, the so-called fixed effects.
- Consistent estimation of the fixed effects is only possible if T is allowed to pass to infinity.
- With fixed T it is not possible to perform valid inference about quantities that require estimates of the fixed effects.
- This is particularly problematic in non-linear models where often the parameter estimates have little meaning and it is more interesting to evaluate partial effects or elasticities.

- Consider a standard linear panel data model of the form

$$\mathrm{E}\left[y_{it}|x_{it},\alpha_i\right] = \alpha_i + \beta x_{it}, \quad i = 1,\ldots,n, \quad t = 1,\ldots,T.$$

- β (but not α_i) can be consistently estimated with fixed T.
- β gives the partial effect of x_{it} on $\mathrm{E}\left[y_{it}|x_{it},\alpha_i\right]$.
- What if we are interested in the semi-elasticity of $\mathrm{E}\left[y_{it}|x_{it},\alpha_i\right]$ with respect to x_{it}?
- For individual i this semi-elasticity is

$$\frac{\partial \ln \mathrm{E}\left[y_{it}|x_{it},\alpha_i\right]}{\partial x_{it}} = \frac{\beta}{\alpha_i + \beta x_{it}},$$

  and therefore it cannot be consistently estimated without a consistent estimate of α_i.

- Let y_{it} be a binary variable such that

$$\mathrm{E}\left[y_{it}|x_{it},\alpha_i\right] = \Pr\left[y_{it}=1|x_{it},\alpha_i\right] = \frac{\exp\left(\alpha_i+\beta x_{it}\right)}{1+\exp\left(\alpha_i+\beta x_{it}\right)}.$$

- It is well known that under suitable regularity conditions (Andersen, 1970, and Chamberlain, 1980) it is possible to estimate β consistently with fixed T.

- β is not particularly meaningful, at least for economists.

  - It can be seen as the partial effect of x_{it} on the log odds ratio (Cramer, 2003, p. 13, Buis, 2010).

  - It is also related to the partial effect on probabilities computed conditionally on $\sum_{i=1}^{T} y_{it}$ (Cameron and Trivedi, 2005, p. 797).

- Some practitioners opt for reporting the partial effects and semi-elasticities evaluated at an arbitrary value $\alpha_i = c$

$$\frac{\partial \Pr\left[y_{it} = 1 \middle| x_{it}, \alpha_i = c\right]}{\partial x_{it}} = \beta \frac{\exp\left(\beta x_{it} + c\right)}{\left(1 + \exp\left(\beta x_{it} + c\right)\right)^2},$$

$$\frac{\partial \ln \Pr\left[y_{it} = 1 \middle| x_{it}, \alpha_i = c\right]}{\partial x_{it}} = \beta \frac{1}{1 + \exp\left(\beta x_{it} + c\right)}$$

  often setting $\alpha_i = 0$.

- These, of course, is not meaningful because the choice of where to evaluate the individual effect is completely arbitrary.

- Wooldridge (2010, p. 622-3) considers an example where labour force participation of married women depends on the number of kids less than 18, on the log of husband's income, and time dummies.

```
. qui xtlogit lfp lhinc kids i.period, fe

. ereturn display, first
```

lfp	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lhinc	-.1842911	.0826019	-2.23	0.026	-.3461878	-.0223943
kids	-.6438386	.1247828	-5.16	0.000	-.8884084	-.3992688
period						
2	-.0928039	.0889937	-1.04	0.297	-.2672283	.0816205
3	-.2247989	.0887976	-2.53	0.011	-.398839	-.0507587
4	-.2479323	.0888953	-2.79	0.005	-.422164	-.0737006
5	-.3563745	.0888354	-4.01	0.000	-.5304886	-.1822604

- Setting $\alpha_i = 0$, the average elasticity of $\Pr\left[y_{it} = 1 | x_{it}, \alpha_i = 0\right]$ with respect to husband's income can be computed using `margins`.

```
. margins, eydx(lhinc)

Average marginal effects                        Number of obs     =      5,275
Model VCE    : OIM

Expression   : Pr(lfp|fixed effect is 0), predict(pu0)
ey/dx w.r.t. : lhinc
```

	Delta-method ey/dx	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lhinc	-.1677164	.0844434	-1.99	0.047	-.3332225	-.0022103

- To illustrate how meaningless this result is, let's repeat the exercise defining husband's income in thousands of dollars.

```
. gen double lhinck=log(hinc/1000)

. qui xtlogit lfp lhinck kids i.period, fe nolog

. ereturn display, first
```

lfp	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lhinck	-.1842911	.0826019	-2.23	0.026	-.3461878	-.0223943
kids	-.6438386	.1247828	-5.16	0.000	-.8884084	-.3992688
period						
2	-.0928039	.0889937	-1.04	0.297	-.2672283	.0816205
3	-.2247989	.0887976	-2.53	0.011	-.398839	-.0507587
4	-.2479323	.0888953	-2.79	0.005	-.422164	-.0737006
5	-.3563745	.0888354	-4.01	0.000	-.5304886	-.1822604

- Using again `margins` to estimate the average elasticity of $\Pr\left[y_{it} = 1 | x_{it}, \alpha_i = 0\right]$ with respect to husband's income we now get a different result.

```
. margins, eydx(lhinck)

Average marginal effects                        Number of obs     =      5,275
Model VCE    : OIM

Expression   : Pr(lfp|fixed effect is 0), predict(pu0)
ey/dx w.r.t. : lhinck
```

	Delta-method ey/dx	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lhinck	-.1389368	.0645866	-2.15	0.031	-.2655242	-.0123495

- The problem, of course, is that changing the scale in which income is measured only changes the values of the fixed effects, which are not estimated.
- Therefore, $\Pr\left[y_{it} = 1 \middle| x_{it}, \alpha_i = 0\right]$ is evaluated at exactly the same parameters, but using different regressors.
- Therefore, partial effects and elasticities evaluated at $\alpha_i = 0$ are not only meaningless, but their value will depend on how the regressors are measured.
- **However, the average elasticity of** $\Pr\left[y_{it} = 1 \middle| x_{it}, \alpha_i\right]$ **with respect to the husband's income can be estimated consistently.**

- Let $x_{it} = \ln(X_{it})$ where X_{it} is the husband's income.
- We want to estimate the average of

$$\mathrm{e}_{it} = \frac{\partial \ln \Pr[y_{it} = 1 | x_{it}, \alpha_i]}{\partial x_{it}} = \beta \frac{1}{1 + \exp(\beta x_{it} + \alpha_i)}$$

- e_{it} obviously depends on α_i and therefore cannot be consistently estimated with fixed T.
- However, to estimate $E[\mathrm{e}_{it}]$ we do not actually need to compute e_{it} because

$$E[\mathrm{e}_{it}] = \beta(1 - E[y_{it}])$$

  which can be consistently estimated by $\hat{\beta}(1 - \bar{y})$, where $\bar{y} = \frac{1}{nT} \sum_{i=1}^{n} \sum_{i=1}^{T} y_{it}$.
- This results was first obtained by Yoshitsugu Kitazawa (2012).

In short

- When $x_{it} = \ln(X_{it})$, $E[e_{it}]$ is the average elasticity with respect to X_{it}.
- Otherwise, $E[e_{it}]$ is the average semi-elasticity with respect to x_{it}.
- If x_{it} is discrete, for small β, $E[e_{it}]$ is approximately the percentage change of $\Pr(y_{it} = 1 | x_{it}, \alpha_i)$ resulting from a unit change in x_{it}.
- Unfortunately, the trick does not apply to the partial effects:
  - The partial effects have the form $\beta \times \mathrm{Var}[y_{it} | x_{it}, \alpha_i]$;
  - $\mathrm{Var}[y_{it} | x_{it}, \alpha_i]$ cannot be estimated without an estimate of α_i, but can be bounded;
  - It is not clear that having bounds on the partial effects is interesting.

- To perform inference about $E\left[\mathrm{e}_{it}\right]$ we need to be able to estimate its variance.
- The computation of such variance is greatly simplified by the fact that $\hat{\beta}$ and $\bar{y}$ are uncorrelated.
- Indeed, conditionally on the value of the regressors, changes in $\bar{y}$ are absorbed by the fixed effects; therefore $\hat{\beta}$ is uncorrelated with $\bar{y}$ because β is estimated by maximizing the conditional likelihood, which does not depend on α_i.
- Hence:

$$\mathrm{Var}\left[\hat{\beta}\left(1-\bar{y}\right)\right]=\mathrm{Var}\left[\hat{\beta}\right]\left(1-\bar{y}\right)^{2}+\mathrm{Var}\left[\bar{y}\right]\hat{\beta}^{2}.$$

- aextlogit is a wrapper for xtlogit which estimates the fixed effects logit and reports estimates of the average (semi-) elasticity of $\Pr\left(y_{it} = 1 | x_{it}, \alpha_i\right)$, and the corresponding standard errors and t-statistics.
- Syntax is standard:

```
aextlogit depvar [indepvars] [if] [in] [iweight] [, options]
```

---

betas: displays the logit estimates
nolog: suppress the display of the iteration log

```
. aextlogit lfp lhinc  kids i.period,  nolog
note: multiple positive outcomes within groups encountered.
note: 4,608 groups (23,040 obs) dropped because of all positive or
      all negative outcomes.

Conditional fixed-effects logistic regression   Number of obs      =      5275
Group variable: id                              Number of groups   =      1055
                                                Obs per group: min =         5
                                                               avg =         5
Log likelihood  = -2003.4184                                   max =         5
```

Average (semi) elasticities of Pr(y=1|x,u)

lfp	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lhinc	-.058623	.0262806	-2.23	0.026	-.110132	-.0071139
kids	-.204805	.0397334	-5.15	0.000	-.282681	-.126929
period						
2	-.0295209	.02831	-1.04	0.297	-.0850076	.0259658
3	-.0715085	.0282534	-2.53	0.011	-.1268841	-.0161329
4	-.0788673	.0282859	-2.79	0.005	-.1343067	-.0234278
5	-.1133627	.0282757	-4.01	0.000	-.1687821	-.0579433

Average of lfp = .68190005 (Number of obs = 28315)

```
. aextlogit lfp lhinck kids i.period,  nolog
note: multiple positive outcomes within groups encountered.
note: 4,608 groups (23,040 obs) dropped because of all positive or
      all negative outcomes.

Conditional fixed-effects logistic regression   Number of obs      =      5275
Group variable: id                              Number of groups   =      1055
                                                Obs per group: min =         5
                                                               avg =         5
Log likelihood  = -2003.4184                                   max =         5
```

Average (semi) elasticities of Pr(y=1|x,u)

lfp	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lhinck	-.058623	.0262806	-2.23	0.026	-.110132	-.0071139
kids	-.204805	.0397334	-5.15	0.000	-.282681	-.126929
period						
2	-.0295209	.02831	-1.04	0.297	-.0850076	.0259658
3	-.0715085	.0282534	-2.53	0.011	-.1268841	-.0161329
4	-.0788673	.0282859	-2.79	0.005	-.1343067	-.0234278
5	-.1133627	.0282757	-4.01	0.000	-.1687821	-.0579433

Average of lfp = .68190005 (Number of obs = 28315)

- A similar results applies to the partial effects in the exponential regression model (Poisson):

$$\mathrm{E}\left[y_{it}|x_{it},\alpha_i\right] = \exp\left(\alpha_i + \beta x_{it}\right), \quad i = 1,\ldots,n, \quad t = 1,\ldots,T.$$

$$\mathrm{E}\left[\frac{\partial \mathrm{E}\left[y_{it}|x_{it},\alpha_i\right]}{\partial x_{it}}\right] = \beta \mathrm{E}\left[\exp\left(\alpha_i + \beta x_{it}\right)\right] = \beta \mathrm{E}\left[y_{it}\right]$$

  which can be consistently estimated by $\hat{\beta}\bar{y}$.

- Maybe `margins` should be disabled after `xtlogit` and `xtpoisson` when the `fe` option is used?

References

- Andersen, E.B. (1970). "Asymptotic properties of conditional maximum likelihood estimators," *Journal of the Royal Statistical Society*, Series B, 32, 283-301.
- Buis, M.L. (2010). "Stata tip 87: Interpretation of interactions in non-linear models," *The Stata Journal*, 10, 305-308.
- Chamberlain, G. (1980). "Analysis of covariance with qualitative data," *Review of Economic Studies*, 47, 225–238.
- Cramer, J.S. (2003). *Logit Models from Economics and Other Fields*. Cambridge: CUP.
- Kitazawa, Y. (2012). "Hyperbolic transformation and average elasticity in the framework of the fixed effects logit model," *Theoretical Economics Letters*, 2, 192-199.
- Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.