Bias in Conditional and Unconditional Fixed Effects Logit Estimation: a Correction*

Tom Coupé

Economics Education and Research Consortium, National University of Kyiv Mohyla Academy

Address: Vul Voloska 10, 04070 Kyiv, Ukraine

Email: tcoupe@eerc.kiev.ua

Tel: 38-044 492-8012

* I thank the referees and Ethan Katz for helpful comments and the Economics Research and Outreach Center (EROC), Kyiv for financial support. The STATA do-file is available on the *Political Analysis* Web site.

Abstract

In a recent paper published in this journal, Katz (2001) compares the bias in conditional and unconditional fixed effects logit estimation using Monte Carlo Simulation. This note shows that while Katz's (2001) specification has "wrong" fixed effects (in the sense that the fixed effects are the same for all individuals), his conclusions still hold if I correct his specification (so that the fixed effects do differ over individuals). This note also illustrates the danger, when using logit, of including dummies when no fixed effects are present.

1. Introduction

In a recent paper published in this journal, Katz (2001) studies the estimation of the binary choice logit model when panel data are available, using Monte Carlo Simulation. He specifies the relationship between the binary variable **y** and **X** as follows

$$(1)\ P(y_{i,t} = 1 | \gamma, \alpha_i, \beta) = \frac{\exp(\gamma + \alpha_i + x_{i,t}\beta)}{1 + \exp(\gamma + \alpha_i + x_{i,t}\beta)}$$

Where γ is a constant, α_i an individual fixed effect and β the coefficient of the explicative variable. Data are available for many individuals i and time periods t, so I have a panel data set-up.

Katz then aims to compare two ways of controlling for fixed effects, using unconditional maximum likelihood but including individual specific dummies and estimating the model by conditional maximum likelihood. Conditional maximum likelihood is known to give consistent estimates of β (Chamberlain, 1980) but does not provide estimates of the individual fixed effects α_i, which are needed if one wants to compute statistics like marginal effects. The unconditional-with-dummies estimator provides estimates of the individual fixed effects α_i but has been proven to lead to inconsistency in β due to the incidental parameter problem – for t=2, Abrevaya (1997) has shown that the unconditional - with - dummies estimator of β equals two times the conditional maximum likelihood estimator of β. Monte Carlo simulations lead Katz to conclude that: "The implication for applied researchers is that the conditional estimator is always safe when T<20 and that the unconditional estimator is safe when 16≤T<20. When 8<T<16, the bias in the unconditional estimator is small and may be acceptable to the researcher but when T is close to two, the bias is substantial. Even in this case, however, the bias takes a predictable direction and magnitude".

Unfortunately, Katz's Monte Carlo simulations are not what they pretend to be. On p. 381, Katz writes: "I controlled p by adjusting the constant term γ, *and I set* α_i *to 1 for all i*[1]." This is a crucial mistake. Indeed, if all α_i**'**s take the same value, there is basically no fixed effect anymore. That is, the subscript i can be removed from α_i and the new constant term $\gamma' = \gamma + \alpha$ brings us to a simple logit model without fixed effects.

$$(2)\ P(y_{i,t} = 1, \gamma', \beta) = \frac{\exp(\gamma' + x_{i,t}\beta)}{1 + \exp(\gamma' + x_{i,t}\beta)}$$

Hence, in Katz's paper the true model has no fixed effects in contrast to what Katz claims[2,3].

In Table 1, I estimate the set-up of Katz's paper using conditional logit and unconditional logit with dummies, and then add a third estimation method, unconditional logit without dummies (that is, simple logit)[4]. Note that the real underlying model is three times a simple logit, but only the third estimation method "realizes" this.

[Table 1 Here]

Results for the conditional logit and the unconditional logit with dummies are similar to Katz's results. Important to note, however, is column 3, a simple logit gives results similar to the conditional estimator even for small t and hence, is much better than the unconditional with dummies estimator for small t. Hence, in Katz's set-up it is never useful to use the unconditional with dummies specification.

Note that Katz's study is an interesting illustration of the fact that in the logit case, including irrelevant variables as regressors can cause biased estimates, even if these irrelevant variables are not correlated with the "true" explicative variables – indeed, including individual specific dummies, which by construction are uncorrelated with the (random) X, in the case where there are NO fixed effects will cause a bias in the estimates of the coefficient of X. I mentioned before that for t=2, Abrevaya (1997) has shown that the unconditional - with - dummies estimator of β equals two times the conditional maximum likelihood estimator of β. In our case, I found exactly the same result, the coefficient of the dummy-variables specification is double the coefficient of the conditional specification – hence, the bias that I cause by including the dummies has nothing to do with the presence of fixed effects but is solely due to the method of estimation.

To correct Katz's study I generated α_i from N(0,1) instead of setting all α_i 's to one. To get fixed effects, I further need a correlation between **X** and the fixed effects. I propose the following specification that models X as a function of α_i[5].

$$(3)\ x_{it} = \frac{\sqrt{2}}{2}(a_i + \varepsilon_{it}), \varepsilon_{it} : N(0,1)$$

In this way, I have a correlation between x_{it} and α_i that is independent of time T, while x_{it} is standard normally distributed[6,7].

[Table 2 Here]

Our conclusion from Table 2 is essentially the same as Katz's conclusions: from t=8, the unconditional logit with dummies estimators gives results that on average are quite close to the true β. But note that in contrast to Table 1, here, the use of simple logit (column 3) is clearly inferior as one would expect since the simple logit completely ignores the fixed effects.

2. Conclusions

This note shows that while Katz's (2001) specification has no meaningful fixed effects, his conclusions still hold if I modify his specification such as to incorporate fixed effects – hence, when one has a panel data set with a relatively large number of time periods, using dummies to take into account the fixed effects is not unreasonable. However, this note also points out the danger, when using logit, of including dummies when no fixed effects are present.

3. References

Abrevaya, Jason. 1997. "The Equivalence of Two Estimators of the Fixed-Effects Logit Model." *Economics Letters* 55:41-43.

Chamberlain, Gary. 1980. "Analysis of Covariance with Qualitative Data." *Review of Economic Studies* 47:225-238.

Greene, William. 2004. "The Behaviour of the Maximum Likelihood Estimator of Limited Dependent Variable Models in the Presence of Fixed Effects." *Econometrics Journal* 7: 98-119

Katz, Ethan. 2001. "Bias in Conditional and Unconditional Fixed Effects Logit Estimation." *Political Analysis* 9:379-384.

4. Endnotes

[1] Italics added.

[2] Of course, one could claim that there are fixed effects, but they just happen to be the same. However, in such case the fixed effects are meaningless.

[3] This mistake is also present in the computer program written by Katz and posted on the Political Analysis website (http://polmeth.wustl.edu/pa/vol9no4.html), though in a slightly different form. In the computer program, the fixed effect of the first individual is dropped, hence the formula is

$$P(y_{i,t} = 1 | \gamma, \alpha_i, \beta) = \frac{\exp(\gamma + \alpha_i + x_{i,t}\beta)}{1 + \exp(\gamma + \alpha_i + x_{i,t}\beta)}$$

with α_i=1 for all i>1 and zero for α_1. Hence in this case, I just make a difference between the "fixed" effect of the first individual (γ) and the "fixed" effects of all other individuals (γ+ α_i).

[4] The parameters are the same as for column 1 in Table 2 of Katz (2001): β=0.5, γ=-1

[5] I could generate α_i from N(0,1) and combine this with a randomly generated **X**. This leads to what economists would call a "random effect" since by construction there is no correlation between the individual specific effects and the explicative variable **X**. The estimation procedure in this case would be to use the random effects estimator, rather than including dummies to capture the "fixed effects".

[6] Instead, one could follow Greene (2004) by generating individual specific effects that are correlated with **X** as follows:

$\alpha_i = \sqrt{T}\bar{x}_i + a_i, a_i N(0,1)$

However, this specification implies that the correlation between the fixed effects and the explicative variable will decrease as T grows since the influence of an individual x_{it} on the average decreases with t. Hence, using this specification one cannot distinguish between the effect of increasing t and decreasing the correlation between the fixed effects and the explicative variable.

[7] Note that this implies that I use a standard normal distribution for X rather than a uniform distribution (which was used by Katz). The distribution of the X generally, however, has no influence on the results (see e.g. Greene (2004)).

5. Tables

Table 1: Katz's specification under different estimation methods

N=100	Conditional Logit	Unconditional Logit with dummies	Unconditional Logit without dummies
t=2	0.472 (0.032)	0.944 (0.064)	0.488 (0.022)
t=4	0.520 (0.017)	0.696 (0.023)	0.514 (0.015)
t=8	0.506 (0.012)	0.579 (0.014)	0.504 (0.011)
t=12	0.505 (0.095)	0.551 (0.010)	0.498 (0.009)
t=16	0.495 (0.082)	0.529 (0.008)	0.500 (0.008)

Based on 500 Monte Carlo simulations. Parameters: p=0.56, β=0.5, γ'=-1. Monte Carlo standard errors in brackets.

Table 2: Fixed Effect Specification under different estimation methods

N=100	Conditional Logit	Unconditional Logit with dummies	Unconditional Logit without dummies
T=2	0. 535 (0.020)	1.070 (0.041)	1.122 (0.009)
T=4	0.502 (0.009)	0.675 (0.013)	1.145 (0.006)
T=8	0.508 (0.005)	0.583 (0.006)	1.030 (0.004)
t=12	0.501 (0.005)	0.549 (0.005)	1.183 (0.003)
t=16	0.496 (0.004)	0.529 (0.004)	1.165 (0.003)

Based on 500 Monte Carlo simulations. Parameters: β=0.5, γ'=-1. Monte Carlo standard errors in brackets.