Do multilevel models ever give different results?

Kelvyn Jones
Kelvyn.jones@bristol.ac.uk

University of Bristol

Abstract
It is sometimes said that the use of multilevel models over OLS regression makes no substantive difference to interpretation and represents something of a fuss over nothing. This short paper demonstrates with a simple example that this is not always the case.

Keywords
Multilevel models, within- and between- regressions

It is sometimes claimed that the use of multilevel modelling over single-level (ordinary least squares) regression makes no substantive difference to model results and interpretation, in that it only affects the standard errors of the coefficients, and not the coefficients themselves. For example (Bickel, 2007, p.xi-pxii) states

> *"There* is *usually little difference between* the unstandardized *regression coefficients estimated with conventional OLS regression and multilevel regression … it sometimes seems that not much is promised beyond marginally improved estimates of standard errors, and identification of the correct numbers of degrees of freedom in inferential tests*."

And again (Bickel, 2007, p.2)

> "*When comparing OLS and multilevel regression results, we may find that differences among coefficient values are inconsequential, and tests of significance may lead to the same decisions. A great deal of effort seems to have yielded precious little gain.*"

Indeed this viewpoint is one reading of the title of Bickel's book. But he is not alone in making these points. For example Gorard (2003a, p.49) writes

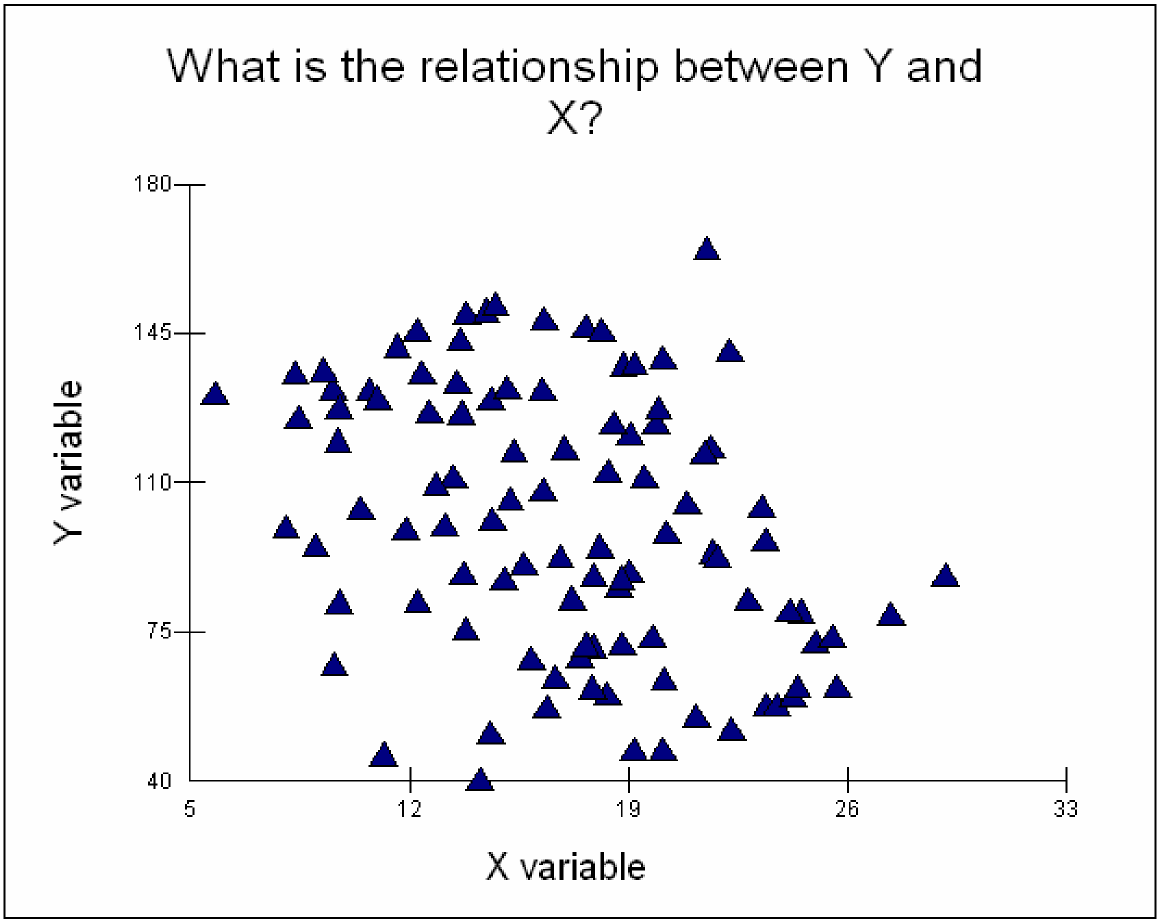
> "*MLM is therefore simply regression that allows the analyst to use both individuals and groups of individuals in the same model to avoid flouting the assumption of independent cases, since the accuracy ('standard errors') of any results can be affected by the clustered nature of the data.*'"

and

> "*Autocorrelation anyway only leads to the loss of power…, and this could be righted more simply by increasing sample size rather than changing methods of analysis.*" *(*Gorard, 2003a, p.60; see also Gorard 2003b)

But here is an example, where that is definitely not the case

We begin with a simple plot of two variables and ask what the relationship between them is.



Our first answer is to fit the simple regression of Y on X (**Xvar**):

$$\mathrm{Y}_i \sim \mathrm{N}(XB,\ \Omega)$$

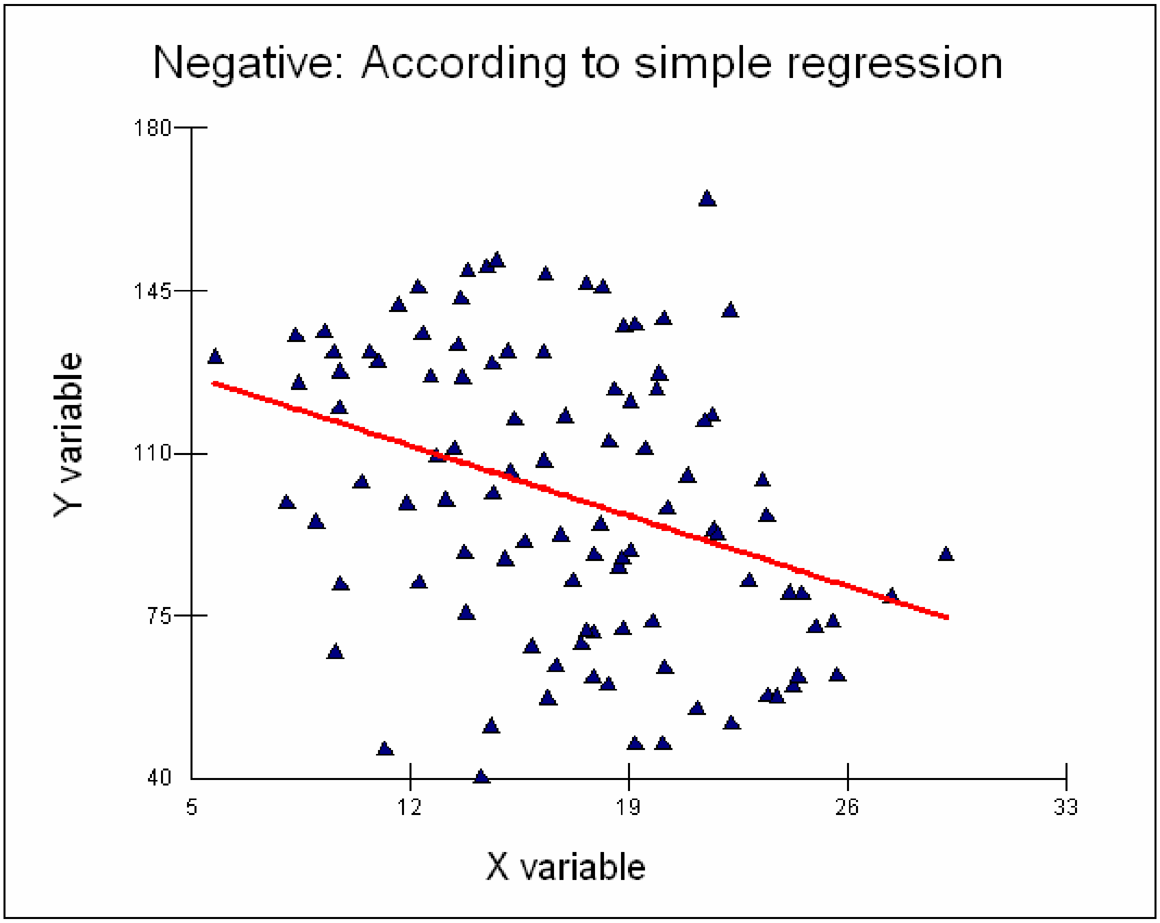
$$\mathrm{Y}_i = \beta_{0i}\mathrm{Cons} + -2.165(0.593)\mathrm{Xvar}_i$$

$$\beta_{0i} = 137.631(10.490) + e_{0i}$$

$$\left[e_{0i}\right] \sim \mathrm{N}(0,\ \Omega_e)\ :\ \Omega_e = \left[844.060(119.368)\right]$$

$$-2{*}loglikelihood(IGLS\ Deviance) = 957.610(100\ \text{of}\ 100\ \text{cases in use})$$

(where Cons =1 for all individuals and its coefficient, β_0, is the intercept). The fitted regression line, shown on the following plot, has a significant (t = 3.61, p< 0.0001) negative slope.



There can be little doubt when viewed as a single-level model, that the relation is a negative one, albeit one with quite a lot of scatter around the line.

Turning now to a random intercept multilevel model (Goldstein 2003), we can recognise that the observations belong to 10 groups.

$$\mathrm{Y}_{ij} \sim \mathrm{N}(XB, \Omega)$$

$$\mathrm{Y}_{ij} = \beta_{0ij}\mathrm{Cons} + 2.050(0.052)\mathrm{Xvar}_{ij}$$

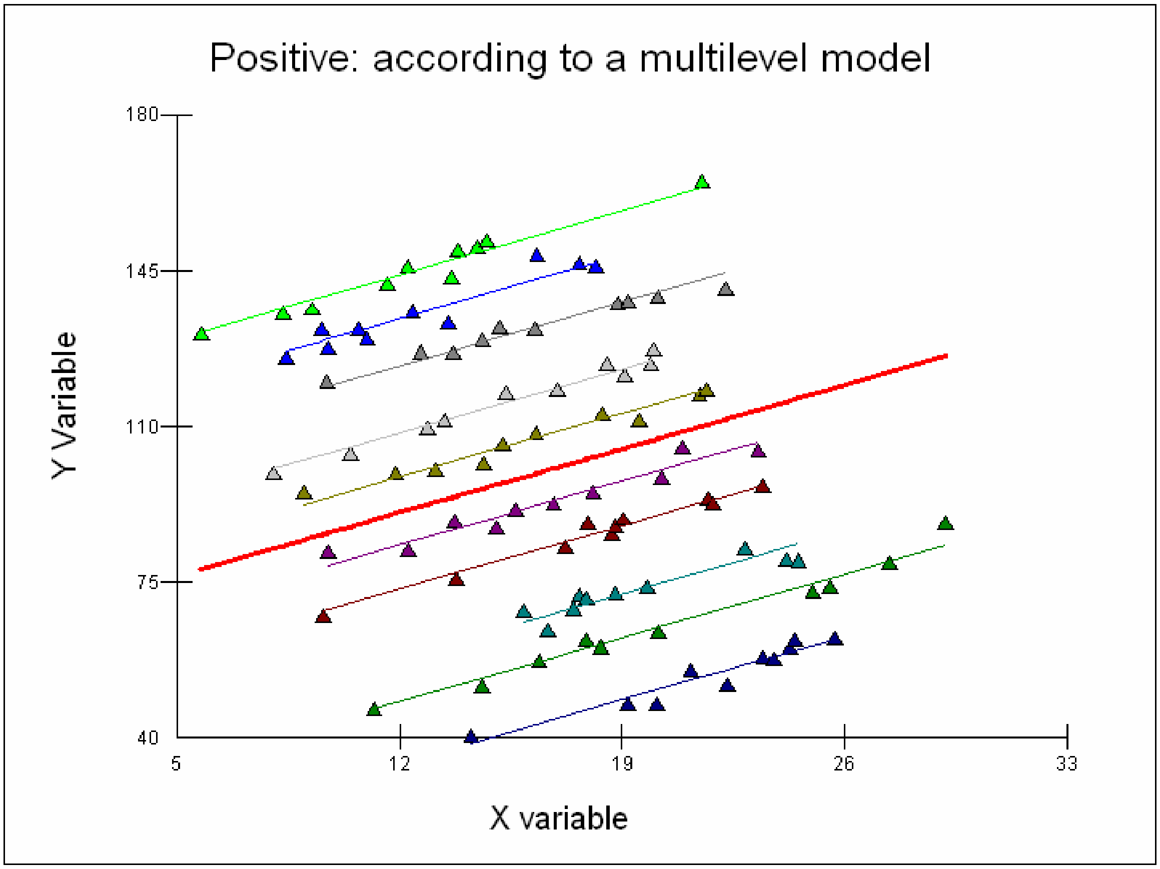
$$\beta_{0ij} = 66.035(11.286) + u_{0j} + e_{0ij}$$

$$\left[u_{0j}\right] \sim \mathrm{N}(0, \Omega_u) \;:\; \Omega_u = \left[1265.495(566.157)\right]$$

$$\left[e_{0ij}\right] \sim \mathrm{N}(0, \Omega_e) \;:\; \Omega_e = \left[4.106(0.612)\right]$$

*-2*loglikelihood(IGLS Deviance)* = 505.372(100 of 100 cases in use)

Quite a different result is found in that the general slope is now positive and significant. We can see what is going on by plotting a fitted line for each and every group, and the overall population average line across all groups (in red). The population average line is what is estimated in a multilevel model (the so-called fixed effect) and is a weighted average of the group-specific lines; weighted to emphasize the most reliably estimated lines (Jones and Bullen 1994), that is ones with the greatest number of observations and the greatest diversity in the Xvar variable.



With single–level regression, it is assumed that the observations are independently and identically distributed, and this gives an overall negative relationship. There is no recognition that within each group the underlying relationship is positive. In contrast the multilevel model allows the intercept of each group to take on a different value from an overall distribution. As the following table shows, the multilevel model is a much better fit to these data, with a considerably smaller deviance. The multilevel model gives a substantially better interpretation of the data. There is no reason why such relationships should not be found in reality.

The effect of taking account of groups in the multilevel model is marked here because the group-specific intercept is negatively related to the mean of X for each group. This behaviour can be elucidated by including the group mean of X in the multilevel model alongside the deviations of X from that mean (Paccagnella, 2006).

$$\mathrm{Y}_{ij} \sim \mathrm{N}(XB,\ \Omega)$$

$$\mathrm{Y}_{ij} = \beta_{0ij}\mathrm{Cons} + \text{-}9.690(0.962)\mathrm{Xbar}_{j} + 2.052(0.052)\mathrm{X\text{-}Xbar}_{ij}$$

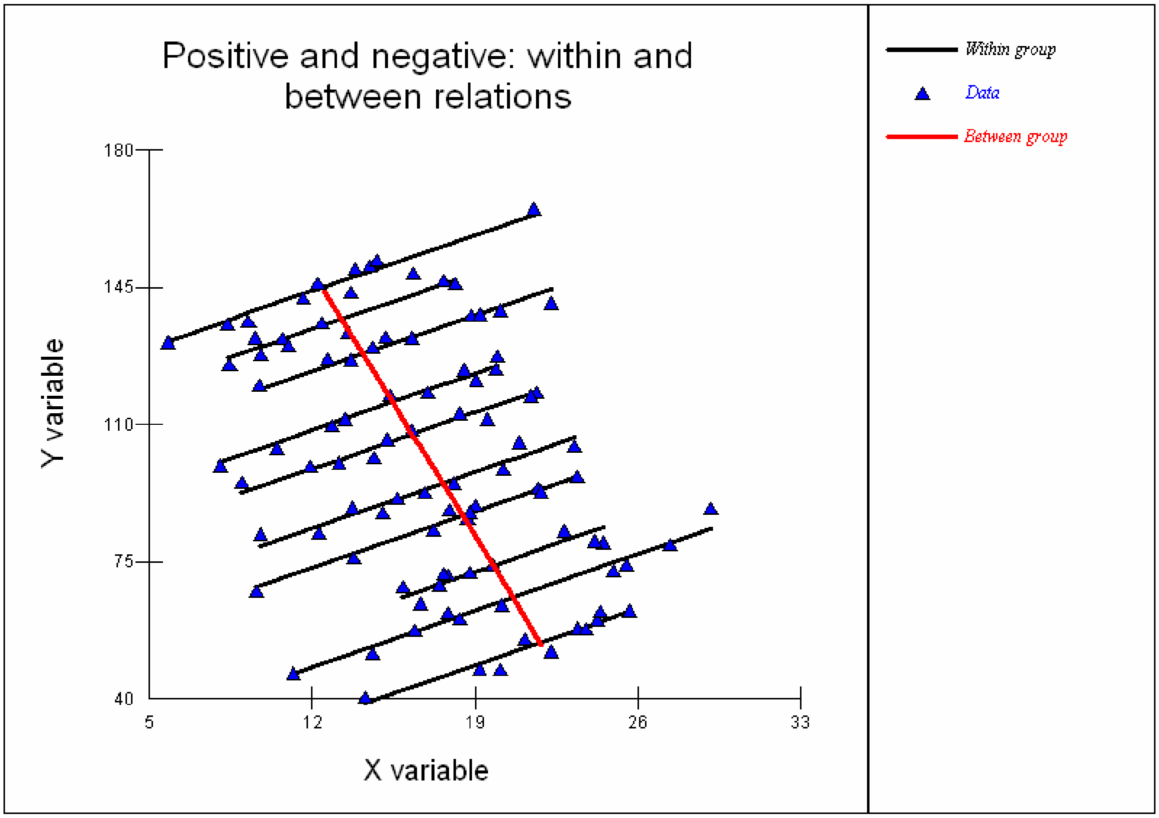
$$\beta_{0ij} = 265.441(16.575) + u_{0j} + e_{0ij}$$

$$\left[u_{0j}\right] \sim \mathrm{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[79.193(35.601)\right]$$

$$\left[e_{0ij}\right] \sim \mathrm{N}(0,\ \Omega_e)\ :\ \Omega_e = \left[4.106(0.612)\right]$$

*-2*loglikelihood(IGLS Deviance)* = 477.702(100 of 100 cases in use)

It is now clear that the between-group relation between Y and Xbar is markedly negative, while the within-group relation with X-Xbar is positive. The true relation between Y and X is only revealed when the within- and between-group relations are considered jointly in a multilevel model.



Comparison of results from single-level (OLS) and multilevel regression of Y on Xvar

	Single-level model		**Multilevel model**	
	Estimate	(Standard error)	Estimate	Standard error
Coefficients				
Cons (β_0)	137.60	(10.49)	66.04	(11.28)
Xvar (β_1)	-2.17	(0.59)	2.05	(0.05)
Variances				
Between-individual	844.06	(119.37)	4.11	(0.61)
Between-group	-	-	1265.49	(566.16)
Deviance	957.61		505.37	
Units: Individuals	100		100	
Units: Groups	-		10	

References

Bickel, Robert. 2007. *Multilevel Analysis for Applied Research: It's Just Regression*. Guilford Press.

Goldstein, Harvey. 2003. *Multilevel Statistical Models*. London: Arnold.

Gorard, Stephen. 2003a. What is multi-level modelling for? *British Journal of Educational Studies* 51(1): 46-63.

Gorard, Stephen. 2003b. *Quantitative Methods in Social Science Research*. Continuum.

Jones, Kelvyn, and Nina Bullen. 1994. Contextual models of urban house prices: A comparison of fixed- and random-coefficient models developed by expansion. *Economic Geography* 70(3): 252-272.

Paccagnella, Omar. 2006. Centering or not centering in multilevel models? The role of the group mean and the assessment of group effects. *Evaluation Review* 30(1): 66-85