SC708: Hierarchical Linear Modeling
Instructor: Natasha Sarkisian
Analyzing an HLM article: Brief Answers

Article: Huffman, Matt L. 2004. More pay, more inequality? The influence of average wage levels and the racial composition of jobs on the Black–White wage gap. Social Science Research 33, 498–520.

1. What are the two levels of this analysis?
Persons and jobs

2. What variables are used on level 1?
See table 1 under individuals

3. What variables are used on level 2?
See table 1 under jobs

4. What centering options are used for each of these variables (mark in the variable lists above)? Given the centering options, what is the meaning of the intercept?

All but race, race composition, and job rank are grand mean centered.
The intercept is the logged earnings for someone White, in an all-White job that has job rank of zero, at means of all other variables (all White job with job rank 0 is probably not a realistic scenario).

5. Which effects are modeled as random (i.e., what coefficients can vary across models)?

Race only – most likely only Black dummy, not other race/ethnicity dummies (they did not test whether any other level 1 variables have significant slope variance; they just use all other variables as controls).

6. How many variance-covariance elements would there be in the tau matrix for this analysis?

Three, two variances and a covariance. By the way, this statement is false: Finally, u_{0j} and u_{1j} are level-2 random effects, assumed to be uncorrelated and with means of zero. The article used HLM software, and HLM software always allow for the two random effects to be correlated with each other (unlike Stata where default is no correlation).

7. Interpret the results presented in Table 2. What are the main findings? Discuss effects of level 1 variables, level 2 variables, and cross-level interactions on wages.

Table 2

Effects of individual and job-level characteristics on earnings (logged): 2-level hierarchical linear regression results

Variable	Model 1 coeff.	Model 2 coeff.	Model 3 coeff.	Model 4 coeff.	Model 5 coeff.
Intercept (β_0)					
Intercept (γ_{00})	2.353*	2.383*	2.397*	1.896*	1.890*
Job % Black (γ_{01})	—	—	−.169*	—	.068*
Job rank, local hierarchy (γ_{02})	—	—	—	.010*	.010*
Job % Black job rank	—	—	—	—	−.001
Black (β_1)					
Intercept (γ_{10})	—	−.051*	−.039*	.030*	.091*
Job % Black (γ_{11})	—	—	−.083*	—	−.299*
Job rank, local hierarchy (γ_{12})	—	—	—	−.002*	−.002*
Job % Black × job rank	—	—	—	—	.004*
Control variables included	None	Level-1	All	All	All
Level-1 R^2	—	.203	.336	.407	.404
Variance components					
Level-1 variance (σ^2)	.269	.237	.237	.236	.237
Intercept (τ_{00})	.154	.100	.044	.015	.015
Intraclass correlation coefficient (ρ)	.364	.297	.157	.060	.060

Notes. $*p < .001$ (two-tailed tests).

Focusing on Model 5: a White person in a job with no Blacks that has a rank of 0 (once again, unrealistic scenario) has log wage value of 1.890. For a White person, every 1 unit increase in %Black adds 0.068 to log wage (we do not say that this is for job rank 0 because job rank by % Black interaction is not significant for Whites). But what is one unit increase in job %Black? We would think that it is 1%, but if we look at Table 1, we see that this variable is not actually % but proportion ranging from 0 to 1 rather than from 0-100. So 1 unit increase means difference between no Blacks to all Blacks. So 0.068 increase in log wages corresponds to a change in a White person's wage if they go from a job with no Blacks to a job with all Blacks—not a very large increase for some a dramatic change. Still, for a White person, %Black in a job appears to benefit them slightly. Also, for a White person, one unit increase in job rank results in 0.01 increase in log wage (job rank is actually measured in percentiles so it ranges from 0-100 so one unit is one percentile). Turning to effects of race, we find that for someone in a job with 0% Black and ranked at 0, being Black rather than White results in 0.091 increase in log wage. This is again an unrealistic scenario because one cannot be Black in a job with 0% Blacks. That effect of being Black is moderated by both % Black and job rank, and the two also interact with each other (so that is a three-way interaction, between Black, %Black in a job and job rank). To calculate the effect of being Black in a job with 100% Black that has rank 0, we add intercept of Black plus coefficient for job % black: 0.091-.299=-0.208. So for an all-Black job at rank 0, we see a negative effect of being Black on wages. Next, we can calculate the effect of being Black in a job with 0% Black that has, say, rank of 50 (50th percentile). We add the intercept for Black (which is the main effect of Black) plus 50 times the cross-level interaction between Black and job rank: 0.091-0.002*50=-0.009. So there is only a slight negative effect (likely non-significant) of being Black in a job with 0% Black at 50th percentile. What if the same job is 100% Black? We have to take both two-level interactions and three-level interaction into account: di 0.091-.299*1 - 0.002*50 +0.004*1*50=-.108. So the negative effect of being Black in a job that is 100% Black at 50th percentile in terms of job rank is more pronounced: in this situation, a Black person's log wage is .108 lower than a White person's.

In this interpretation, we focused on Black coefficient and used the other variables as moderators; we could reverse it and focus on level 2 variables and use Black as a moderator. For example, the effect of job % Black at rank 0 for Whites is 0.068 (that is, as a job goes from 0% Black to 100% Black, a White person's wage increases by 0.068). The effect of job % Black at rank 0 for a Black person, though, is 0.068 – .299=-.231 (so as a job goes from 0% Black to 100% Black, a Black person's wage decreases by .231). So the same cross-level interaction (-.299) can be used either to modify the coefficient for Black if that is our main focus, or the coefficient for job % Black if that's the main focus. But really, this is all just interpreted the same way as we interpret interactions in regular OLS – we should just realize that the intercept for Black here is actually the main effect of Black and each of the coefficients in that section for Black is actually an interaction coefficient (Black x job % Black, etc.).

Equation for predicted values:

Logged earnings=1.89+.068*Job % Black + .01*Job rank -.001*Job%Black * Job rank + .091*Black - .299*Black*Job % Black-.002*Black*Job rank +.004*Black*Job % Black*Job rank

Effect of Job % Black: .068 -.001* Job rank -.299*Black+.004*Black*Job rank

Effect of Job rank: .01 -.001*Job%Black -.002*Black +.004*Black*Job % Black

Difference between Black and White: .091-.299*Job % Black-.002*Job rank +.004*Job % Black*Job rank

Effects of Job% Black and Job Rank depending on the levels of both variables:

Effect	**White**	**Black**	**Condition**
Effect of Job% Black (from 0 to 100%=1 unit)	.068	.068-.299=-.231	0 Job rank
	.068-.001*100=-.032	.068-.299-.001*100+.004*100=.069	100 Job rank
Effect of Job rank (from 0 to 100 rank=100 units)	.01*100=1.0	(.01-.002)*100=.8	0% Black
	(.01-.001)*100=0.9	(.01-.001-.002+.004)*100=1.1	100% Black

Difference of Black vs White:

	0 Job rank	**100 Job rank**
0% Black job	.091	.091-.002*100=-.109
100% Black job	.091-.299=-.208	.091-.299-.002*100+.004*100=-.008

8. Interpret variance component information presented in the table.

Level 1 unexplained variance only decreases between model 1 and model 2 because that is where we introduce all the level 1 variables. Level 1 variables can explain both level 1 and level 2 variance, while level 2 variables can only explain level 2 variance. So after Model 2, we only add level 2 variables and therefore we only explain level 2 variance. We can only see here how well we are doing interpreting intercept variance; we are also explaining slope variance but that is not included in the table. Level 2 intercept variance decreases from to .015 in the final model; so we explained most of it. We can't say, however, whether that variance component was significant to start with and whether what is remaining in the last model is statistically significant. We can also see that ICC decreases across models – that's because we continue explaining level 2 variance while level 1 variance stays pretty much the same, so level 2 variance becomes smaller and smaller proportion of the unexplained variance.

9. How are the intraclass correlations calculated?

See footnote 15

10. How did the author calculate the R squared presented in the table?

See footnote 14; that is in fact not just level 1 R squared (as it is reported in the table) but overall R2 because it assesses change in both components. Also, note that they left out the slope component and the covariance. It would be better for them to reestimate the model with fixed slope and use level 1 and level 2 variance from that model to calculate R squared. Perhaps they did that but that's not clear from their description, so we don't know what happened to the slope variance.

11. What information about variance components is missing from the table? What additional insights would that information provide us?

Missing components: no significance tests, no slope variance, no covariance component. If we had that, we could see whether remaining variance is still significant, as well as see how well we do in explaining slope variance when we add cross-level interactions. We could also see how the slope of Black is related to the size of the earnings for Whites across jobs.

12. Using Stata, how would you generate a graph like the one presented in Figure 1?

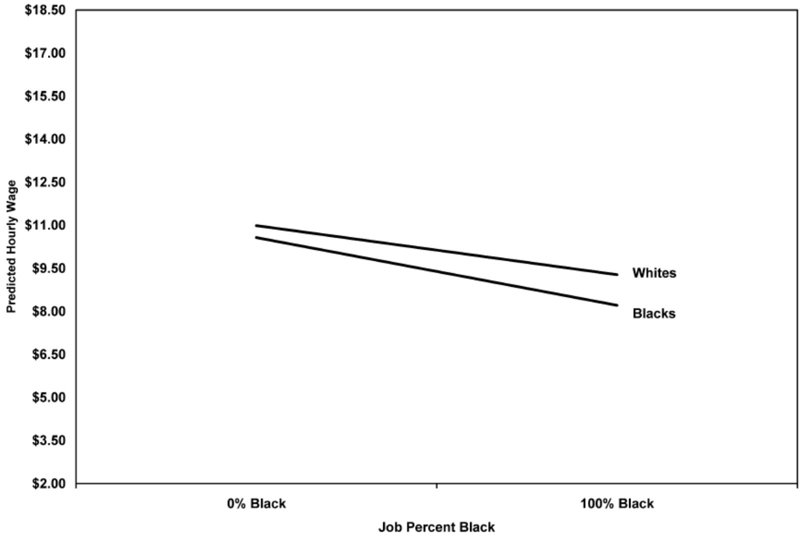

Fig. 1. Association between job percent Black and wages, by race. *Note*. Predictions based on Model 3 of Table 2.

margins, at(Black=(0 1) percBlack=(0 1)) atmeans

marginsplot, x(percBlack)

We would need to tinker further to get the appearance of the graph correct but I wanted to just focus on the substantive shape.