Let's clarify how to interpret RD estimates where the slopes differ around the discontinuity. Recall that we fit the following three models:

$$READSCORE_i = \beta_0 + \beta_1 COHORTSIZE_i + \beta_2 SMALLCLASS_i + \epsilon_i \tag{1}$$

$$\begin{aligned} READSCORE_i = &\beta_0 + \beta_1 COHORTSIZE_i + \beta_2 SMALLCLASS_i + \\ &\beta_3 COHORTSIZE \times SMALLCLASS_i + \epsilon_i \end{aligned} \tag{2}$$

$$\begin{aligned} READSCORE_i = &\beta_0 + \beta_1 COHORTSIZE_i + \beta_2 \text{COHORTSIZE}_i^2 + \\ &\beta_3 SMALLCLASS_i + \epsilon_i \end{aligned} \tag{3}$$

These were our results:

Table 1: Regression Discontinuity estimates of the effects of class size on reading scores

	(1)	(2)	(3)
Constant	75.825*** (4.202)	96.046*** (8.621)	68.334*** (1.516)
size	−0.139 (0.119)	−0.725** (0.248)	
poly(size, 2)1			−33.987 (17.666)
poly(size, 2)2			22.716* (10.154)
I(size >40.5)	3.953* (1.800)	−24.346* (10.708)	5.894** (1.990)
size:I(size >40.5)		0.757** (0.282)	
Observations	423	423	423
R^2	0.015	0.031	0.026
Note:	*p<0.05; **p<0.01; ***p<0.001		

We know from our graphical observations that the effect of an offer of small class size is not negative at the discontinuity (see Figure 1), but the coefficient on class size is a whopping -24.346. So what gives?

We need to project the fitted values that our regression results predict *at the discontinuity*. The most straightforward way is to plug values in for Eq. 2 for grade cohorts that are just under- and over the threshold for being divided in two by Maimonides' Rule using the estimated coefficients in Table 1.

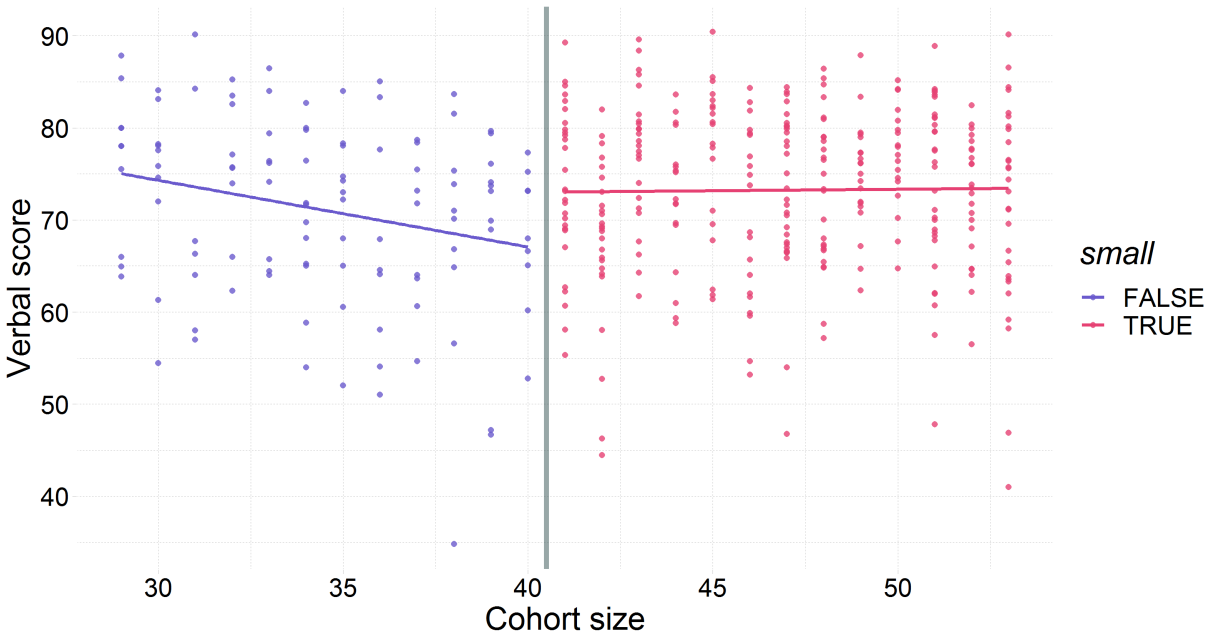

Figure 1: Relationship between reading scores and cohort size, allowing relationship to vary on either side of discontinuity

So:

Big Class, grade cohort=40:

$$\begin{aligned} \hat{READSCORE_i} =& 96.046 + (-0.725)(40) + (-24.346)(0) + (0.757)(40)(0) \\ & 96.046 + (-29) + 0 + 0 \\ & 67.046 \end{aligned} \tag{4}$$

Small Class, grade cohort=41:

$$\begin{aligned} \hat{READSCORE_i} =& 96.046 + (-0.725)(41) + (-24.346)(1) + (0.757)(41)(1) \\ & 96.046 + (-29.725) + (-24.346) + 31.037 \\ & 73.012 \end{aligned} \tag{5}$$

So the predicted effect of being assigned to receive a smaller class when we allow the slopes to vary around the discontinuity is $73.012 - 67.046 = 5.9766$, or slightly larger than either the linear, constant slope or the quadratic specifications.