Supplementary Materials: Online Appendix for Regression Discontinuity Applications with Rounding Errors in the Running Variable

Yingying Dong*
University of California Irvine

Revised April 2013

Appendix

This online supplementary appendix contains two main sections:

Appendix A.1 derives the bias corrected RD treatment effect estimand when the outcome model is a fourth order polynomial, i.e., equations (4) and (5)

Appendix A.2 gives the estimated biases along with standard errors for the two empirical applications in the paper.

Appendix A.1: Recall that the regression of the outcome Y_i on the underlying continuous running variable X_i^* can be written as the regression in equation (1), where the true treatment effect is given by $\tau = b_0$. Although one cannot estimate this regression, by Corollary 1, $B = M^{-1}C$, where B is the vector of coefficients b_0, b_1, ..., b_J and C is the vector of coefficients c_0, c_1, ..., c_J.

In the case of $J = 4$, $B = M^{-1}C$ has the form

$$\begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} = \begin{pmatrix} 1 & \mu_1 & \mu_2 & \mu_3 & \mu_4 \\ 0 & 1 & 2\mu_1 & 3\mu_2 & 4\mu_3 \\ 0 & 0 & 1 & 3\mu_1 & 6\mu_2 \\ 0 & 0 & 0 & 1 & 4\mu_1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} \tag{1}$$

Notice that for each $j = 0, ..., J$, the coefficients in column $j+1$ of M are just the binomial coefficients of $(X+e)^j = \sum_{k=0}^{j} \binom{j}{k} X^{j-k} e^k$. The true treatment effect $\tau = b_0$ is given by the first element of the right-hand side of equation (1), i.e., $\tau = c_0 - \frac{1}{2}c_1 + \frac{1}{6}c_2 - \frac{1}{30}c_4$, which is equation (4).

For the case where e has a uniform distribution, $\mu_k = 1/(k+1)$ and M becomes

$$M = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & \frac{3}{2} & 2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{so} \quad \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{6} & 0 & -\frac{1}{30} \\ 0 & 1 & -1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & -\frac{3}{2} & 1 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix},$$

so $\tau = c_0 - \mu_1 c_1 + \left(2\mu_1^2 - \mu_2\right) c_2 + \left(-6\mu_1^3 + 6\mu_2\mu_1 - \mu_3\right) c_3 + \left(24\mu_1^4 - 36\mu_1^2\mu_2 + 8\mu_3\mu_1 + 6\mu_2^2 - \mu_4\right) c_4$, which then gives equation (5).

*Correspondence: Department of Economics, 3151 Social Science Plaza, University of California Irvine, Irvine, CA 92697-5100, USA. Phone: (949)-824-4422. Email: yyd@uci.edu. http://yingyingdong.com/.

Appendix A.2

Table A1 Biases in the RD estimates of the health insurance rate increase at age 65 using age in years

	3rd order polynomial			4th order polynomial		
	(1)	(2)	(3)	(1)	(2)	(3)
[-6, +6)	0.010	0.008	0.007	0.006	-0.006	-0.006
	(0.007)	(0.007)	(0.007)	(0.023)	(0.023)	(0.023)
[-9, +9)	0.010	0.010	0.009	0.008	0.005	0.004
	(0.003)***	(0.003)***	(0.003)***	(0.008)	(0.008)	(0.008)
[-12, +12)	0.008	0.008	0.008	0.007	0.005	0.005
	(0.002)***	(0.002)***	(0.002)***	(0.004)	(0.004)	(0.004)
[-15, +15)	0.007	0.007	0.007	0.010	0.010	0.010
	(0.002)***	(0.002)***	(0.002)***	(0.003)***	(0.003)***	(0.003)***

Note: Estimates are based on HRS 1992-2008; (1) does not control for covariates; (2) controls for year dummies; (3) controls for year dummies and additional demographic variables. *Significant at the 10% level; ** Significant at the 5% level; ***Significant at the 1% level.

Table A2 Biases in the RD estimates of the retirement effect on food consumption

	(1)		(2)	
[-6,+6)	0.080	(0.020)***	0.079	(0.019)***
[-10,+10)	0.075	(0.026)***	0.067	(0.026)***
[-15,+15)	0.068	(0.009)***	0.064	(0.009)***

Note: Estimates are based on male household heads, UHS 1997-2006; (1) controls for year dummies family size, family size squared, and education levels; (2) only controls for year dummies. Boostrapped standard errors are in the parentheses; * significant at the 10% level; ** significant at the 5% level, *** significant at the 1% level.