MIT (14.32) **Spring 2021**
J. Angrist **(2023 practice)**

FINAL EXAM

1 True/False/Uncertain (2 points each)

Indicate whether the following are true, false, or uncertain, and provide a brief explanation.

1. Suppose you're interested in the effects of class size on student achievement. You propose to explore this by running an experiment that randomizes class size in a fixed number of schools. You'll analyze the experimental data using regressions of the form

$$Y_{ic} = \alpha + \beta X_{ic} + \gamma S_c + \varepsilon_{ic},$$

where Y_{ic} is the 6th grade math score of student i enrolled in class c, X_{ic} is a control for students' 3rd grade test scores, S_c is the size of class c and ε_{ic} is a regression residual.

- Your research budget allows you to collect data on n_c students in each of K classes per school. To increase statistical power, it's usually better to increase K (classes sampled in each school) than to increase n_c (students sampled in each class).

TRUE (1 pt). Because the data are clustered, the effective smpl size is #clusters rather than #students(1 pt)

2. Suppose you're interested in the causal effect of completing an Economics major (denoted by dummy variable M_i) on post-college wages. You'd like to estimate this by regressing graduates' wages on M_i with controls for parental education (P_i) and student academic ability as measured by SAT scores (A_i).

- You have data on P_i but not A_i. When estimated with controls for P_i alone, a regression estimate of the effect of an Economics major on wages is likely to be too large relative to the one you want.

UNCERTAIN (1pt). sign of OVB is sign of effect of A_i (presumed positive) times the regression of A_i on M_i, which is uncertain, econ students may have higher or lower SATs (1 pt)
TRUE (1pt) sign of OVB is sign of effect of A_i (presumed positive) times the regression of A_i on M_i, which is positive since econ students are smarter(1 pt)

3. Female college graduates are said to be less likely than men to land highly paid positions in the tech sector. Imagine investigating this by using a sample of recent male and female MIT grads to estimate the regression model:

$$Y_i = \alpha + b_1 C_i + b_2 F_i + b_3 (C_i \times F_i) + e_i,$$

where Y_i is log wages five years after graduation, C_i is a dummy for being a Computer Science major, and F_i is a dummy for female.

- Parameter b_3 measures the economic returns to a Computer Science major for women.

FALSE (1 pt). Effect of CS major on women is $b_1 + b_3$ (1 pt)

4. Suppose you're interested in the regression of Y_i on X_i. Alas, you don't observe X_i, rather, you observe only $W_i = X_i + u_i$, where u_i is measurement error that's uncorrelated with X_i and Y_i.

- The slope from a regression of Y_i on W_i tends to be too small relative to the slope you seek.

TRUE (1 pt). The slope you get is $r\beta$, where β is slope you want and r is reliability, a value in $[0, 1]$

5. The covariance between regression residuals and fitted values is approximately zero.
FALSE (1 pt). Cov (resids, fits) is exactly zero (1 pt).

2 Short Questions (12 points each) [JA review]

1. The Current Population Survey collects data on American's smoking habits. Consider a regression of a dummy for smoking on a dummy for having a college degree.

   (a) [6 points] Interpret the college coefficient in this regression in terms of conditional probabilities of being a smoker given college completion.

   (b) [6 points] College might affect smoking rates differently for women and men. Write down a simple regression model you can use to test this hypothesis and explan the test.

   (a) The slope coefficient measures P[smokes|went to college]-P[smokes|no college]. Half credit for E instead of P. Full credit for a correct verbal statement like "the slope measures the difference in the conditional prob of smoking between those with and w/o college"
   (b) Look for a model like that in T/F Q3 where Y=smoke, C=college, F=female. Enough to notice this is the same as T/F 3. Loose 2 points for not saying: the test is a t-test for the college=female interaction"

2. The 2020 pandemic-induced economic crisis prompted many governments to borrow and spend in a manner never seen before. This "fiscal stimulus" has prompted debate over the implications of large government deficits. Imagine exploring this using time series data to estimate a regression of the following sort:

$$Y_t = \alpha + \beta U_{t-1} + \beta D_t + \epsilon_t, \tag{1}$$

   where Y_t is American GDP growth in year t, U_{t-1} is the unemployment rate last year, and D_t is the size of the current government debt as a percent of last year's GDP.

   (a) [6 points] Briefly explain why conventional OLS standard errors for this regression are likely to be misleading. Are they too big or too small?

   (b) [6 points] Briefly explain two ways to address this problem.

   (a) time series resids likely (positively) serially correlated; this makes concventional SEs too small (half-credit for writing only "too small"; half-credit for serial corr but "too big"
   (b) method 1: HAC SEs; method 2: GLS or quasi-differencing (half-credit if only one)

3. Recall our in-class experiment analyzing the effects of in-person attendance on midterm grades. This experiment randomized the opportunity to attend in-person.

   (a) [2 points] Ignoring the experiment, we can simply regress midterm grades on in-person attendance in the first half of term. A regression of midterm grades on number of classes attended generates a marginally significant estimate of 1.4. Explain in a few words why this might be a misleading measure of the causal effect of in-class attendance.

   (b) [4 points] Using information from the attendance lottery, we saw that:

      i. Among students eligible to attend, those randomly offered the chance to attend in-person indeed attended 7.1 more days in-person than those not offered the chance to attend. In an instrumental variables (IV) framework, what's this relationship called?

      ii. Among students eligible to attend, those randomly offered the chance to attend in-person earned midterm grades 9.2 points higher than those not offer. In an IV framework, what's this relationship called?

   (c) [2 points] briefly explain how to use the estimates noted in (b) to construct an IV estimate of the effects of in-person days attended on midterm grades

(d) [4 points] Let Y_i denote student i's midterm grade; let X_i indicate students eligible for the in-person attendance lottery; let Z_i indicate eligible students offered the chance to attend in-person, and let D_i denote in-person days actually attended.

i. [2 points] Explain how to use this data to compute "manual two-stage least squares" (2SLS) estimates of the effect D_i on Y_i using a pair of OLS regressions.

ii. [2 points] Explain briefly how you'd get 2SLS estimates in practice – why is the procedure in (i) unsatisfactory?

(a) likely selection bias or OVB (key terms) (full credit for use of terms or for saying "those coming to class more enthusiastic"
(b) (i) first-stage (ii) reduced form (2 points for each)
(c) divide reduced form by first stage
(d) (i) regress D on X and Z, save fits, regress Y on X and D_hat (lose a point if forget X) (ii) use Stata ivregress (1 pt); SEs for manual no good (1 pt)

4. In April 1980, over 125,000 mostly low-skill immigrants arrived in a flotilla from Mariel Bay, Cuba. This episode increased the Miami labor force by 8 percent in a matter of weeks. Card (1990) compares changes in employment and wages in Miami with analogous changes in four control cites (Atlanta, Houston, Los Angeles, and Tampa-St. Petersburg). The idea behind this is that the comparison cities make a good control group in a differences-in differences (DD) setup. The table below shows black unemployment rates by city and year with standard errors in parentheses:

Group	Year		
	1979	1981	1981-1979
Miami	8.3 (1.7)	9.6 (1.8)	1.3 (2.5)
Comparison cities	10.3 (0.8)	12.6 (0.9)	2.3 (1.2)

(a) [6 points] Use these statistics to compute DD estimates and their standard errors (for the latter, assume data for different cities are independent). What do these (DD) estimates suggest about the consequences of large-scale immigration for native workers?

(b) [6 points] In the summer of 1994, tens of thousands of Cubans again boarded boats destined for Miami. This time, however, the US Navy intercepted the would-be immigrants. Only a small fraction reached U.S. shores. Angrist and Krueger (1999) report differences-in-differences estimates using 1993-95 data analogous to those in Card (1990). These appear below:

Group	Year		
	1993	1995	1995-1993
Miami	10.1 (2.1)	13.7 (2.8)	3.6 (3.5)
Comparison cities	11.5 (0.9)	8.8 (0.8)	-2.7 (1.2)

What concerns does this table raise in regard to the Card (1990) findings?

(a) compute DD correctly (3 pts); compute SE correctly assuming city independence (3 pts)
(b) compute DD correctly (2 points); note a large marginally sig DD for a non-event (2 pts), this suggests DD estimates need not be causal (2 pts)

3 Empirical questions (25 points each)

1. Abdulkadiroglu, Pathak, and Walters (2016; APW16) study the Louisiana Scholarship Program (LSP), which awarded vouchers covering the cost of private school attendance using a lottery. LSP vouchers

are awarded to low-income students who've been attending traditional public schools. Schools wishing to enroll voucher recipients are authorized to offer a certain number of seats to applicants with vouchers. Schools with more voucher applicants than authorized seats offer seats first to applicants with already-enrolled siblings, then to students living nearby, and finally to applicants not in these two categories previously enrolled in a low-performing public school. These groups are called priorities: after seats in the first two priority groups are filled, offers are randomly assigned to applicants in the third group. In practice, almost all applicants in the first two priority groups who want a voucher can get one.

(a) [4 points] APW16 consider applicants to be randomly assigned if they're in Priority Group 3 at their first-choice school. Columns 3-5 in APW16 Table 1 (below) compare the demographic characteristics of all LSP applicants in grades 3-8, those subject to random assignment, and those who use a voucher to enroll in an LSP-participating school (ignore the row labeled NSECD). Why is it worth comparing columns 3 and 4?

TABLE 1—DESCRIPTIVE STATISTICS FOR STUDENTS

			Louisiana Scholarship Program		
	Louisiana (1)	RSD (2)	All applicants (3)	Randomized applicants (4)	Enrollees (5)
Female	0.487	0.473	0.489	0.487	0.539
Black	0.451	0.939	0.861	0.885	0.805
Hispanic	0.044	0.031	0.031	0.033	0.039
White	0.468	0.010	0.086	0.058	0.131
NSECD	–	–	0.004	0.005	0.006
Household income: mean	–	–	15,471	15,535	17,400
25th percentile			1,300	1,455	1,452
Median			12,000	12,840	15,000
75th percentile			24,781	24,864	28,032
Observations	715,012	14,689	3,723	1,412	1,019

Notes: Columns 1 and 2 show statistics for students enrolled in Louisiana and Recovery School District (RSD) public schools in grades 3–8 in the 2012–2013 school year. These statistics are obtained from the Louisiana Department of Education website. Column 3 shows statistics for first-time applicants to Louisiana Scholarship Program (LSP) schools in grades 3–8 for 2012–2013. Column 4 shows statistics for LSP applicants subject to first choice random assignment. Column 5 shows statistics for LSP enrollees.

A. Note that fewer than half of applicants are randomized (2 pts). We'd like them to be similar to full applicant pool (2 pts)

(b) [6 points] APW16 computes 2SLS estimates of the effect of using an LSP voucher for private school on test scores, Y_i. The causal relationship of interest is:

$$Y_i = \beta P_i + \gamma_0 d_i + \sum_{j}^{J} \gamma_j S_{ij} + X_i'\delta + \varepsilon_i, \tag{2}$$

where LSP voucher use is indicated by dummy variable P_i. The paper uses a lottery offer dummy (Z_i) as an instrument for P_i. The first-stage equation is:

$$P_i = \pi Z_i + \lambda_0 d + \sum_{j=1}^{J} \lambda_j S_{ij} + X_i'\theta + \eta_i. \tag{3}$$

Equations (2) and (3) control for a dummy (denoted d_i) indicating applicants' who are in priority group 3 at their first-choice school and therefore subject to random assignment. They also control for a set of J dummy variables, denoted S_{ij} for school j, indicating the identify of the applicant's first choice school. For example, an applicant whose first-choice school is school number 1 has $S_{i1} = 1$ and the other school dummies equal to zero. Finally, these equations control for a vector of additional controls, denoted X_i, containing dummies for gender, race, and family income quartiles.

i. [2 points] Why are OLS estimates of equation (2) unlikely to be unsatisfactory?

ii. [2 points] Which, if any, of the controls included in (2) and (3) are likely to be *necessary* for the 2SLS procedure to consistently estimate causal effects? Why?

iii. [2 points] Briefly state the two conditions needed for this 2SLS procedure to reliably estimate the causal effects of LSP voucher use on test scores.

A. (i) selection bias or OVB [full credit for these terms or for saying "families using vouchers may be more motivated"
(ii) Control for first-choice school and priority necessary as these determine assignment risk
(iii) $\pi \neq 0$ in first-stage, eq 3; Z_i uncorr w/u_i in causal model, eq 2. Half credit for correct-sounding verbal statement.

(c) [5 points] APW16 Table 3 (below) compares the demographic characteristics and household income of LSP voucher winners and losers. Briefly explain why this table is important. As an empirical matter, what does it show?

TABLE 3—COVARIATE BALANCE

	Non-offered mean (1)	Offer differential: Full sample (2)	Offer differential: With follow-up (3)
Female	0.474	0.012	0.008
		(0.033)	(0.035)
Black	0.900	−0.034	−0.028
		(0.021)	(0.022)
Hispanic	0.030	0.003	0.001
		(0.012)	(0.013)
White	0.050	0.019	0.018
		(0.015)	(0.016)
NSECD	0.004	−0.001	−0.002
		(0.006)	(0.006)
Household income	15,410	1,636	1,025
		(1,097)	(1,118)
Income below p_{25}	0.254	−0.007	0.000
		(0.029)	(0.030)
Income below p_{50}	0.503	−0.030	−0.017
		(0.035)	(0.036)
Income below p_{75}	0.753	−0.048	−0.028
		(0.034)	(0.035)
Joint p-value	–	0.659	0.932
Observations		1,412	1,248

Notes: This table compares characteristics of offered and non-offered applicants to Louisiana Scholarship Program schools for grades 3–8 in the 2012–2013 school year. The sample is restricted to first-time applicants subject to first choice random assignment. Column 1 reports mean characteristics for applicants not offered a seat, while columns 2 and 3 report differences between offered and non-offered applicants. These differences come from regressions that control for risk set indicators. The sample in column 3 is restricted to applicants with follow-up test scores. p_{25}, p_{50}, and p_{75} refer to the twenty-fifth, fiftieth, and seventy-fifth percentiles of household income in the non-offered group. The last row shows p-values from tests that all differentials equal zero. Standard errors, clustered by risk set, are in parentheses.

A. Why important? Cov balance makes it more likely that the lottery offer is a good instrument for LSP participation [3 points] The table shows that winners ad losers are comparable in the sense of similar demos (2 points);

(d) [5 points] APW16 Table 4 (below) reports first-stage, reduced form, 2SLS, and OLS of the effects of using an LSP voucher. How are the estimates in columns 1-3 of this table related?

Table 4—Two-Stage Least Squares Estimates of Voucher Effects on Test Scores

Subject	First stage (1)	Reduced form (2)	2SLS (3)	OLS (4)
Math	0.679	−0.281	−0.413	−0.386
	(0.029)	(0.061)	(0.091)	(0.066)
Observations		1,247		
ELA	0.679	−0.055	−0.081	−0.120
	(0.029)	(0.053)	(0.079)	(0.056)
Observations		1,248		
Science	0.689	−0.181	−0.263	−0.282
	(0.030)	(0.066)	(0.095)	(0.065)
Observations		1,221		
Social studies	0.690	−0.229	−0.331	−0.270
	(0.030)	(0.060)	(0.089)	(0.059)
Observations		1,220		

Notes: This table reports estimates of the effects of attendance at Louisiana Scholarship Program (LSP) voucher schools on LEAP/iLEAP test scores. The sample includes first-time voucher applicants subject to first choice random assignment applying to grades 3–8 in 2012–2013. Column 1 reports first-stage effects of voucher offers on attendance at an LSP school, while column 2 reports reduced form effects of offers on test scores. Column 3 reports two-stage least squares estimates of the effects of LSP participation, and column 4 reports corresponding ordinary least squares estimates. All models control for risk set indicators and baseline demographics (sex, race, NSECD, and indicators for household income quartiles). Standard errors, clustered by risk set, are in parentheses.

A. within rows, 2SLS=RF/First

(e) [5 points] APW16 argues that negative effects of LSP participation are explained by the low quality of LSP-participating *schools*. What evidence does APW16 Table 2 (below) offer in support of this

Table 2—Descriptive Statistics for Private Schools

	All Louisiana private schools			Matched city sample		
	LSP voucher schools (1)	Oversubscribed LSP schools (2)	Other private schools (3)	LSP voucher schools (4)	Oversubscribed LSP schools (5)	Other private schools (6)
Enrollment in 2012	311	243	323	323	239	349
Enrollment growth, 2000–2012	−12.4%	−16.1%	2.8%	−7.7%	−10.4%	1.9%
Tuition	\$4,898	\$4,653	\$5,760	\$5,115	\$4,740	\$6,430
Fraction black	0.327	0.433	0.158	0.387	0.517	0.188
Fraction Hispanic	0.020	0.021	0.037	0.021	0.021	0.041
Fraction white	0.622	0.517	0.752	0.564	0.433	0.714
Catholic school	0.645	0.679	0.391	0.594	0.619	0.367
Other religious affiliation	0.274	0.304	0.421	0.313	0.357	0.430
Student/teacher ratio	13.5	12.7	11.5	13.3	12.3	10.9
Days in school year	178.6	178.9	177.9	178.8	178.9	177.7
Hours in school day	6.8	6.8	6.7	6.8	6.7	6.7
Observations	124	56	235	96	42	158

Notes: This table reports characteristics of private schools in Louisiana using data from the Private School Universe Survey (PSS). Column 1 shows statistics for schools eligible for Louisiana Scholarship Program vouchers at any time through 2012–2013. Column 2 shows statistics for voucher schools with applicants subject to random assignment in 2012–2013. Column 3 shows statistics for non-LSP private schools. Columns 4, 5, and 6 report statistics for schools in cities with both LSP and non-LSP private schools. The second row reports average enrollment growth between 2000 and 2012 for schools with available data in both years. The third row measures tuition in the most recent available year, usually 2015–2016. Tuition is available for 94 percent of voucher schools and 92 percent of other private schools.

A. Compare characteristics of LSP voucher schools in cols 1-2 to those of other private schools in col 3 (likewise, compare 4-5 to 6) [1 points for thi idea]. We see that the voucher participants (i) suffer declining enrollment growth, while other private enrollment is increasing (ii) are cheaper. both suggest particpating schools are lower quality. [2 points for each observation]

2. Almond, et al. (2010; ADKW) uses a regression discontinuity design to estimate the marginal returns to medical spending. Specifically, this study focuses on the health effects of medical spending on very low birthweight (VLBW) newborns. Low birthweight is highly predictive of elevated infant mortality. VLBW newborns, defined as those weighing less than 1500 grams, are deemed to be at especially high risk of premature death. Hospitals therefore allocate substantial extra care to VLBW babies, and insurers usually cover the cost of this.

   (a) [3 points] An analysis of the of the effects of medical spending on infant mortality might begin with OLS estimates of an equation like:

   $$M_i = X_i'\beta + \rho S_i + \varepsilon_i, \tag{4}$$

   where X_i is a vector of controls for mother's family background, and S_i is the cost of medical care for newborn i. The ADKW dependent variables are indicators of mortality at different horizons (like one-year mortality, meaning the infant dies in the first year after birth).

   - Briefly explain why OLS estimates of (4) are unlikely to be a good measure of the mortality consequences of medical spending (a complete answer should note two problems with an OLS analysis)
     A. selection bias or OVB [1 pt]; higher-income families may get more care and heave healthier babies; babies in very poor health dies early, so spend less [1 each]

   (b) [2 points] ADKW Figure III (below) plots hospital charges and length of stay agains birthweight

   - What does this figure suggest about the relationship between treatment intensity and VLBW status? Why is this important?
     A. Spending seems to jump to the left of the cut-off; this suggests VLBW status indeed boost charges for babies near the cutoff.

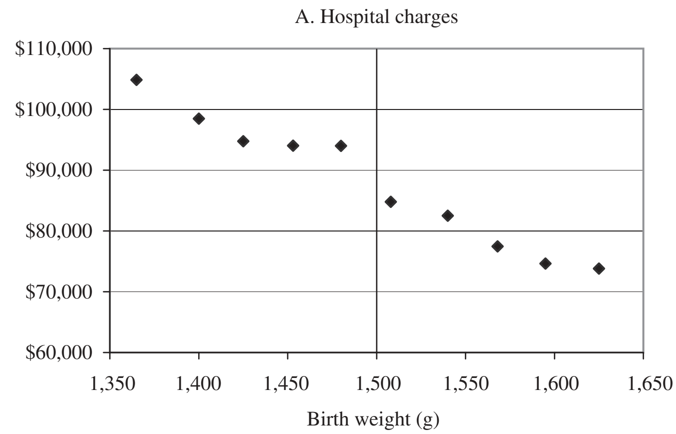

FIGURE III
Summary Treatment Measures around 1,500 g
Data are all births in the five-state sample (AZ, CA, MD, NY, and NJ), as described in the text. Charges are in 2006 dollars. Points represent gram-equivalents of ounce intervals, with births grouped into one-ounce bins radiating from 1,500 g; the estimates are plotted at the median birth weight in each bin.

(c) [5 points] ADKW Table II (below) reports regression-discontinuity estimates to go with the figure. Specifically, the top row of the table reports estimates of coefficient π in the model:

$$S_i = \alpha + \pi VLBW_i + \delta_0(1 - VLBW_i)(g_i - 1500) + \delta_1 VLBW_i(g_i - 1500) + \eta_i, \tag{5}$$

where S_i measures the amount of medical care (measured by hospital charges billed) and the length of stay in the hospital for infant i,, $VLBW_i$ indicates newborns weighing less than 1500 grams, and g_i is infant i's weight in grams. Estimates labeled "local linear model" are computed while weighting observations near 1500 more heavily; other estimates are unweighted. Some models include additional controls listed at the bottom of the table. (For purposes of this question, you can ignore the clustered standard errors in brackets.)

- What does the table suggest about the relationship between VLBW status and neonatal treatment intensity? (Use the mean of the dependent variables reported in the table to interpret magnitudes)

  A. [3 pts] VLBW status indeed boost charges for babies near the cutoff by a large amount, compare 9k to 81K mean; length of stay also goes up, compare 2 extra days to mean of 25, *both significant [full credit requires interpretation of magnitudes as well as stat sig]*

- What's the relationship between the estimates in the table and the points in Figure III?
  A. [2 points] π estimates the size of the jump in the fig

TABLE II

SUMMARY TREATMENT MEASURES BY VERY-LOW-BIRTH-WEIGHT STATUS, FIVE-STATE SAMPLE, 1991–2006

	Hospital charges				Length of stay			
	Local linear model	OLS	OLS	OLS	Local linear model	OLS	OLS	OLS
Birth weight < 1,500 g	9,450 (2,710)**	9,022 (2,448)** [3,538]*	8,205 (2,416)** [3,174]*	9,065 (2,297)** [5,094]	1.97 (0.451)**	1.7768 (0.4165)** [1.0024]	1.7600 (0.4166)** [0.9775]	1.4635 (0.4107)** [0.7928]
Birth weight < 1,500 g × grams from cutoff (100s)		−1,728 (3,700) [8,930]	−3,176 (3,647) [7,937]	617.4876 (3,463) [8,447]		−0.1012 (0.6482) [1.9397]	−0.1356 (0.6467) [1.8419]	−0.5766 (0.6366) [1.4858]
Birth weight ≥ 1,500 g × grams from cutoff (100s)		−7,331 (3,018)* [5,022]	−8,684 (2,978)** [4,337]*	−7,951 (2,823)** [7,562]		−2.3130 (0.5245)** [1.4366]	−2.3779 (0.5250)** [1.4117]	−2.5993 (0.5174)** [1.1464]*
Year controls		No	Yes	Yes		No	Yes	Yes
Main controls		No	No	Yes		No	No	Yes
Mean of dependent variable above cutoff	81,566				24.68			
Observations	28,928				30,935			

Notes. Local linear regressions use a bandwidth of 3 ounces (85 g). OLS models are estimated on a sample within three ounces above and below the VLBW threshold. Five states include AZ, CA, MD, NY, and NJ (various years). Charges are in 2006 dollars. "Main controls" are listed in Online Appendix Table A5, as well as indicators for each year. Some observations have missing charges, as described in the text. Local linear models report asymptotic standard errors. OLS models report heteroscedastic-robust standard errors in parentheses and standard errors clustered at the gram level in brackets.

(d) [5 points] ADKW Figure II (below) shows how mortality varies across the VLBW (1500 gram) threshold; Table I (below the figure) reports the corresponding estimates. These were computed by estimating equation (5) with mortality at various horizons as dependent variable.

- Interpret the estimates in Table I - does VLBW status raise or lower death rates? By how much?
  A. [3 points] VLBW status reduces mortality by 3/4-1 percentage point, a statistically signifnicant drop. [2 points] a very large improvement measured against a mean death rate of 4-5 percent

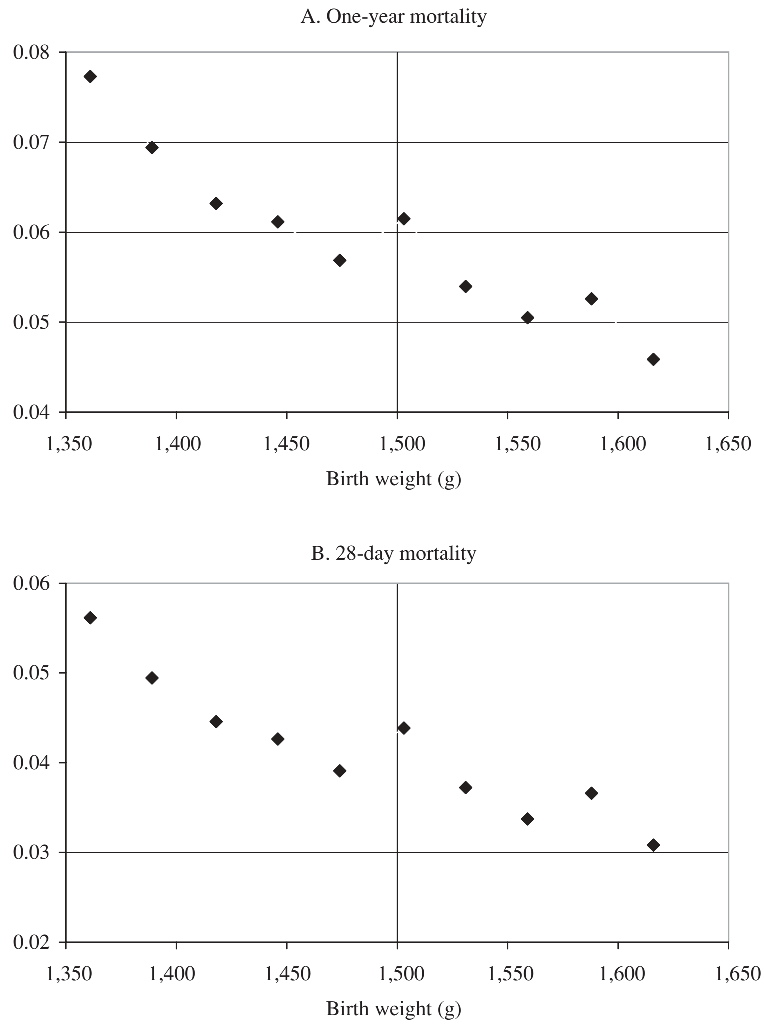

FIGURE II
One-Year and 28-Day Mortality around 1,500 g

NCHS birth cohort–linked birth/infant death files, 1983–1991 and 1995–2003, as described in the text. Points represent gram-equivalents of ounce intervals, with births grouped into one-ounce bins radiating from 1,500 g; the estimates are plotted at the median birth weight in each bin.

TABLE I
INFANT MORTALITY BY VERY-LOW-BIRTH-WEIGHT STATUS, NATIONAL DATA, 1983–2002 (AVAILABLE YEARS)

	One-year mortality				28-day mortality			
	Local linear model	OLS	OLS	OLS	Local linear model	OLS	OLS	OLS
Birth weight < 1,500 g	−0.0121 (0.0023)**	−0.0095 (0.0022)** [0.0048]*	−0.0067 (0.0022)** [0.0040]	−0.0072 (0.0022)** [0.0040]	−0.0107 (0.0019)**	−0.0088 (0.0018)** [0.0038]*	−0.0074 (0.0018)** [0.0031]*	−0.0073 (0.0018)** [0.0031]*
Birth weight < 1,500 g × grams from cutoff (100s)		−0.0136 (0.0032)** [0.0062]*	−0.0119 (0.0032)** [0.0024]**	−0.0111 (0.0032)** [0.0018]**		−0.0114 (0.0027)** [0.0055]*	−0.0102 (0.0027)** [0.0027]**	−0.0097 (0.0027)** [0.0022]**
Birth weight ≥ 1,500 g × grams from cutoff (100s)		−0.0224 (0.0029)** [0.0081]**	−0.0196 (0.0029)** [0.0074]**	−0.0184 (0.0029)** [0.0074]*		−0.0199 (0.0024)** [0.0060]**	−0.0179 (0.0024)** [0.0056]**	0.0172 (0.0024)** [0.0055]**
Year controls		No	Yes	Yes		No	Yes	Yes
Main controls		No	No	Yes		No	No	Yes
Mean of dependent variable above cutoff	0.0553				0.0383			

(e) [6 points] ADKW reports estimates of the effects of marginal spending on mortality using a fuzzy regression discontinuity strategy that builds on the estimates in Tables I and II. Specifically, VLBW classification is used to instrument for spending (S_i) in a causal mortality model that can be written:

$$M_i = \alpha + \pi S_i + \gamma_0(1 - VLBW_i)(g_i - 1500) + \gamma_1 VLBW_i(g_i - 1500) + \upsilon_i \tag{6}$$

- [3 points = 1 point each] What's the instrument for this model? Which figure shows the first stage? Which figure plots the reduced form?
- [3 points for doing the division] Use the estimates in the tables to compute your own rough-and-read fuzzy RD estimate of effect of VLBW-induced spending on mortality

(f) [4 points] In a controversial critique of the ADKW study, Barreca et al (2011) argues that manipulation of reported birthweight invalidates the ADKW research design. Specifically, Barreca et al. claim that the elevated mortality seen just to the right of the VLBW cutoff reflects selection bias: sophisticated doctors, nurses, and hospitals *know* that babies recorded as weighing below 1500g are eligible for lots of extra care. They argue that only a poorly trained or naive care provider would record a weight of 1501 grams, when by reporting 1499, the newborn qualifies for lots of extra free care.

- In the table reproduced below, Barreca et al (2011) report a series of "donut RD" estimates discarding data on newborns exactly at 1500 grams or within 1-3 grams of 1500. Panel A replicates the VLBW effects on mortality reported in columns 4 and 8 of ADKW Table I (showing larger, clustered standard errors). Panels B-D report the donut estimates. Beyond the fact that the clustered standard errors are larger, do the findings here tend to undermine the ADKW causal claims?
  A. yes: Donut estimates are mostly smaller, some insig. Maybe points near cutoff are contaminated [full credit requires observation that donut estimates are smaller, less often significant, and a brief note as to why this matters]. But give 2/4 partial credit for arguing coherently the other way: the logic of doing RD near the cutoff means we shouldnt discard the closest obs, these might be the most useful (a weak counter but insightful nonetheless)

from Barecca et al. 2011

TABLE I

REPLICATION OF ADKW'S MAIN RESULTS ALONG WITH DONUT-RD ESTIMATES

Mortality Outcome	One-year (1)	28-Day (2)	7-Day (3)	24-Hour (4)
Panel A: Our replication of ADKW's estimates				
Weight < 1,500 g	−0.0071	−0.0071*	−0.0046	−0.0033
	(0.0041)	(0.0032)	(0.0028)	(0.0020)
Observations	202,078	202,078	202,078	202,078
Clusters	171	171	171	171
Panel B: Donut RD dropping those at 1,500 g				
Weight < 1,500 g	−0.0033*	−0.0042**	−0.0023	−0.0018
	(0.0014)	(0.0013)	(0.0013)	(0.0010)
Observations	198,534	198,534	198,534	198,534
Clusters	170	170	170	170
Panel C: Donut RD dropping those within 1 g of 1,500-g cutoff				
Weight < 1,500 g	−0.0035*	−0.0043**	−0.0024	−0.0018
	(0.0014)	(0.0012)	(0.0013)	(0.0010)
Observations	198,334	198,334	198,334	198,334
Clusters	168	168	168	168
Panel D: Donut RD dropping those within 2 g of 1,500-g cutoff				
Weight < 1,500 g	−0.0027*	−0.0037**	−0.0019	−0.0013
	(0.0014)	(0.0012)	(0.0012)	(0.0009)
Observations	197,135	197,135	197,135	197,135
Clusters	166	166	166	166
Panel E: Donut RD dropping those within 3 g of 1,500-g cutoff				
Weight < 1,500 g	−0.0018	−0.0026	−0.0018	−0.0011
	(0.0019)	(0.0015)	(0.0015)	(0.0014)
Observations	175,108	175,108	175,108	175,108
Clusters	164	164	164	164

Notes. Results are based on Vital Statistics Linked Birth and Infant Death Data, United States, 1983–2002 (not including 1992–1994). Estimates use a bandwidth of 85 g and rectangular kernel weights, standard errors are clustered at the gram level, and all models include a linear trend in birth weights that is flexible on either side of the cutoff. All estimates include controls for prenatal care, mother's age, mother's education, father's age, child gender, gestational age, mother's race, plurality of birth, birth order, and year. * significant at 5%; ** significant at 1%.