The Long-Run Effect of Income During Infancy: Evidence from the EITC (Preliminary Work)

Andrew Barr[1] and Jonathan Eggleston[2] and Alex Smith[3]

[1]Texas A&M University, [2]U.S. Census Bureau, [3]West Point

May 1, 2019

Census Disclaimer: *This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on statistical issues are those of the authors and not necessarily those of the U.S. Census Bureau. The Census Bureaus Disclosure Review Board and Disclosure Avoidance Officers have reviewed this product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. CBDRB-FY18-315 and CBDRB-FY19-323.*

Army and Department of Defense Disclaimer: *The opinions expressed herein reflect our personal views and not those of the U.S. Army or the Department of Defense.*

Family Income and Child Outcomes

- Well-documented strong correlation between family income and child outcomes
- Early childhood may be a particularly important period (Duncan et al. 2010)
- The primary driver of this correlation is unclear.
  - Direct effect of income itself
  - Environmental factors correlated with income (e.g. school/neighborhood quality or parenting style)

Research Question

What is the long-run causal effect of an increase in family income during infancy on a child's later outcomes?

Prior Work on Causal Effects of Family Income on Child Outcomes

- Welfare-to-work experiments
  - Gennetian and Miller (2002), Morris and Gennetian (2003), Hill et al. (2001), Clark-Kauffman et al.(2003)
- Natural experiments generating variation in local resources
  - Loken (2010), Loken et al. (2012), Akee et al. (2010)
- EITC schedule changes
  - Evans and Garthwaite (2014), Strully et al. (2010), Dahl and Lochner (2012), Bastion and Michelmore (2018)
- Benefit provision during childhood
  - *Cash:* Aizer et al. (2016)
  - *In-Kind:* Hoynes et al. (2016), Barr and Smith (2018), Anders et al. (2018), Heckman et al. 2010, Olds et al. (1998)

This Paper

- Leverage variation in family income due to birth timing with a Regression Discontinuity design
  - Jan 1 birthdate discontinuity in family's EITC amount
  - Child born prior to Jan 1 $\rightarrow$ family income $\uparrow$ in 1st year of life (vs. after Jan 1).
- Detailed data on all students in NC Public Schools admin data
- Universe of tax records (selected years) with child-parent links

Contributions

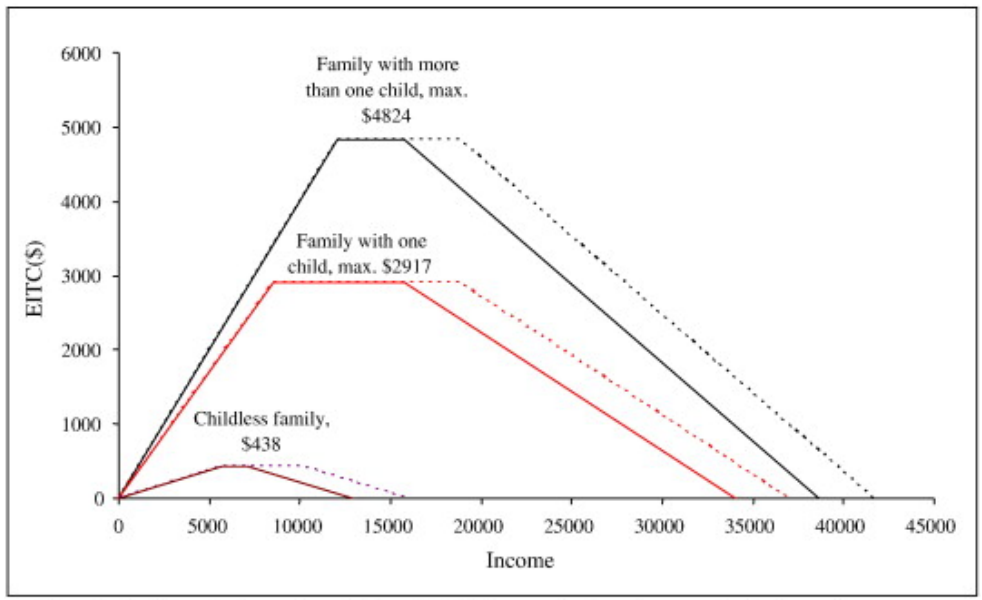
- Causal estimates of the long-run effects of income provided in the 1st year of life.
- Modern causal estimates of the long-term effects of resources during early childhood
- Estimates of the effects on children of a *pure income transfer* generated by the EITC

Preview of Results

- Academic outcomes from NC Public Schools admin data
  - 0.05 SD ↑ in index of student outcomes for likely EITC-eligibile students
    - ★ HS Graduation ↑ 1.9 pp
    - ★ Test scores ↑ 0.057 SD
    - ★ Likelihood of suspension ↓ 2.2 pp
  - Roughly \$1,000 ↑ family income → 10% of FRL/non-FRL gap
  - Effects driven by white children
- Earnings and other outcomes from Census (preliminary)
  - Universe of tax records with child-parent links
  - 0.3 to 0.4 percentile increase in earnings (age 23-25)
  - \$335 to \$890 increase in earnings (age 23-25)

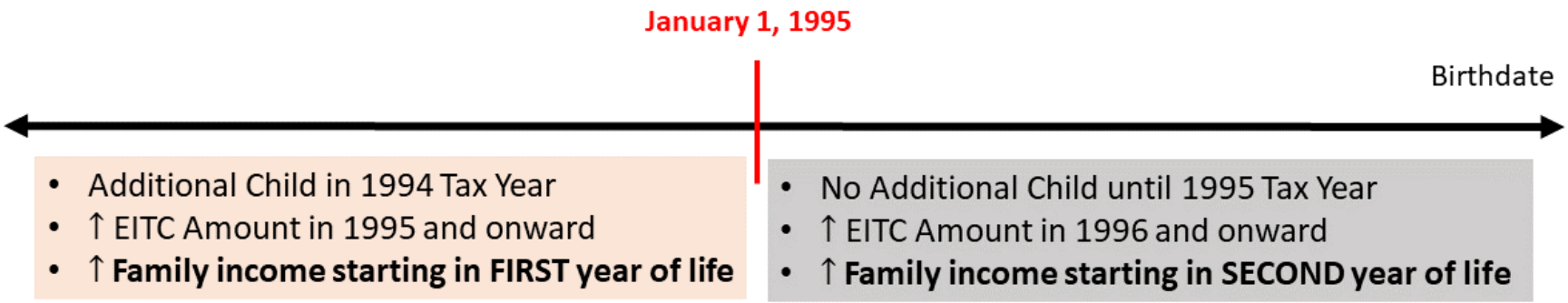
Earned Income Tax Credit (EITC)

- Largest cash transfer program to poor families in the U.S.
  - 27 million families received 65 billion dollars (2017)
  - Lifted 5 million children out of povery (2017)
- Average credit for 1993-1998 was approx $2,000 (2015 dollars)
- Large increase credit amount for 1st child (smaller increase for 2nd)



Note: Dashed lines indicate the additional $3000 increase in the phase-out beginning and ending family income for married couples filing jointly
Source: Urban-Brookings Tax Policy Center, 2007

Family Income Variation at Birthdate Cutoff (Example)



Outline of Presentation

- Take advantage of Jan 1 birth discontinuity in EITC and exact DOB
  - Address concerns about manipulation of birth timing
- Schooling Data from North Carolina
  - Admin data for all students in NC Public Schools
  - Math and verbal test scores, suspensions, graduation
  - Use FRL status as proxy for EITC eligibility
- Adult Outcomes from Tax Records and Census Surveys
  - Universe of tax records with child-parent links, linked to ACS
  - Educational attainment, earnings, and employment
  - Use actual or imputed 1040 income in year prior to determine eligibility

Outline of Presentation

- Take advantage of Jan 1 birth discontinuity in EITC and exact DOB
  - Address concerns about manipulation of birth timing

- Schooling Data from North Carolina
  - Admin data for all students in NC Public Schools
  - Math and verbal test scores, suspensions, graduation
  - Use FRL status as proxy for EITC eligibility

- Adult Outcomes from Tax Records and Census Surveys
  - Universe of tax records with child-parent links, linked to ACS
  - Educational attainment, earnings, and employment
  - Use actual or imputed 1040 income in year prior to determine eligibility

Regression Discontinuity

$$Y_{ibt} = \beta_0 + \beta_1 D_b + \beta_2 z_b + \beta_3 (D_b * z_b) + \theta_t + \epsilon_{ibt},$$

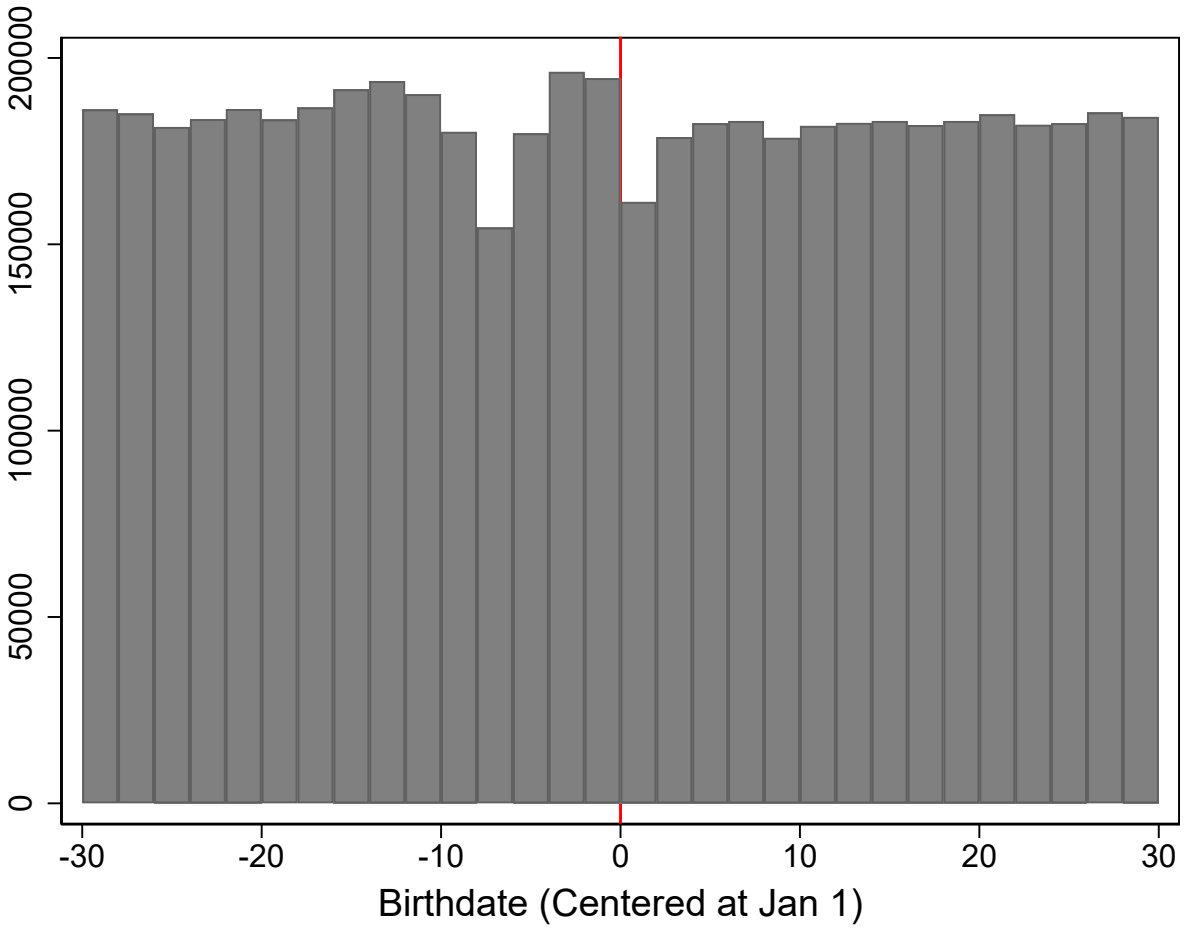
Where:

- Y_{ibt} – outcome for child i born on month-day b in birth year t (re-centered around Jan 1).
- D_b – indicator equal to 1 if birthdate is prior to Jan 1.
- z_b – difference between birthdate and Jan 1 ("assignment" variable).
- θ_t – Birth year (re-centered) fixed effects.

Identifying Assumption: Treatment assignment is "as good as random" at the the threshold.

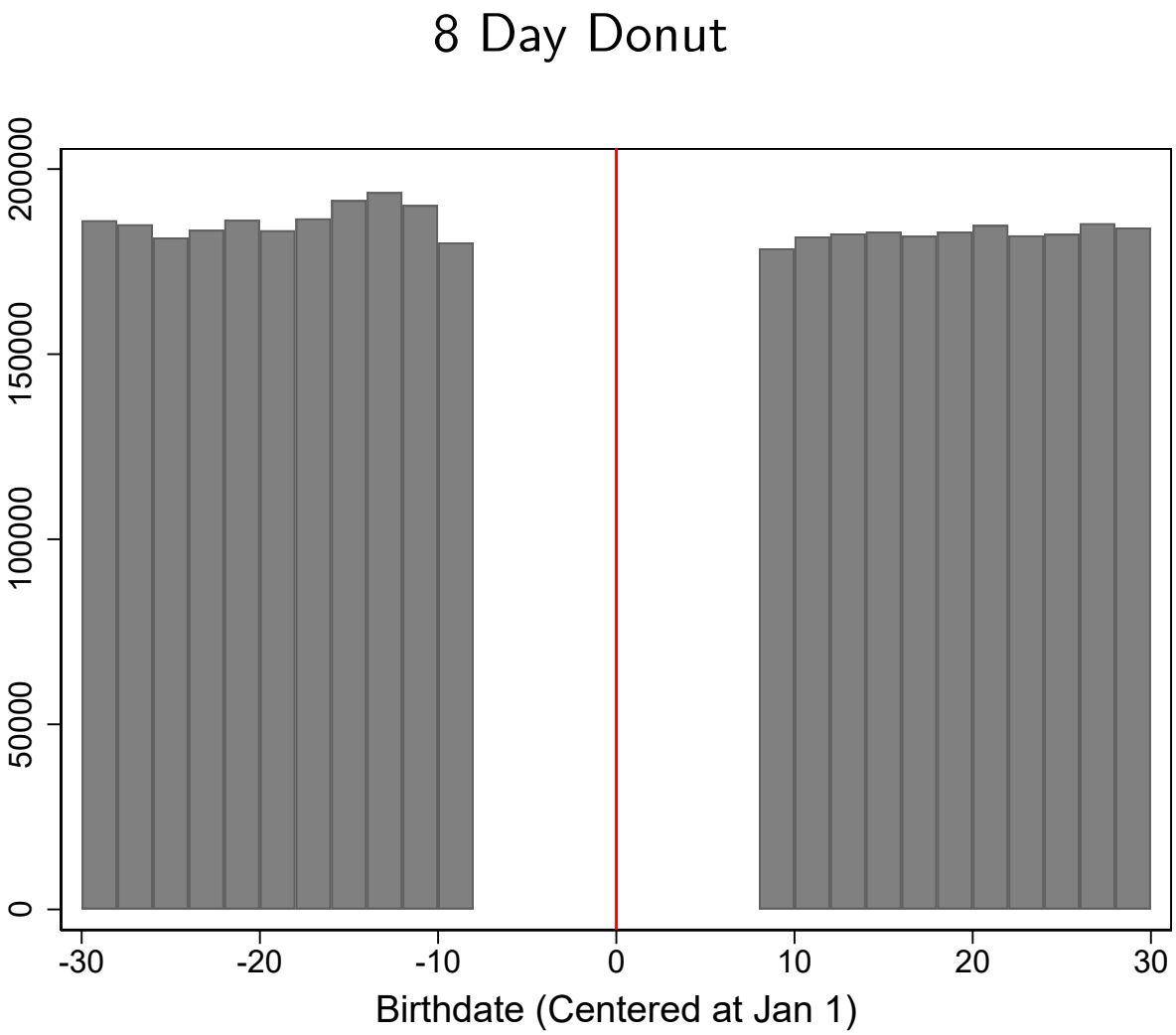
Main Concern: Manipulation of the timing of birth.

Distribution of Births

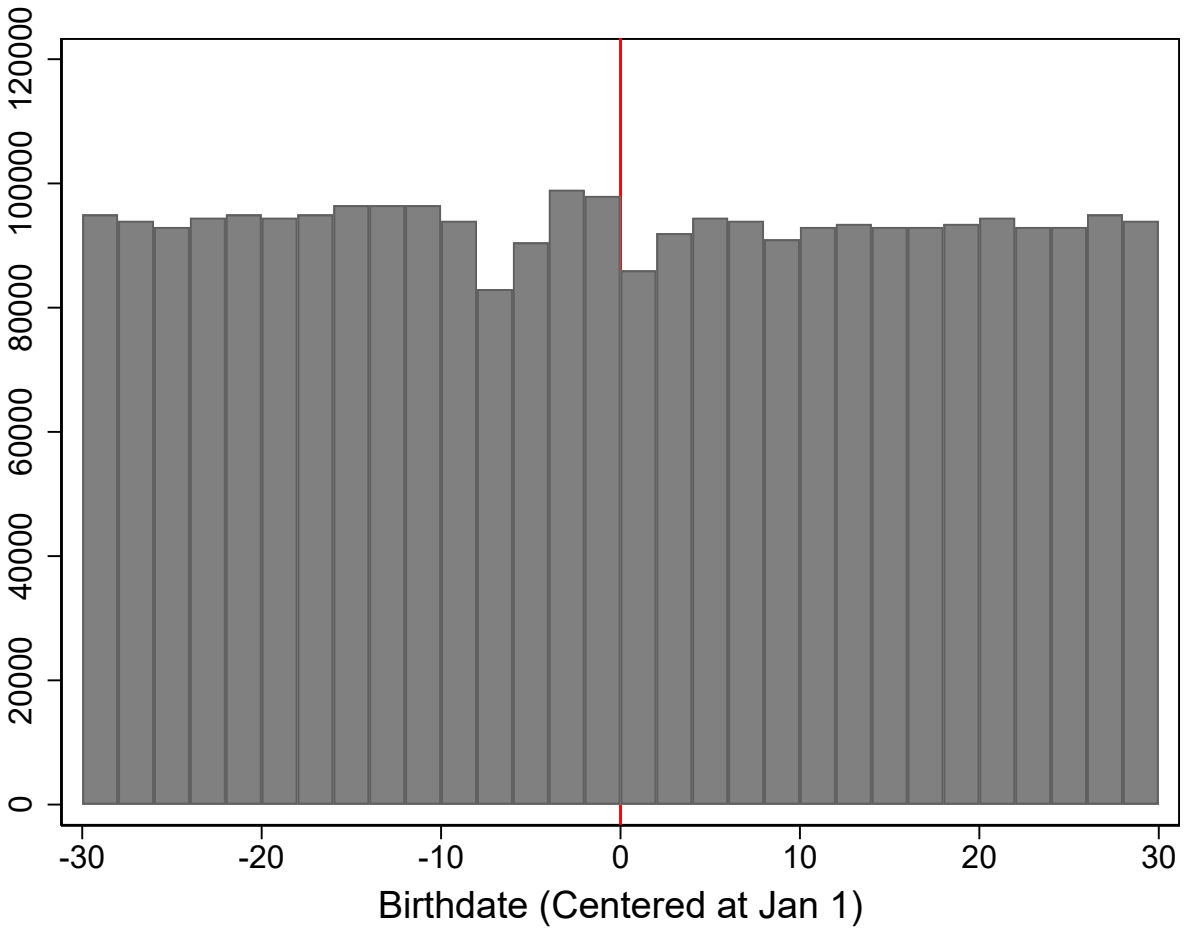


All births in tax records 1979-1992

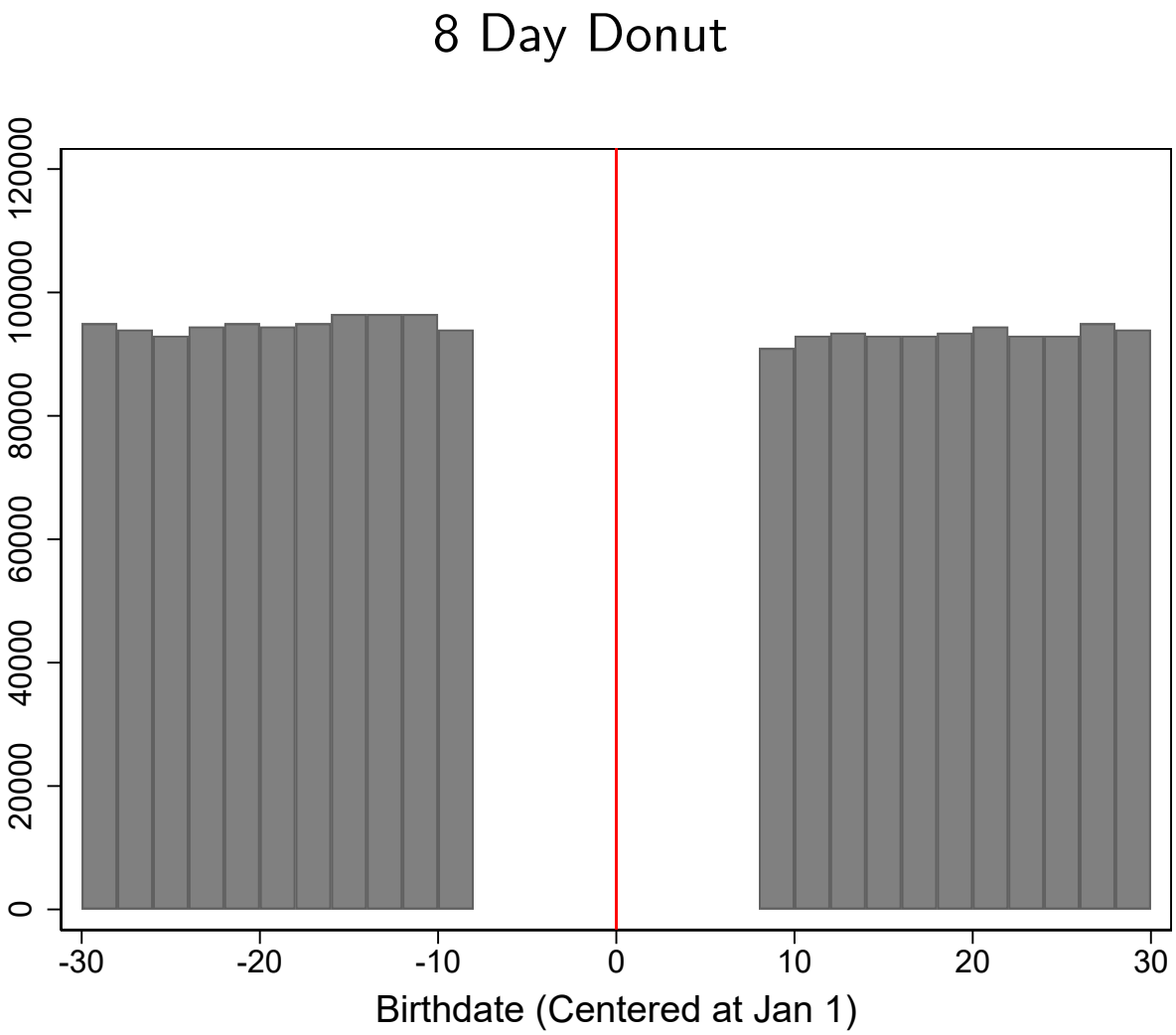
Distribution of Births



Distribution of Births: EITC Eligible - First Born

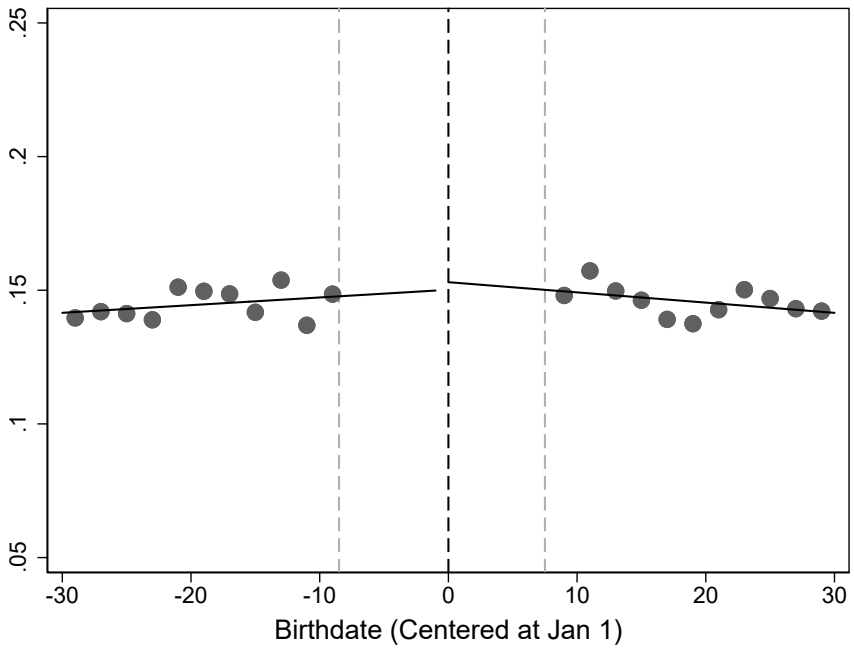


Distribution of Births: EITC Eligible - First Born



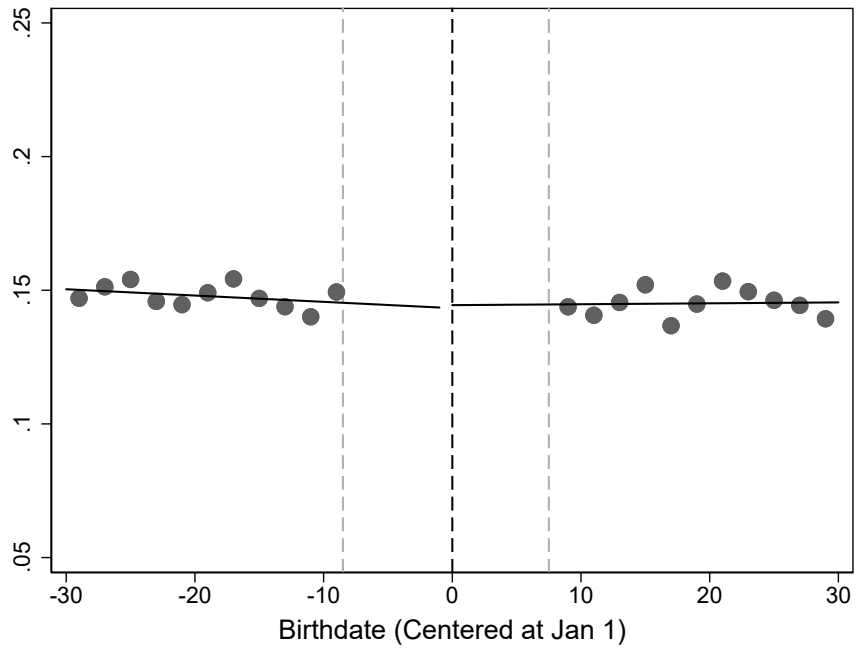
Covariate Balance: EITC Eligible - First Born

Black



Estimate: -0.005 (0.007)

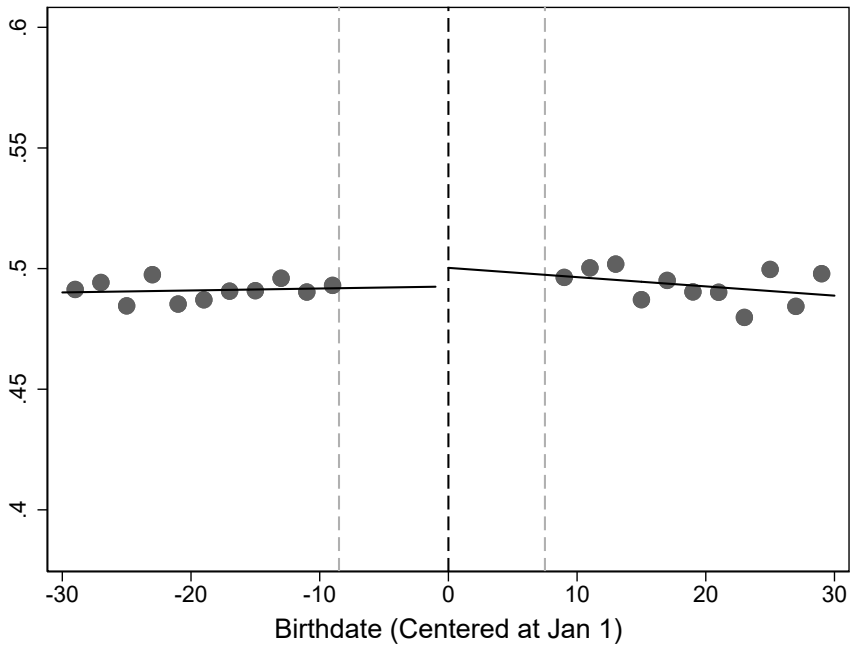
Hispanic



Estimate: 0.000 (0.006)

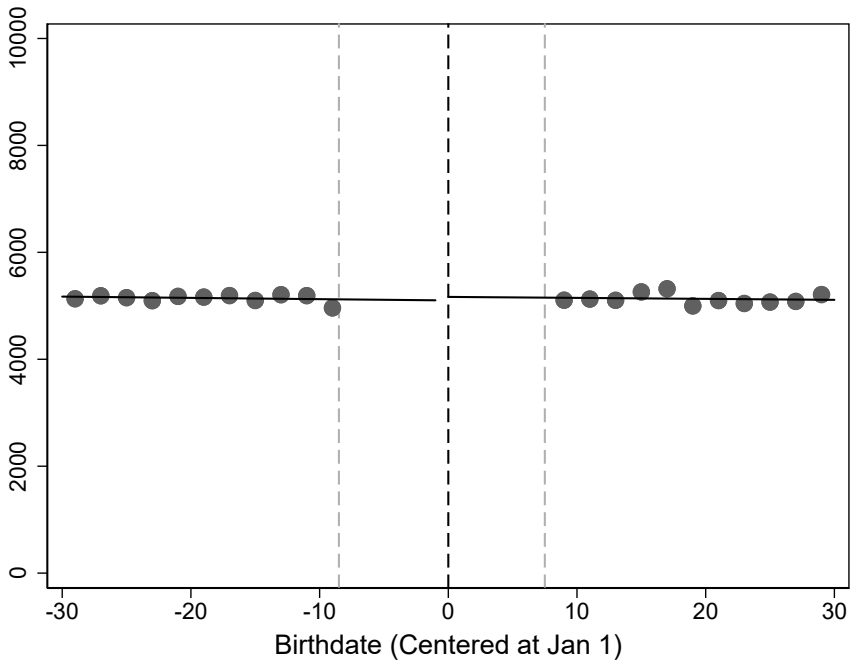
Covariate Balance: EITC Eligible - First Born

Male



Estimate: 0.004 (-0.005)

Parents' Predicted AGI



Addressing Potential Balance Issues

- Manipulation and imbalance are generally minimal in our sample
- Exclude "donut" of days around the cutoff
- Other Approaches:
  - Falsification groups with equal or greater manipulation, but no EITC eligibility
  - Exclude areas with high rates of C-section births (ongoing)
  - Isolate other subgroups with no manipulation (ongoing)

Outline of Presentation

- Take advantage of Jan 1 birth discontinuity in EITC and exact DOB
  - Address concerns about manipulation of birth timing
- Schooling Data from North Carolina
  - Admin data for all students in NC Public Schools
  - Math and verbal test scores, suspensions, graduation
  - Use FRL status as proxy for EITC eligibility
- Adult Outcomes from Tax Records and Census Surveys
  - Universe of tax records with child-parent links, linked to ACS
  - Educational attainment, earnings, and employment
  - Use actual or imputed 1040 income in year prior to determine eligibility

NC Public School Data

- Administrative data for all students in NC Public Schools that contains:
  - Math and reading test scores for grades 3-8.
  - Suspensions
  - HS graduation
  - Student characteristics (i.e. FRL status, race, gender)
  - Exact Birthdate
- Primary Sample:
  - Born within 1 month of January 1 in 1993-1998.
  - Observations: 142,094

Student Outcome Index Construction

- We construct an index using a weighted sum of z-scores from three student outcomes to improve statistical power (similar to Kling et al. 2007):
  - Mean of normalized math and reading test scores for grade 3-8 (50%).
  - Normalized high school graduation (25%).
  - Normalized "ever suspended" with sign reversed (25%).

Free/Reduced Lunch (FRL) Status and EITC Eligibility

- We focus on FRL students as the EITC "treated" group.
  - Ever show up as FRL eligible
- Similar income thresholds
- Ex: for family of three with one child in 2000,
  - FRL Eligibility: Less than $25,600
  - EITC Eligibility: Less than $27,400

Free/Reduced Lunch (FRL) Status and EITC Eligibility

- We focus on FRL students as the EITC "treated" group.
  - Ever show up as FRL eligible
- Similar income thresholds
- Ex: for family of three with one child in 2000,
  - FRL Eligibility: Less than $25,600
  - EITC Eligibility: Less than $27,400
- Preliminary estimates from the ASEC supplement to the CPS indicate roughly a $1,000 increase in income across the threshold for similar samples
  - In near future, we will use tax data to estimate first-stage among FRL students in North Carolina

Summary Statistics

	FRL (1)	Non-FRL (2)
Student Outcome Index	-0.08	0.40
Test Score Index	-0.03	0.63
HS Graduation	0.73	0.85
Any Suspension	0.17	0.07
Black	0.40	0.13
Limited English Proficiency	0.10	0.02
Male	0.52	0.51
Observations	56,400	38,668

Regression Discontinuity

$$Y_{ibt} = \beta_0 + \beta_1 D_b + \beta_2 z_b + \beta_3 (D_b * z_b) + \theta_t + \epsilon_{ibt},$$

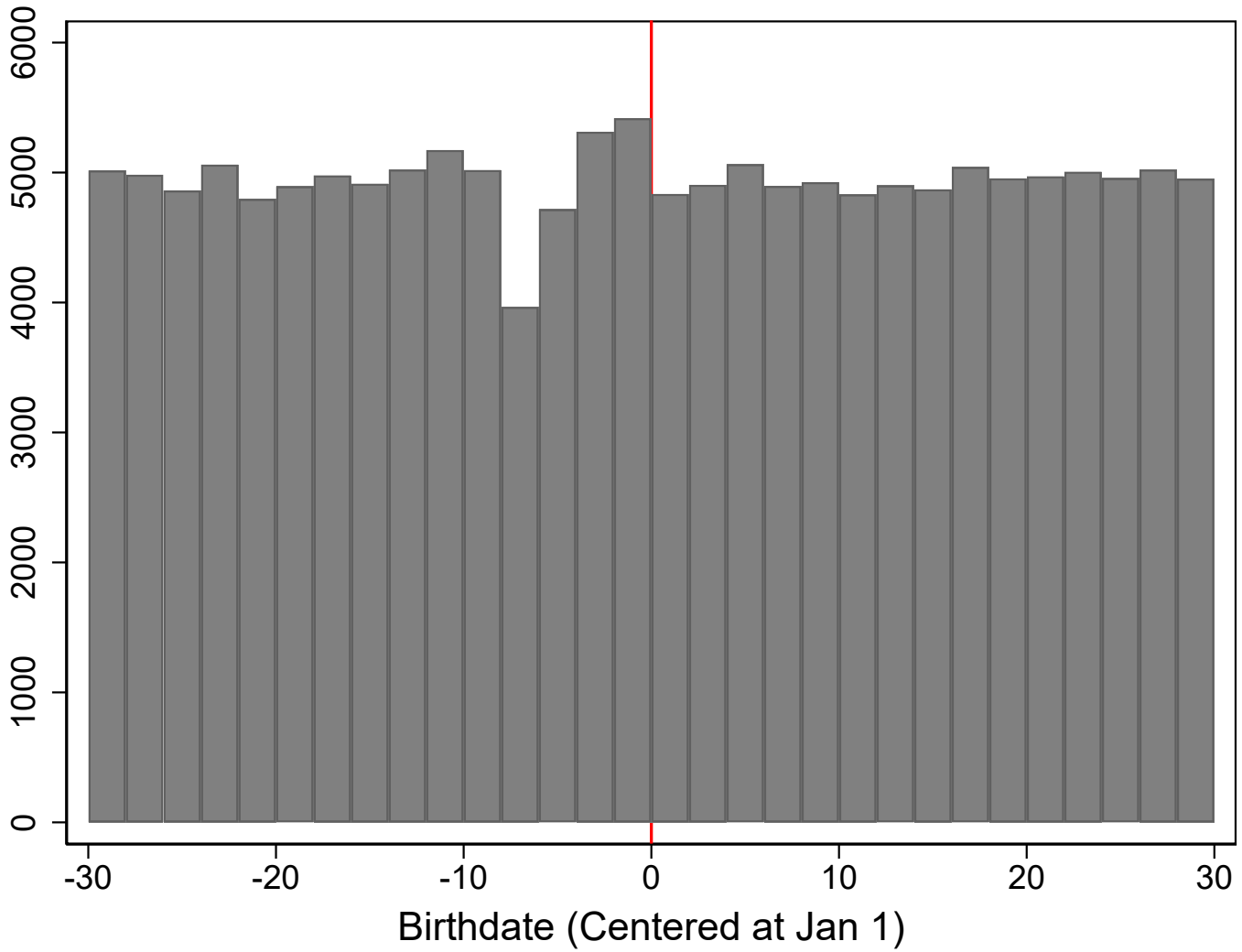
Where:

- Y_{ibt} – student outcome for child i born on month-day b in birth year t (re-centered around Jan 1).
- D_b – indicator equal to 1 if birthdate is prior to Jan 1.
- z_b – difference between birthdate and Jan 1 ("assignment" variable).
- θ_t – Birth year fixed effects (re-centered around Jan 1).

Identifying Assumption: Treatment assignment is "as good as random" at the the threshold.

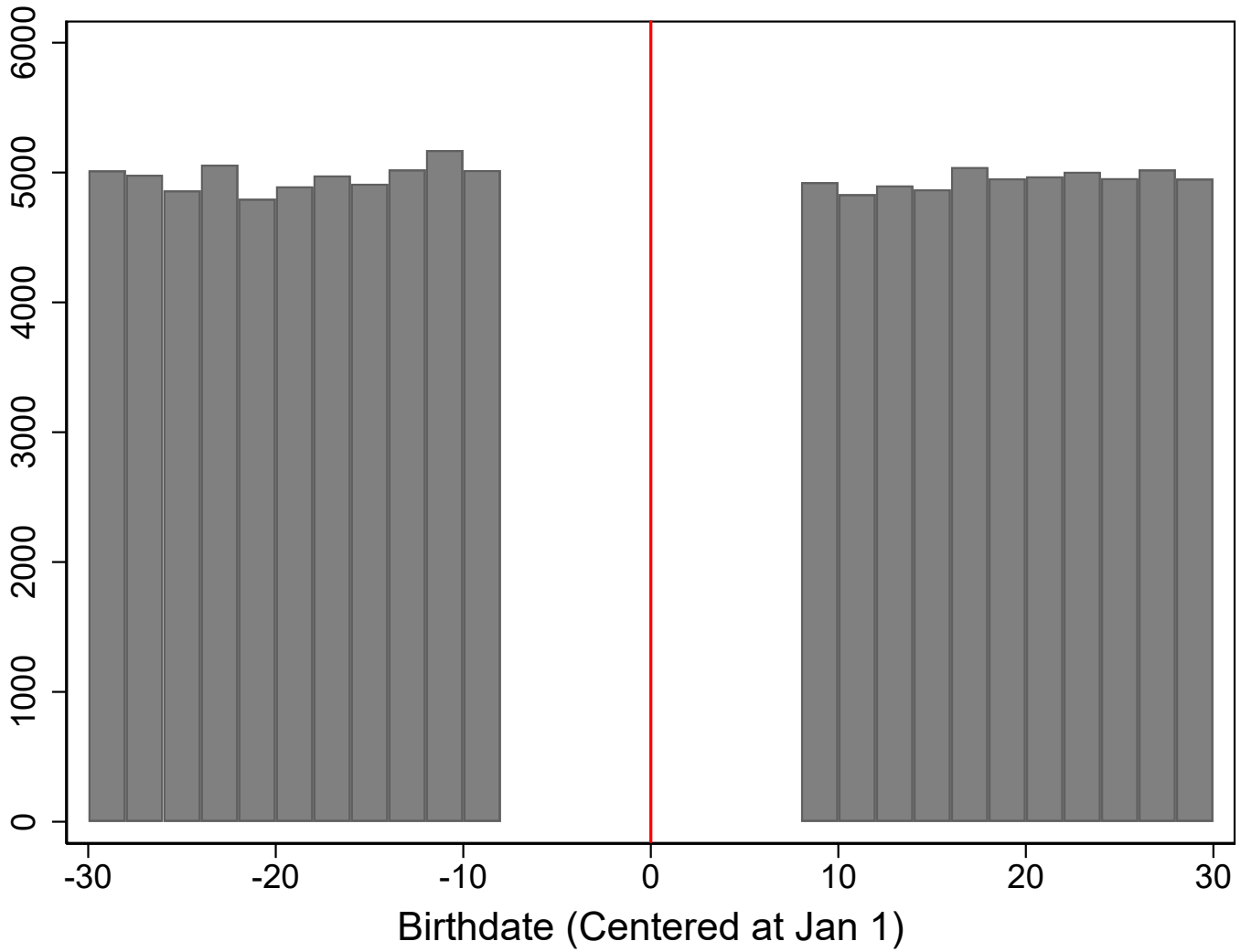
Main Concern: Manipulation of the timing of birth.

Distribution of Student Births

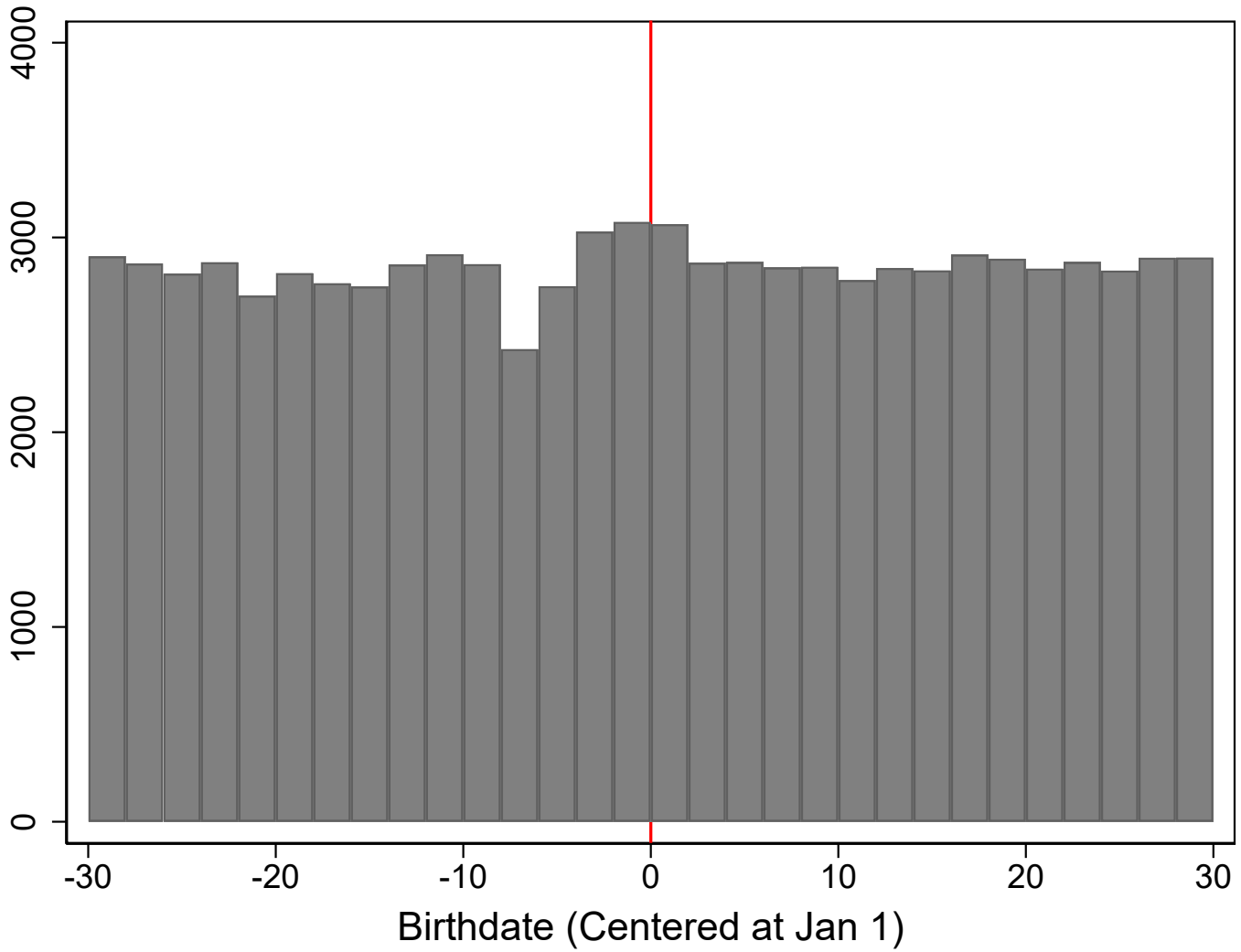


All births in North Carolina schooling data (1993-1998)

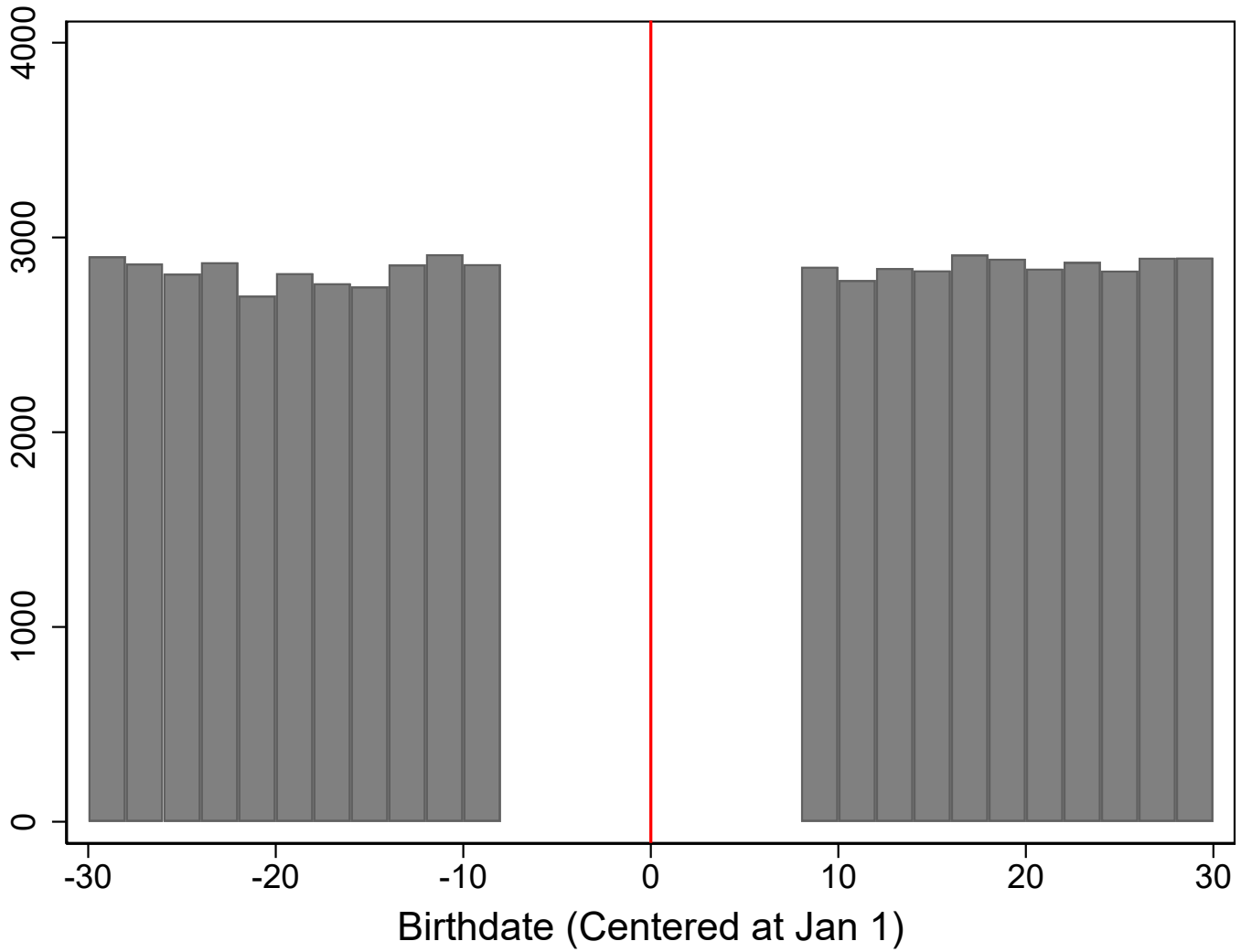
Distribution of Student Births



Distribution of Free/Reduced Lunch (FRL) Student Births

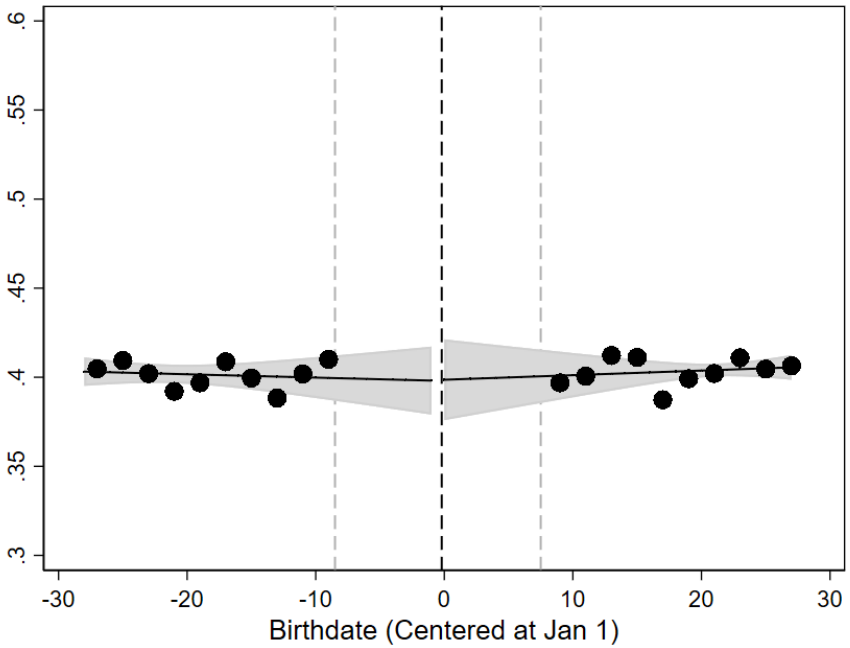


Distribution of Free/Reduced Lunch (FRL) Student Births



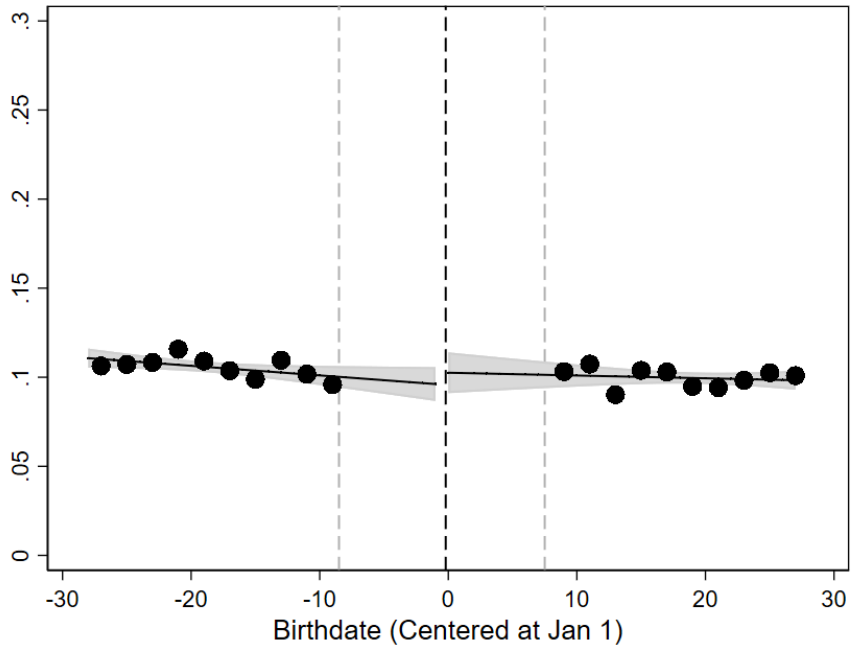
Covariate Balance: FRL Students

Black



Estimate: -0.001 (0.012)

Limited English Proficiency



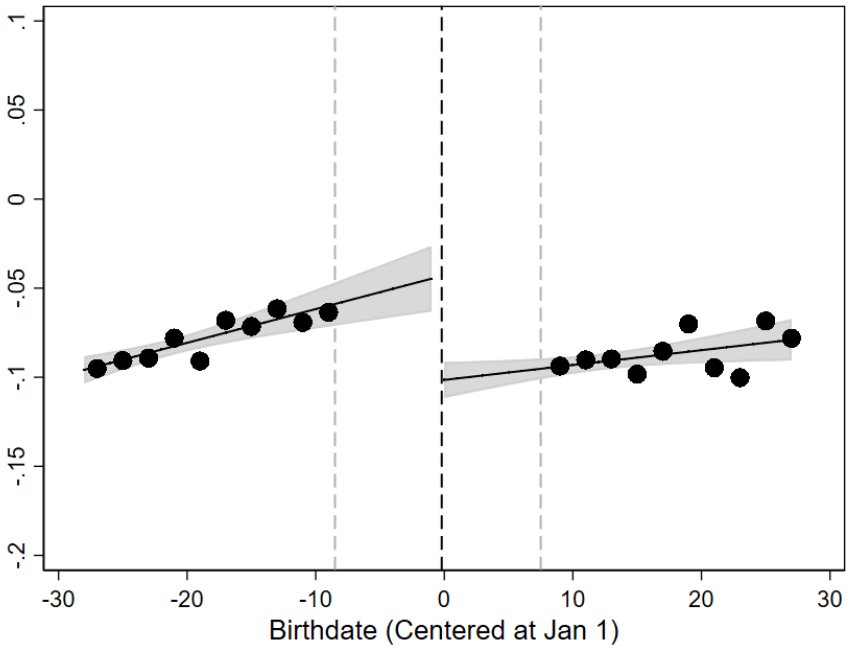
Estimate: 0.004 (0.006)

Covariate Balance (FRL Only)

	(1)	(2)	(3)
Black	-0.019**	-0.018*	-0.001
	(0.009)	(0.010)	(0.012)
Mean	*0.406*	*0.406*	*0.405*
LEP	-0.029	0.001	0.004
	(0.026)	(0.006)	(0.006)
Mean	*0.106*	*0.101*	*0.102*
Male	-0.014*	-0.013**	-0.008
	(0.007)	(0.006)	(0.009)
Mean	*0.517*	*0.516*	*0.516*
Donut Size (Days)	0	+/-1	+/-8

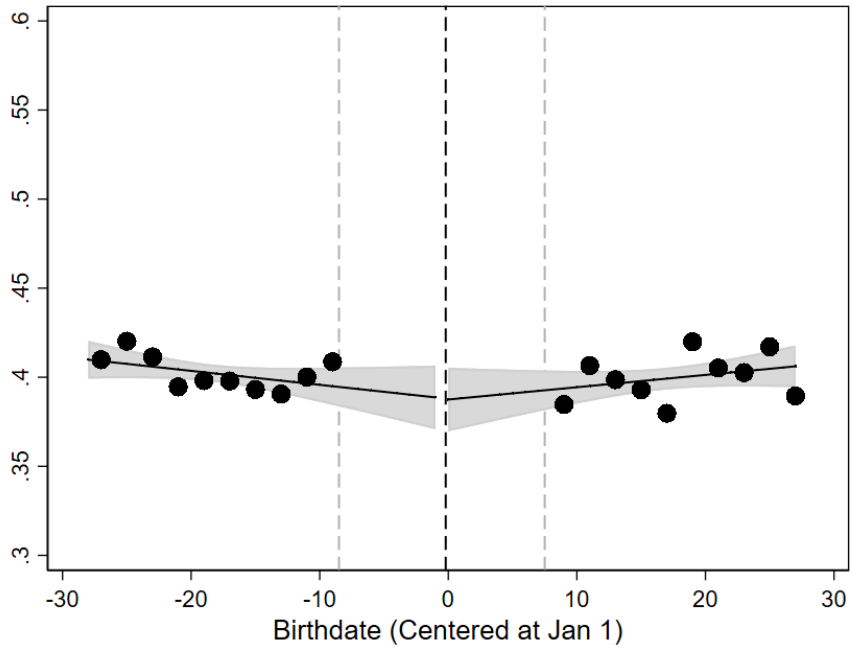
Student Outcome Index

FRL Students



Estimate: 0.055 (0.012)

Non-FRL Students

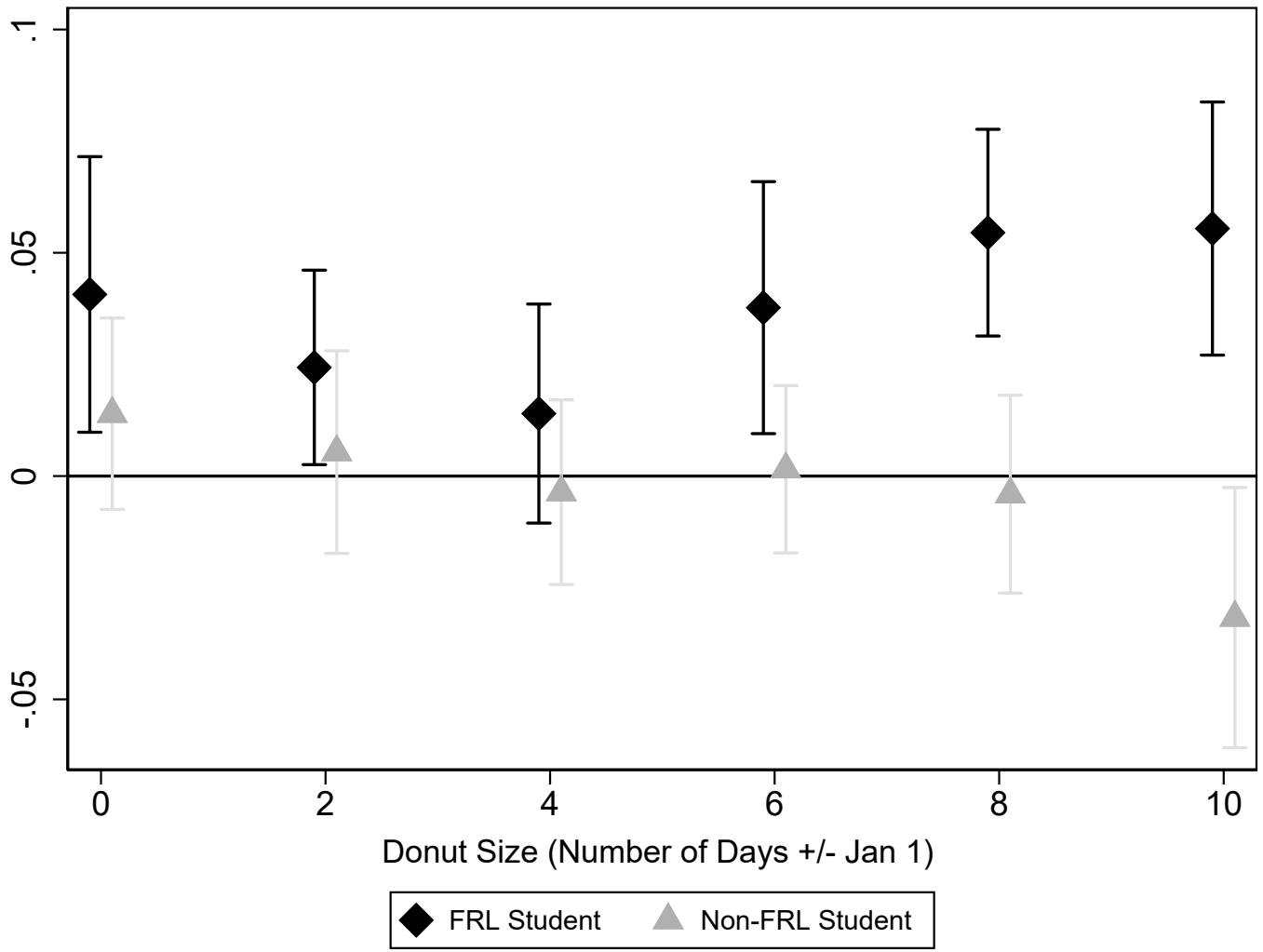


Estimate: -0.004 (0.011)

Student Outcome Index

	(1)	(3)	(5)
FRL Student	0.041***	0.026**	0.055***
	(0.016)	(0.012)	(0.012)
Obs	*56,604*	*56,138*	*56,400*
Mean	*-0.084*	*-0.082*	*-0.082*
Non-FRL Student	0.014	0.011	-0.004
	(0.011)	(0.011)	(0.011)
Obs	*37,970*	*38,051*	*38,668*
Mean	*0.396*	*0.398*	*0.401*
Cutoff Year Fixed Effect	X	X	X
Day-of-Week Fixed Effect			
Demographic Controls			
Donut Size (Days)	0	+/-1	+/-8

RD Estimate for Student Outcome Index (by Donut Size)

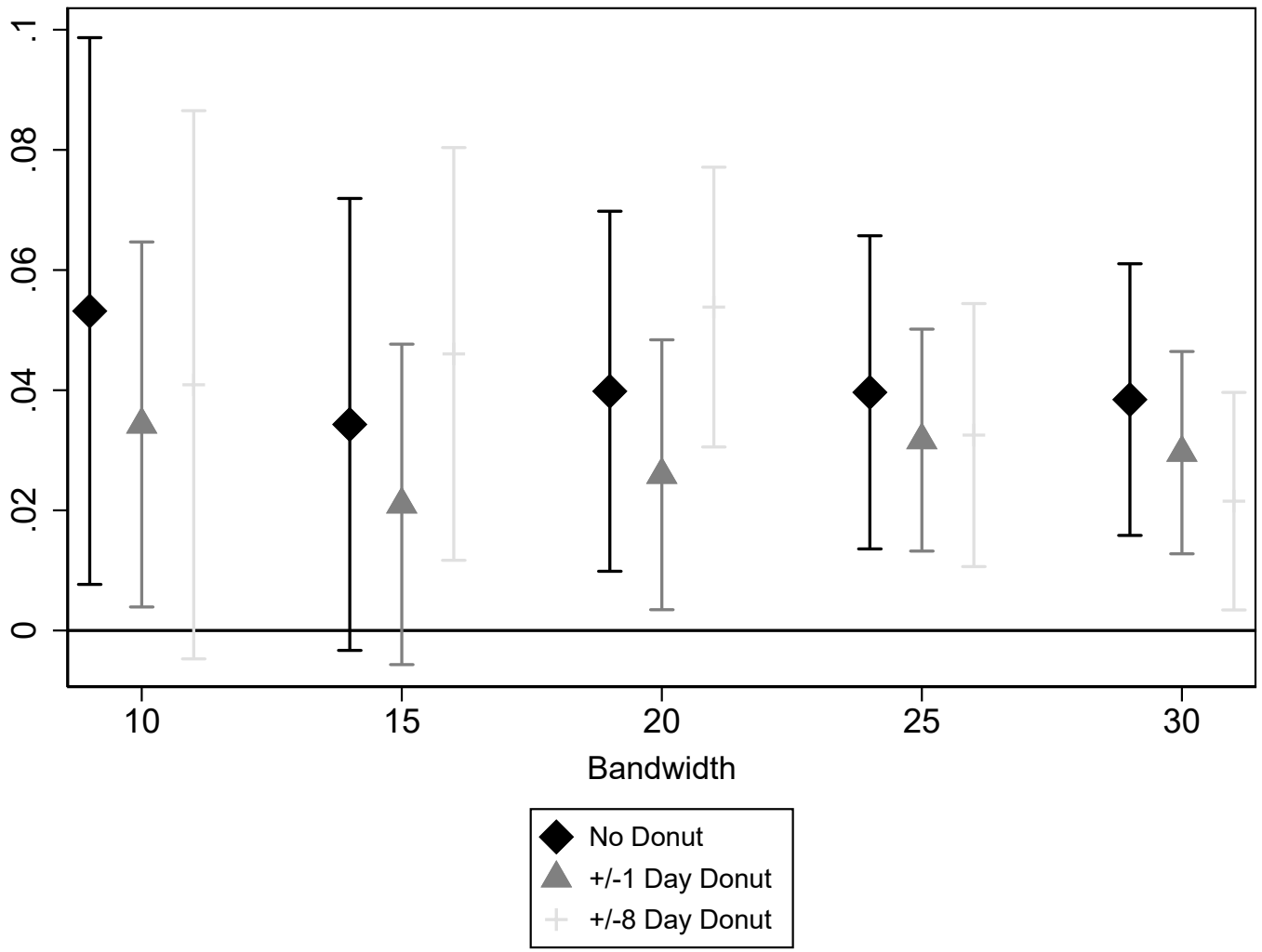


Student Outcome Index

	(1)	(2)	(3)
FRL Student	0.055***	0.054***	0.051***
	(0.012)	(0.012)	(0.011)
Obs	*56,400*	*56,400*	*56,400*
Mean	*-0.082*	*-0.082*	*-0.082*
Non-FRL Student	-0.004	-0.004	-0.005
	(0.011)	(0.011)	(0.012)
Obs	*38,668*	*38,668*	*38,668*
Mean	*0.401*	*0.401*	*0.401*
Cutoff Year Fixed Effect	X	X	X
Day-of-Week Fixed Effect		X	X
Demographic Controls			X

Donut Size =+/-8

RD Estimate for Student Outcome Index by Bandwidth/Donut Size (FRL Only)



Single Outcomes

	FRL		Non-FRL	
	(1)	(2)	(3)	(4)
Test Score Index	0.022	0.057***	0.016	-0.003
	(0.017)	(0.020)	(0.014)	(0.025)
Obs	*50,886*	*51,111*	*31,922*	*32,454*
Mean	*-0.030*	*-0.030*	*0.626*	*0.634*
HS Graduation	0.013*	0.019**	0.002	-0.012
	(0.007)	(0.010)	(0.014)	(0.015)
Obs	*45,623*	*45,986*	*30,874*	*31,476*
Mean	*0.727*	*0.725*	*0.846*	*0.846*
Any Suspension	-0.008**	-0.022**	0.001	-0.002
	(0.004)	(0.010)	(0.006)	(0.008)
Obs	*52,813*	*53,122*	*34,604*	*35,170*
Mean	*0.174*	*0.173*	*0.069*	*0.069*
Donut Size (Days)	+/-1	+/-8	+/-1	+/-8

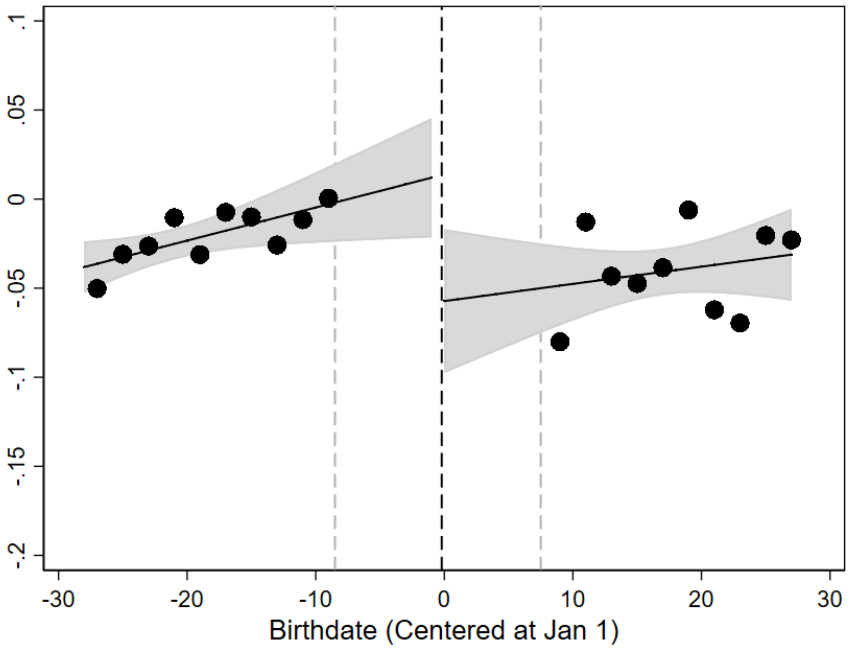
*Samples: HS Graduation – observed in NC in 7th grade. Any Suspension– observed in NC in Middle or High School.

RD Estimate for Student Outcome Index (by Subgroups)

	White	Black	Boy	Girl
	(1)	(2)	(3)	(4)
FRL Student	0.096***	0.024	0.052**	0.053***
	(0.019)	(0.022)	(0.020)	(0.018)
Obs	*21,507*	*22,740*	*29,169*	*27,231*
Mean	*0.015*	*-0.171*	*-0.147*	*-0.013*
Non-FRL Student	-0.002	0.012	-0.023	0.017
	(0.016)	(0.037)	(0.017)	(0.019)
Obs	*29,886*	*4,857*	*19,739*	*18,929*
Mean	*0.440*	*0.211*	*0.363*	*0.440*

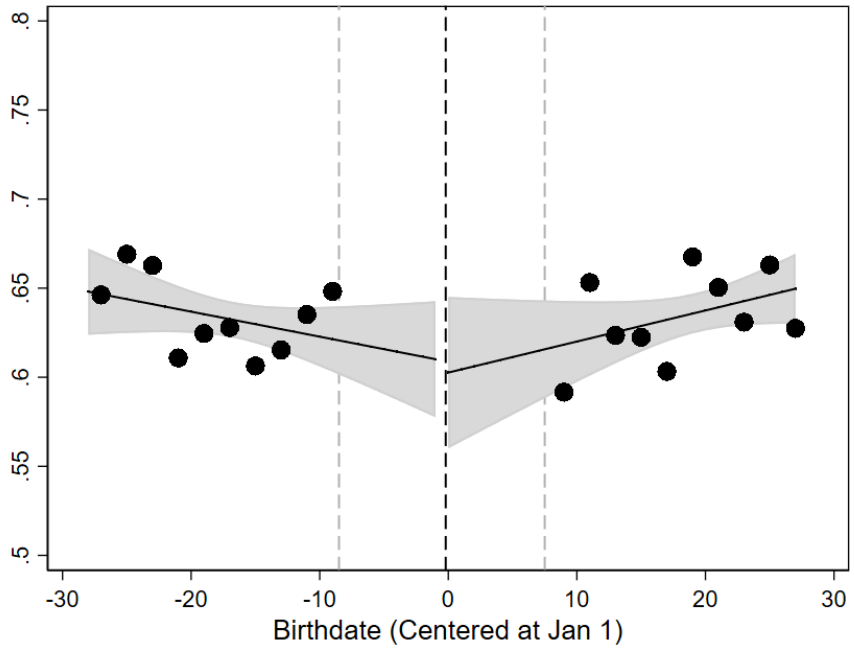
Index Components: Test Score Only Index

FRL Students



Estimate: 0.057 (0.020)

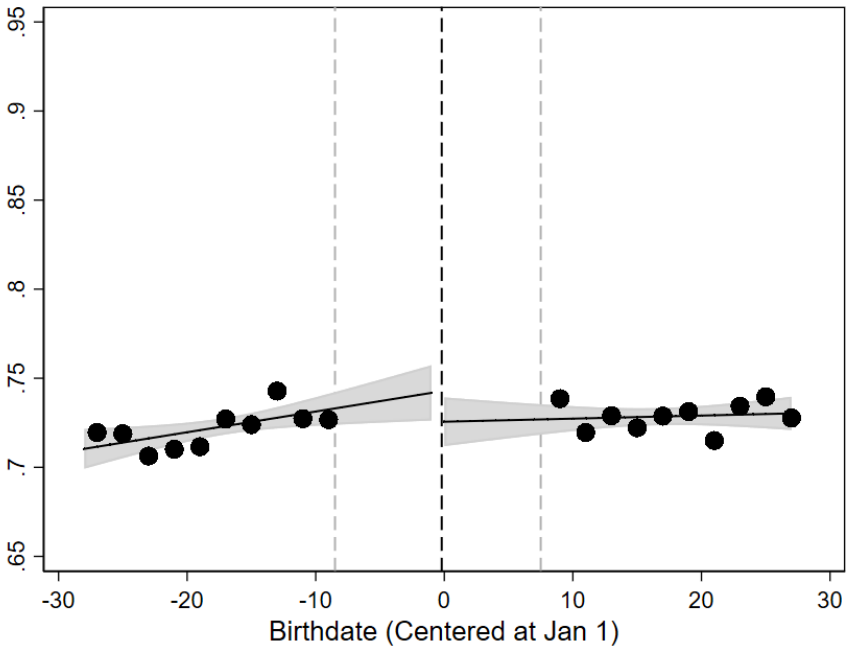
Non-FRL Students



Estimate: -0.003 (0.025)

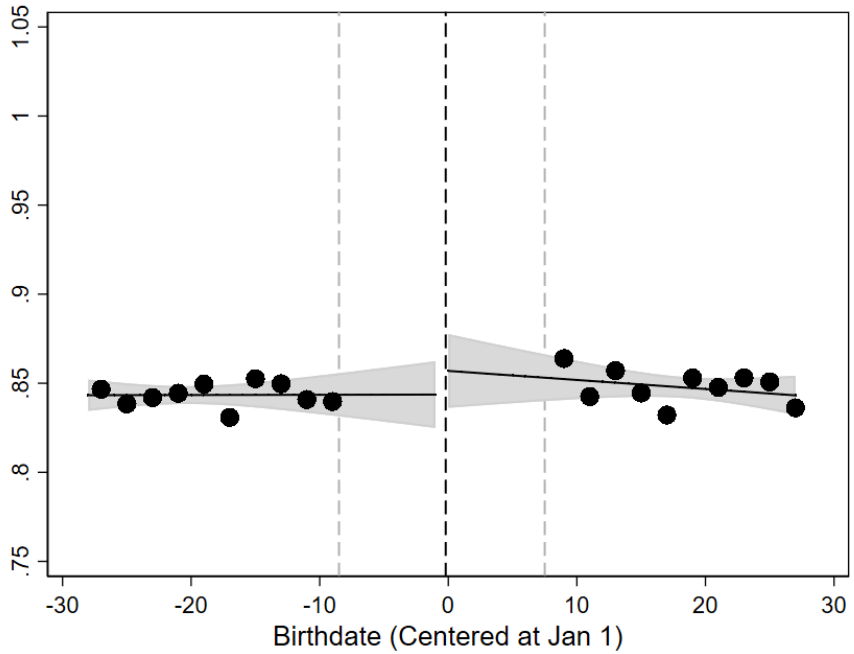
Index Components: HS Graduation

FRL Students



Estimate: 0.013 (0.007)

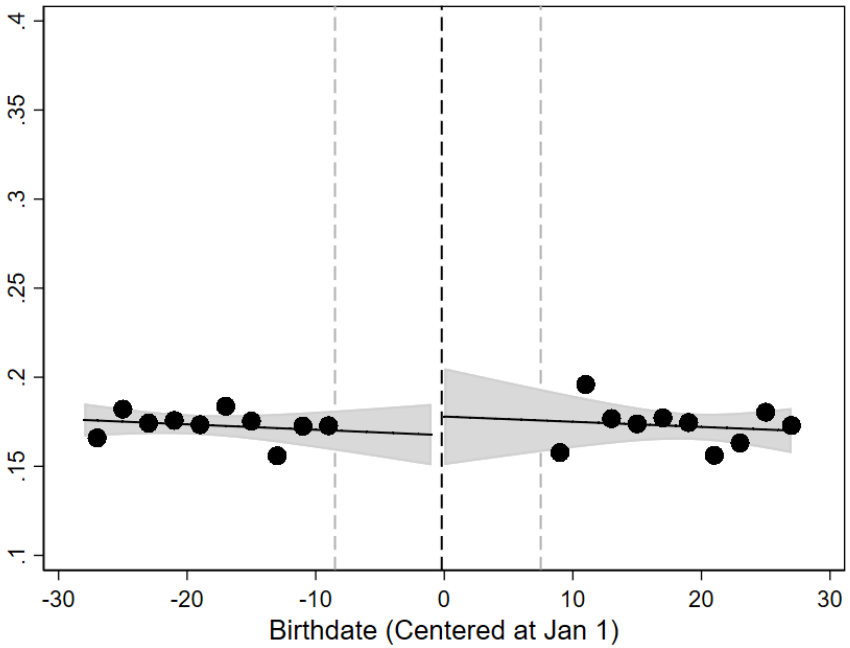
Non-FRL Students



Estimate: -0.012 (0.015)

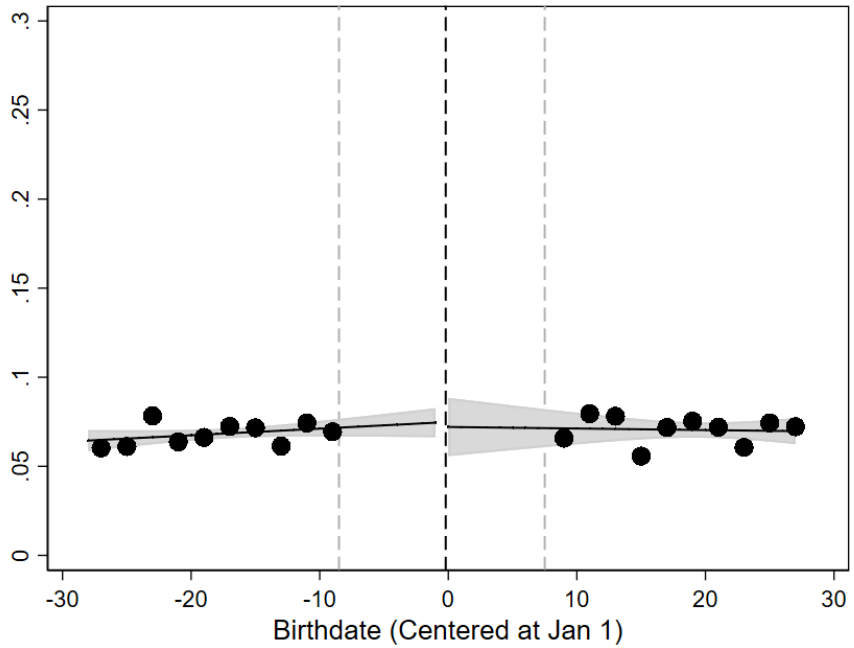
Index Components: Ever Suspended

FRL Students



Estimate: -0.022 (0.010)

Non-FRL Students



Estimate: -0.002 (0.008)

Summary of Schooling Results

- Results:
  - 0.02 SD $\uparrow$ in index of student outcomes for likely EITC-eligibile students
  - Roughly \$1,000 (5-10%) $\uparrow$ family income in infancy $\rightarrow$ %4 of FRL/Non-FRL gap
  - No effect on student unlikely to have been eligible

Outline of Presentation

- Take advantage of Jan 1 birth discontinuity in EITC and exact DOB
  - Address concerns about manipulation of birth timing
- Schooling Data from North Carolina
  - Admin data for all students in NC Public Schools
  - Math and verbal test scores, suspensions, graduation
  - Use FRL status as proxy for EITC eligibility
- Adult Outcomes from Tax Records and Census Surveys
  - Universe of tax records with child-parent links, linked to ACS
  - Educational attainment, earnings, and employment
  - Use actual or imputed 1040 income in year prior to determine eligibility

Adult Outcome Data (Tax-Tax and Tax-ACS)

- Universe of tax records with child-parent links [Tax-Tax]
  - Tax years 1979, 1984, 1989, 1994-1995, and 1998-2016
  - AGI (1040) and wage income (W-2)
  - Exact birthdate and birth place (Numident)
  - Birth order (1040)
- ACS (2005-2016) linked with tax records [Tax-ACS]
  - Educational attainment
  - Teen parenthood (not today)
  - Government assistance (not today)
  - Sex and race
- Primary Sample:
  - Born within 1 months of January 1 in 1980, 1985, 1990.
  - Ever observed in tax data as dependent
  - Donut of +/-8 Days
  - Two analytical samples (# individuals):
    - ★ Tax-Tax : ˜750,000
    - ★ Tax-ACS :˜50,000

Key Outcomes

- Universe of tax records with child-parent links
  - 3-Year Average Wage Income
    - ★ Ages 23-25 or Ages 26-28
    - ★ Includes zeroes
  - Wage Percentile within Cohort
- ACS linked with tax records
  - BA degree
    - ★ Ages 23-30

"Treatment" Group

- EITC Eligible
  - Based on Family's AGI from 1040 in year $t-1$.
  - Independent of child's *actual* birthdate.
  - Eligible - AGI qualifying for $EITC > 0$ *if child born prior to Jan 1* (or No 1040).

Regression Discontinuity

$$Y_{ibt} = \beta_0 + \beta_1 D_b + \beta_2 z_b + \beta_3 (D_b * z_b) + \theta_t + \epsilon_{ibt},$$

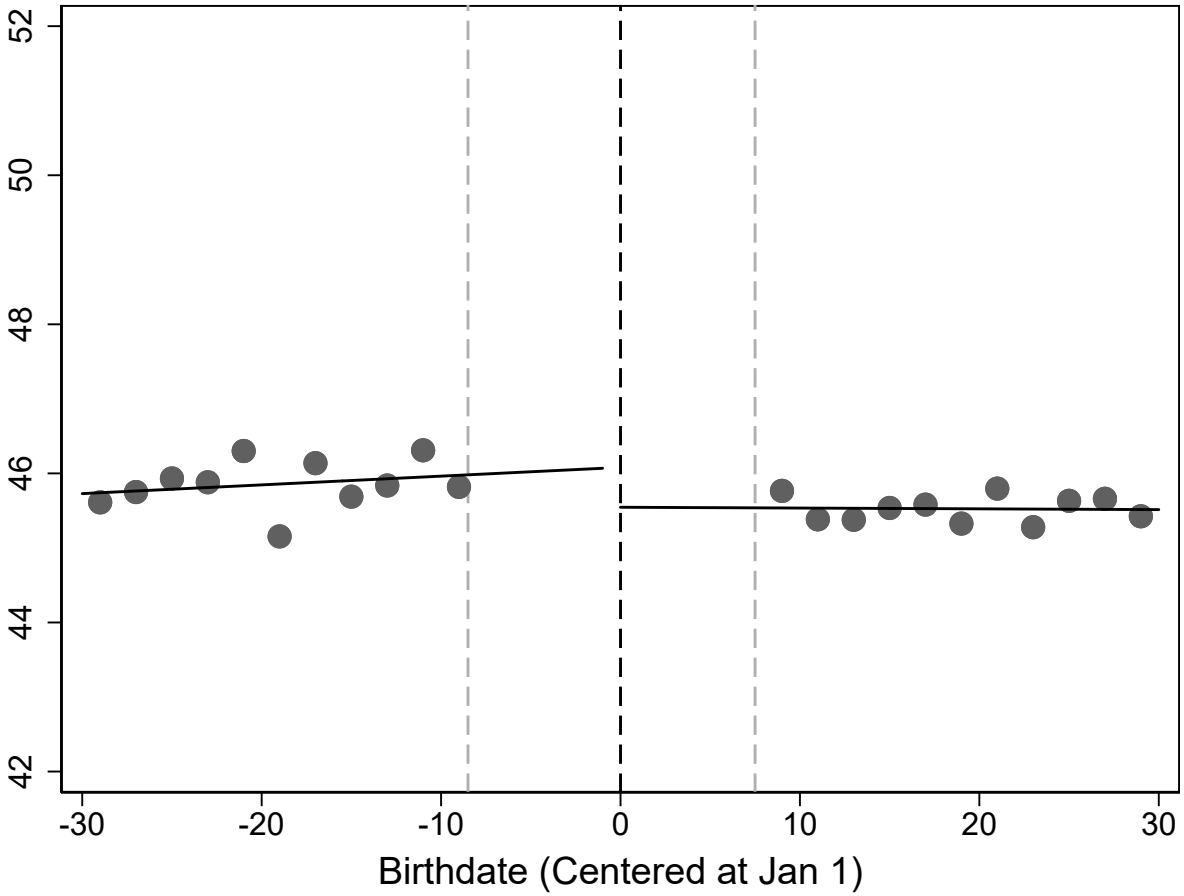
Where:

- Y_{ibt} – adult outcome for child i born on month-day b in birth year t (re-centered around Jan 1).
- D_b – indicator equal to 1 if birthdate is prior to Jan 1.
- z_b – difference between birthdate and Jan 1 ("assignment" variable).
- θ_t – Birth year fixed effects (re-centered around Jan 1).

Identifying Assumption: Treatment assignment is "as good as random" at the the threshold.

Main Concern: Manipulation of the timing of birth.

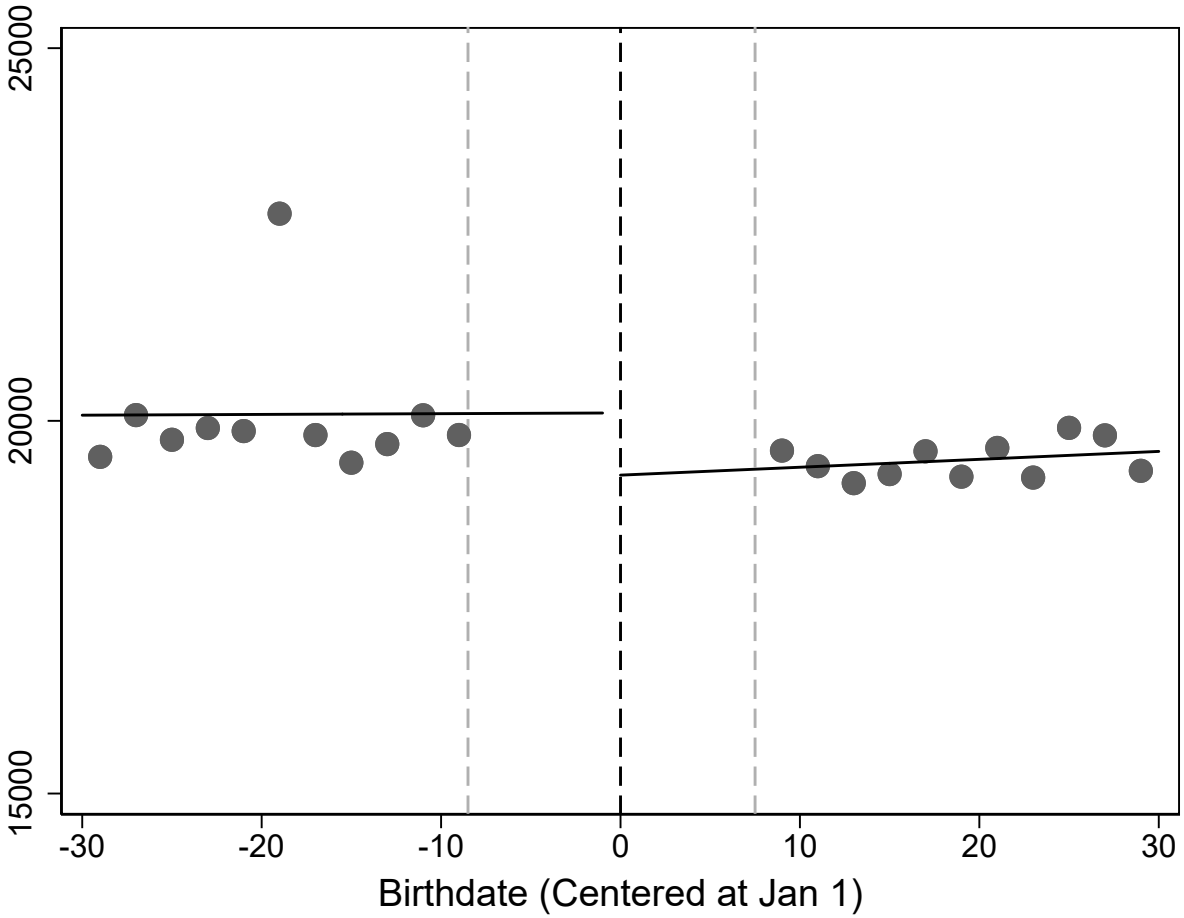
Wage Percentile within Cohort - EITC Eligible



Estimate: 0.437 (0.260)

Age 23-25

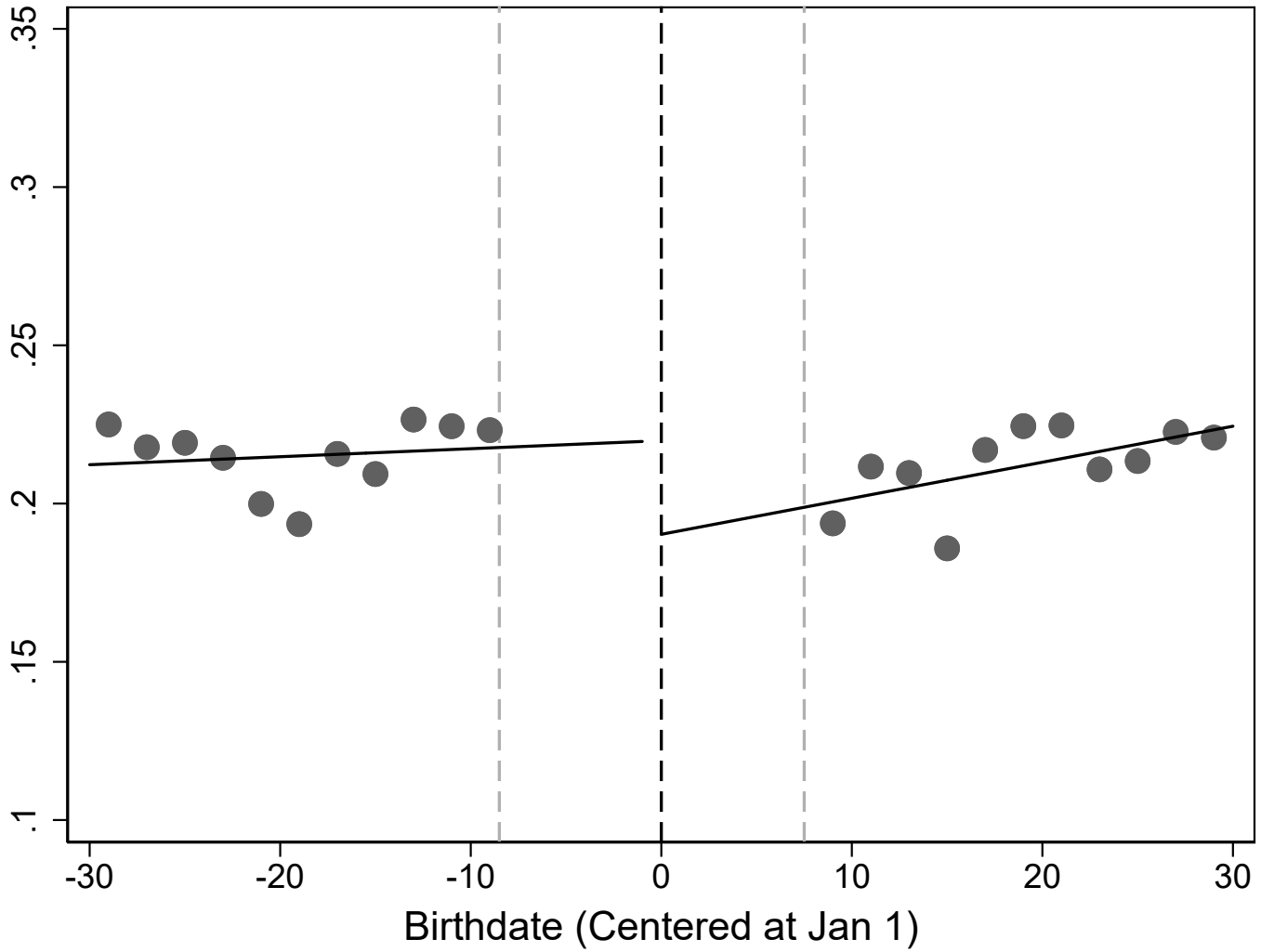
Wages - EITC Eligible



Estimate: 890.6 (514.8)

Age 23-25

Bachelor's Degree - EITC Eligible



Estimate: 0.024 (0.012)

Wage Outcomes (Age 23)

	Focal Years (Actual AGI Available)		All Years (Imputed AGI for Missing Years)	
	(1)	(2)	(3)	(4)
Wage Percentile	0.437*	1.182**	0.284**	0.566**
	(0.260)	(0.486)	(0.132)	(0.224)
Obs	*378,000*	*142,000*	*2,165,000*	*1,016,000*
Mean	*45.70*	*45.57*	*47.83*	*47.98*
Wage	890.6*	2125.0	335.3**	626.4**
	(514.8)	(1374.0)	(151.9)	(252.4)
Obs	*378,000*	*142,000*	*2,165,000*	*1,016,000*
Mean	*19770*	*18790*	*20550*	*19370*
height				
Birthyears	1979-1992	1987-1992	1979-1992	1987-1992

Threats to Internal Validity

- Manipulation of birth timing
  - Limited evidence of manipulation among first-born EITC eligible
  - Covariate balance

Threats to Internal Validity

- Manipulation of birth timing
  - Limited evidence of manipulation among first-born EITC eligible
  - Covariate balance
- Other things that shift across the January 1 threshold
  - School start dates (states dropped from analyses)
  - Dependency status for financial aid (Denning 2018)
  - ???

Threats to Internal Validity

- Manipulation of birth timing
  - Limited evidence of manipulation among first-born EITC eligible
  - Covariate balance
- Other things that shift across the January 1 threshold
  - School start dates (states dropped from analyses)
  - Dependency status for financial aid (Denning 2018)
  - ???

  Falsification Exercises: Examine presence of discontinuity for similar groups that are not eligible for substantial income boost
  - Ineligible families
  - Kid 2 in eligible families (minor or no incremental credit for our sample)

Comparison Groups

- Ineligible
  - Based on Family's AGI from 1040 in year $t-1$.
  - Independent of child's *actual* birthdate.
  - Ineligible - 150-200% of maximum EITC-eligible income *if child born prior to Jan 1* (e.g., 2015 and 1 kid $39,000-$59,000)
- Later births
  - Later births - No family income boost regardless of eligibility/birth timing (prior to 1991).

Comparison Groups

- Ineligible
  - Based on Family's AGI from 1040 in year $t-1$.
  - Independent of child's *actual* birthdate.
  - Ineligible - 150-200% of maximum EITC-eligible income *if child born prior to Jan 1* (e.g., 2015 and 1 kid $39,000-$59,000)
- Later births
  - Later births - No family income boost regardless of eligibility/birth timing (prior to 1991).

No discontinuities in these samples (generally negative point estimates)

Summary of Adult Outcome Results (preliminary)

- Earnings ↑ at age 23
  - 0.4 to 1.2 percentiles (1-2.6%)
  - $900-$2,000 (4.4-10%) increase
- Larger effects at age 26
- Positive effects on BA completion
- No effect (negative point estimates) for comparison groups
  - Children in ineligible families
  - 2nd child in eligible families

Thinking about Magnitudes

- Results:
  - First-stage estimates imply increase in income of roughly $1,000
- Comparing with existing estimates:
  - Mother's Pension: $1,528 (13%)
  - Food Stamps*: $1,460 per year of exposure (15%, insig.)
  - Head Start: $9,012 (20%)

Conclusion

- Results indicate that income during infancy can have profound and long-lasting effects
- Next Steps:
  - Getting at mechanisms:
    - ★ Effects on expenditures (Consumer expenditure survey)
    - ★ Effects on migration
    - ★ Effects on parental employment patterns
  - Income at other ages:
    - ★ Effect of income provided by second child
  - Use additional variation
    - ★ Size of credit based on income
    - ★ Awareness of EITC across geography
  - Heterogeneity by school or neighborhood quality

Thanks!