

Micronutrient and Long-run Human Capital: Evidence from Folic Acid Fortification

Wenjie Zhan*

Abstract

How does prenatal nutrition affect children's long-run outcomes? This paper presents new empirical evidence from the 1996 folic acid fortification mandate for enriched grain products. I proxy birth state fortification exposure with baseline rates of birth defects tied to folate deficiency and define exposed cohorts by whether the first trimester occurred after fortification. Comparing exposed and unexposed cohorts across states with different levels of baseline folate deficiency, I find that in-utero fortification exposure (1) is associated with higher test scores in mathematics among 4th- and 8th-graders; (2) shifts the time of college-age adults (19-22 years old) from work to schooling; (3) improves the graduate school enrollment among young adults over college ages (23-30 years old) without negative impacts on their labor supply. Excluding long-run human capital benefits could understate the net benefits of fortification by \$2.85–\$9.75 million per year. (JEL I18, J22, J24, N32, N52, Q18)

*Wenjie Zhan is a Ph.D. candidate at the Department of Agricultural and Resource Economics, University of California, Davis (email: wjzhan@ucdavis.edu). I thank Marianne Bitler, Timothy Beatty, Stephen Vosti, Richard Sexton, Ashish Shenoy, Rachel Soloveichik, Debora Mazetto, and seminar participants at the 2024 AAEA Annual Meeting, the 2025 ASSA Annual Meeting and the 2025 ASHEcon Annual Meeting for their helpful comments.

1 Introduction

Early-life nutrition can play a critical role in long-run human capital formation by supporting both physical and cognitive development ([Ampaabeng and Tan, 2013](#); [Portrait, Van Wingerden and Deeg, 2017](#)). While evidence is mounting that early-life nutrition shocks have long-run effects on human capital, most economic studies focus on overall nutrition intake or macronutrients ([Almond, Currie and Duque, 2018](#)). Evidence on the long-run effects of micronutrient deficiencies or inadequacies remains limited, despite their essential role in supporting key biological functions. This paper contributes new empirical evidence on the long-run effects of prenatal nutrition on human capital outcomes by studying the folic acid fortification of enriched grain products in the late 90s in the U.S.

Fortification is a cost-effective strategy to enhance access to micronutrients. The U.S. has a long history of fortifying foods with iodine, iron, and various vitamins. Folic acid fortification was the most recent effort to combat maternal deficiency in folate, a critical nutrient for child mortality and neurodevelopment. Maternal folate deficiency, particularly concerning during pregnancy, can lead to severe birth defects and cognitive impairments in children ([Roth et al., 2011](#); [Irvine et al., 2022](#)). To reduce these risks, the U.S. Food and Drug Administration (FDA) mandated the fortification of 40 μ g/100g of folic acid, i.e., the synthetic form of folate, in enriched grain products starting March 5, 1996. While public health literature widely recognizes the immediate benefits of folic acid fortification in reducing birth defects and improving infant health, its long-term effects on human capital remain underexplored.

I leverage geographical variation in pre-fortification birth defects tied to folate deficiency and the timing of folic acid fortification of grain products to assess the program's effect. Folic acid fortification effectively reduced folate deficiency ([Wald et al., 2001](#)), with greater benefits observed in states with higher baseline deficiencies (Section 5). Folate is crucial for neural tube formation during the first trimester of pregnancy, and neurological damage during this stage is often irreversible. Thus, the effects of maternal exposure to folic acid fortification may manifest in later life stages. If fortification is effective, we should observe significant improvements in the outcomes of individuals exposed to folic acid fortification during early fetal development, particularly in states with higher baseline folate deficiency. Due to the lack of large-scale data on maternal folate deficiency, I use the pre-fortification prevalence of birth defects tied to folate deficiency to capture maternal exposure to folic acid fortification.

I begin by documenting several first-stage facts: (i) measured folate content rose across a wide range of foods following fortification; (ii) dietary folate intake and serum folate concen-

trations increased; and (iii) the prevalence of folate-sensitive congenital anomalies declined, with larger reductions in places with higher baseline rates.

To explore how in-utero fortification exposure affects test scores of K-12 students, I construct a migration-adjusted fortification exposure at the state-of-school level and link it to state-average test scores of 4th- and 8th-graders from the public-use National Assessment of Education Progress (NAEP). I find positive associations in mathematics: fortification exposure is linked to a 2.28 point increase in average score (on a 0 to 500 scale) and to higher achievement shares, with 1.67% percentage more students at or above the NAEP Basic level and 2.67% more at or above the NAEP Proficient level.

I combine the spatial variation in baseline fortification exposure and the timing of fortification with the microdata on young adults' educational and labor outcomes from the American Community Survey (ACS), and estimate long-run effects of in-utero exposure to fortification using a cohort difference-in-differences (cohort-DiD) approach. I find that in-utero exposure to fortification reallocated time from work toward schooling among college-age adults: college enrollment increased by 1.81–6.17 percentage points (3.70–12.66% of the sample mean), full-time employment fell by 1.17–4.00 p.p. (3.80–12.97% of the sample mean), and annual earnings declined by \$457–\$1,562 (4.06–13.86% of the sample mean). These patterns are consistent with more time spent in school. Among young adults beyond college age, in-utero exposure to fortification leads to a 0.69–2.34 p.p. increase in graduate-school enrollment (11.30–38.59%), supporting greater human capital investment among the exposed group.

Benchmarking against other nutrition policies, the implied human capital effects are comparable to those attributed to salt iodization and smaller than those estimated for the Food Stamps Program, a substantially larger intervention. To my knowledge, this is the first study to quantify the long-run human capital benefits of folic acid fortification. My back-of-the-envelope calculation indicates that excluding these long-run gains from cost–benefit analyses could understate the program's net benefits by \$2.85–\$9.75 million per year.

This paper contributes to both economic and public health literature on food fortification. Most economic studies have examined the long-term benefits of salt iodization on cognitive development, health, and socioeconomic outcomes—such as improved cognitive performance and higher earnings ([Feyrer, Politi and Weil, 2017](#); [Serena, 2019](#); [Adhvaryu et al., 2020](#); [Huang, Liu and Zhou, 2020](#); [Deng and Lindeboom, 2022a](#); [Tafesse, 2022](#)). One exception is [Niemesh \(2015\)](#), which finds that iron fortification of bread increased working-age adults' incomes and school enrollment and raised their children's long-run wages. In contrast, evidence on the

human capital effects of folic acid fortification is scarce. Biologically, the case for folate is strong: unlike iodine and iron deficiencies—which primarily impair thyroid function and oxygen transport, respectively—folate deficiency directly disrupts neural development and can have severe and lasting consequences. Folic acid fortification might therefore exert stronger effects on cognitive development and downstream outcomes such as schooling and earnings. Adoption has also been far less widespread globally than salt iodization or iron supplementation, particularly in developing countries, so empirical evidence from the United States can inform policy design elsewhere. The public health literature has emphasized short-run health effects of folic acid supplementation (e.g., [Wald et al., 2001](#); [Quinlivan et al., 2002](#); [Kancherla et al., 2022](#)) and cost–benefit analyses of fortification (e.g., [Grosse et al., 2005](#); [Bentley et al., 2009](#); [Llanos et al., 2007](#)). By examining the long-run human capital outcomes, this study extends that work and provides new empirical evidence on the broader developmental consequences of folic acid fortification.

This paper also relates to the fetal origins literature. A large body of research shows that prenatal and early childhood nutritional conditions have enduring consequences. Adverse shocks, such as famine ([Meng and Qian, 2006](#); [Almond et al., 2007](#); [Chen and Zhou, 2007](#); [Meng and Qian, 2009](#); [Lindeboom, Portrait and Van den Berg, 2010](#); [Scholte, Van Den Berg and Lindeboom, 2015](#); [Deng and Lindeboom, 2022b](#)) and Ramadan fasting ([Almond and Mazumder, 2011](#); [Almond, Mazumder and Van Ewijk, 2015](#); [Majid, 2015](#); [Greve, Schultz-Nielsen and Tekin, 2017](#)), have been linked to poorer health and labor market outcomes in adulthood. In contrast, positive interventions, such as breastfeeding ([Fitzsimons and Vera-Hernández, 2022](#)), iodine supplementation ([Field, Robles and Torero, 2009](#); [Araújo, Carrillo and Sampaio, 2021](#)), and food assistance programs like WIC and food stamps ([Hoynes, Page and Stevens, 2011](#); [Rossin-Slater, 2013](#); [Hoynes, Schanzenbach and Almond, 2016](#); [Bailey et al., 2024](#)), have been shown to support cognitive development and improve long-term socioeconomic outcomes. This study extends this literature by examining the long-run impacts of prenatal exposure to a previously understudied intervention.

The paper is organized as follows: Section 2 provides the policy backgrounds; Section 3 describes the data; Section 4 outlines the research design and discusses identifying assumptions; Section 5 presents first-stage evidence; Section 6 presents results; Section 7 discusses robustness; and Section 8 concludes.

2 Background

2.1 Folate deficiency disorder and associated birth defects

Folate deficiency is a major cause of neural tube defects (NTDs), the most common congenital anomalies of the central nervous system (CNS) in newborns (Smithells et al., 1983). Severe NTDs, such as anencephaly, are typically fatal, with most affected infants dying before or shortly after birth.¹ Infants with less severe NTDs, like spina bifida, can survive into adulthood but often carry a high risk of lifelong physical and mental disabilities (Yi et al., 2011).² In the early 1990s, approximately 4,000 fetuses in the U.S. (about 1 in 1,000) were affected by NTDs annually, with one-third lost due to selective or spontaneous abortions (Cragan et al., 1995; Mersereau et al., 2004). Folate deficiency can also lead to other congenital CNS anomalies, such as hydrocephaly (Naz et al., 2016; Liu et al., 2018). These birth defects can develop as early as the first month of pregnancy when the neural tube begins to form, and failure to close the neural tube by the end of the first trimester can cause irreversible damage to the central nervous system (Obeid, Holzgreve and Pietrzik, 2013). While in-utero surgery may offer some palliative benefits, this neurological damage remains irreversible (Greene and Copp, 2014). Timely medical intervention is often difficult, as most ultrasound screenings occur in the second trimester—when anomalies are easier to detect—and many pregnant women in the U.S. lack adequate prenatal care (Blumenfeld, Siegler and Bronshtein, 1993).

2.2 Sources of folate

Folate exists naturally in foods such as beef liver, dark green leafy vegetables, beans, peas, nuts, and a variety of fruits and fruit juices. However, meeting the recommended intake during pregnancy through diet alone is challenging (Czeizel, 2000). Data from the National Health and Nutrition Examination Surveys (NHANES) III (1988–1994) show that women ages 15–49 consumed an average of 233.68 µg of folate per day, well below the 400 µg recommended by the U.S. Public Health Service for pregnant women. One reason dietary intake falls short is that natural food folate is unstable under typical cooking conditions, which can substantially reduce the amount ultimately absorbed, making it a less reliable way to improve folate status during pregnancy (McNulty and Pentieva, 2004).

Folate is also available from nutritional supplements, including over-the-counter folic acid tablets and multivitamins sold in pharmacies. Folic acid supplements are often prescribed dur-

¹Infants with anencephaly are born without parts of the skull and brain.

²The backbone of infants with spina bifida does not close properly, leaving a section of the spinal cord and spinal nerves exposed to the outside without the protection of the backbone.

ing prenatal visits ([Ray, Singh and Burrows, 2004](#)). A key challenge, however, is low awareness of and adherence to supplementation recommendations ([Toivonen et al., 2018](#)). According to CDC guidance,³ folic acid should be taken starting at least one month before conception. Yet about 50% of U.S. pregnancies are unintended ([Finer and Zolna, 2016](#)). From 1995 to 1998, only about 30% of U.S. women reported taking a daily vitamin containing folic acid, and fewer than 10% knew it should be taken before pregnancy ([Petrini, Damus and Johnston, 1999](#)). Access and affordability also pose barriers, particularly for low-income women ([Czeizel, 2000](#)). These constraints point to the need for a low-cost, preferably passive approach to ensure adequate folic acid intake among women who may become pregnant.

2.3 Folic acid fortification and other fortifications in the U.S.

The United States has a long history of using food fortification to improve public health. Salt iodization began in the 1920s, vitamin D fortification of milk followed in the 1930s, and flour and bread were enriched with B vitamins and iron in the 1930s and 1940s. The most recent effort, folic acid fortification of grain products, started in the 1990s. The first wave of grain fortification in the 1940s followed the identification of specific nutrient deficiency disorders in the U.S. In the early 1940s, the FDA issued the first standard of identity for enriched flour, requiring the addition of iron and B vitamins (niacin, thiamin, and riboflavin). By the 1950s, these standards extended to other cereal grain products, including bread, rice, macaroni, and noodles ([Hutt, 1984](#); [Committee on Use of Dietary Reference Intakes in Nutrition Labeling, 2004](#)). Folic acid fortification is the most recent amendment to the standard of identity for enriched grain products and is widely regarded as one of the most successful public health initiatives in recent decades ([Berry, Mulinare and Hamner, 2010](#)).

As with earlier fortification campaigns, the folic acid policy was driven by accumulating evidence that folic acid prevents neural tube defects (NTDs). In October 1990, as part of the Nutrition Labeling and Education Act, Congress directed the FDA to evaluate the link between folic acid and NTDs and to develop a plan for adding folic acid to foods ([Wright, 2003](#)). On September 14, 1992, the United States Public Health Service (USPHS) recommended that all women of childbearing age consume 400 μg of folic acid daily to prevent NTDs. In response, on March 5, 1996, the FDA amended the standard of identity to require 140 $\mu\text{g}/100 \text{ g}$ of folic acid in enriched grain products by January 1, 1998 ([Food and Drug Administration, 1996](#)). In practice, fortification was largely completed by mid-1997 ([Jacques et al., 1999](#)). I therefore define event time as March 1996—the month the FDA authorized folic acid fortification. Because

³See <https://www.cdc.gov/nccddd/folicacid/recommendations.html> (accessed on 05/20/2022).

enriched wheat flour is used in many processed foods, fortification extended beyond breads and pastas; for example, some chips contain folic acid (Figure 1). Before the mandate, voluntary addition of folic acid was prohibited in standardized foods⁴ and discouraged in other products to avoid overfortification and nutrient imbalances in the population ([Food and Drug Administration, 1996, 2015](#)).



FIGURE 1: CHIPS WITH ENRICHED WHEAT FLOUR AS AN INGREDIENT

3 Data

3.1 Birth certificate data

The Vital Statistics Natality files cover all U.S. live births and report detailed birth outcomes (birth month/year, state of birth, birth weight, gestational age in weeks, and congenital anomalies) and maternal characteristics (age, race, Hispanic origin, education, and prenatal care use) ([National Center for Health Statistics, 2003](#)). I use these data for two purposes. First, I proxy baseline folate deficiency with the pre-fortification prevalence of folate-sensitive congenital anomalies and, combining gestational age with the policy's authorization and rollout dates, assign in-utero exposure at the cohort level (Section 4.1). Second, I test whether exposure to fortification changed the distribution of infant and maternal characteristics, to rule out compositional shifts as an explanation for the main results.

3.2 Aggregate test score data

I use state test-score profiles from the public National Assessment of Educational Progress (NAEP, the Nation's Report Card) to examine how fortification exposure relates to outcomes for 4th- and 8th-graders. Restricted-use NAEP microdata are ordinarily accessible by application through the Institute of Education Sciences at the U.S. Department of Education, but

⁴"Standardized foods" are products with a federal standard of identity, such as enriched grain products.

access is currently paused.⁵

The public NAEP reports state-level averages and the shares of students at or above the NAEP achievement levels (Basic or Proficient levels)⁶ for grades 4, 8, and 12 in mathematics, reading, science, and writing at one- to two-year intervals since 1990. The state average scores are calculated by averaging the scaled scores of a representative sample of students in each state (between 0 and 500). I focus on 4th- and 8th-grade math and reading because the public series for science, writing, and all 12th grade tests are available for far fewer years. My sample includes: 4th-grade math in 2000, 2003, 2005, 2007, 2009, and 2011; 4th-grade reading in 2002, 2003, 2005, 2007, 2009, and 2011; 8th-grade math in 2000, 2003, 2005, 2007, 2009, 2011, and 2015; and 8th-grade reading in 2002, 2003, 2005, 2007, 2009, 2011, and 2015. These cohorts align closely with those in the subsequent young-adult analysis.

3.3 Young adult outcome data

Outcome data are drawn from the American Community Survey Public-Use Microdata Sample (ACS PUMS), 2016–2023 ([Ruggles and Williams., 2025](#)). I focus on young adults because the earliest fully exposed cohorts, who were conceived after the March 1996 authorization and born in the fourth quarter of 1996, are in their twenties during these waves. The analysis sample includes individuals ages 19–29. I study two domains of human capital. For education, I examine high school completion (diploma or GED) and current enrollment in post-secondary education (college or graduate/professional school). For labor markets, I consider labor force participation, employment, full-time status, and annual earnings.

3.4 Other data

I compile baseline state characteristics from multiple sources. Demographic data on race, gender, age, and total population come from the Intercensal Population Estimates ([US Census Bureau, 1990](#)). Birth and death rates, the unemployment rate, the value of products sold per farm, and average farm size come from the County and City Data Book (1988) ([US Census Bureau, 2009](#)). Transfer payments are from the Bureau of Economic Analysis's Regional Economic Information System (REIS) ([Bureau of Economic Analysis, 1988](#)). These data are combined with exposure data from birth certificates for balance test (Table A1).

In the baseline specification, I control for local economic conditions using a Bartik-style un-

⁵I applied for restricted-use student-level data, but on April 2, 2025 I was informed that all applications were paused. I do not expect access to resume soon.

⁶See <https://nces.ed.gov/nationsreportcard/mathematics/achieve.aspx> for the details about the NAEP achievement levels

employment measure, following [Ganong and Liebman \(2018\)](#) and [East \(2020\)](#). For each state of birth, I interact pre-policy sectoral employment shares from the BLS Quarterly Census of Employment and Wages (QCEW) with annual national changes in sectoral unemployment rates and sum across sectors to obtain a predicted state unemployment rate ([US Bureau of Labor Statistics, 1989-2002](#)). This approach mitigates concerns that fortification could mechanically influence the contemporaneous unemployment rate. I also include time-varying controls for potentially confounding policies: (i) Medicaid/SCHIP eligibility for pregnant women, as estimated by [Hoynes and Luttmer \(2011\)](#); (ii) mental health parity laws; (iii) an indicator for the first major AFDC waiver; and (iv) an indicator for the implementation of TANF. I first assign these exposures at the birth record level using the Natality files and then aggregate to the state-cohort level to align with the definition of fortification exposure.

4 Methods

An ideal empirical strategy would be a randomized controlled trial assigning pregnant women to receive folic acid supplements and following their children into adulthood to compare outcomes. This approach is not feasible at scale. Instead, I use the 1996 U.S. folic acid fortification of grain products as a natural experiment to estimate the long-run human capital effects of prenatal folic acid supplementation. Specifically, I use a cohort DiD design to compare cohorts exposed and unexposed to the fortification in-utero in states with high and low baseline CNS anomaly rates.

My approach parallels studies that leverage baseline regional disease prevalence to estimate the benefits of health interventions. For example, researchers have used baseline hookworm infection rates to study hookworm eradication ([Bleakley, 2007](#)), malaria prevalence to evaluate malaria eradication ([Bleakley, 2010; Kuecken, Thuilliez and Valfort, 2021](#)), measles incidence to assess vaccination ([Atwood, 2022](#)), pneumonia rates to examine the introduction of sulfa antibiotics ([Lazuka, 2020](#)), and goiter prevalence to analyze salt iodization ([Feyrer, Politi and Weil, 2017; Adhvaryu et al., 2020](#)).

4.1 Exposure measure

I proxy exposure to folic acid fortification using baseline shares of infants diagnosed with central nervous system (CNS) anomalies. Birth certificates record five categories of CNS anomalies: spina bifida, anencephaly, hydrocephaly, microcephaly, and “other” CNS anomalies. Folate deficiency is a leading cause of neural tube defects (NTDs) ([Wald et al., 2001](#)). Spina bifida and anencephaly are the most common NTDs, and other NTD subtypes are grouped under

“other” CNS anomalies. Folate deficiency can also contribute to hydrocephaly and microcephaly, directly or indirectly through NTDs (Abdel-Salam and Czeizel, 2000; Naz et al., 2016; Liu et al., 2018). Overall, the medical literature indicates that folic acid supplementation substantially reduces the risk of CNS anomalies.

I define the baseline period as January 1989–June 1993. Most states began reporting congenital anomalies in 1989; the exceptions are Louisiana (1990), Nebraska (1990), Oklahoma (1991), New York (1993), and New Mexico (not reported during the study period). This window maximizes state coverage—all states and the District of Columbia except New Mexico. Limiting the baseline to the first half of 1993 ensures that cohorts born afterward have at least four pre-periods for the event-study analysis. Figure 2a presents the baseline CNS anomaly rates across states. I define high-exposure states as those in the top quartile of baseline CNS anomaly rates. This yields 14 states: IN, IA, KS, MD, MN, NE, NJ, NY, ND, RI, SD, TN, TX, and VT. I test the robustness of using alternative thresholds for high-exposure state in Section 6.6.

I determine exposure timing using weeks of gestation recorded on birth certificates. An infant is classified as exposed if the first trimester ends after March 1996 (the month fortification was authorized), since neural tube closure occurs by the first trimester and folic acid reduces the risk of CNS anomalies by helping the neural tube close properly. Next, I aggregate the birth-level exposure indicator to the quarter-year level. As shown in Figure 2b, the share exposed during the first trimester rises sharply starting with births in 1996 Q4. I therefore define individuals born from 1996 Q4 onward as the exposed group. This pre–post timing, combined with spatial variation in baseline CNS anomaly rates, provides the key identifying variation in my empirical strategy.

I validate the exposure measure with the following two exercises. First, using NHANES III, I map state-level pre-fortification CNS anomaly rates to biomarkers of folate status and find a negative association with both serum and RBC folate (Table 1). Second, I show that trends in CNS anomalies diverge by baseline risk: high-exposure states exhibit larger post-fortification declines (Figure 5b; see Section 5).

A number of prior studies model treatment intensity using continuous, dose–response DiD designs. In this setting, that approach raises two concerns. First, the relationship between folate deficiency and cognitive development may be nonlinear, so interpreting a single slope from a continuous-dose model relies on functional-form and monotonicity assumptions that may not hold. Second, identification in dose–response DiD requires stronger parallel-trends

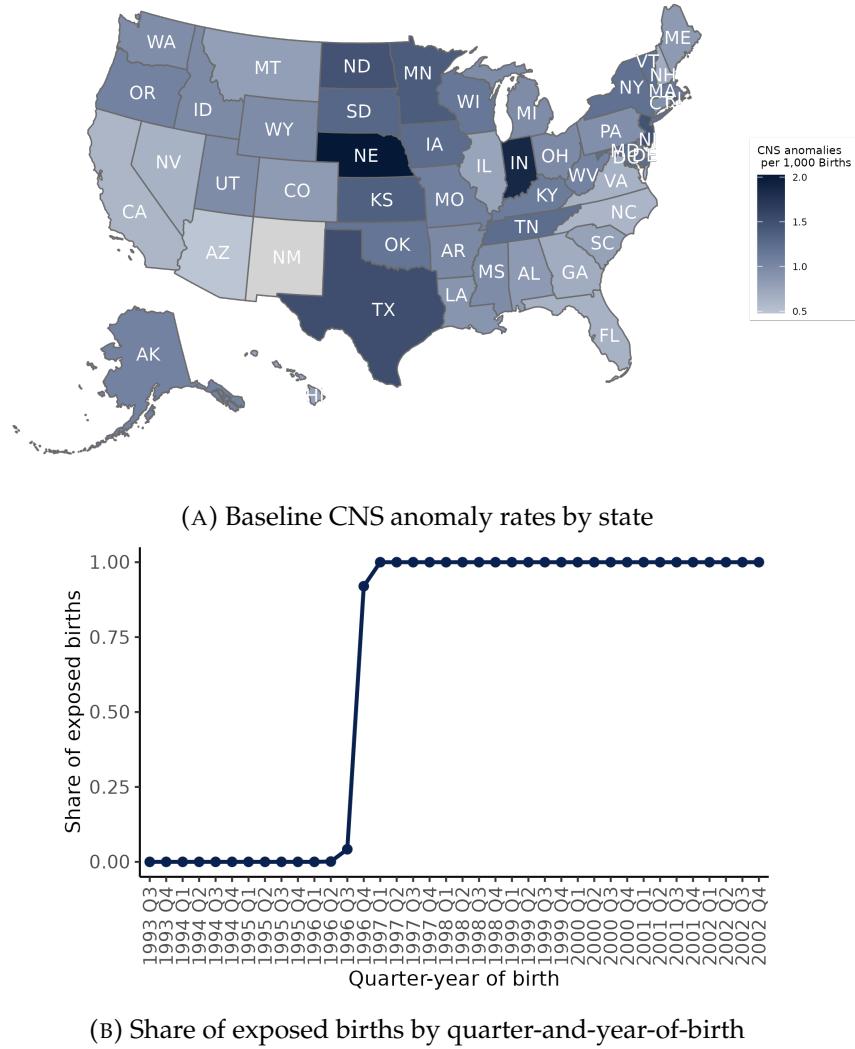


FIGURE 2: SPATIAL AND TEMPORAL VARIATION IN FORTIFICATION EXPOSURE

Notes: In Figure 2a, I compute baseline CNS anomaly rates from the restricted-use Natality files and aggregate them to the state of birth. The baseline window is January 1989–June 1993. In Figure 2b, a birth is classified as exposed if its first trimester ends after March 1996, when folic acid fortification was authorized. I then aggregate the birth-level exposure indicator to county-quarter-year cell averages.

conditions, i.e., parallel trends within each dose level and stable composition across dose cells, which are demanding here (Callaway, Goodman-Bacon and Sant’Anna, 2024). Therefore, I use a binary design contrasting high- and low-baseline-risk states (defined by pre-fortification prevalence of folate-sensitive anomalies) below and report dose–response estimates in Section 6.4 as a robustness check; they are directionally consistent with the binary results but less precise, consistent with the concerns discussed above.

4.2 Empirical strategy for test scores of 4th- and 8th-graders

Since the public aggregate NAEP data do not report students’ state and year of birth but only state of school, survey year, and grade, I map fortification exposure from state of birth to state

TABLE 1: CORRELATION BETWEEN BASELINE CNS ANOMALY RATE AND FOLATE BIOMARKERS

	Serum foalte (1)	RBC folate (2)	Serum foalte (3)	RBC folate (4)
CNS anomaly rate	-0.5375** (0.2607)	-11.41** (4.572)		
$\mathbf{1}\{\text{CNSA top quartile}\}$			-0.6059*** (0.1854)	-10.55*** (3.171)
Observations	10,842	10,913	10,842	10,913
R ²	8e-04	0.0014	0.0021	0.0025
Est./Dep. var. mean	106.91%	105.42%	102.36%	101.52%

Notes: Dependent variables are individual-level folate measure. In parentheses are heteroskasticity-robust standard errors. Regressions are weighted by MEC final examination sample weights. Data source is public-use NHANES iii. Geographical identifiers that are not suppressed include 35 counties from 13 states. CNS anomaly rate is measured at state level.

of school by constructing a migration-adjusted measure that weights birth-state exposure by the observed composition of students' birth places within each state-grade-year cell using ACS microdata from the same survey years as the NAEP files. Let j index state of school, $g \in \{4, 8\}$ grade, c survey year, and s state of birth or foreign countries (I group all foreign countries together). Define $\pi_{s|j,g,c}$ as the share of students in (j, g, c) who were born in s (including a "foreign born" category), and let $(\mathbf{1}\{\text{CNSA top quartile}\}_s)$ indicate whether birth state s is in the top quartile of baseline CNS anomaly rates. The exposure used for NAEP cell (j, g, c) is

$$P(\text{CNSA top quartile})_{j,g,t} = \sum_s \pi_{s|j,g,t} \times \mathbf{1}\{\text{CNSA top quartile}\}_s,$$

with the indicator set to 0 for foreign born. This construction produces a state-of-school exposure measure that accounts for observed migration patterns among 4th- and 8th-graders.

I assign representative ages of 10 (grade 4) and 14 (grade 8) to back out birth cohorts and define exposed cohorts as those born after 1996. Figures A1a and A1b show that roughly 90.8% of 4th graders are ages 9–10 and roughly 89.3% of 8th graders are ages 13–14. Therefore, these choices likely bias exposure downward, so the resulting estimates are conservative. The estimates are consistent when I reassign representative ages to 9 (grade 4) and 13 (grade 8) (Table B1).

The regression model to analyze the relationship between in-utero fortification exposure and test scores of 4th- and 8th-graders is:

$$Y_{jgc} = \beta \cdot P(\text{CNSA top quartile})_{j,g,c} \times \mathbf{1}\{\text{exposed cohorts}\}_{gc} + \zeta_j + \delta_{gc} + \xi_g + C_{jgc} + \varepsilon_{jgc}, \quad (1)$$

where Y_{jgt} represents the average test score outcomes of state j in survey year c , by grade g ,

$P(\text{CNSA top quartile})_{j,g,c}$ is the fortification exposure adjusted for migration pattern, $\mathbf{1}\{\text{exposed cohorts}\}_{gc}$ is a dummy for birth cohort born after 1996, ζ_j is state-of-school fixed effects to account for cohort-invariant unobserved heterogeneity, δ_{gc} is assigned-year-of-birth fixed effects to control for cohort-specific shocks, ξ_g is grade fixed effects to control for unobservables related to grade, C_{jgc} is a set of migration-adjusted exposure measures of confounding policies including Medicaid/CHIP expansion, welfare reform, and state mental health parity laws (Section 3.4) and a Bartik-style measure of state-by-year unemployment rate to control for local economic conditions, each use the same migration-adjusted shares as in $P(\text{CNSA top quartile})_{j,g,c}$. I pool fourth and eighth graders and focus on a pre versus post comparison rather than dynamic effects due to a limited number of state by grade by year cells in aggregate data.

4.3 Empirical strategy for young adults' outcomes

My baseline regression model to examine effects of in-utero fortification exposure on young adults' outcomes is:

$$Y_{istc} = \beta \cdot \mathbf{1}\{\text{CNSA top quartile}\}_s \times \mathbf{1}\{\text{exposed cohorts}\}_i + \mu_s + \lambda_t + \gamma_c + C_{istc} + \varepsilon_{istc}, \quad (2)$$

where Y_{istc} represents the outcome for individual i who born in state s and quarter-and-year t recorded in survey year c , $\mathbf{1}\{\text{CNSA top quartile}\}_s$ is an indicator equal to 1 for the states with the highest quarter of baseline CNS anomaly rates and 0 otherwise, $\mathbf{1}\{\text{exposed cohorts}\}_i$ is an indicator equal to 1 for individuals born on or after 1996 Q4 and 0 otherwise, μ_s is state-of-birth fixed effects to account for cohort-invariant unobserved heterogeneity, λ_t is quarter-and-year-of-birth fixed effects to control for cohort-specific shocks, γ_c is survey year fixed effects to control for unobservables related to age,⁷ C_{ist} is a set of control variables including (i) individual characteristics, including gender, race dummies, and Hispanic origin, (ii) exposure measures of confounding policies including Medicaid/CHIP expansion, welfare reform, and state mental health parity laws (Section 3.4), and (iii) a Bartik-style measure of state-by-year unemployment rate to control for local economic conditions.

The coefficient β is the primary parameters of interest, which capture the effects of in-utero exposure to folic acid fortification across cohorts. Because my outcome data does not identify who is folate-deficient, β_τ should be interpreted as intent-to-treat (ITT) effects. Below, I approximate treatment-on-the-treated (TOT) effects by scaling the ITT using the estimated pre-fortification share of folate-deficient women ages 19–45 from NHANES.

⁷Because we can obtain age from survey year and year of birth, controlling for age fixed effects is equivalent to controlling for survey year fixed effects.

Because the baseline exposure threshold at which fortification has no effect is unknown, it is plausible that residents of non-top-quartile states also benefited. The difference-in-differences estimate therefore captures a contrast: the effect in top-quartile states relative to the effect in all other states. If the effect in other states is positive, this contrast understates the true impact for top-quartile states and does not recover the program's overall average effect. The estimates should be interpreted as a lower bound for high-baseline states rather than a nationwide average.

4.4 Identifying assumptions

The validity of this empirical strategy relies on two assumptions: parallel trends and no anticipation. Parallel trends require that, absent folic acid fortification, average outcomes for high- and low-exposure groups would have evolved similarly across birth cohorts. Although this is not directly testable because we do not observe the counterfactual, I conduct partial checks of its plausibility. First, I examine whether exposure is correlated with pre-1989 county- or state-level characteristics. A violation could arise if prenatal high-exposure areas systematically differ in factors that also influence long-run outcomes. For example, if poorer states happened to have higher CNS anomaly rates and economic conditions at birth independently affect adult outcomes, estimates might simply reflect underlying economic conditions rather than the effect of fortification. I then regress the exposure measure, $\mathbf{1}\{\text{CNSA top quartile}\}_s$, on these baseline characteristics. Table A1 shows that while some covariates are individually correlated with high exposure, as a whole they explain only about 30% of cross-state variation and less than 10% of cross-county variation, suggesting the high-exposure designation is largely orthogonal to observables and consistent with the parallel trends assumption. In baseline model, I also include rich controls for individual characteristics and confounding policy and economic variables wherever possible.

Second, I present dynamic effects using an event-study design to examine the presence of any pre-treatment trends for young adults' outcomes. In most cases, we do not observe evidence of such trends. In Section 6.6, I add time-varying covariates to relax the unconditional parallel trends assumption underlying the baseline results. With these covariates included, the parallel trends assumption needs only to hold conditional on them.

The no-anticipation assumption requires (i) that mothers in high-exposure areas did not alter behavior in advance of fortification, and (ii) that food manufacturers did not begin fortifying before March 1996. In this setting, anticipatory responses are unlikely. ([Petrini, Damus and Johnston, 1999](#)) indicates low awareness of folic acid among women of childbearing age.

Because the fortification mandate was motivated by scientific evidence and directed at food manufacturers, it likely had low salience for the general public. On the supply side, voluntary folic acid fortification was prohibited for standardized foods and discouraged elsewhere due to concerns about overfortification and nutrient imbalances ([Food and Drug Administration, 1996, 2015](#)). Consistent with these arguments, I do not observe any evidence for anticipatory behavior in my event-study results.

5 Descriptive First-Stage Evidence

5.1 Folate content in foods increased after folic acid fortification

First, folic acid content in foods rose markedly after fortification. Using the 1994–1996 and 1998 waves of the Continuing Survey of Food Intakes by Individuals (CSFII), I compare per-serving folic acid in the same foods before and after the mandate, based on USDA's recipe-based nutrient calculations. Because the CSFII records the reason for composition changes, including enrichment/fortification, reformulation, agricultural or processing modifications, and implementation of the Nutrition Labeling and Education Act. I can isolate changes attributable to fortification. As shown in Figure 3, folic acid increased across a wide range of products, from white bread to snack/cookie bars; in total, more than 350 basic food items show higher folic acid due to fortification ([Anderson et al., 2001](#)).

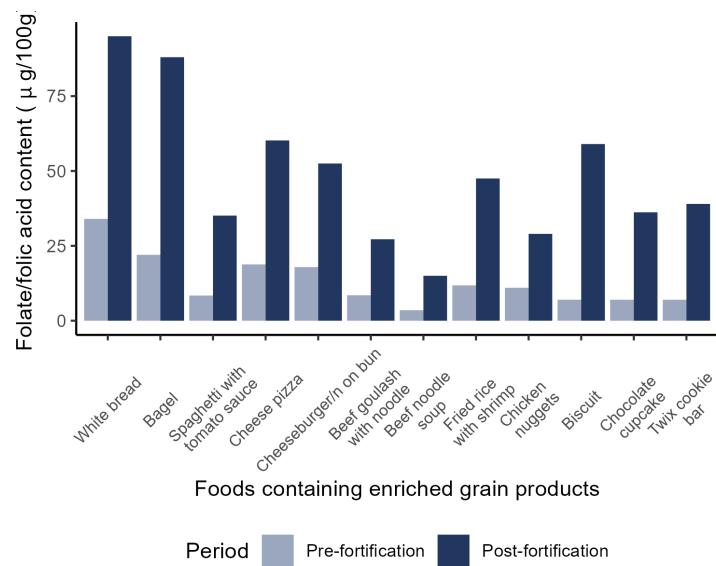


FIGURE 3: CHANGES IN FOLATE CONTENTS IN SELECTED FOODS DUE TO FORTIFICATION

Notes: Data on food folate content is from USDA Continuing Survey of Food Intakes by Individuals (CSFII) 1994–1996 and 1998. Folate content is estimated by USDA based on the recipe. Changes in folate content in this graph are solely due to fortification.

5.2 Dietary folate intake and blood folate increased after folic acid fortification

Second, using NHANES, I document sharp increases in dietary folate intake and blood folate among women ages 19–45 following fortification. Mean dietary intake rose by nearly 50% (Figure 4a). The share with intake below 400 $\mu\text{g}/\text{day}$ fell from 98.65% to 69.87% ($\Delta = -28.78$ percentage points). These intake measures exclude folic acid from supplements and medications (Ahluwalia et al., 2016). Biomarkers show parallel gains: serum folate more than doubled and red blood cell (RBC) folate rose by nearly 50% (Figure 4b), indicating sustained improvements in folate status. Blood folate measurement methods were stable over 1999–2006 (Pfeiffer et al., 2012).

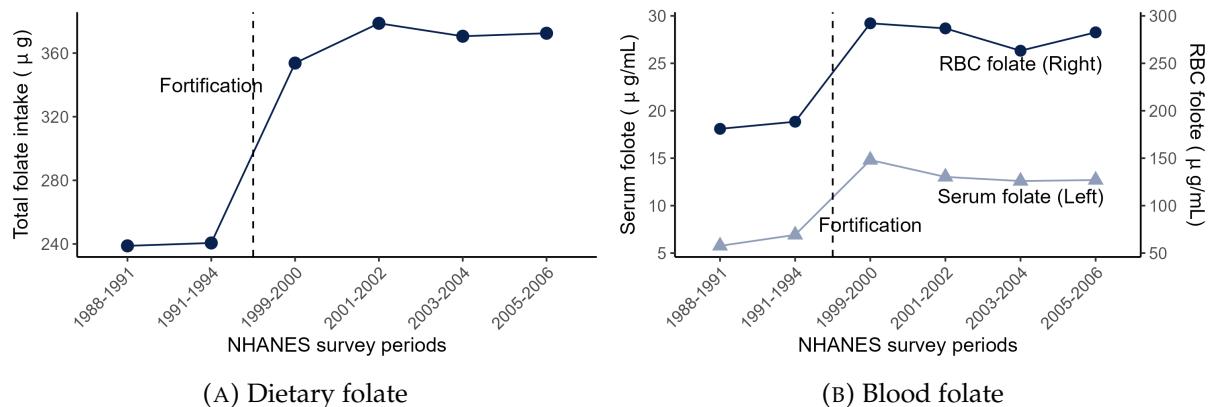


FIGURE 4: DIETARY AND BLOOD FOLATE CONCENTRATIONS BEFORE AND AFTER FORTIFICATION

Notes: Data is from harmonized NHANES data cleaned by Nguyen et al. (2023) to ensure comparability of folate measures across survey periods. Mobile examination center (MEC) final examination sample weights are used for all folate measures in all survey periods.

5.3 Congenital anomalies declined after folic acid fortification

Finally, Figure 5a shows that as folate intake and absorption rose, CNS anomaly rates declined. After a flat period from 1992 to 1996, incidence fell sharply following fortification. This pattern is unlikely to reflect broader healthcare improvements, as rates of other congenital anomalies remained stable over the same period. Figure 5b further compares trends by exposure (defining high exposure as the top quartile of baseline CNS anomaly rates): rates fell in both high- and low-exposure states, with a larger decline in high-exposure areas, supporting the validity of the exposure measure used in my research design.

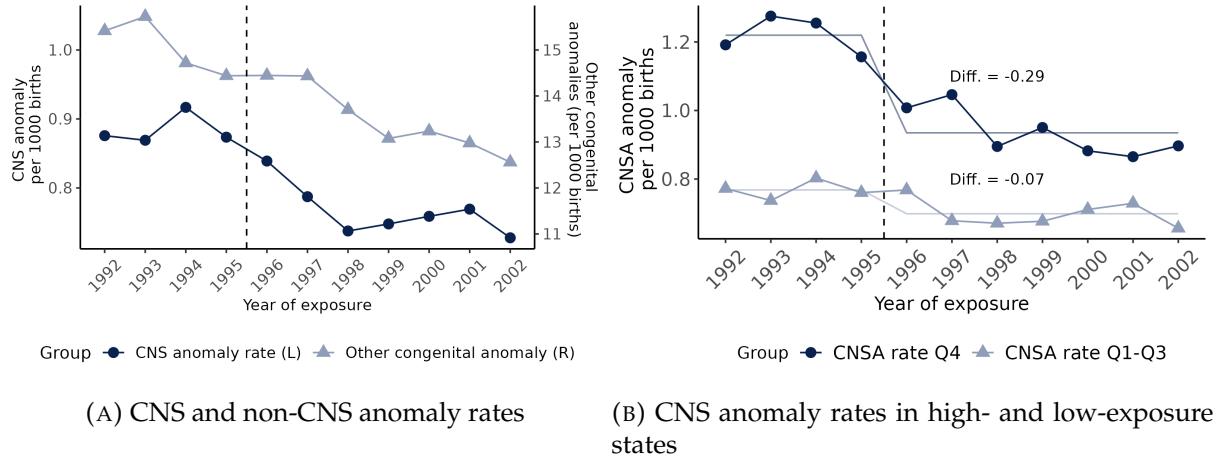


FIGURE 5: TRENDS IN CONGENITAL ANOMALY RATES

Notes: The unit of CNS Anomaly rate is cases per 1,000 births. High exposure is defined as the top quartile baseline CNS anomaly rate.

6 Long-run Effects of folic acid fortification

6.1 Effects on 4th- and 8th graders

Table 2 reports the effects of in-utero exposure to folic acid fortification on NAEP performance among 4th- and 8th-graders assuming 4th graders are 10 years old and 8th-graders are 14 years old. Columns (1)-(3) show positive associations between fortification exposure and performance in mathematics: average math scores rise by 2.28 (0.88% at the mean), and the shares of students scoring at or above the NAEP Basic level increase by 1.67%, and the shares at or above the NAEP Proficient level increase by 2.67%. I do not find noticeable changes in average reading scores or in the share of students at or above NAEP Basic. Fortification exposure is linked to lower shares at or above NAEP Proficient, but this pattern is not consistent across specifications (see Appendix B).

Results for 4th- and 8th-grade mathematics are broadly consistent when I assume fourth graders are nine and eighth graders are thirteen (Table B1); when I use state of birth exposure as a proxy for state of school exposure (Table B2); when I omit controls for confounding policies and the unemployment rate (Table B3); when I include state of school by grade fixed effects and grade effects (Table B4); and when I include state of school by year of birth and grade fixed effects (Table B5).

6.2 Effects on college-age adults (19–22 years old)

Table 3 reports the effects of in utero exposure to folic acid fortification on education and labor outcomes for college-age young adults (19–22). Post-secondary school enrollment among the

TABLE 2: LONG-RUN EFFECTS OF FOLIC ACID FORTIFICATION ON TEST SCORES OF 4TH- AND 8TH-GRADERS

	Math			Reading		
	Average score (1)	% \geq Basic (2)	% \geq Proficient (3)	Average score (4)	% \geq Basic (5)	% \geq Proficient
$P(\text{CNSA top quartile}) \times \text{exposed cohorts}$	2.28*** (0.78)	1.67** (0.77)	2.67*** (0.79)	-1.28 (1.02)	0.09 (1.35)	-1.32** (0.61)
Observations	459	459	459	459	459	459
R ²	0.99	0.92	0.91	0.99	0.93	0.92
Dep. var. mean	257.72	76.15	34.82	238.93	68.96	31.35
Est./Dep. var. mean	0.88%	2.20%	7.67%	-0.54%	0.13%	-4.20%
State-of-school FE	✓	✓	✓	✓	✓	✓
Year-of-birth FE	✓	✓	✓	✓	✓	✓
Grade FE	✓	✓	✓	✓	✓	✓
Confounder controls	✓	✓	✓	✓	✓	✓

Notes: The table presents cohort DD estimates and their standard errors. $P(\text{CNSA top quartile})$ is the fortification exposure adjusted for migration. It varies across state of school, survey year, and grade. The exposed cohorts are the one with assigned year of birth > 1996 . I control for state-of-school, year-of-birth, grade fixed effects, Medicaid eligibility, exposure to mental-health parity laws and welfare reforms, and the local unemployment rate. The unit of observation is state-by-year-by-grade cells. Regressions and dependent-variable means are weighted by the number of students by state-of-school, survey year, and grade from ACS. Standard errors are clustered by state of school. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

high-exposure group rises by 1.39 p.p. (2.77% at the sample mean), driven by higher college enrollment of 1.79 p.p. (3.67% at the mean); both increases are marginally significant. There is no detectable change in labor-force participation, but high-exposure individuals are 1.16 p.p. less likely to work full time (3.76% at the mean) and earn \$452.9 less per year (4.02% at the mean). Taken together, these patterns suggest that in utero exposure to fortification shifts time toward schooling over work at college ages. Figures 6a–6e present cohort-by-cohort dynamic effects of in utero exposure to folic acid fortification. I find no noticeable pre-trends for most outcomes.

In Section 5, I use NHANES to estimate that 98.65% of women aged 19–45 were folate-deficient before fortification, and that the share deficient fell by 28.78 percentage points afterward. I use these figures to bound treatment-on-the-treated (TOT) effects from the reduced-form ITT estimates. Let $p_{\text{def}}^{\text{pre}} = 0.99$ and $\Delta p_{\text{def}} = 0.29$. If all previously deficient women benefited, the implied treatment rate is $p_{\text{def}}^{\text{pre}}$, yielding a lower bound $\text{TOT}_L = \text{ITT}/p_{\text{def}}^{\text{pre}}$. If only those whose deficiency was reversed benefited, the relevant treatment rate is Δp_{def} , yielding an upper bound $\text{TOT}_U = \text{ITT}/\Delta p_{\text{def}}$. Thus, $\text{TOT} \in [\text{ITT}/0.99, \text{ITT}/0.29]$, with the lower (upper) bound corresponding to broader (narrower) treatment incidence assumptions. Therefore, the treatment-on-the-treated (TOT) effects are as follows: post-secondary enrollment increases by

TABLE 3: LONG-RUN EFFECTS OF FOLIC ACID FORTIFICATION ON COLLEGE-AGE ADULTS (19-22 YEARS OLD)

	School enrollment		Labor outcomes		
	post-secondary school enrollment	College enrollment	Being in the labor force	Working full-time	Annual wage (\$1,000)
		(1)	(2)	(3)	(4)
1{CNSA top quartile} × exposed cohorts	0.0139* (0.0077)	0.0179* (0.0093)	-0.0047 (0.0057)	-0.0116*** (0.0037)	-0.4529*** (0.1481)
Observations	993,911	993,911	993,911	993,911	993,911
R ²	0.0640	0.0686	0.0335	0.0565	0.0621
Dep. var. mean	0.5023	0.4881	0.6916	0.3091	11.2703
Est./Dep. var. mean	2.77%	3.67%	-0.68%	-3.76%	-4.02%
State-of-birth FE	✓	✓	✓	✓	✓
Quarter-year-of-birth FE	✓	✓	✓	✓	✓
Survey year FE	✓	✓	✓	✓	✓
Individual char controls	✓	✓	✓	✓	✓
Confounder controls	✓	✓	✓	✓	✓

Notes: The table presents cohort DD estimates and their standard errors. The treated group consists of births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top quartile of baseline CNS anomaly rates. I control for state-of-birth fixed effects; quarter-by-year-of-birth and survey-year fixed effects; and controls for gender, race, Hispanic origin, Medicaid eligibility, exposure to mental-health parity laws and welfare reforms, and the local unemployment rate. The unit of observation is the individual. Regressions and dependent-variable means are weighted by the IPUMS person weight; percentile calculations are weighted by the number of births. Standard errors are clustered by state of birth. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

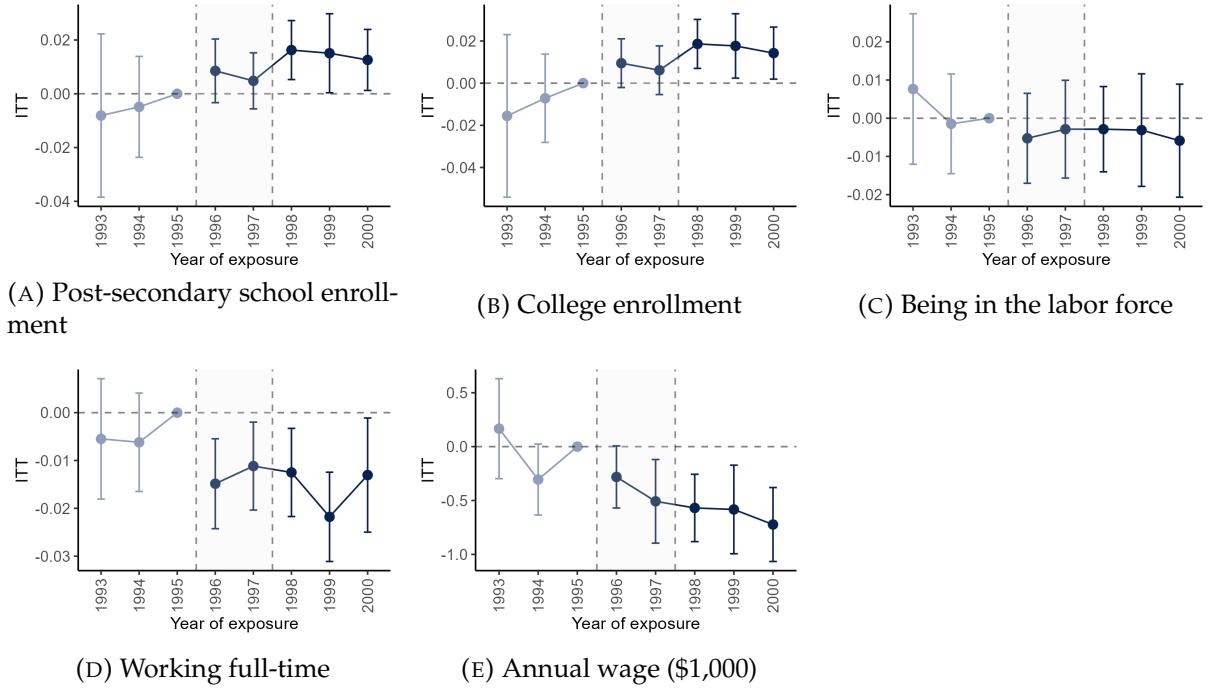
1.40-4.79 p.p. (2.80-9.55% at the mean) and college enrollment by 1.81-6.17 p.p. (3.70-12.66%), while the likelihood of working full time decreases by 1.17-4.00 p.p. (3.80-12.97%), and annual wages decline by \$457-\$1,562 (4.06-13.86%).

6.3 Effects on young adults over college ages (23-30 years old)

Table 4 reports the effects of in utero exposure to folic acid fortification on education and labor outcomes for young adults beyond college age (23-30). In the high-exposure group, graduate-school enrollment increases by 0.68 p.p. (11.19% at the mean), the likelihood of earning a bachelor's degree rises by 0.78 p.p. (2.23% at the mean), and the likelihood of earning a STEM degree increases by 0.54% (4.22% at the mean). At the same time, labor outcomes are unchanged: labor-force participation, full-time work, and annual wage are similar to the comparison group.

Figures 7a-7f show cohort-by-cohort dynamic effects of in utero exposure to folic acid fortification for young adults beyond college age (23-30). The patterns align with the overall estimates in Table 4 and show no noticeable pre-trends. That said, the dynamic effects on the

FIGURE 6: LONG-RUN EFFECTS OF FOLIC ACID FORTIFICATION ON COLLEGE-AGE YOUNG ADULTS (19-22 YEARS OLD)



Notes: The figure plots cohort-specific (dynamic) estimates with 95% confidence intervals. The treated group is the births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top quartile of baseline CNS anomaly rates. I control for state-of-birth fixed effects; quarter-by-year-of-birth and survey-year fixed effects; and controls for gender, race, Hispanic origin, Medicaid eligibility, exposure to mental-health parity laws and welfare reforms, and a Bartik-style measure of local unemployment rate. The shaded region denotes cohorts with partial exposure. The unit of observation is individuals. Regressions and dependent-variable means are weighted by the IPUMS person weight. Standard errors are clustered by state of birth.

likelihood of earning a Bachelor's degree and a STEM degree are imprecise, so these estimates should be interpreted with caution.

Similarly, using the pre-fortification share of folate-deficient women (99%) and the post-fortification reduction (29%) as treatment rates, the implied TOTs indicate that in-utero fortification exposure increases the probability of enrolling in graduate school by 0.69-2.34 p.p. (11.3-38.59% at the mean).

For transparency, I report the results of fortification exposure on the full distribution of grade attendance and educational attainment for age-appropriate groups in Tables A2 and A3. Table A2 shows that enrollment rises among exposed young adults, driven primarily by college enrollment; effects on attendance at grades below college are near zero, and the corresponding sample means are similar across exposure groups. The sample mean for graduate-school attendance declines from 0.34 in the age-appropriate group to 0.04 in the full 19-31 sample, and the estimated effects on graduate-school attendance for 19-31-year-olds are not

TABLE 4: LONG-RUN EFFECTS OF FOLIC ACID FORTIFICATION ON YOUNG ADULTS OVER COLLEGE AGES (23-30 YEARS OLD)

	Educational outcomes			Labor outcomes		
	Graduate school enrollment (1)	Bachelor's degree (2)	STEM degree (3)	Being in the labor force (4)	Working full-time (5)	Annual wage (\$1,000) (6)
1{CNSA top quartile} × Exposed cohorts	0.0068*** (0.0019)	0.0078** (0.0032)	0.0054** (0.0020)	-0.0021 (0.0038)	-0.0034 (0.0039)	-0.4263 (0.3591)
Observations	982,727	982,727	982,727	982,727	982,727	982,727
R ²	0.0129	0.0787	0.0286	0.0106	0.0397	0.0924
Dep. var. mean	0.0605	0.3507	0.1290	0.8366	0.5919	30.1348
Est./Dep. var. mean	11.19%	2.23%	4.22%	-0.25%	-0.57%	-1.41%
State-of-birth FE	✓	✓	✓	✓	✓	✓
Quarter-year-of-birth FE	✓	✓	✓	✓	✓	✓
Survey year FE	✓	✓	✓	✓	✓	✓
Individual char controls	✓	✓	✓	✓	✓	✓
Confounder controls	✓	✓	✓	✓	✓	✓

Notes: The table presents cohort DD estimates and their standard errors. The treated group consists of births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top quartile of baseline CNS anomaly rates. I control for state-of-birth fixed effects; quarter-by-year-of-birth and survey-year fixed effects; and controls for gender, race, Hispanic origin, Medicaid eligibility, exposure to mental-health parity laws and welfare reforms, and the local unemployment rate. The unit of observation is the individual. Regressions and dependent-variable means are weighted by the IPUMS person weight; percentile calculations are weighted by the number of births. Standard errors are clustered by state of birth. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

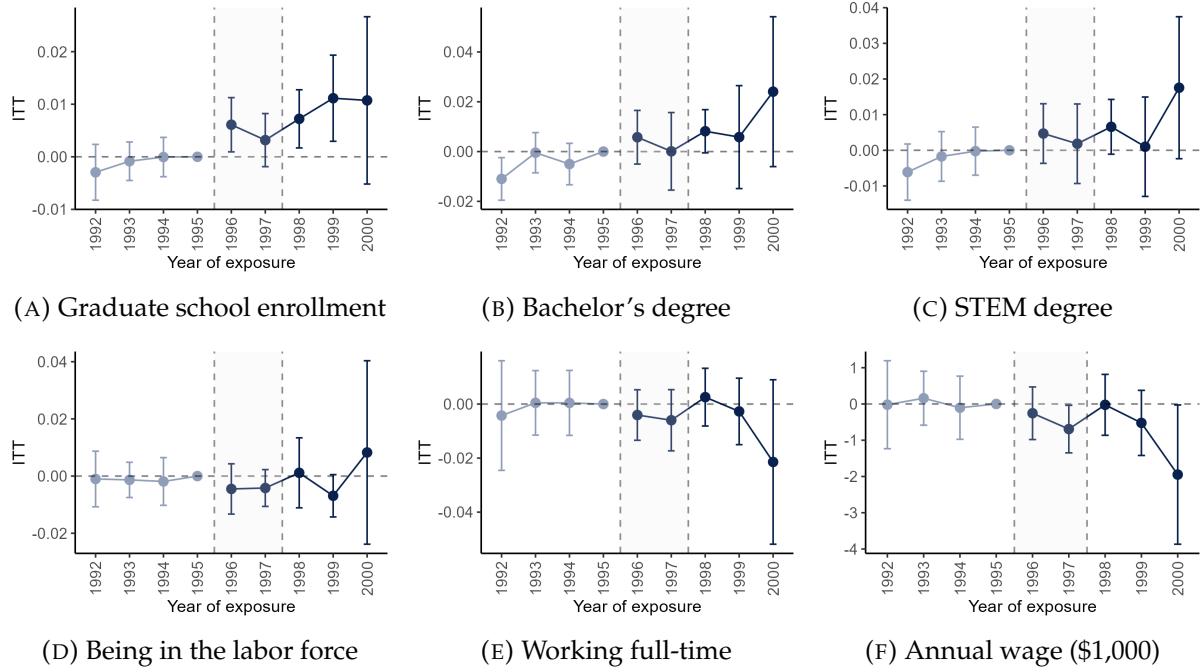
statistically different from zero, likely because many individuals in their early 20s are still in college.

Table A3 shows no differences between exposed and unexposed young adults in completion of lower schooling (kindergarten, K–12) or in having some college without a degree. This is expected, as these measures are relatively insensitive to potential changes in cognitive ability associated with fortification exposure. Among those over 22, graduate-degree completion also does not differ, likely because many graduate degrees are completed well after age 22. Overall, I do not find that the full-distribution results are inconsistent with the main findings.

6.4 Dose response

Figures A2a–A2k re-estimate the models using the continuous baseline CNS anomaly rate as the exposure proxy, tracing out a (potentially nonlinear) dose–response between exposure intensity and outcomes. The patterns largely mirror the main results: among 19–22-year-olds, in utero exposure is associated with higher post-secondary enrollment and lower full-time work and wages, though the estimates are noisier; among 23–30-year-olds, exposure is associated with higher graduate-school enrollment. The added imprecision likely reflects (i) measure-

FIGURE 7: LONG-RUN EFFECTS OF FOLIC ACID FORTIFICATION ON YOUNG ADULTS OVER COLLEGE AGES (23-30 YEARS OLD)



Notes: The figure plots cohort-specific (dynamic) estimates with 95% confidence intervals. The treated group is the births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top quartile of baseline CNS anomaly rates. I control for state-of-birth fixed effects; quarter-by-year-of-birth and survey-year fixed effects; and controls for gender, race, Hispanic origin, Medicaid eligibility, exposure to mental-health parity laws and welfare reforms, and a Bartik-style measure of local unemployment rate. The shaded region denotes cohorts with partial exposure. The unit of observation is individuals. Regressions and dependent-variable means are weighted by the IPUMS person weight. Standard errors are clustered by state of birth.

ment error in the continuous proxy, which attenuates estimates toward zero, and (ii) the possibility that very small baseline-deficiency differences have limited behavioral salience. Consistent with this, the main specification—Q4 versus the rest—produces cleaner, more precise estimates.

6.5 Fertility selection

This section evaluates whether the long-run effects of fortification reflect fertility selection. Tables A4 and A5 show no detectable impact of fortification exposure on birth outcomes or on the distribution of most maternal characteristics. The one exception is an increase in the share of mothers aged ≤ 22 . Figure A3, however, indicates that this rise follows a broader upward trend rather than a discrete change at fortification. If anything, a larger share of younger mothers would bias our estimates toward zero, given their higher baseline risk of adverse outcomes. The persistence of our main effects in the presence of this compositional shift suggests the true impact of in-utero fortification exposure is likely understated. Overall, these results argue against fertility selection as the primary driver of my findings.

6.6 Robustness

This section shows that the estimated effects on post-secondary school enrollment, full-time employment, and annual earnings are robust to multiple checks. A placebo randomization further indicates that the main estimates are unlikely to be driven by random noise.

(i) *Alternative model specifications.* I begin by testing robustness to alternative model specifications. Figures C1a–C1l compare the baseline with four variants: (i) a parsimonious model with only state-of-birth, quarter-by-year-of-birth, and survey-year fixed effects (omitting demographics and other time-varying covariates); (ii) the baseline excluding time-varying covariates; (iii) the baseline replacing separate state and survey-year effects with state-of-birth-by-survey-year fixed effects; and (iv) the baseline adding state-of-residence fixed effects. Conclusions are unchanged across specifications.

(ii) *Alternative exposure thresholds.* To assess sensitivity to the exposure definition, I vary the baseline-risk cutoff used to define treated states. The baseline uses the top quartile (top 25%) of pre-existing CNS anomaly rates. I re-estimate the models using alternative thresholds—top 30% and top 20%—holding all other specification choices fixed. As shown in Figures C2a–C2l, the point estimates remain stable in sign and magnitude, and the confidence intervals largely overlap the baseline estimates. This pattern indicates the results are not driven by an arbitrary cutoff. Precision moves as expected with the number of treated states (slightly tighter at 30%, slightly looser at 20%), but the qualitative conclusions are unchanged.

(iii) *Sharper comparison.* I compare young adults born in states in the top versus bottom quartiles of baseline CNS anomaly rates. Excluding the middle quartiles sharpens the exposure contrast, reduces attenuation from measurement error, and sidesteps functional-form concerns associated with pooling medium- and low-exposure states in the baseline. Figures C3a–C3k show larger, more precisely estimated effects than in the baseline. Event-study pre-trends are flat in this extreme-groups comparison, further supporting parallel trends.

(iv) *Excluding the 2020 data.* I exclude 2020 to mitigate bias from the pandemic-related spike in nonresponse. ACS response rates were 94.7% (2016), 93.7% (2017), 92.0% (2018), 86.0% (2019), 71.2% (2020), 85.3% (2021), 84.4% (2022), and 84.7% (2023). The 2020 rate is substantially lower than in other years, and post-pandemic rates remain below pre-pandemic levels but are stable through 2023. Figures C4a–C4k show that excluding 2020 does not alter my conclusions.

(v) *Randomization test.* Finally, to test robustness to random noise, I run a randomization (placebo) exercise. I recompute ATT estimates after randomly reassigning treatment status 500

times while preserving its empirical distribution across states. Figures C5c–C5g show that the main estimates lie well into the tails of the placebo distributions, suggesting that they are unlikely to be driven by chance.

7 Magnitudes, economic significance, and policy comparison

7.1 Benchmarking magnitudes

To gauge magnitude, I benchmark the long-run education effects of folic acid fortification against two micronutrient policies (salt iodization and iron fortification of bread) and the Food Stamp Program. Because these policies differ in target populations, timing, exposure definitions, and outcomes, the exercise is illustrative rather than a strict apples-to-apples comparison.

For salt iodization, I draw on [Adhvaryu et al. \(2020\)](#), which estimates in-utero exposure using a continuous proxy (baseline goiter prevalence) and report effects on years of schooling and income for adults aged 39–60. Moving exposure from the 25th to the 75th percentile increases years of schooling by 0.0712 years for women (about 0.63% of the mean; ITT) and 0.0313 years for men (about 0.27% of the mean; ITT) (Table 6 in [Adhvaryu et al. \(2020\)](#)). They also find income increases of 14.9% for women (ITT) and 2.88% for men (ITT) following salt iodization (Table 4 in [Adhvaryu et al. \(2020\)](#)).

For iron fortification, [Niemesh \(2015\)](#) estimate that moving from zero to a full 19 years of exposure at a one-standard-deviation difference in iron consumption implies a 0.05-year increase in schooling (ITT) among ages 22–50 (imprecisely estimated), and a 2.9% increase in total income controlling for years of schooling (Table 7 in [Niemesh \(2015\)](#)).

For the Food Stamp Program, [Bailey et al. \(2024\)](#) estimate that full exposure—from conception through age five—increases years of schooling by 0.2294 years (TOT), and raises labor income by 7.125% (TOT).

For folic acid fortification, I translate the estimated impact of fortification on post-secondary enrollment into years of schooling using the college enrollment effect for ages 19–22 from Table 3 and the graduate-school enrollment effect for those over 22 from Table 4. I conservatively assume a 60% completion rate for both levels; non-completers attend one year; a completed bachelor’s degree takes four years; and a completed graduate program averages two years. Under these assumptions, an additional college enrollee contributes $0.40 \times 1 + 0.60 \times 4 = 2.8$ years of schooling on average, and an additional graduate-school enrollee contributes

$0.40 \times 1 + 0.60 \times 2 = 1.6$ years. Multiplying by the ITT effects (0.0179 for college and 0.0068 for graduate school) yields an ITT gain of $2.8 \times 0.0179 + 1.6 \times 0.0068 = 0.0610$ years. To bound the treatment-on-the-treated (TOT), I scale by the first-stage exposure range from Section 5, i.e, pre-fortification deficiency share = 0.99 and post-fortification decline = 0.29, which implies $\text{TOT} \in [\frac{0.061}{0.99}, \frac{0.061}{0.29}] \approx [0.0616, 0.2103]$ years. Standard errors scale by the same factors. Starting from the enrollment standard error 0.0093 for college enrollment and 0.0019 for graduate school enrollment, the ITT years standard error is $0.0093 \times 2.8 + 0.0019 \times 1.6 = 0.0291$. For the bounds, $0.0291/0.99 \approx 0.0294$ and $0.0291/0.29 \approx 0.1003$.

Figure 8 summarizes the comparisons. For years of schooling, folic acid fortification's effect is comparable to salt iodization for women and smaller than the Food Stamp Program. Yet fortification is far more cost-effective: it yields roughly 26.9–91.8% of FSP's schooling gains at $\approx 0.01\%$ of FSP's annual cost ($\approx \$3$ million (Grosse et al., 2005) vs $\approx \$29$ billion (Food and Nutrition Service, 2005)).

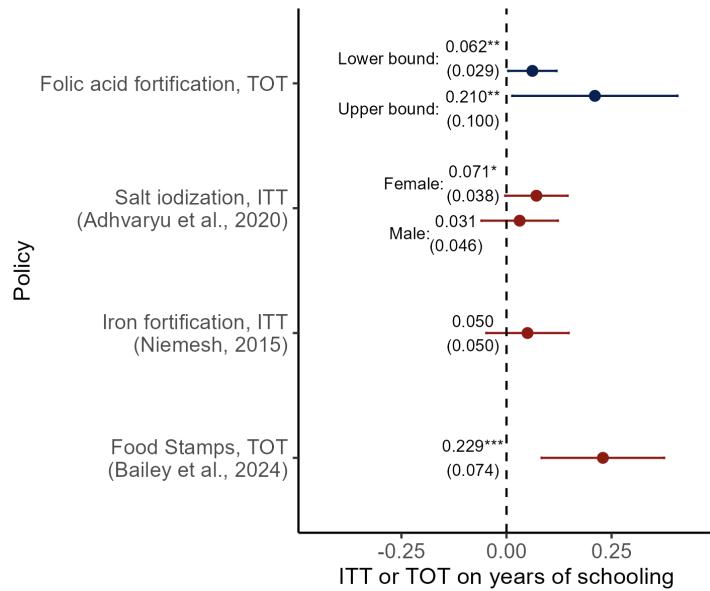


FIGURE 8: COMPARING LONG-RUN EFFECTS ON YEARS OF SCHOOLING ACROSS NUTRITION INTERVENTIONS

Notes: The figure plots point estimates for each nutrition intervention's effect on years of schooling and log(income), with 95% confidence intervals. The folic acid estimate converts its effect on post-secondary school enrollment into years of schooling and log(income); all other values are regression coefficients taken directly from the cited studies. Adhvaryu et al. (2020) estimate the effect of prenatal exposure to salt iodization using baseline goiter prevalence as a continuous proxy for iodine deficiency; Niemesh (2015) estimate the effect of prenatal exposure to iron fortification of bread using estimated iron consumption as a proxy for iron deficiency. For comparability, both are rescaled to reflect a shift from the 25th to the 75th percentile of the corresponding exposure measure. Bailey et al. (2024) report the effect of exposure to Food Stamp from conception to age five.

7.2 Monetizing the long-run human capital gains

How much does the shift from work to school in early adulthood translate to longer-run human capital gains? Using a conservative 20% return to college education ([Hoekstra, 2009](#); [Zimmerman, 2014](#)), the implied increase in later adult income operating through schooling alone is $\text{TOT} \in [\frac{0.0610 \times 0.20}{0.99}, \frac{0.0610 \times 0.20}{0.29}] \approx [1.23\%, 4.21\%]$. Because this excludes non-education pathways, it should be viewed as a lower bound. For example, [Bailey et al. \(2024\)](#) conclude that schooling accounts for roughly one-third of the income gains from early-life nutrition; health and other channels plausibly explain the remainder.

I then take a benchmark for prime-age annual earnings of \$53,996 (full-time workers aged 30-50, estimated using ACS IPUMS 2016-2023) and a 40-year working horizon with a 7% discount rate ([US Office of Management and Budget, 2003](#)). Based on my estimate for earnings via the schooling channel, i.e., 1.23% to 4.21%, The present value (PV) annuity factor is $\Phi(40, 0.07) = \frac{1 - (1 + 0.07)^{-40}}{0.07} \approx 13.33$. Thus the PV gain per treated person from the schooling channel is $\Delta \text{PV}_{\text{HC}} \approx (1.23\% \text{ to } 4.21\%) \times \$53,996 \times 13.33 \approx \$8,853 \text{ to } \$30,302$.

To express this at a cohort scale, consider 1,109,368 births in high-exposure states (the average number of births in the 14 high-exposure states from 1996 to 2002). Using the first-stage exposure change of 29% ([Section 5](#)) as a treated share, about 321,717 births are effectively exposed at the TOT margin. The cohort PV from the schooling channel is then $\text{Benefit} \approx 321,717 \times (\$8,853 \text{ to } \$30,302) \approx \$1.85 \text{ to } \$6.25 \text{ million}$. This calculation suggests that the net benefits of fortification would be underestimated by \$2.85–\$9.75 million per year if omitting its long-run human capital benefits.

8 Conclusion

This paper provides the first evidence on the long-run human-capital effects of folic acid fortification of enriched grains. I exploit the 1996 authorization of fortification and cross-state differences in baseline folate deficiency, using pre-fortification CNS anomaly rates as a proxy, to measure in-utero exposure to folic acid fortification. I find positive associations between fortification exposure and performance in mathematics of 4th- and 8th-graders. I also find a consistent reallocation from work to school among exposed college-age adults (19-22 years old): exposed individuals are more likely to enroll in college, less likely to work full time, and earn slightly less in the short run. For young adults beyond college age (23-30 years old), I find that exposed cohorts are more likely to enroll in graduate school. Estimates are stable across a wide set of alternative specifications, including control sets, exposure thresholds, and sample

selection.

I benchmark these estimates against other nutrition interventions (salt iodization, iron fortification) and the Food Stamp Program. The years of schooling impacts are comparable to those documented for salt iodization and smaller than those for the Food Stamp Program, a much larger program. However, because folic acid fortification costs a tiny fraction of the Food Stamp Program's budget, it is much more cost-effective.

The findings of this paper align with the broader fetal-origins literature: well-timed early-life interventions can generate long-run human capital gains. Given the low unit cost of fortification and its minimal reliance on behavior change, similar or larger benefits are plausible in settings where folate deficiency is more prevalent and access to supplements is limited, including many low- and middle-income countries. For policymakers operating under tight budget constraints, targeting severe micronutrient deficiencies via fortification can deliver favorable benefit-cost ratios relative to broad food subsidies.

The findings of this paper provide strong suggestive evidence of long-run human capital effects from in utero fortification exposure. However, further evidence is needed to establish a more compelling causal relationship when proper data is available in the future. First, we need to broaden the outcome space for K-12 children and young adults, such as IQ test results, credits earned, and professional licensure pass rates, as well as non-cognitive and behavioral outcomes such as discipline records, arrests, and crime. Second, we also need more evidence from both children at different ages and older adults. Tracking human capital trajectories over the life course would provide a more complete assessment of fortification's effects. At present, few public datasets jointly include comparable human capital measures across states, birth cohorts straddling fortification, and place of birth. Building linked data—connecting birth records to administrative education, earnings, and health files or to longitudinal surveys—would help fill this gap and enable stronger tests of persistence. Finally, extending the analysis to fortification implementation in low- and middle-income countries would offer a stronger test of the external validity of this paper, given that the baseline folate deficiency and program coverage in these countries are often substantially different from the U.S. ([McLean, de Benoist and Allen, 2008](#)).

References

- Abdel-Salam, Ghada, and Andrew E Czeizel.** 2000. "A case-control etiologic study of microcephaly." *American Journal of Medical Genetics*, 39(1): 1–10.
- Adhvaryu, Achyuta, Steven Bednar, Teresa Molina, Quynh Nguyen, and Anant Nyshadham.** 2020. "When it rains it pours: The long-run economic impacts of salt iodization in the United States." *Review of Economics and Statistics*, 102(2): 395–407.
- Ahluwalia, Namanjeet, Johanna Dwyer, Ana Terry, Alanna Moshfegh, and Clifford Johnson.** 2016. "Update on NHANES dietary data: focus on collection, release, analytical considerations, and uses to inform public policy." *Advances in Nutrition*, 7(1): 121–134.
- Almond, Douglas, and Bhashkar Mazumder.** 2011. "Health capital and the prenatal environment: the effect of Ramadan observance during pregnancy." *American Economic Journal: Applied Economics*, 3(4): 56–85.
- Almond, Douglas, Bhashkar Mazumder, and Reyn Van Ewijk.** 2015. "In utero Ramadan exposure and children's academic performance." *The Economic Journal*, 125(589): 1501–1533.
- Almond, Douglas, Janet Currie, and Valentina Duque.** 2018. "Childhood circumstances and adult outcomes: Act II." *Journal of Economic Literature*, 56(4): 1360–1446.
- Almond, Douglas, Lena Edlund, Hongbin Li, and Junsen Zhang.** 2007. "Long-term effects of the 1959–1961 China famine: Mainland China and Hong Kong."
- Ampaabeng, Samuel K, and Chih Ming Tan.** 2013. "The long-term cognitive consequences of early childhood malnutrition: the case of famine in Ghana." *Journal of Health Economics*, 32(6): 1013–1027.
- Anderson, Ellen, Betty Perloff, Jaspreet KC Ahuja, and Nancy Raper.** 2001. "Tracking nutrient changes for trends analysis in the United States." *Journal of Food Composition and Analysis*, 14(3): 287–294.
- Araújo, Daniel, Bladimir Carrillo, and Breno Sampaio.** 2021. "The long-run economic consequences of iodine supplementation." *Journal of Health Economics*, 79: 102490.
- Atwood, Alicia.** 2022. "The long-term effects of measles vaccination on earnings and employment." *American Economic Journal: Economic Policy*, 14(2): 34–60.
- Bailey, Martha J, Hilary Hoynes, Maya Rossin-Slater, and Reed Walker.** 2024. "Is the social safety net a long-term investment? Large-scale evidence from the food stamps program." *Review of Economic Studies*, 91(3): 1291–1330.
- Bentley, Tanya GK, Milton C Weinstein, Walter C Willett, and Karen M Kuntz.** 2009. "A cost-effectiveness analysis of folic acid fortification policy in the United States." *Public Health Nutrition*, 12(4): 455–467.

- Berry, RJ, J Mulinare, and HC Hamner.** 2010. "Folic acid fortification: neural tube defect risk reduction—a global perspective." *Folate in Health and Disease*, 2.
- Bleakley, Hoyt.** 2007. "Disease and development: evidence from hookworm eradication in the American South." *Quarterly Journal of Economics*, 122(1): 73–117.
- Bleakley, Hoyt.** 2010. "Malaria eradication in the Americas: A retrospective analysis of childhood exposure." *American Economic Journal: Applied Economics*, 2(2): 1–45.
- Blumenfeld, Zeev, Ephraim Siegler, and Moshe Bronshtein.** 1993. "The early diagnosis of neural tube defects." *Prenatal Diagnosis*, 13(9): 863–871.
- Bureau of Economic Analysis.** 1988. "Regional Economic Information System (REIS): Historical Data: Local Area Personal Income and Employment: Personal Current Transfer Receipts, 1988."
- Callaway, Brantly, Andrew Goodman-Bacon, and Pedro HC Sant'Anna.** 2024. "Difference-in-differences with a continuous treatment." National Bureau of Economic Research.
- Chen, Yuyu, and Li-An Zhou.** 2007. "The long-term health and economic consequences of the 1959–1961 famine in China." *Journal of Health Economics*, 26(4): 659–681.
- Committee on Use of Dietary Reference Intakes in Nutrition Labeling.** 2004. *Dietary reference intakes: guiding principles for nutrition labeling and fortification*. National Academies Press.
- Cragan, Janet D, Helen E Roberts, Larry D Edmonds, Muin J Khoury, Russell S Kirby, Gary M Shaw, Ellen M Velie, Ruth D Merz, Mathias B Forrester, Roger A Williamson, et al.** 1995. "Surveillance for anencephaly and spina bifida and the impact of prenatal diagnosis—United States, 1985–1994." *MMWR. CDC Surveillance Summaries: Morbidity and Mortality Weekly Report. CDC Surveillance Summaries*, 44(4): 1–13.
- Czeizel, Andrew E.** 2000. "Primary prevention of neural-tube defects and some other major congenital abnormalities: recommendations for the appropriate use of folic acid during pregnancy." *Paediatric Drugs*, 2: 437–449.
- Deng, Zichen, and Maarten Lindeboom.** 2022a. "A bit of salt, a trace of life: Gender norms and the impact of a salt iodization program on human capital formation of school aged children." *Journal of Health Economics*, 83: 102614.
- Deng, Zichen, and Maarten Lindeboom.** 2022b. "Early-life famine exposure, hunger recall, and later-life health." *Journal of Applied Econometrics*, 37(4): 771–787.
- East, Chloe N.** 2020. "The effect of food stamps on children's health evidence from immigrants' changing eligibility." *Journal of Human Resources*, 55(2): 387–427.
- Feyrer, James, Dimitra Politi, and David N Weil.** 2017. "The cognitive effects of micronutrient deficiency: Evidence from salt iodization in the United States." *Journal of the European Economic Association*, 15(2): 355–387.

- Field, Erica, Omar Robles, and Maximo Torero.** 2009. "Iodine deficiency and schooling attainment in Tanzania." *American Economic Journal: Applied Economics*, 1(4): 140–69.
- Finer, Lawrence B, and Mia R Zolna.** 2016. "Declines in unintended pregnancy in the United States, 2008–2011." *New England Journal of Medicine*, 374(9): 843–852.
- Fitzsimons, Emla, and Marcos Vera-Hernández.** 2022. "Breastfeeding and child development." *American Economic Journal: Applied Economics*, 14(3): 329–66.
- Food and Drug Administration.** 1996. "Food standards: Amendment of standards of identity for enriched grain products to require addition of folic acid; final rule (21 CFR Parts 136, 137, and 139)." *Federal Register*, 61: 8781–8797.
- Food and Drug Administration.** 2015. "Questions and answers on FDA's fortification policy: guidance for industry."
- Food and Nutrition Service.** 2005. "SNAP Data Tables." <https://fns-prod.azureedge.us/sites/default/files/resource-files/snap-annualsummary-11.pdf> (Accessed November 29, 2024).
- Ganong, Peter, and Jeffrey B Liebman.** 2018. "The decline, rebound, and further rise in SNAP enrollment: Disentangling business cycle fluctuations and policy changes." *American Economic Journal: Economic Policy*, 10(4): 153–176.
- Greene, Nicholas DE, and Andrew J Copp.** 2014. "Neural tube defects." *Annual Review of Neuroscience*, 37: 221–242.
- Greve, Jane, Marie Louise Schultz-Nielsen, and Erdal Tekin.** 2017. "Fetal malnutrition and academic success: Evidence from Muslim immigrants in Denmark." *Economics of Education Review*, 60: 20–35.
- Grosse, Scott D, Norman J Waitzman, Patrick S Romano, and Joseph Mulinare.** 2005. "Reevaluating the benefits of folic acid fortification in the United States: economic analysis, regulation, and public health." *American Journal of Public Health*, 95(11): 1917–1922.
- Hoekstra, Mark.** 2009. "The effect of attending the flagship state university on earnings: A discontinuity-based approach." *The Review of Economics and Statistics*, 91(4): 717–724.
- Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond.** 2016. "Long-run impacts of childhood access to the safety net." *American Economic Review*, 106(4): 903–934.
- Hoynes, Hilary, Marianne Page, and Ann Huff Stevens.** 2011. "Can targeted transfers improve birth outcomes?: Evidence from the introduction of the WIC program." *Journal of Public Economics*, 95(7-8): 813–827.
- Hoynes, Hilary W, and Erzo FP Luttmer.** 2011. "The insurance value of state tax-and-transfer programs." *Journal of Public Economics*, 95(11-12): 1466–1484.

- Huang, Qingyang, Chang Liu, and Li-An Zhou.** 2020. "Farewell to the God of Plague: Estimating the effects of China's Universal Salt Iodization on educational outcomes." *Journal of Comparative Economics*, 48(1): 20–36.
- Hutt, Peter Barton.** 1984. "Government regulation of the integrity of the food supply." *Annual Review of Nutrition*, 4(1): 1–21.
- Irvine, Nathalie, Gillian England-Mason, Catherine J Field, Deborah Dewey, and Fariba Aghajafari.** 2022. "Prenatal folate and choline levels and brain and cognitive development in children: a critical narrative review." *Nutrients*, 14(2): 364.
- Jacques, Paul F, Jacob Selhub, Andrew G Bostom, Peter WF Wilson, and Irwin H Rosenberg.** 1999. "The effect of folic acid fortification on plasma folate and total homocysteine concentrations." *New England Journal of Medicine*, 340(19): 1449–1454.
- Kancherla, Vijaya, Lorenzo D Botto, Laura A Rowe, Nathan A Shlobin, Adrian Caceres, Anastasia Arychnyna-Smith, Kathrin Zimmerman, Jeffrey Blount, Zewdie Kibruyisfaw, Kemel A Ghotme, et al.** 2022. "Preventing birth defects, saving lives, and promoting health equity: an urgent call to action for universal mandatory food fortification with folic acid." *The Lancet Global Health*, 10(7): e1053–e1057.
- Kuecken, Maria, Josselin Thuilliez, and Marie-Anne Valfort.** 2021. "Disease and human capital accumulation: Evidence from the Roll Back Malaria partnership in Africa." *The Economic Journal*, 131(637): 2171–2202.
- Lazuka, Volha.** 2020. "Infant Health and Later-Life Labor Market Outcomes Evidence from the Introduction of Sulfa Antibiotics in Sweden." *Journal of Human Resources*, 55(2): 660–698.
- Lindeboom, Maarten, France Portrait, and Gerard J Van den Berg.** 2010. "Long-run effects on longevity of a nutritional shock early in life: the Dutch Potato famine of 1846–1847." *Journal of Health Economics*, 29(5): 617–629.
- Liu, Jufen, Lei Jin, Zhiwen Li, Yali Zhang, Le Zhang, Linlin Wang, and Aiguo Ren.** 2018. "Prevalence and trend of isolated and complicated congenital hydrocephalus and preventive effect of folic acid in northern China, 2005–2015." *Metabolic Brain Disease*, 33: 837–842.
- Llanos, Adolfo, Eva Hertrampf, Fanny Cortes, Andrea Pardo, Scott D Grosse, and Ricardo Uauy.** 2007. "Cost-effectiveness of a folic acid fortification program in Chile." *Health Policy*, 83(2-3): 295–303.
- Majid, Muhammad Farhan.** 2015. "The persistent effects of in utero nutrition shocks over the life cycle: Evidence from Ramadan fasting." *Journal of Development Economics*, 117: 48–57.
- McLean, Erin, Bruno de Benoist, and Lindsay H Allen.** 2008. "Review of the magnitude of folate and vitamin B12 deficiencies worldwide." *Food and Nutrition Bulletin*, 29(2_suppl1): S38–S51.

McNulty, Helene, and Kristina Pentieva. 2004. "Folate bioavailability." *Proceedings of the Nutrition Society*, 63(4): 529–536.

Meng, Xin, and Nancy Qian. 2006. "The long run health and economic consequences of famine on survivors: Evidence from China's Great Famine."

Meng, Xin, and Nancy Qian. 2009. "The long term consequences of famine on survivors: evidence from a unique natural experiment using China's great famine." National Bureau of Economic Research.

Mersereau, P, K Kilker, H Carter, E Fassett, J Williams, A Flores, C Prue, L Williams, C Mai, and J Mulinare. 2004. "Spina Bifida and Anencephaly Before and After Folic Acid Mandate—United States, 1995–1996 and 1999–2000." *MMWR: Morbidity & Mortality Weekly Report*, 53(17).

National Center for Health Statistics. 2003. "All-county natality files for 1989-2003." as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program.

Naz, Naila, Alicia Requena Jimenez, Anna Sanjuan-Vilaplana, Megan Gurney, and Jaleel Miyan. 2016. "Neonatal hydrocephalus is a result of a block in folate handling and metabolism involving 10-formyltetrahydrofolate dehydrogenase." *Journal of Neurochemistry*, 138(4): 610–623.

Nguyen, Vy Kim, Lauren YM Middleton, Lei Huang, Neil Zhao, Eliseu Verly Jr, Jacob Kvasnicka, Luke Sagers, Chirag J Patel, Justin Colacino, and Olivier Jolliet. 2023. "Harmonized US National Health and Nutrition Examination Survey 1988-2018 for high throughput exposome-health discovery." *MedRxiv*.

Niemesh, Gregory T. 2015. "Ironing out deficiencies: evidence from the united states on the economic effects of iron deficiency." *Journal of Human Resources*, 50(4): 910–958.

Obeid, Rima, Wolfgang Holzgreve, and Klaus Pietrzik. 2013. "Is 5-methyltetrahydrofolate an alternative to folic acid for the prevention of neural tube defects?" *Journal of Perinatal Medicine*, 41(5): 469–483.

Petrini, JR, K Damus, and RB Johnston. 1999. "Knowledge and Use of Folic Acid by Women of Childbearing Age—United States, 1995 and 1998." *MMWR: Morbidity & Mortality Weekly Report*, 48(16): 325–327.

Pfeiffer, Christine M, Jeffery P Hughes, Ramon A Durazo-Arvizu, David A Lacher, Christopher T Sempos, Mindy Zhang, Elizabeth A Yetley, and Clifford L Johnson. 2012. "Changes in measurement procedure from a radioassay to a microbiologic assay necessitate adjustment of serum and RBC folate concentrations in the US population from the NHANES 1988–2010." *The Journal of Nutrition*, 142(5): 894–900.

- Portrait, France RM, TF Van Wingerden, and DJH Deeg.** 2017. "Early life undernutrition and adult height: the Dutch famine of 1944–45." *Economics & Human Biology*, 27: 339–348.
- Quinlivan, EP, J McPartlin, H McNulty, M Ward, JJ Strain, DG Weir, and JM Scott.** 2002. "Importance of both folic acid and vitamin B12 in reduction of risk of vascular disease." *The Lancet*, 359(9302): 227–228.
- Ray, Joel G, Gita Singh, and Robert F Burrows.** 2004. "Evidence for suboptimal use of peri-conceptional folic acid supplements globally." *BJOG: An International Journal of Obstetrics & Gynaecology*, 111(5): 399–408.
- Rossin-Slater, Maya.** 2013. "WIC in your neighborhood: New evidence on the impacts of geographic access to clinics." *Journal of Public Economics*, 102: 51–69.
- Roth, Christine, Per Magnus, Synnve Schjølberg, Camilla Stoltenberg, Pål Surén, Ian W McKeague, George Davey Smith, Ted Reichborn-Kjennerud, and Ezra Susser.** 2011. "Folic acid supplements in pregnancy and severe language delay in children." *JAMA*, 306(14): 1566–1573.
- Ruggles, Steven, Sarah Flood Matthew Sobek Daniel Backman Grace Cooper Julia A. Rivera Drew Stephanie Richards Renae Rodgers Jonathan Schroeder, and Kari C.W. Williams.** 2025. "Ipums usa: Version 16.0 [dataset]." (No Title).
- Scholte, Robert S, Gerard J Van Den Berg, and Maarten Lindeboom.** 2015. "Long-run effects of gestation during the Dutch Hunger Winter famine on labor market and hospitalization outcomes." *Journal of Health Economics*, 39: 17–30.
- Serena, Benjamin Ly.** 2019. "Cognitive consequences of iodine deficiency in adolescence: evidence from salt iodization in Denmark." *The Scandinavian Journal of Economics*.
- Smithells, RW, MJ Seller, R Harris, DW Fielding, CJ Schorah, NC Nevin, S Sheppard, AP Read, S Walker, and J Wild.** 1983. "Further experience of vitamin supplementation for prevention of neural tube defect recurrences." *The Lancet*, 321(8332): 1027–1031.
- Tafesse, Wiktoria.** 2022. "The effect of Universal Salt Iodization on cognitive test scores in rural India." *World Development*, 152: 105796.
- Toivonen, KI, E Lacroix, M Flynn, PE Ronksley, KA Oinonen, A Metcalfe, and TS Campbell.** 2018. "Folic acid supplementation during the preconception period: a systematic review and meta-analysis." *Preventive Medicine*, 114: 1–17.
- US Bureau of Labor Statistics.** 1989-2002. "Quarterly Census of Employment and Wages: 1989-2002."
- US Census Bureau.** 1990. "Intercensal Population Estimates: 1980-1990."
- US Census Bureau.** 2009. "County and City Data Book [United States], 1988."

US Office of Management and Budget. 2003. "Circular A-4." https://obamawhitehouse.archives.gov/omb/circulars_a004_a-4#2 (accessed on 20 August 2025).

Wald, NJ, MR Law, JK Morris, and DS Wald. 2001. "Quantifying the effect of folic acid." *The Lancet*, 358(9298): 2069–2073.

Wright, Kevin. 2003. "Folic Acid and the American Food Supply: A historical account of the FDA's creation of the current folic acid regulations."

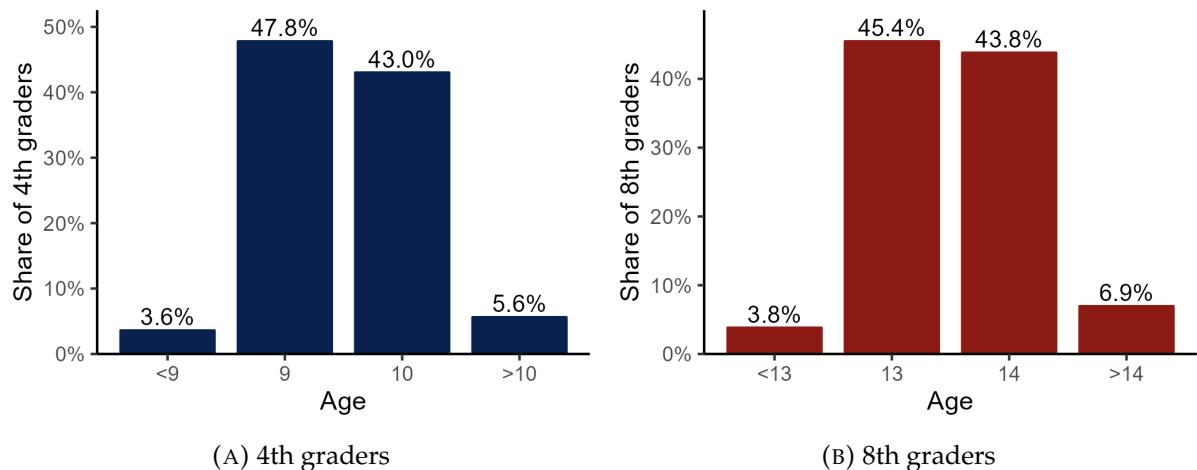
Yi, Yunni, Marion Lindemann, Antje Colligs, and Claire Snowball. 2011. "Economic burden of neural tube defects and impact of prevention with folic acid: a literature review." *European Journal of Pediatrics*, 170(11): 1391–1400.

Zimmerman, Seth D. 2014. "The returns to college admission for academically marginal students." *Journal of Labor Economics*, 32(4): 711–754.

Appendix

A Figures and tables

FIGURE A1: AGE DISTRIBUTION OF 4TH- AND 8TH-GRADERS



Notes: The data is from the ACS from the same survey year as the public NAEP data.

TABLE A1: BALANCE TEST

	$\mathbf{1}\{\text{CNSA top quartile}\}_s$		
	State-level	County Level	
	(1)	(2))	(3)
<i>Demographic features</i>			
Share of black (%), 1988	-0.0023 (0.0153)	-0.0007 (0.0009)	-0.0007 (0.0036)
Share of female (%), 1988	0.0145 (0.2684)	0.0341*** (0.0084)	0.0341 (0.0237)
Share of under 5 (%), 1988	0.2604 (0.2354)	0.0812*** (0.0124)	0.0812 (0.0660)
Share of over 65 (%), 1988	-0.0061 (0.1058)	-0.0281*** (0.0057)	-0.0281 (0.0317)
Birth rate (%), 1988	-0.0223 (0.0585)	-0.0129*** (0.0028)	-0.0129 (0.0187)
Death rate (%), 1988	0.1621 (0.2671)	0.0566*** (0.0114)	0.0566* (0.0301)
Log population, 1988	0.0930 (0.1198)	-0.0620*** (0.0076)	-0.0620* (0.0340)
<i>Economic conditions</i>			
Transfer income p.p. (million \$), 1988	-0.3067 (0.5091)	0.0288 (0.0321)	0.0288 (0.2770)
Income p.p. (million \$), 1985	107.9 (110.4)	36.61*** (5.999)	36.61 (22.93)
Federal funds p.p. (million \$), 1986	-234.2* (135.3)	-7.814* (4.400)	-7.814 (9.496)
Unemployment rate (%), 1986	-0.0555 (0.0662)	-0.0088** (0.0037)	-0.0088 (0.0202)
<i>Agriculture</i>			
Value of produces sold per farm (million \$), 1987	-4.560 (3.019)	-0.5773*** (0.1024)	-0.5773 (0.4570)
Average farm size (million acres), 1987	152.6 (181.6)	-3.731 (9.547)	-3.731 (15.16)
State FEs			Y
Observations	49	3,000	3,000
R ²	0.3209	0.0849	0.0849
Dep. var. mean	0.2941	0.2933	0.2933

Notes: Regressions are weighted by population of 1988. Data on share of black, share of female, share of under 5, share of the over 65, and population are from County Intercensal Estimates; data on birth rate, death rate, value of produces sold per farm, and average farm size are from County Databook 1988; data on transfers is from Bureau of Economic Analysis, Regional Economic Information System (REIS); unemployment data is from Bureau of Labor Statistics. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

TABLE A2: LONG RUN EFFECTS OF FOLIC ACID FORTIFICATION ON FULL DISTRIBUTION OF GRADE ATTENDING

	Respondent was attending					
	Any school (1)	Grade 6 and above (2)	Grade 7 and above (3)	Grade 12 and above (4)	College and above (5)	Graduate school (6)
Panel A: All young adults						
1{CNSA top quartile} × Exposed cohorts	0.0188* (0.0108)	0.0188* (0.0108)	0.0187* (0.0108)	0.0187* (0.0108)	0.0189* (0.0112)	0.0004 (0.0017)
Observations	1,976,638	1,976,638	1,976,638	1,976,638	1,976,638	1,976,638
R ²	0.1590	0.1590	0.1589	0.1570	0.1444	0.0183
Dep. var. mean	0.3504	0.3504	0.3500	0.3477	0.3369	0.0381
Est./Dep. var. mean	5.36%	5.36%	5.33%	5.37%	5.61%	1.08%
State-of-birth FE	✓	✓	✓	✓	✓	✓
Quarter-year-of-birth FE	✓	✓	✓	✓	✓	✓
Survey year FE	✓	✓	✓	✓	✓	✓
Individual char controls	✓	✓	✓	✓	✓	✓
Confounder controls	✓	✓	✓	✓	✓	✓

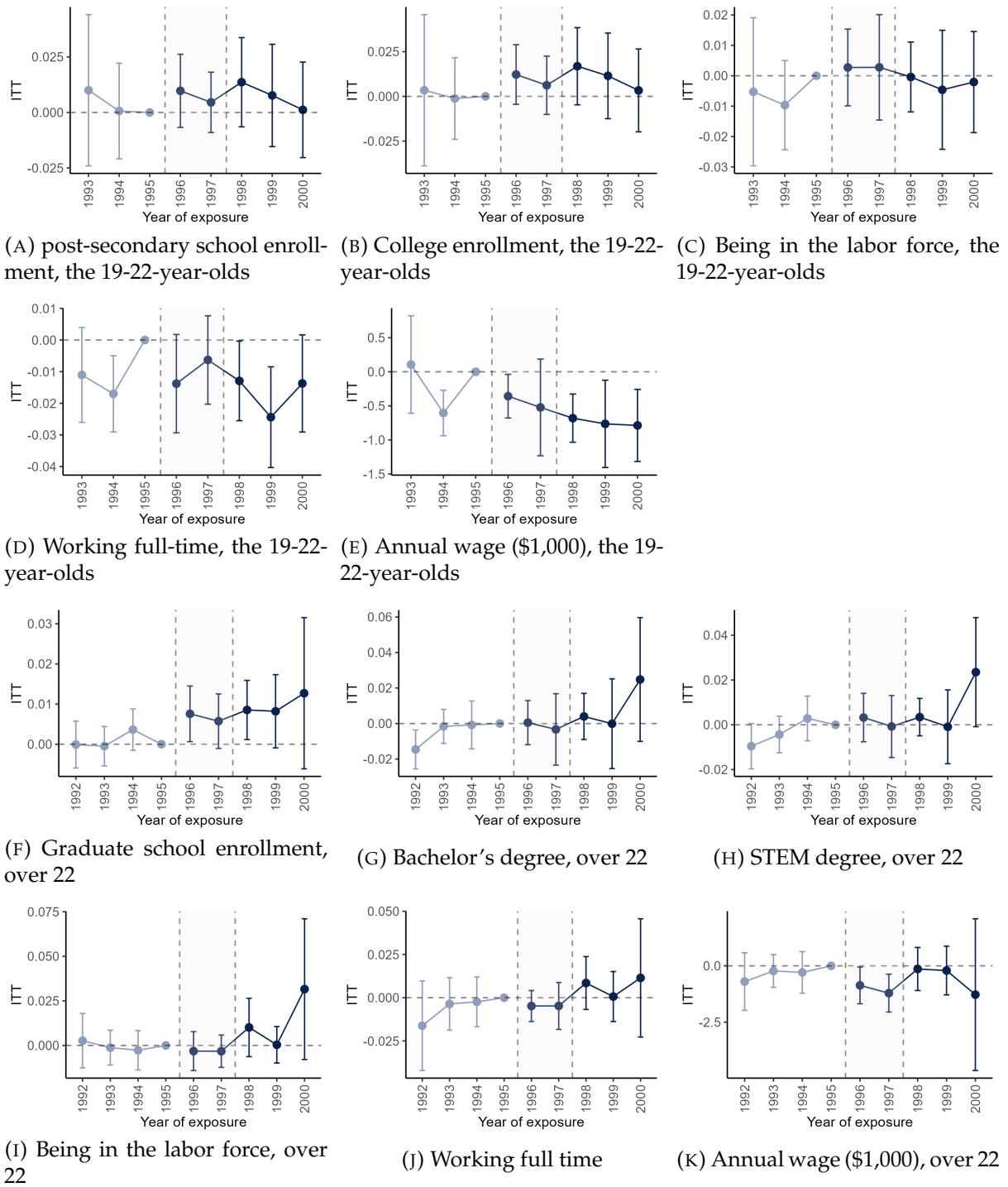
Notes: The table presents standard DD estimates and their standard errors. The treated group consists of births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top quartile of baseline CNS anomaly rates. I control for state-of-birth fixed effects; quarter-by-year-of-birth and survey-year fixed effects; and controls for gender, race, Hispanic origin, Medicaid eligibility, exposure to mental-health parity laws and welfare reforms, and the local unemployment rate. The unit of observation is the individual. Regressions and dependent-variable means are weighted by the IPUMS person weight; percentile calculations are weighted by the number of births. Standard errors are clustered by state of birth. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

TABLE A3: LONG RUN EFFECTS OF FOLIC ACID FORTIFICATION ON FULL DISTRIBUTION OF SCHOOL COMPLETION

	Respondent completed more than					
	Any school					
		Kindergarten	Grade 6	Grade 9	Grade 12	High school diploma or GED
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All young adults						
$\mathbf{1}\{\text{CNSA top quartile}\} \times \text{Exposed cohorts}$	-0.0002 (0.0003)	-0.0003 (0.0003)	-0.0004 (0.0004)	-0.0007 (0.0006)	-0.0019 (0.0014)	-0.0024 (0.0020)
Observations	1,976,638	1,976,638	1,976,638	1,976,638	1,976,638	1,976,638
R ²	0.0014	0.0014	0.0014	0.0021	0.0107	0.0124
Dep. var. mean	0.9938	0.9936	0.9926	0.9872	0.9511	0.9332
Est./Dep. var. mean	-0.02%	-0.03%	-0.04%	-0.07%	-0.20%	-0.25%
	Some college	College for at least one year	Associate degree	Bachelor's degree	Graduate degree	
	(7)	(8)	(9)	(10)	(11)	
Panel B: Young adults over 22						
$\mathbf{1}\{\text{CNSA top quartile}\} \times \text{Exposed cohorts}$	0.0047 (0.0036)	0.0051 (0.0034)	0.0092*** (0.0032)	0.0078** (0.0032)	-0.0010 (0.0026)	
Observations	982,727	982,727	982,727	982,727	982,727	
R ²	0.0433	0.0485	0.0761	0.0787	0.0361	
Dep. var. mean	0.6694	0.6038	0.4364	0.3507	0.0593	
Est./Dep. var. mean	0.71%	0.85%	2.10%	2.23%	-1.65%	
State-of-birth FE	✓	✓	✓	✓	✓	✓
Quarter-year-of-birth FE	✓	✓	✓	✓	✓	✓
Survey year FE	✓	✓	✓	✓	✓	✓
Individual char controls	✓	✓	✓	✓	✓	✓
Confounder controls	✓	✓	✓	✓	✓	✓

Notes: The table presents standard DD estimates and their standard errors. The treated group consists of births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top quartile of baseline CNS anomaly rates. I control for state-of-birth fixed effects; quarter-by-year-of-birth and survey-year fixed effects; and controls for gender, race, Hispanic origin, Medicaid eligibility, exposure to mental-health parity laws and welfare reforms, and the local unemployment rate. The unit of observation is the individual. Regressions and dependent-variable means are weighted by the IPUMS person weight; percentile calculations are weighted by the number of births. Standard errors are clustered by state of birth. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

FIGURE A2: DOSE-RESPONSE RESULTS



Notes: The figure plots event-study estimates with 95% confidence intervals. The shaded region denotes cohorts with partial exposure. The unit of observation is the individual. Regressions and dependent-variable means are weighted by the IPUMS person weight; percentile calculations are weighted by the number of births. Standard errors are clustered by state of birth.

TABLE A4: IMPACTS OF EXPOSURE TO FOLIC ACID FORTIFICATION ON BIRTH OUTCOMES

	Birth weight (grams)	Low birth weight (1 if birth weight < 2500)	Gestation weeks	Preterm (1 if gestation weeks < 37)
	(1)	(2)	(3)	(4)
1{CNSA top quartile} × Exposed cohorts	-0.3562 (2.956)	0.0002 (0.0006)	0.0129 (0.0179)	0.0013 (0.0014)
Observations	84,017	84,017	84,017	84,017
R ²	0.8532	0.7881	0.7817	0.7270
Dep. var. mean	3320	0.0748	38.8453	0.1141
Est./Dep. var. mean	-0.01%	0.24%	0.03%	1.12%
State-of-birth FE	✓	✓	✓	✓
Quarter-year-of-birth FE	✓	✓	✓	✓
Individual char controls	✓	✓	✓	✓

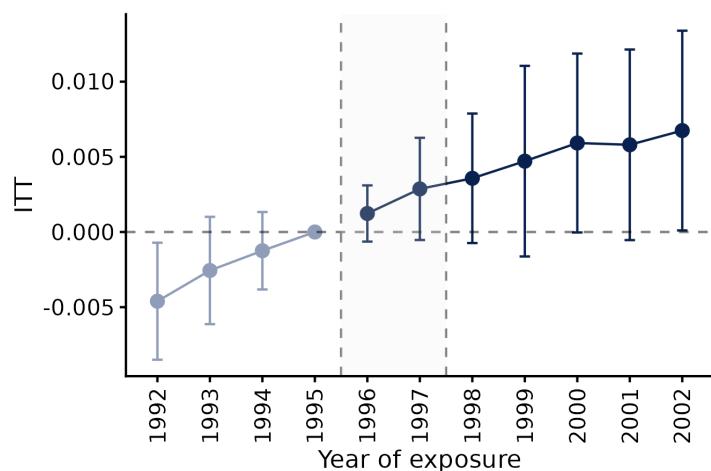
Notes: The table presents standard DD estimates and their standard errors. I aggregate natality records to county-by-quarter-of-birth cells and merge a state-level exposure measure to each cell. The treated group consists of births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top quartile of baseline CNS anomaly rates. I control for county-of-birth fixed effects, quarter-by-year-of-birth fixed effects, and time-varying covariates capturing Medicaid eligibility, exposure to mental-health parity laws and welfare reforms, and the local unemployment rate. Regressions and dependent-variable means are weighted by the number of birth in each cell. Standard errors are clustered by state of birth. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

TABLE A5: IMPACTS OF EXPOSURE TO FOLIC ACID FORTIFICATION ON COMPOSITION OF MATERNAL CHARACTERISTICS

	Black	Age ≤ 22	23 ≤ Age ≤ 29	Education < college	Unmarried	Inadequate prenatal care
	(1)	(2)	(3)	(4)	(5)	(6)
1{CNSA top quartile} × Exposed cohorts	-0.0006 (0.0032)	0.0057** (0.0024)	-0.0047 (0.0055)	0.0021 (0.0063)	0.0037 (0.0096)	0.0146 (0.0126)
Observations	84,017	84,017	84,017	83,905	84,017	84,017
R ²	0.9871	0.9308	0.6128	0.9477	0.9122	0.7810
Dep. var. mean	0.1508	0.2697	0.3745	0.5486	0.3262	0.2379
Est./Dep. var. mean	-0.43%	2.10%	-1.26%	0.38%	1.13%	6.12%
State-of-birth FE	✓	✓	✓	✓	✓	✓
Quarter-year-of-birth FE	✓	✓	✓	✓	✓	✓
Individual char controls	✓	✓	✓	✓	✓	✓
Confounder controls	✓	✓	✓	✓	✓	✓

Notes: The table presents standard DD estimates and their standard errors. I aggregate natality records to county-by-quarter-of-birth cells and merge a state-level exposure measure to each cell. The treated group consists of births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top quartile of baseline CNS anomaly rates. I control for county-of-birth fixed effects, quarter-by-year-of-birth fixed effects, and time-varying covariates capturing Medicaid eligibility, exposure to mental-health parity laws and welfare reforms, and the local unemployment rate. Regressions and dependent-variable means are weighted by the number of birth in each cell. Standard errors are clustered by state of birth. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

FIGURE A3: DYNAMIC EFFECTS OF EXPOSURE TO FOLIC ACID FORTIFICATION ON THE SHARE OF MOTHERS ≤ 22



Notes: The table presents standard DD estimates and their standard errors. I aggregate natality records to county-by-quarter-of-birth cells and merge a state-level exposure measure to each cell. The treated group consists of births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top quartile of baseline CNS anomaly rates. I control for county-of-birth fixed effects, quarter-by-year-of-birth fixed effects, and time-varying covariates capturing Medicaid eligibility, exposure to mental-health parity laws and welfare reforms, and the local unemployment rate. Regressions and dependent-variable means are weighted by the number of birth in each cell. Standard errors are clustered by state of birth. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The shaded region denotes cohorts with partial exposure.

B Robustness checks and falsification tests: young adults' outcomes

TABLE B1: LONG-RUN EFFECTS OF FOLIC ACID FORTIFICATION ON TEST SCORES OF 4TH AND 8TH GRADERS, ASSIGNING SMALLER AGES

	Math			Reading		
	Average score (1)	% \geq basic (2)	% \geq proficient (3)	Average score (4)	% \geq basic (5)	% \geq proficient
$P(\text{CNSA top quartile}) \times \text{exposed cohorts}$	2.22*** (0.81)	1.61* (0.82)	2.62*** (0.82)	-1.32 (1.07)	0.05 (1.38)	-1.35** (0.65)
Observations	459	459	459	459	459	459
R ²	0.99	0.92	0.91	0.99	0.93	0.92
Dep. var. mean	257.72	76.15	34.82	238.93	68.96	31.35
Est./Dep. var. mean	0.86%	2.11%	7.52%	-0.55%	0.07%	-4.31%
State-of-school FEes	✓	✓	✓	✓	✓	✓
Year-of-birth FEes	✓	✓	✓	✓	✓	✓
Grade FEes	✓	✓	✓	✓	✓	✓
Confounder controls	✓	✓	✓	✓	✓	✓

Notes: The table presents cohort DD estimates and their standard errors. $P(\text{CNSA top quartile})$ is the fortification exposure adjusted for migration. It varies across state of school, survey year, and grade. The exposed cohorts are the one with assigned year of birth > 1996 . The unit of observation is state-by-year-by-grade cells. Regressions and dependent-variable means are weighted by the number of students by state-of-school, survey year, and grade from ACS. Standard errors are clustered by state of school. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

TABLE B2: LONG-RUN EFFECTS OF FOLIC ACID FORTIFICATION ON TEST SCORES OF 4TH AND 8TH GRADERS, USING STATE-OF-BIRTH EXPOSURE AS A ROUGH PROXY FOR STATE-OF-SCHOOL EXPOSURE

	Math			Reading		
	Average score (1)	% \geq basic (2)	% \geq proficient (3)	Average score (4)	% \geq basic (5)	% \geq proficient
$P(\text{CNSA top quartile}) \times \text{exposed cohorts}$	1.81*** (0.66)	1.31** (0.64)	1.99*** (0.65)	-0.88 (0.85)	0.17 (1.08)	-0.85 (0.54)
Observations	450	450	450	450	450	450
R ²	0.99	0.92	0.91	0.99	0.93	0.92
Dep. var. mean	257.80	76.22	34.91	239.00	69.03	31.42
Est./Dep. var. mean	0.70%	1.72%	5.71%	-0.37%	0.24%	-2.71%
State-of-school FEes	✓	✓	✓	✓	✓	✓
Year-of-birth FEes	✓	✓	✓	✓	✓	✓
Grade FEes	✓	✓	✓	✓	✓	✓
Confounder controls	✓	✓	✓	✓	✓	✓

Notes: The table presents cohort DD estimates and their standard errors. The exposed cohorts are the one with assigned year of birth > 1996 . The unit of observation is state-by-year-by-grade cells. Regressions and dependent-variable means are weighted by the number of students by state-of-school, survey year, and grade from ACS. Standard errors are clustered by state of school. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

TABLE B3: LONG-RUN EFFECTS OF FOLIC ACID FORTIFICATION ON TEST SCORES OF 4TH- AND 8TH-GRADERS, NOT CONTROLLING FOR CONFOUNDERS

	Math			Reading		
	Average score (1)	% \geq Basic (2)	% \geq Proficient (3)	Average score (4)	% \geq Basic (5)	% \geq Proficient (0.79)
$P(\text{CNSA top quartile}) \times \text{exposed cohorts}$	2.00** (0.90)	1.39 (0.92)	2.43** (1.00)	-1.53 (1.16)	-0.21 (1.42)	-1.53* (0.79)
Observations	459	459	459	459	459	459
R ²	0.99	0.92	0.90	0.99	0.93	0.92
Dep. var. mean	257.72	76.15	34.82	238.93	68.96	31.35
Est./Dep. var. mean	0.78%	1.82%	6.98%	-0.64%	-0.30%	-4.89%
State-of-school FEs	✓	✓	✓	✓	✓	✓
Year-of-birth FEs	✓	✓	✓	✓	✓	✓
Grade FEs	✓	✓	✓	✓	✓	✓
Confounder controls						

Notes: The table presents cohort DD estimates and their standard errors. $P(\text{CNSA top quartile})$ is the fortification exposure adjusted for migration. It varies across state of school, survey year, and grade. The exposed cohorts are the one with assigned year of birth > 1996 . The unit of observation is state-by-year-by-grade cells. Regressions and dependent-variable means are weighted by the number of students by state-of-school, survey year, and grade from ACS. Standard errors are clustered by state of school. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

TABLE B4: LONG-RUN EFFECTS OF FOLIC ACID FORTIFICATION ON TEST SCORES OF 4TH AND 8TH GRADERS, CONTROLLING FOR STATE-OF-SCHOOL-BY-YEAR-OF-BIRTH AND GRADE FIXED EFFECTS

	Math			Reading		
	Average score (1)	% \geq basic (2)	% \geq proficient (3)	Average score (4)	% \geq basic (5)	% \geq proficient (0.61)
$P(\text{CNSA top quartile}) \times \text{exposed cohorts}$	4.53*** (1.12)	3.69*** (1.10)	4.66*** (0.89)	-0.82 (1.10)	1.15 (1.48)	-1.00 (0.61)
Observations	459	459	459	459	459	459
R ²	1.00	0.95	0.95	1.00	0.96	0.95
Dep. var. mean	257.72	76.15	34.82	238.93	68.96	31.35
Est./Dep. var. mean	1.76%	4.85%	13.38%	-0.34%	1.66%	-3.20%
State-of-school-by-year-of-birth FEs	✓	✓	✓	✓	✓	✓
Grade FEs	✓	✓	✓	✓	✓	✓
Confounder controls	✓	✓	✓	✓	✓	✓

Notes: The table presents cohort DD estimates and their standard errors. The exposed cohorts are the one with assigned year of birth > 1996 . The unit of observation is state-by-year-by-grade cells. Regressions and dependent-variable means are weighted by the number of students by state-of-school, survey year, and grade from ACS. Standard errors are clustered by state of school. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

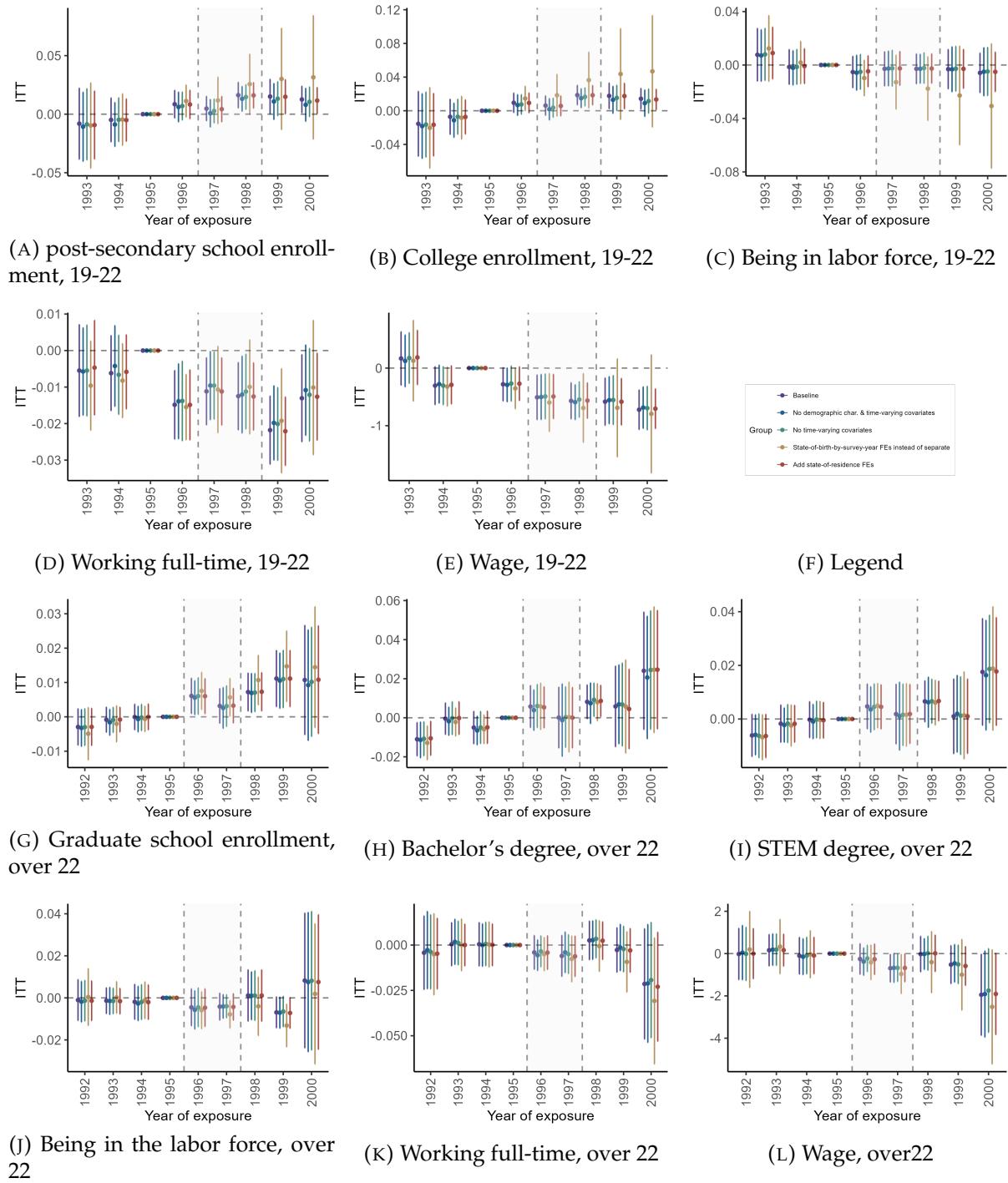
TABLE B5: LONG-RUN EFFECTS OF FOLIC ACID FORTIFICATION ON TEST SCORES OF 4TH AND 8TH GRADERS, CONTROLLING FOR STATE-OF-SCHOOL-BY-GRADE AND SURVEY YEAR FIXED EFFECTS

	Math			Reading		
	Average score (1)	% \geq basic (2)	% \geq proficient (3)	Average score (4)	% \geq basic (5)	% \geq proficient (6)
P(CNSA top quartile) \times exposed cohorts	3.72* (2.00)	3.04* (1.52)	4.98** (2.07)	2.72 (2.18)	3.27* (1.81)	3.08** (1.53)
Observations	459	459	459	459	459	459
R ²	0.98	0.90	0.87	0.98	0.92	0.86
Dep. var. mean	257.72	76.15	34.82	238.93	68.96	31.35
Est./Dep. var. mean	1.44%	4.00%	14.29%	1.14%	4.74%	9.83%
State-of-school-by-grade FE	✓	✓	✓	✓	✓	✓
Survey year FE	✓	✓	✓	✓	✓	✓
Confounder controls	✓	✓	✓	✓	✓	✓

Notes: The table presents cohort DD estimates and their standard errors. The exposed cohorts are the one with assigned year of birth > 1996 . The unit of observation is state-by-year-by-grade cells. Regressions and dependent-variable means are weighted by the number of students by state-of-school, survey year, and grade from ACS. Standard errors are clustered by state of school. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

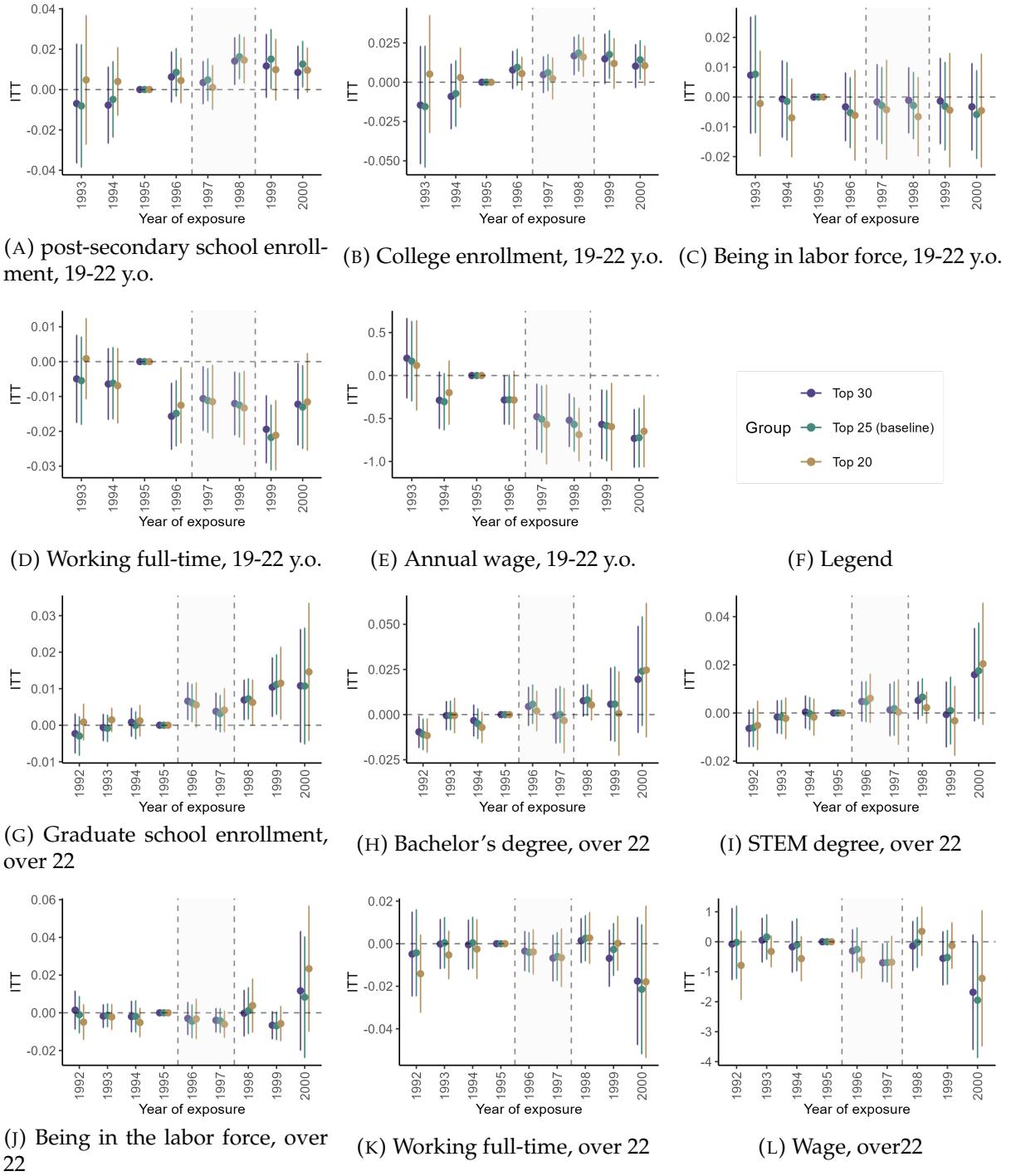
C Robustness checks and falsification tests: young adults' outcomes

FIGURE C1: ROBUSTNESS TO ALTERNATIVE MODEL SPECIFICATIONS



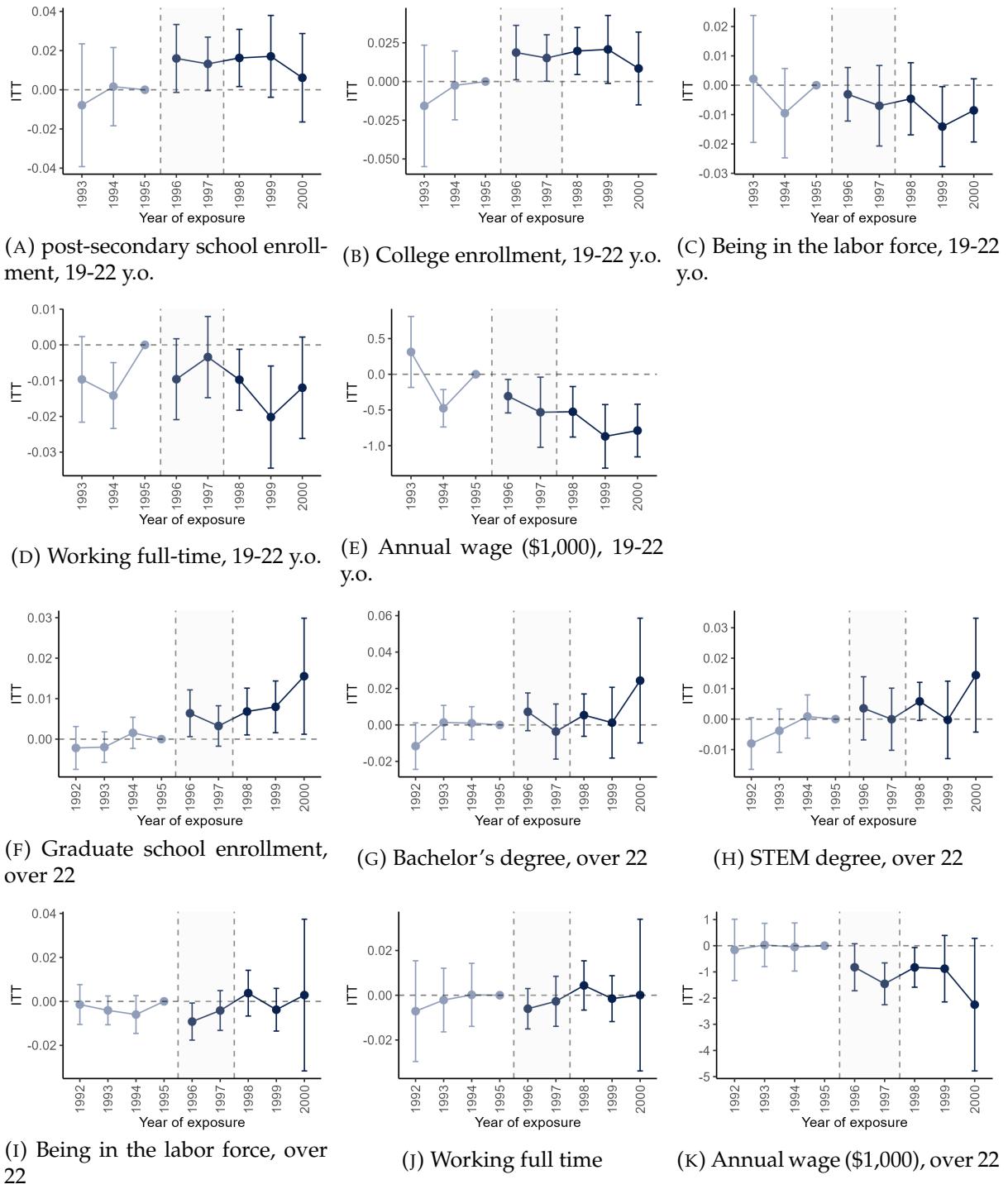
Notes: The figure plots event-study estimates with 95% confidence intervals. The treated group consists of births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top quartile of baseline CNS anomaly rates. The shaded region denotes cohorts with partial exposure. The unit of observation is the individual. Regressions and dependent-variable means are weighted by the IPUMS person weight; percentile calculations are weighted by the number of births. Standard errors are clustered by state of birth.

FIGURE C2: ROBUSTNESS TO ALTERNATIVE EXPOSURE THRESHOLDS



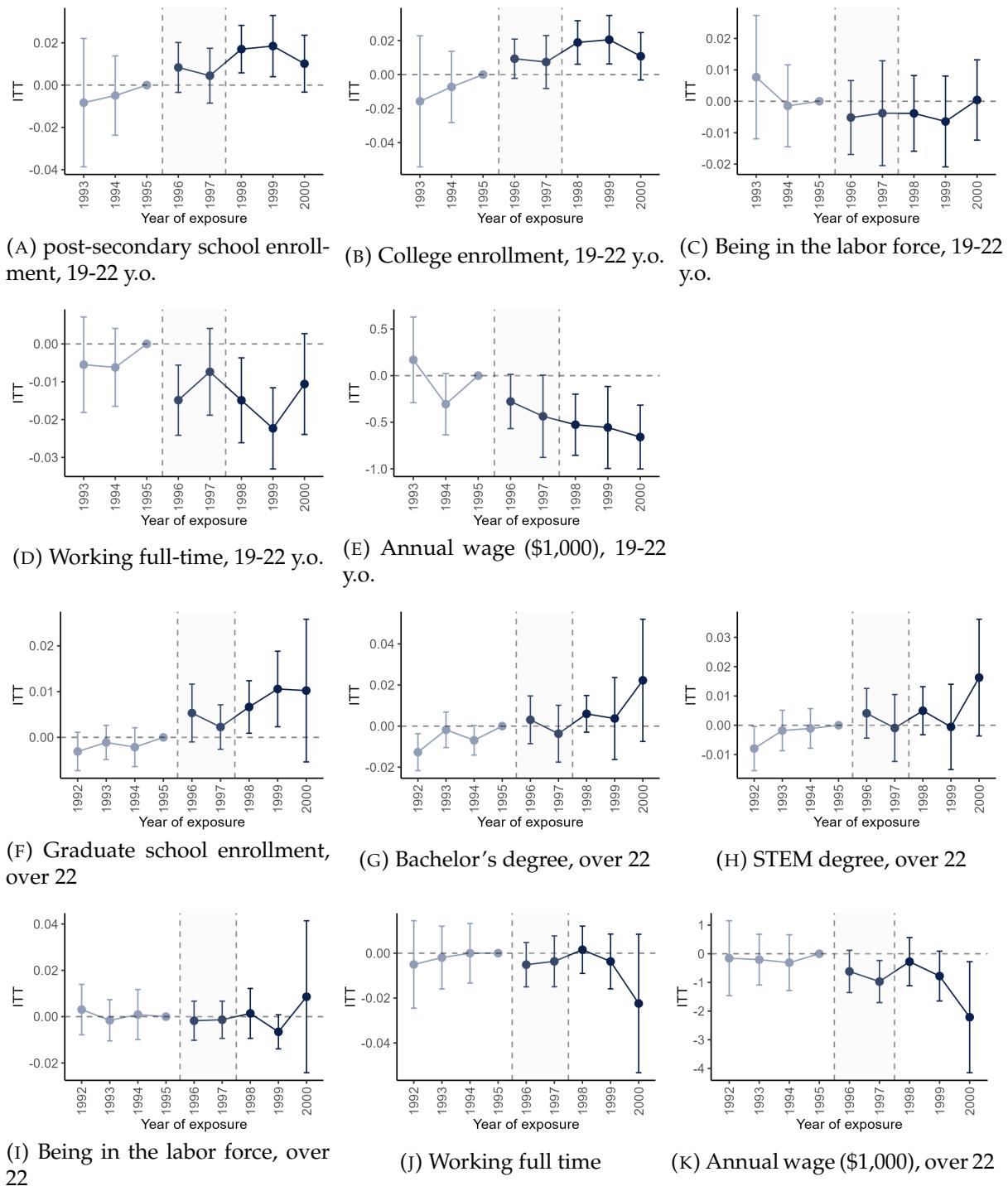
Notes: The figure plots event-study estimates with 95% confidence intervals. The treated group consists of births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top 30%, top quartile (25%), or top 20% of baseline CNS anomaly rates. The shaded region denotes cohorts with partial exposure. The unit of observation is the individual. Regressions and dependent-variable means are weighted by the IPUMS person weight; percentile calculations are weighted by the number of births. Standard errors are clustered by state of birth.

FIGURE C3: STATES WITH Q4 VS. Q1 EXPOSURE



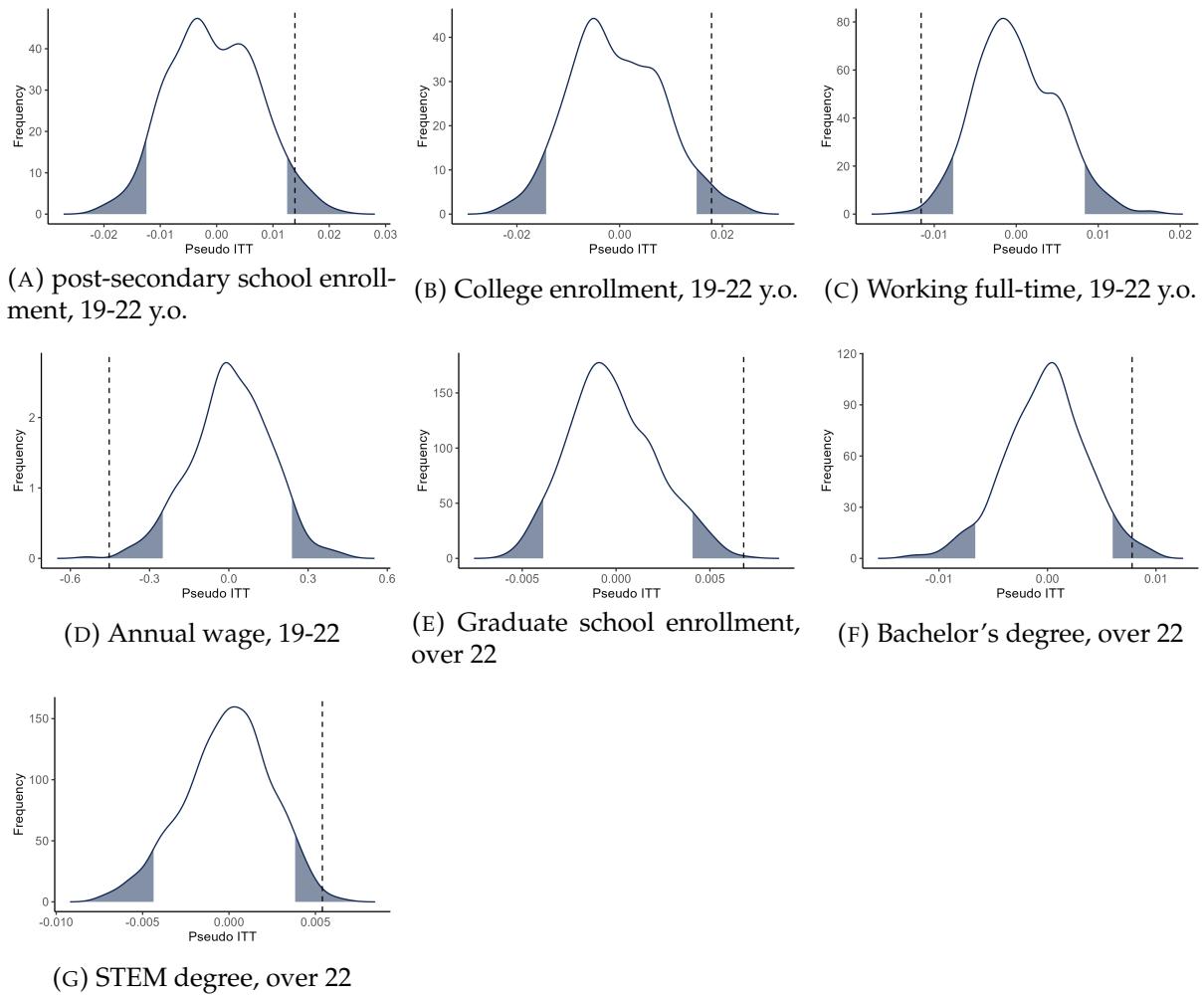
Notes: The figure plots event-study estimates with 95% confidence intervals. The treated group consists of births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top quartile of baseline CNS anomaly rates. The shaded region denotes cohorts with partial exposure. The unit of observation is the individual. Regressions and dependent-variable means are weighted by the IPUMS person weight; percentile calculations are weighted by the number of births. Standard errors are clustered by state of birth.

FIGURE C4: DROP ACS IPUMS 2020



Notes: The figure plots event-study estimates with 95% confidence intervals. The treated group consists of births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top quartile of baseline CNS anomaly rates. The shaded region denotes cohorts with partial exposure. The unit of observation is the individual. Regressions and dependent-variable means are weighted by the IPUMS person weight; percentile calculations are weighted by the number of births. Standard errors are clustered by state of birth.

FIGURE C5: RANDOMIZATION TEST



Notes: The figure plots estimates of overall effects. The treated group consists of births whose first trimester ended after the March 1996 authorization of folic acid fortification and who were born in states in the top quartile of baseline CNS anomaly rates. The unit of observation is the individual. Regressions are weighted by the IPUMS person weight; percentile calculations are weighted by the number of births. Shaded areas represent ≤ 5 th and ≥ 95 th percentiles of our simulated null distribution.