Long-Run Human Capital Effects of Fortifying Grain Products with Folic Acid

Wenjie Zhan*

Abstract

This paper examines the long-run human capital effects of folic acid fortification. I exploit cross-state variation in baseline birth defect rates linked to folate deficiency and the timing of the first trimester relative to fortification. In utero exposure shifts young adults' time toward schooling: graduate school enrollment rises by 0.8 percentage points, accompanied by a 0.7-percentage-point decline in labor force participation. These findings are most consistent with improved cognitive ability rather than fertility selection or labor market conditions. Back-of-the-envelope calculations suggest the resulting human capital gains could generate tens of millions of dollars in additional tax revenue. (JEL I18, J22, J24, Q18)

1   Introduction

Early-life nutrition plays an important 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 it is well documented that early-life nutrition has long-run economic effects, most studies focus on overall nutrition intake or macronutrients (Almond, Currie and Duque, 2018). Evidence on the long-run effects of micronutrient deficiencies 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 late-1990s U.S. fortification of enriched grain products with folic acid.

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 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, on March 5, 1996, the U.S. Food and Drug Administration (FDA) mandated that enriched grain products, including wheat flour, bread, rice, macaroni, and noodles, be fortified with 40µg/100g of folic acid, the synthetic form of folate. Manufacturers were required to comply by January 1, 1998. While public health research has documented the immediate health benefits of folic acid fortification, its long-run effects on human capital remain underexplored.

I begin by documenting several first-stage facts following the fortification mandate: (i) measured folate content increased across a wide range of foods, (ii) dietary folate intake and serum folate concentrations rose, and (iii) the prevalence of folate-sensitive congenital anomalies declined, with larger reductions in states with higher baseline rates.

I then leverage cross-state variation in pre-fortification birth defect rates linked to folate deficiency, calculated from birth certificate data, and match these rates to young adults' outcomes in the American Community Survey (ACS) Public Use Microdata Sample (PUMS) by state and year of birth. I validate the baseline birth defect rate as an exposure measure by showing that it is a plausible pre-period marker of folate deficiency (Sections 4.1 and 5.1). I define exposed cohorts based on whether the first trimester occurred after fortification, since this is the period when folic acid affects folate-sensitive birth defects. I use the mandate authorization date

(March 5, 1996) rather than the full implementation date (January 1, 1998) as the event time because prior work shows that fortification was largely completed before the implementation deadline (Jacques et al., 1999), and the first-stage evidence in Section 5.1 is consistent with this timing. Using a cohort difference-in-differences design, I compare exposed cohorts born in states with the highest baseline prevalence of these defects to unexposed cohorts and to cohorts born in other states to identify the mandate's effects.

I find that in utero exposure to fortification raises graduate school enrollment among young adults aged 23–30 by 0.8 percentage points (14% at the sample mean), accompanied by a 0.7-percentage-point decline in labor force participation (0.84% at the mean). Graduate school enrollment increases across demographic groups, while the decline in labor force participation is strongest among white and non-Hispanic individuals. Women also adjust labor supply on the intensive margin, with exposed women less likely to work full time. Effects for Hispanic individuals are smaller and less precisely estimated, consistent with the fact that the 1996 mandate did not cover some staples commonly consumed in Hispanic households.

I discuss three potential mechanisms behind this shift toward schooling. First, fertility selection is unlikely to explain the results because birth outcomes and maternal characteristics do not respond to in utero fortification exposure, aside from a small increase in the share of births to younger mothers after fortification. Second, I find no evidence that changes in local labor market conditions are a driver: the effects are not diminished by place-of-residence-by-year fixed effects, and in utero fortification exposure does not affect the likelihood of unemployment or occupational quality among those employed. Finally, consistent with improved cognitive ability as a plausible mechanism, I find suggestive evidence that fortification exposure is associated with higher math scores among 4th and 8th graders.

Converting the graduate school enrollment estimates into years of schooling, I find that the education gain from folic acid fortification is comparable to that from salt iodization for women and is substantially smaller than that from the Food Stamp Program, a much larger intervention. My back-of-the-envelope calculation suggests that the resulting increase in tax revenue from additional schooling—through longer-term income gains—is about $15.5 million per year, far exceeding the program's reported annual cost of $3 million.

This paper contributes to the economics literature on the human capital effects of food fortification. Existing work focuses mainly on the long-run benefits of salt iodization for cog-

nitive development, health, and socioeconomic outcomes (Feyrer, Politi and Weil, 2017; Serena, 2019; Adhvaryu et al., 2020; Huang, Liu and Zhou, 2020; Deng and Lindeboom, 2022*a*; Tafesse, 2022). One exception is Niemesh (2015), which finds that iron fortification of bread increased working-age adults' incomes and raised their children's school enrollment and long-run wages. In contrast, evidence on the human capital effects of folic acid fortification remains scarce. Folate deficiency can cause irreversible damage to early neural development, potentially resulting in lower cognitive ability later in life (Irvine et al., 2022). A natural question is whether early-life folate deficiency affects downstream educational and labor outcomes, both of which depend on cognitive skills. Folic acid fortification has also been adopted far less widely than salt iodization or iron supplementation, particularly in developing countries, so evidence from the United States can inform policy design elsewhere.

This paper also contributes to the public health literature on folic acid fortification. Prior work has emphasized the 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 long-run human capital outcomes, this study extends that work and provides new empirical evidence on the broader developmental consequences of folic acid fortification.

More broadly, this paper 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 (Almond et al., 2007; Chen and Zhou, 2007; Lindeboom, Portrait and Van den Berg, 2010; Scholte, Van Den Berg and Lindeboom, 2015; Deng and Lindeboom, 2022*b*) and Ramadan fasting (Almond and Mazumder, 2011; Almond, Mazumder and Van Ewijk, 2015; Majid, 2015), have been linked to poorer health and labor 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 policy background. Section 3 de-

4

scribes the data. Section 4 outlines the empirical methods. Section 5 presents the results. Section 6 discusses mechanisms. Section 7 interprets the magnitudes of the estimates and presents a back-of-the-envelope calculation of program benefits. 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 central nervous system anomalies (CNSAs) 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, such as spina bifida, can survive into adulthood but often face a high risk of lifelong physical and cognitive disabilities (Yi et al., 2011).[2] In the early 1990s, roughly 4,000 fetuses in the U.S. (about 1 in 1,000) were affected by NTDs each year, with about one third lost to selective or spontaneous abortion (Cragan et al., 1995; Mersereau et al., 2004). Folate deficiency is also associated with other congenital CNSAs, such as hydrocephaly (Naz et al., 2016; Liu et al., 2018). These defects can arise as early as the first month of pregnancy, when the neural tube begins to form, and failure of the neural tube to close 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 provide some palliative benefits, it cannot reverse the underlying neurological damage (Greene and Copp, 2014). Timely medical intervention is often difficult because routine screening for neural tube defects typically relies on the second-trimester anatomic ultrasound at 18 to 22 weeks, and detection rates are much lower in the first trimester (American College of Obstetricians and Gynecologists, 2017). Many pregnant women in the U.S. also lack adequate prenatal care (Blumenfeld, Siegler and Bronshtein, 1993).

2.2   Sources of Folate

Folate occurs naturally in foods such as beef liver, dark green leafy vegetables, beans, peas, nuts, and many fruits and fruit juices. However, meeting recommended intake during pregnancy through diet alone is difficult (Czeizel, 2000). Data from the National Health and Nutri-

---

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

[2]In spina bifida, the backbone does not close properly, leaving a section of the spinal cord and spinal nerves exposed without full protection.

tion Examination Survey (NHANES) III (1988–1994) show that women aged 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 often falls short is that naturally occurring food folate is unstable under typical cooking conditions, which can substantially reduce the amount ultimately absorbed, making diet alone a less reliable way to improve folate status during pregnancy (McNulty and Pentieva, 2004).

Folate is also available through supplements, including over-the-counter folic acid tablets and multivitamins. Folic acid is often recommended during prenatal visits (Ray, Singh and Burrows, 2004). A key challenge, however, is low awareness of and adherence to supplementation guidance (Toivonen et al., 2018). CDC recommendations emphasize starting folic acid at least one month before conception.[3] 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 can also be barriers, particularly for low-income women (Czeizel, 2000). These constraints highlight the need for a low-cost, preferably passive approach to ensuring 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, began in the 1990s. The first wave of grain fortification in the 1940s followed the identification of specific nutrient deficiency disorders among the U.S. population. 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 was 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).

---

[3]See https://www.cdc.gov/ncbddd/folicacid/recommendations.html (accessed on 05/20/2022).

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 μg/100 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 the event time as March 1996, when the FDA authorized folic acid fortification. Because enriched grain products are ingredients in many processed foods, the added folic acid shows up in products beyond the enriched grain category; for example, some chips contain folic acid (Figure 1). Before the mandate, voluntary addition of folic acid was prohibited in standardized foods[4] and discouraged in other products, primarily to avoid overfortification and nutrient imbalances (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 (birth certificate data) cover all U.S. live births and report detailed information on birth outcomes, including birth month and year, state of birth, birth weight, gestational age in weeks, and congenital anomalies, as well as maternal characteristics such as

---

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

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 using the pre-fortification prevalence of folate-sensitive congenital anomalies and assign cohort-level in utero exposure using gestational age and the policy's authorization date (Section 4.1). Second, I test whether fortification exposure affects birth outcomes or shifts the distribution of maternal characteristics, which helps assess whether fertility selection could be driving the main results.

3.2 Young Adult Outcome Data

Outcome data for young adults come from the American Community Survey Public-Use Microdata Sample (ACS PUMS), 2019–2024 (Ruggles and Williams., 2025). For educational outcomes, I study graduate school enrollment (including master's, professional, and doctoral programs) and educational attainment, including completion of high school (diploma or GED), bachelor's degrees, and STEM degrees. For labor market outcomes, I examine labor force participation, employment, full-time status, and income. These outcomes are standard measures of human capital and are commonly examined in related work (Niemesh, 2015; Adhvaryu et al., 2020).

I use the 2019–2024 PUMSs so that each survey year includes both exposed and unexposed cohorts. I focus on young adults because the oldest exposed cohorts are only 28 years old during this period. The main analysis sample includes individuals aged 23–30, which captures early-career outcomes for young adults beyond typical college ages. I discuss alternative sample definitions in Section 5.3.

3.3 Other Data

I examine how the exposure measure relates to baseline state characteristics in Section 4.1. I compile these characteristics from multiple sources. Demographic measures, including 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). Measures of transfer payments come from the Bureau of Economic Analysis's Regional Economic Information System (REIS) (Bureau of Economic Analysis, 1988).

In Section 5.1, I examine trends in dietary folate intake and blood folate concentrations

using publicly available data from NHANES III (1988–1994) and the continuous NHANES (1999–2006). NHANES III is a nationally representative survey of the civilian, noninstitutionalized U.S. population aged 2 months and older. Beginning in 1999, NHANES shifted to a continuous design, collecting data annually and releasing them in two-year cycles. Both NHANES III and the continuous NHANES combine an in-home interview with a standardized physical examination in mobile examination centers, along with laboratory testing of biospecimens. The surveys provide detailed information on demographics, socioeconomic characteristics, health conditions, and diet, as well as anthropometrics and biomarkers. For folate status, NHANES reports dietary folate intake based on 24-hour dietary recalls and measures serum and red blood cell (RBC) folate in the laboratory component. I restrict the continuous NHANES sample to 1999–2006 because folate laboratory methods changed after 2006, limiting comparability across periods (Yetley et al., 2011).

In Section 5.3, as a robustness check, I control for a Bartik-style predicted unemployment measure to proxy for economic conditions at birth. Baseline sectoral employment shares come from the Bureau of Labor Statistics' Quarterly Census of Employment and Wages (QCEW) (US Bureau of Labor Statistics, 1989-2002). Following East (2020), for each state of birth I interact these baseline employment shares with annual national changes in sector-specific unemployment rates and then sum across sectors to construct a predicted state unemployment rate.

In Section 5.4, I examine the association between fortification exposure and academic performance in elementary and middle school. Test score outcomes come from the National Assessment of Educational Progress (NAEP) public state-level aggregate data for 4th and 8th graders. The public NAEP reports state average scores and the shares of students at or above the NAEP achievement levels (Basic and Proficient)[5] in mathematics, reading, science, and writing, typically at intervals of one or two years since 1990. State average scores are computed by averaging scaled scores from a representative sample of students in each state, with scores ranging from 0 to 500. I focus on 4th and 8th grade math and reading because public series for science, writing, and all 12th grade tests are available for far fewer years. The analysis 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,

---

[5]See https://nces.ed.gov/nationsreportcard/mathematics/achieve.aspx for details on the NAEP achievement levels.

2011, and 2015; and 8th grade reading in 2002, 2003, 2005, 2007, 2009, 2011, and 2015.

4   Methods

To study long-run responses to in utero exposure to fortification, I exploit two sources of varia-
tion: (i) cross-state differences in the pre-fortification prevalence of birth defects linked to folate
deficiency, and (ii) the timing of the first trimester relative to fortification. Because folate de-
ficiency during the first trimester can cause irreversible damage, any benefits from improved
maternal folate status may emerge later in life. This approach is valid if baseline pre fortification
defect prevalence captures residents' potential to benefit from fortification, and if fortification
actually increased folate intake rather than remaining a policy change on paper. Section 4.1 dis-
cusses the validity of using birth defects to capture cross-state variation in exposure, and Section
5.1 presents evidence that fortification increased folate intake and reduced birth defects. Com-
pared to a randomized trial that assigns folic acid supplements to mothers and follows children
into adulthood, this design is more feasible at scale and avoids ethical concerns.

This design, which uses baseline regional disease prevalence to infer the benefits of a health
intervention, is common in health economics. For example, researchers use 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

Using the birth certificate data, I calculate each state's pre-fortification prevalence of central
nervous system anomalies (CNSA), including spina bifida, anencephaly, hydrocephaly, mi-
crocephaly, and "other" CNSAs. 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 reported under "other" CNSAs. Folate deficiency may also contribute
to hydrocephaly and microcephaly, either directly or indirectly through NTDs (Abdel-Salam
and Czeizel, 2000; Naz et al., 2016; Liu et al., 2018).

I define the baseline period as January 1989 through June 1993. Most states began reporting

10

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 coverage, including all states and the District of Columbia except New Mexico. I end the baseline in the first half of 1993 so that cohorts born afterward have at least four pre-periods available for the event-study analysis. Figure 2a shows baseline CNSA rates across states. I define high-exposure states as those in the top quartile of baseline CNSA rates, yielding 14 states: IN, IA, KS, MD, MN, NE, NJ, NY, ND, RI, SD, TN, TX, and VT. I test robustness to alternative high-exposure thresholds and to using the continuous CNSA rate as the exposure measure in Section 5.3.

To validate the exposure measure, I link state-level pre-fortification CNSA rates to biomarkers of folate status from NHANES III. Table 1 shows a negative association: on average, individuals living in high-baseline-CNSA states have lower blood folate concentrations. In Section 5.1, I also show that high-baseline-CNSA states experienced larger post-fortification declines in CNSA rates (Figure 4d), which further supports the validity of the exposure measure.

Table 1: Correlation between baseline CNSA rate and folate biomarkers

	Serum folate (1)	RBC folate (2)	Serum folate (3)	RBC folate (4)
Continuous CNSA rate	-0.538** (0.261)	-11.410** (4.572)		
High CNSA			-0.606*** (0.185)	-10.554*** (3.171)
Observations	10,842	10,913	10,842	10,913
R^2	0.001	0.001	0.002	0.003
Dep. var. mean	7.332	198.186	7.332	198.186

Notes: Blood folate is measured in μg/mL. Robust standard errors are in parentheses. Regressions are weighted by MEC examination weights. Source: NHANES III public-use data covering 35 counties across 13 states. CNSA rate is measured at state level. ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

Another concern is that baseline CNSA rates are not randomly assigned and may simply capture other state characteristics, such as differences in demographics, economic conditions, or local policy environments. For example, if states with higher baseline CNSA rates also tend to be poorer, and early-life economic conditions independently shape adult outcomes, the estimates could reflect underlying economic hardship rather than fortification. To address this concern, Table A1 regresses the high-baseline-CNSA indicator on pre-1989 local characteristics, including demographic, economic, and agricultural measures at both the state and county

levels. Column (1) shows that individual state characteristics have imprecise coefficients, with the full set of state-level observables explaining about 30% of cross-state variation. Column (2) shows that when all regressors are defined at the county level, these same observables explain much less—only about 5%. Adding state fixed effects raises the R^2 to roughly 30% (Column (3)). These results suggest that high-baseline-CNSA status does not simply proxy for a single observable factor or a combination of local characteristics.

I determine exposure timing using gestational weeks recorded on birth certificates and then map exposure to quarter of birth, since PUMS does not report gestational length. Specifically, I classify an infant as exposed if the first trimester ends after March 1996, because neural tube closure occurs by the end of the first trimester and folic acid reduces CNSAs by supporting proper closure. I then aggregate this birth-level exposure indicator to the quarter-by-year level to match the PUMS data. As shown in Figure 2b, the share exposed rises sharply starting with births in 1996 Q4. I therefore define individuals born in 1996 Q4 or later as the exposed cohorts. This pre–post timing, combined with cross-state variation in baseline CNSA rates, provides the key identifying variation in my empirical strategy.

(a) Baseline CNSA rates by state

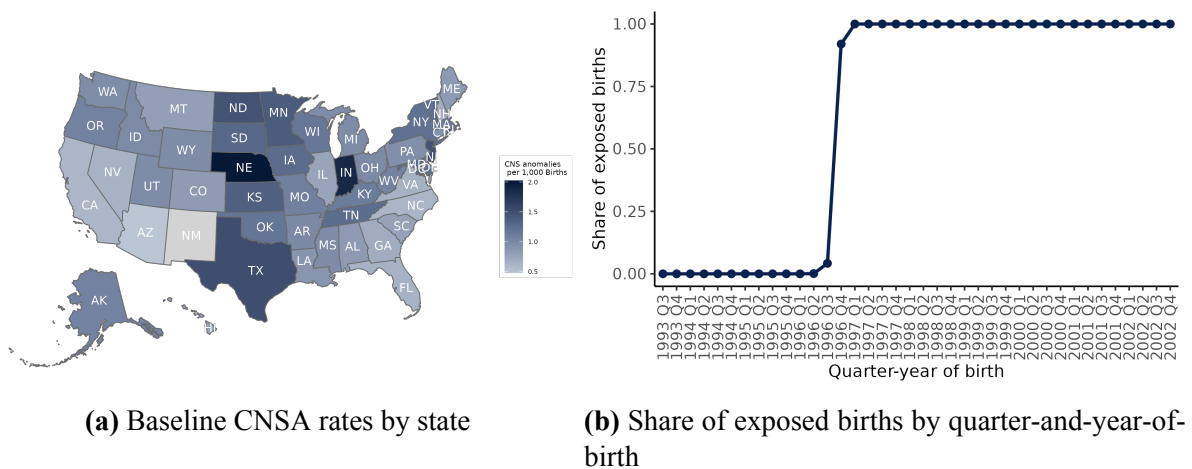(b) Share of exposed births by quarter-and-year-of-birth

Figure 2: Spatial and temporal variation in fortification exposure

Notes: Figure 2a maps baseline CNSA rates (January 1989–June 1993) by state of birth. Figure 2b defines a birth as exposed if the first trimester ends after March 1996, when fortification was authorized, and then aggregates this exposure indicator to county-by-quarter-by-year averages.

4.2   The Estimation Method

I estimate the following cohort difference-in-differences (DiD) model to identify the long-run effects of in utero exposure to fortification:

$$Y_{istc} = \beta\big(\text{High CNSA}_s \times \text{Exposed cohorts}_t\big) + \gamma_{sc} + \delta_t + X'_{istc}\theta + C'_{sc}\pi + \varepsilon_{istc}, \quad (1)$$

where Y_{istc} denotes the outcome for individual i born in state s and birth cohort t (quarter-by-year), observed in ACS survey year c. High CNSA$_s$ is an indicator equal to 1 for states in the top quartile of the baseline CNSA rate and 0 otherwise. Exposed cohorts$_t$ is an indicator equal to 1 for cohorts born in 1996 Q4 or later and 0 otherwise. γ_{sc} are state-of-birth-by-survey-year fixed effects, which absorb time-varying state factors common to all individuals born in the same state and observed in the same survey year. δ_t are quarter-by-year-of-birth fixed effects that control for cohort-specific shocks. X_{istc} are individual controls (gender, race indicators for White, Black, and other, and an indicator for Hispanic). C_{sc} includes controls for potentially confounding policies measured at the state-by-cohort level, including Medicaid and CHIP eligibility expansions, welfare reform, and state mental health parity laws. The error term is ε_{istc}. The coefficient β is the primary parameter of interest and captures differential changes in outcomes for exposed cohorts born in high-baseline-CNSA states relative to other cohorts and states.

My preferred specification compares high- versus high-baseline-CNSA states rather than estimating a dose–response relationship using continuous CNSA rates, because the latter is harder to interpret and requires stronger identifying assumptions (Callaway, Goodman-Bacon and Sant'Anna, 2024). The trade-off is that a binary design discards some variation and may be sensitive to the choice of threshold. To address these concerns, Section 5.3 tests robustness to alternative exposure thresholds and to dose–response specifications.

Because the threshold at which fortification would have zero effect is unknown, low-baseline-CNSA states may also have benefited. As a result, the cohort-DiD design does not compare treated and untreated groups. Instead, it identifies a relative effect: the impact in top-quartile states relative to the average impact in all other states. If fortification improved outcomes outside the top quartile, this contrast understates the effect in high-baseline states and does not recover the nationwide average effect. The estimates should therefore be interpreted as a lower

13

bound for high-baseline states.

The validity of this empirical strategy relies on two assumptions: parallel trends and no anticipation. The parallel trends assumption requires that, absent fortification, average outcomes for young adults in high- and low-baseline-CNSA states would have evolved similarly across birth cohorts. To assess pre-treatment trends, I estimate dynamic effects for key outcomes using an event-study design. Specifically, I replace $\beta\big(\text{High CNSA}s \times \text{Exposed cohorts}t\big)$ with $\sum_\tau \beta\tau,, \tau \neq 1995\big(\text{High CNSA}_s \times \mathbf{1}\text{Cohort} = \tau\big)$, where τ indexes exposure year and I construct exposure years by grouping cohorts into four-quarter bins after fortification. For example, exposure year 1996 includes cohorts born in 1996Q4–1997Q3. Section 5.3 further assesses the credibility of these event-study results by examining sensitivity to hypothesized violations of the parallel trends assumption following Rambachan and Roth (2023).

The no-anticipation assumption requires that (i) prospective mothers did not change behavior in anticipation of fortification, and (ii) food manufacturers did not begin fortifying before March 1996. In this setting, anticipatory responses are unlikely. Petrini, Damus and Johnston (1999) documents low awareness of folic acid among women of childbearing age, and because the 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 over-fortification and nutrient imbalances (Food and Drug Administration, 1996, 2015). Consistent with these arguments, the event-study results show no evidence of anticipatory behavior.

5 Results

5.1 The First Stage

This section presents first-stage evidence on the effects of fortification. Descriptive evidence suggests that, after fortification, folic acid content in foods increased, dietary folate intake and blood folate concentration rose, and the CNSA rate declined.

Folate content in foods increased after folic acid fortification. First, folic acid content in foods increased sharply after fortification. Using the 1996 and 1998 Continuing Survey of Food Intakes by Individuals (CSFII), Lewis et al. (1999) compare per-serving folic acid in the same foods before and after the mandate, based on USDA 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, they 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 and cookie bars; in total, more than 350 basic food items show higher folic acid due to fortification (Anderson et al., 2001).
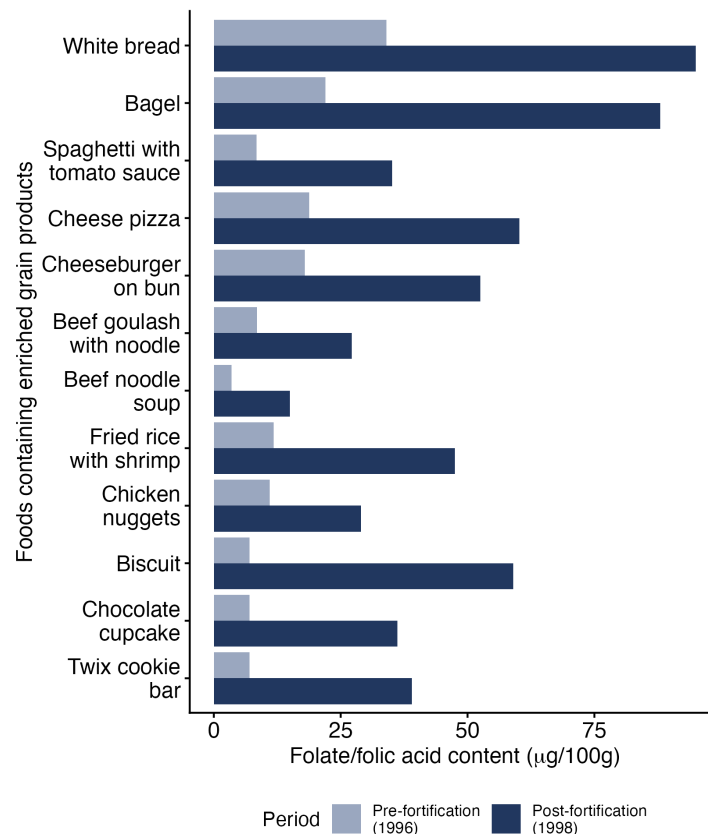

Figure 3: Changes in folate contents in selected foods attributable to fortification

Notes: This figure uses estimates from Lewis et al. (1999). Food folate content data come from the USDA Continuing Survey of Food Intakes by Individuals (CSFII) 1996 and 1998. Folate content is estimated by the USDA using recipe-based calculations. The changes shown reflect only the portion attributable to fortification.

Dietary folate intake and blood folate increased after folic acid fortification. Second, using NHANES, I document sharp post-fortification increases in dietary folate intake and blood folate among women aged 19–45. Mean dietary intake rose by nearly 50% (Figure 4a). The share with intake below 400 μg/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 similar gains: serum folate more than doubled and red blood cell (RBC) folate rose by nearly 50% (Figure 4b), indicating sustained improvements

in folate status.[6]

Congenital anomalies declined after folic acid fortification. Finally, Figure 4c shows
that CNSA rates declined as folate intake and absorption increased. After a relatively flat pe-
riod from 1992 to 1996, incidence fell substantially following fortification. That rates of other
congenital anomalies were stable over the same period and did not begin to decline until after
1997 suggests this pattern is unlikely to reflect broad improvements in health. Figure 4d shows
larger post-fortification declines in states with higher baseline CNSA rates, supporting the use
of the high-baseline-CNSA indicator as the exposure measure in my research design.

(a) Dietary folate

(b) Blood folate

(c) CNSA and non-CNSA rates

(d) CNSA rates in high- and low-baseline-CNSA
states

Figure 4: Trends in dietary folate, blood folate concentration, and CNSA rate

Notes: Data in Figures 4a and 4b come from the harmonized NHANES files constructed by Nguyen et al. (2023),
which ensure comparability of folate measures across survey periods. All folate measures use Mobile Examination
Center (MEC) final examination sample weights. Data in Figures 4c and 4d come from the Vital Statistics Natality
files (birth certificate data).

---

[6]Laboratory methods for blood folate were stable from 1999 to 2006 but changed after 2006 (Yetley et al.,
2011; Pfeiffer et al., 2012). Because public-use NHANES does not report state identifiers, I cannot separately plot
trends for high- and low-baseline-CNSA states.

5.2 Main Results

Table 2 reports the effects of in utero exposure to folic acid fortification on educational and labor outcomes for young adults aged 23–30. The results suggest that exposed young adults allocate more time to schooling rather than work.

Columns (1)–(4) show that graduate school enrollment increases by 0.8 percentage points (14% at the mean), while completed educational attainment changes little, including high school diploma or GED completion, bachelor's degree attainment, and holding a STEM degree. This suggests that in utero exposure primarily affects individuals who would have completed high school and earned a bachelor's degree regardless, but who otherwise would not have pursued graduate education. These results are consistent with Niemesh (2015), who find null effects of iron fortification of bread on high school completion and college attendance among adults aged 22–50.

Consistent with higher graduate school enrollment, Column (5) shows a 0.7-percentage-point decline in labor force participation (0.84% at the mean). The magnitudes of the enrollment and labor supply effects are similar. The likelihood of being employed, full-time work (more than 30 hours per week), and income also decline, but these estimates are less precise (Columns (6)–(8)).

Table 2: Long-run effects of folic acid fortification on young adults' outcomes

	Educational outcomes				Labor outcomes			
	Grad school enrollment	HS diploma or GED	Bachelor's degree	STEM degree	Labor force participation	Employed	Hours worked per week ≥ 30	Income ($100,000)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High CNSA×Exposed cohorts	0.008***	-0.001	0.004	0.002	-0.007**	-0.007*	-0.009	-0.012*
	(0.003)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)	(0.006)	(0.006)
Observations	1,065,355	1,065,355	1,065,355	1,065,355	1,065,355	1,065,355	1,065,355	1,065,355
R^2	0.011	0.010	0.066	0.020	0.011	0.019	0.025	0.087
Dep. var. mean	0.059	0.943	0.369	0.142	0.843	0.791	0.741	0.362
Est./Dep. var. mean	14.441%	-0.157%	1.116%	1.720%	-0.841%	-0.909%	-1.146%	-3.436%
State-of-birth×Survey-year FEs	✓	✓	✓	✓	✓	✓	✓	✓
Quarter-year-of-birth FEs	✓	✓	✓	✓	✓	✓	✓	✓
Control variables	✓	✓	✓	✓	✓	✓	✓	✓

Notes: This table presents cohort-DiD estimates with standard errors clustered at the state of birth. High CNSA is an indicator for states in the 4th quartile of baseline CNSA rates. Exposed cohorts are those whose first trimester ended after the March 1996 authorization of folic acid fortification. Controls include state-of-birth-by-survey-year and quarter-by-year-of-birth fixed effects, along with gender, race, ethnicity, Medicaid eligibility, exposure to mental-health parity laws, and welfare reforms. Regressions and dependent-variable means are weighted by IPUMS person weights. ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

Figures 5a and 5b plot event-study estimates for graduate school enrollment and labor force participation, the two outcomes that provide the clearest evidence on the impacts of folic acid fortification. Neither outcome shows noticeable differential pre-trends prior to fortification. Event-study estimates for the remaining outcomes appear in Figures A1a–A1f and are broadly consistent with the DD estimates in Table 2. Section 5.3 further examines the sensitivity of these results to potential violations of the parallel trends assumption.
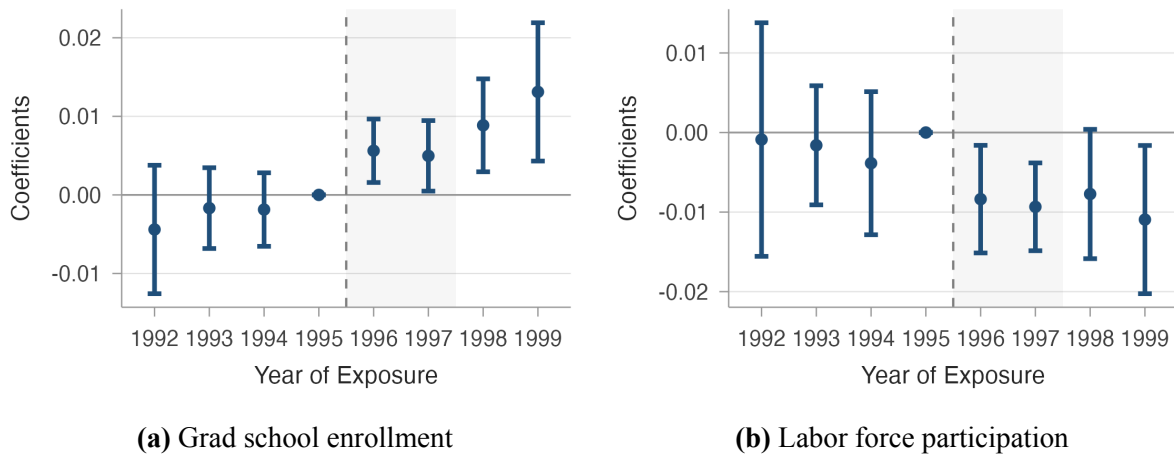

(a) Grad school enrollment

(b) Labor force participation

Figure 5: Long-run effects of folic acid fortification on young adults' outcomes, event-study plots

Notes: These figures plot event-study estimates with 95% confidence intervals. Standard errors are clustered by state of birth. Exposure year groups cohorts into four-quarter bins; for example, 1996 includes cohorts born in 1996Q4–1997Q3. Controls include state-of-birth-by-survey-year and quarter-by-year-of-birth fixed effects, along with gender, race, ethnicity, Medicaid eligibility, exposure to mental-health parity laws, and welfare reforms. The shaded region denotes partially exposed cohorts (exposed before January 1, 1998). Regressions are weighted by IPUMS person weights.

Figures A2a–A2b report results by gender, race (white versus non–white), and ethnicity (Hispanic versus non-Hispanic). Graduate school enrollment increases across all groups, with the largest gains among women. The decline in labor force participation is strongest for white and non-Hispanic individuals. Women also adjust labor supply on the intensive margin: alongside a less precisely estimated decline in labor force participation, exposed women are less likely to work full time (30 or more hours per week). Estimates for Hispanic individuals are generally noisy, consistent with the fact that the 1996 mandate did not cover staple foods widely consumed in Hispanic households, such as corn masa flour and tortillas (Redpath, Kancherla and Oakley, 2018).[7]

All regression estimates are intent-to-treat effects (ITTs). Interpreting the treated group as

---

[7]The FDA did not permit voluntary fortification of corn masa flour until 2016 (Food and Drug Administration, 2016).

individuals who avoid CNSA because of fortification, I convert ITTs to treatment effects on the treated (TOTs) by scaling the reduced-form estimates by the fortification-induced change in the CNSA rate. Figure 4d shows that the CNSA rate falls by 0.24 per 1,000 births in high-baseline-CNSA states and by 0.07 per 1,000 births in low-baseline-CNSA states, implying a first-stage difference of $0.24 - 0.07 = 0.17$. The implied TOTs are therefore $0.008/0.17 = 0.047$ for graduate school enrollment and $-0.007/0.17 = -0.041$ for labor force participation.

5.3 Robustness

This section shows that the estimates for labor force participation and graduate school enrollment appear reasonably robust.

Model specifications. I begin by assessing robustness to alternative model specifications. Online Appendix Figures B1a and B1b show that the results are robust across the following specifications: (i) a parsimonious model that includes only state-of-birth-by-survey-year and quarter-by-year-of-birth fixed effects; (ii) the preferred specification but without controls for confounding policies; (iii) the preferred specification, but with fixed effects entered separately as state-of-birth, survey-year, and quarter-by-year-of-birth fixed effects; and (iv) the preferred specification with an added Bartik-style state-by-year unemployment rate to capture economic conditions at birth.[8]

Exposure thresholds and dose response. To assess sensitivity to the exposure definition, I vary the cutoff used to define high-baseline-CNSA states, using the top 30% and top 20%, holding all other specification choices fixed. Online Appendix Figures B2a and B2b show that the results remain robust. I also estimate dose–response effects by replacing the binary high-baseline-CNSA indicator with a continuous measure of baseline CNSA rates. The results are consistent with the main estimates, although the dose–response estimates for labor force participation are less precise.

Sample selection. Online Appendix Figures B3a and B3b show that the main results are robust to alternative sample choices. I reestimate the models using: (i) PUMS 2015–2024; (ii) PUMS 2019–2024 excluding 2020; and (iii) PUMS 2019–2024 with a wider age range of 22 to 32. For (i), the 2015–2018 PUMS contributes only unexposed cohorts within the age window, which increases the number of pre-treatment cohorts. For (ii), ACS 2020 has a substantially

---

[8]The construction of this Bartik-style unemployment rate follows East and Velásquez (2024).

lower response rate than other years, and response rates remain below pre-pandemic levels but are relatively stable through 2024.[9] For (iii), I test whether the results are sensitive to the age window by including younger individuals who may not yet have entered graduate school and older individuals at a different life stage.

Randomization test. To assess robustness to random noise, I estimate pseudo effects by randomly reassigning the fortification exposure 1,000 times while preserving its empirical distribution across states. Online Appendix Figures B4a and B4b show that the main estimates fall well into the tails of the pseudo effect distributions, suggesting they are unlikely to be driven by chance.

Sensitivity to parallel trend violation. Finally, I evaluate how sensitive the main results are to potential violations of the parallel trends assumption using the approach in Rambachan and Roth (2023). For the first exposed cohort, I estimate breakdown values, defined as the size of a deviation in pre-trends that would make the estimates statistically insignificant at the 90% confidence level. The breakdown values are approximately 0.6 for both graduate school enrollment and labor force participation (Online Appendix Figures B5a and B5b). While such deviations cannot be completely ruled out, they seem unlikely given the observed pre-treatment dynamics. Overall, the estimates appear reasonably robust to hypothesized departures from parallel trends between individuals born in high- and low-baseline-CNSA states.

5.4   Potential Mechanisms

This section examines three potential mechanisms and presents suggestive evidence that improved cognitive ability, rather than fertility selection or changes in labor market conditions, is most consistent with the shift towarding schooling.

Fertility selection. Because fortification reduces birth defects, it may also change the composition of live births, which could contribute to the shift toward schooling. The shift could reflect healthier infants who gain an early advantage in human capital accumulation, or a shift toward births to more advantaged mothers who place greater value on education. To examine this channel, I use birth certificate data to estimate the effects of fortification on birth outcomes and maternal characteristics. Birth outcomes include birth weight, an indicator for low birth weight (birth weight $< 2,500$ grams), gestational weeks, and an indicator for preterm birth

---

[9]ACS response rates are 94.7% (2016), 93.7% (2017), 92.0% (2018), 86.0% (2019), 71.2% (2020), 85.3% (2021), 84.4% (2022), 84.7% (2023), and 82.9% (2024).

(gestational weeks < 37). Maternal characteristics include indicators for being non-white, being Hispanic, age 22 or younger, having less than a college education, being unmarried, and having inadequate prenatal care.

Table A2 shows no evidence that birth outcomes respond to fortification exposure. Because birth certificate data cover only live births, these results do not imply that fortification has no effect on infant health. Table A3 instead shows an increase in the share of mothers aged 22 or younger, consistent with a shift toward more disadvantaged mothers among live births. If fortification increases survival for infants born to disadvantaged mothers, average birth outcomes among live births can remain unchanged even as outcomes improve for births that would have occurred regardless. Overall, these results suggest that the main findings are unlikely to be driven by healthier infants or by a shift toward more advantaged mothers. If anything, the compositional change could attenuate the estimated effects on later outcomes, since young adults are more likely to have a mother aged 22 or younger.

Local labor market conditions. Another possibility is that the shift reflects labor market conditions that coincided with fortification exposure. If exposed cohorts entered adulthood during weaker job markets, they may have been more likely to stay in school, including in graduate programs, to delay labor market entry when returns to work were temporarily low. To assess this channel, I reestimate the main results while adding state- or PUMA-of-residence-by-survey-year fixed effects,[10] which absorb time-varying differences in local labor market conditions. If local labor markets were driving the results, the estimated effects should shrink once these fixed effects are included. Table A4 shows that the estimates are essentially unchanged, indicating that local labor market conditions explain little of the shift.

Table A5 supports this conclusion. If weak job markets were the main driver, we would expect higher unemployment among those in the labor force and/or declines in job quality among those employed. However, I find no effect on unemployment conditional on labor force participation and no clear changes in occupational quality indices that capture earnings, socioeconomic status, prestige, or the education level associated with the occupation. Overall, the evidence provides little support for local labor market conditions as the mechanism behind the shift toward graduate schooling.

Cognitive ability. Improvements in cognitive ability may increase educational investment

---

[10]PUMA: Public Use Microdata Area, the smallest geographic unit in the ACS PUMS. In the sample, over 95% of respondents work in their state of residence.

21

([Murnane, Willett and Levy, 1995](#); [Bowles, Gintis and Osborne, 2001](#); [Heckman, Stixrud and Urzua, 2006](#)). The medical literature suggests that folic acid exposure can improve cognitive development (see [Irvine et al. (2022)](#) for a review). To test this channel, I use state-level test score data from the National Assessment of Educational Progress (NAEP) to provide suggestive evidence.[11] Consistent with this mechanism, higher fortification exposure is associated with higher math scores among 4th and 8th graders.

I begin by mapping fortification exposure from state of birth to state of school, since the public NAEP data report only state of school, survey year, and grade. I construct a migration-adjusted exposure measure that weights birth-state exposure by the observed composition of student birth places within each state-grade-year cell using ACS PUMS 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 (with all foreign countries grouped together). Define $\pi_{s|j,g,c}$ as the share of students in (j, g, c) who were born in s, and let $\mathbf{1}\{\text{High CNSA}\}_s$ indicate whether birth state s is in the top quartile of baseline CNSA rates. The exposure used for NAEP cell (j, g, c) is

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

with the indicator set to 0 for foreign born. This yields a state of school exposure measure that accounts for observed migration patterns among 4th and 8th graders, and captures the likelihood that students in state j were born in a high-baseline-CNSA state.

I assign representative ages of 10 (grade 4) and 14 (grade 8) to back out likely birth cohorts and define exposed cohorts as those born after 1996. Online Appendix Figures [B6a](#) and [B6b](#) show that roughly 90.8% of 4th graders are ages 9 to 10 and roughly 89.3% of 8th graders are ages 13 to 14. This assignment likely biases exposure downward, so the estimates are conservative. The results are similar when I instead assign ages of 9 (grade 4) and 13 (grade 8) (Online Appendix Table [B1](#)).

The regression model is:

$$Y_{jgc} = \zeta \big(P(\text{High CNSA})_{jgc} \times \text{Exposed cohorts}_{gc} \big) + \eta_{jc} + \theta_g + Z_{jgc} + \varepsilon_{jgc}, \qquad (2)$$

---

[11]Restricted-use NAEP microdata are ordinarily accessible by application through the Institute of Education Sciences at the U.S. Department of Education, but access is currently paused. I 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.

where Y_{jgc} is the average test score outcome for state j in survey year c and grade g; $P(\text{High CNSA})_{jgc}$ is the migration-adjusted exposure; and Exposed cohorts$_{gc}$ is an indicator for likely birth cohorts after 1996.[12] η_{jc} includes state-of-school-by-likely-year-of-birth fixed effects that absorb unobserved factors common to students who attend school in the same state and belong to the same cohort, including time-varying differences in school policies, resources, and local economic conditions. θ_g are grade fixed effects. Z_{jgc} includes migration-adjusted exposure measures for potentially confounding policies, including Medicaid and CHIP expansions, welfare reform, and state mental health parity laws, constructed using the same weights as in $P(\text{High CNSA})_{jgc}$. I pool 4th and 8th graders and focus on a pre-versus-post comparison because the aggregate data include a limited number of state-by-grade-by-year cells.

Table 3 reports the estimated relationship between in utero exposure to folic acid fortification and NAEP performance among 4th and 8th graders. Columns (1) to (3) show positive associations in mathematics: when the likelihood of being born in a high-baseline-CNSA state moves from 0 to 1, average math scores rise by 4.51 points (1.73% at the mean), the share scoring at or above NAEP Basic increases by 3.66 percentage points, and the share at or above NAEP Proficient increases by 4.63 percentage points. I find no noticeable changes in reading scores or in the share at or above NAEP Basic. Taken together with the earlier evidence, these results are most consistent with improved cognitive ability as a mechanism behind the shift toward graduate schooling.

---

[12]Because quarter of birth is not observed in NAEP, I define exposed cohorts as students likely born after 1996.

23

Table 3: In utero fortification exposure and test scores of 4th and 8th graders

	Math			Reading		
	Average score (1)	% ≥ Basic (2)	% ≥ Proficient (3)	Average score (4)	% ≥ Basic (5)	% ≥ Proficient (6)
P(High CNSA)×Exposed cohorts	4.508***	3.661***	4.632***	-0.822	1.140	-1.009
	(1.130)	(1.104)	(0.902)	(1.095)	(1.478)	(0.610)
Observations	408	408	408	408	408	408
R^2	0.995	0.945	0.938	0.995	0.951	0.945
Dep. var. mean	259.991	75.502	34.361	241.256	69.351	31.305
Est./Dep. var. mean	1.734%	4.849%	13.480%	-0.341%	1.644%	-3.224%
State-of-school-by-(likely)-year-of-birth FEs	✓	✓	✓	✓	✓	✓
Grade FEs	✓	✓	✓	✓	✓	✓
Control variables	✓	✓	✓	✓	✓	✓

Notes: This table presents cohort-DiD estimates with standard errors clustered at state of school. P(High CNSA) is fortification exposure adjusted for migration. Exposed cohorts are those with likely year of birth after 1996, determined by assigning age 10 to grade 4 and age 14 to grade 8. Controls include state-of-school-by-(likely)-year-of-birth, grade fixed effects, Medicaid eligibility, exposure to mental-health parity laws, and welfare reforms. The unit of observation is state-by-year-by-grade cells. Regressions are weighted by student counts from ACS PUMS. ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

6 Magnitudes and Economic Importance

This section compares the long-run effects of folic acid fortification with salt iodization, iron fortification of bread, and the Food Stamp Program (later SNAP). I focus on educational outcomes because labor force participation among young adults is not directly comparable to estimates for older adults, given substitution between work and schooling. Most older adults, however, completed their education in young adulthood.

To facilitate comparison, I convert the estimated effect of folic acid fortification on graduate school enrollment into additional years of schooling by assuming that graduate enrollment corresponds to 1.78 additional years of schooling.[13] Multiplying by the ITT effect on graduate school enrollment yields an ITT gain of $0.008 \times 1.78 = 0.014$ years, and multiplying by the TOT effect yields a TOT of $0.047 \times 1.78 = 0.084$ years. Standard errors scale by the same factors.[14]

---

[13] I use 1.78, which is calculated by multiplying the share of each degree among all graduate degrees by the typical minimum years needed to complete the degree. The shares are calculated from PUMS by restricting the sample to non-Hispanic individuals over 35 years old. The shares of each graduate degree and their typical minimum time to complete are Master's degrees (0.72, 1 year), professional degrees (0.17, 3 years), and Ph.D. degrees (0.11, 5 years). $(0.72 \times 1 + 0.17 \times 3 + 0.11 \times 5) \times 0.6 = 1.78$.

[14] Starting from the enrollment standard error of 0.003, the ITT years standard error is $0.003 \times 1.78 = 0.005$.

For salt iodization, I draw on Adhvaryu et al. (2020), who proxy in utero exposure using baseline goiter prevalence and report effects on years of schooling and income for adults aged 39 to 60. Moving exposure from the 25th to the 75th percentile increases years of schooling by 0.071 years for women (about 0.63% of the mean, ITT) and 0.031 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 nineteen years of exposure, at a one standard deviation difference in iron consumption, implies a 0.05-year increase in schooling (ITT) among adults aged 22 to 50, although the estimate is imprecise. He also finds a 2.9% increase in total income when 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 0.229 years (TOT) and raises labor income by 7.13% (TOT).

Figure 6 summarizes these comparisons. Although this comparison is illustrative rather than a strict like-for-like comparison due to different target populations, timing, exposure definitions, and outcomes across studies, we can see that the effect of folic acid fortification is on par with that of salt iodization for women and substantially smaller than the Food Stamp Program, a much larger program. Yet fortification appears far more cost-effective. It delivers 29% of schooling gains at about 0.01% of the Food Stamp Program's annual cost—about $3 million (Grosse et al., 2005) versus about $29 billion (Food and Nutrition Service, 2005).

How much does the shift from work to schooling in early adulthood translate into long-run human capital gains? Using an estimated increase in schooling of 0.084 years (TOT), the income return to one additional year of schooling of 9% from Barrow and Rouse (2006), and prime-age annual income of $67,218 (full-time workers aged 35 to 50, estimated using IPUMS 2019–2024), the implied increase in annual income at prime age, discounted to year 0 by 35 years at a 7% discount rate (US Office of Management and Budget, 2003), is $0.084 \times 0.090 \times$ $67,218 \times \frac{1}{(1+0.07)^{35}} \approx$ $47.6 per exposed individual. To scale the schooling channel to a cohort, consider 1,088,000 infants in the 14 high-exposure states—the average annual number of births from 1996 to 1999, the cohort present value from one year of additional income is then $1,088,000 \times$ $47.6 \approx$ $51.7 million. Applying an effective marginal tax rate of 30%[15]

Scaling to the TOT using the first-stage difference of 0.17 implies a TOT years standard error of $0.003/0.17 \times 1.78 = 0.031$.

[15]The effective marginal tax rate is the percentage of an additional dollar of earnings that is unavailable to a
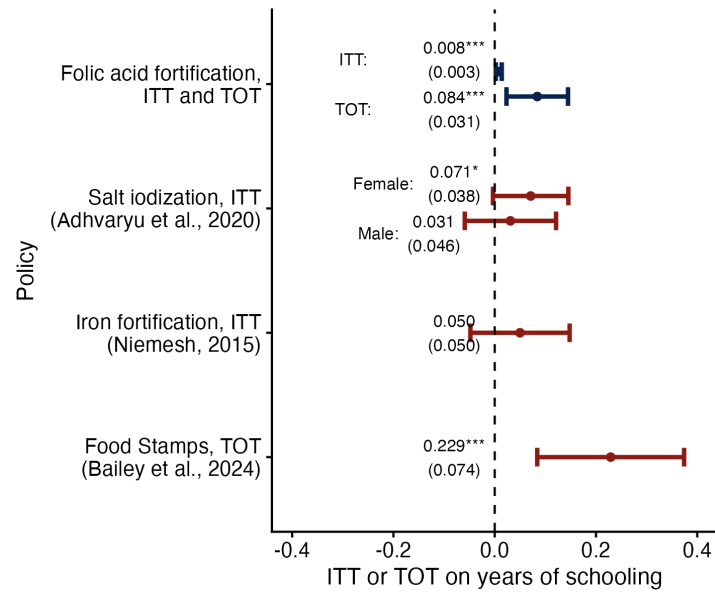
Figure 6: Long-run effects on years of schooling by nutrition policies

Notes: This figure plots point estimates for each nutrition intervention's effect on years of schooling with 95% confidence intervals. The folic acid estimate converts its effect on graduate school enrollment into years of schooling; all other values are regression coefficients from the cited studies. Adhvaryu et al. (2020) use baseline goiter prevalence as a continuous proxy for iodine deficiency; Niemesh (2015) use estimated iron consumption as a proxy for iron deficiency. Bailey et al. (2024) report the effect of Food Stamps exposure from conception to age five.

(Congressional Budget Office, 2012), the implied increase in tax revenue is $51.7 \times 0.30 \approx$ $15.5 million.

Because this calculation captures only the schooling channel, total benefits could be larger. For example, Bailey et al. (2024) find that schooling explains roughly one third of the income gains from early-life nutrition, with health and other channels plausibly accounting for the remainder. Overall, my calculation suggests that folic acid fortification is likely to break even given its roughly $3 million annual cost, and that omitting long-run human capital benefits can materially understate the benefits of fortification by tens of millions of dollars.

7   Conclusion

This paper provides the first evidence on the long-run human capital effects of folic acid fortification. Leveraging cross-state variation in baseline rates of central nervous system anomalies (CNSA)—birth defects linked to folate deficiency—and the timing of the first trimester relative to fortification, I document a shift toward schooling in early adulthood: among young adults aged 23–30, graduate school enrollment rises by 0.8 percentage points while labor force

worker because it is paid in taxes or offset by reductions in benefits from social safety net programs.

26

participation falls by 0.7 percentage points. The schooling gains appear across demographic groups and are largest for women. Labor supply responses are concentrated among White and non-Hispanic individuals, and exposed women are less likely to work full time. Estimates for Hispanic individuals are smaller and noisier, consistent with weaker exposure because the 1996 mandate did not cover staples widely consumed in Hispanic households. Evidence on mechanisms points away from fertility selection and changes in local labor market conditions. Instead, fortification exposure is associated with higher math scores among 4th and 8th graders, consistent with improved cognitive ability as a plausible driver of the shift toward schooling.

The implied long-run human capital gains are comparable to those from salt iodization and substantially smaller than those from the Food Stamp Program, a much larger intervention. Back-of-the-envelope calculations suggest that additional annual tax revenue from schooling gains alone can cover the program's reported annual cost. Cost-benefit analyses that ignore these long-run effects may understate program benefits by tens of millions of dollars.

This study has several limitations. First, the exposure measure is constructed from live births rather than all pregnancies. Data on the full universe of pregnancies are not available; even fetal death files primarily cover fetuses at 20 weeks or later, and many states began reporting congenital anomalies in these data later than in birth certificates. Second, exposed cohorts are still too young to study downstream adult outcomes such as prime-age earnings or family formation. Following these cohorts over the life course would provide a more complete assessment of fortification's effects. Third, because the design compares higher- and lower-exposure groups rather than treated and untreated groups, estimates capture relative rather than absolute effects. Despite these limitations, the results inform how prenatal nutrition can shape downstream human capital, aligning with the fetal-origins literature in showing that early-life interventions can generate long-run gains.

Compared with large programs such as the Food Stamp Program, which broadly subsidizes food consumption, fortification targets specific micronutrient deficiencies and is low-cost yet delivers meaningful benefits. This contrast highlights potential efficiency gains from more targeted nutrition policy—focusing resources on nutrients and foods with the highest marginal returns rather than subsidizing healthy foods more generally. Such targeting may be especially attractive in an era of tightening welfare budgets.

This paper also suggests scope for larger gains from folic acid fortification in low- and

middle-income countries, where folate deficiency is often more severe (McLean, de Benoist and Allen, 2008). The U.S. experience is encouraging, but more evidence is needed on how fortification performs in other settings, particularly where diets, staple foods, and the feasibility and quality of fortification technologies may differ.

References

Abdel-Salam, Ghada, and Andrew E Czeizel. 2000. "A case-control etiologic study of microcephaly."

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."

American College of Obstetricians and Gynecologists. 2017. "Practice Bulletin No. 187: Neural Tube Defects." *Obstetrics & Gynecology*, 130(6): e279–e290.

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.

Barrow, Lisa, and Cecilia E Rouse. 2006. "The economic value of education by race and ethnicity." *Economic Perspectives, Second Quarter*.

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 Dagnosis*, 13(9): 863–871.

Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001. "The determinants of earnings: A behavioral approach." *Journal of Economic Literature*, 39(4): 1137–1176.

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.

Congressional Budget Office. 2012. "Effective Marginal Tax Rates for Low- and Moderate-Income Workers." https://www.cbo.gov/sites/default/files/112th-congress-2011-2012/reports/marginaltaxratesone-column.pdf.

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. 2022*a*. "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. 2022*b*. "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.

East, Chloe N, and Andrea Velásquez. 2024. "Unintended consequences of immigration enforcement: household services and high-educated mothers' work." *Journal of Human Resources*, 59(5): 1458–1502.

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 Drug Administration. 2016. "Food additives permitted for direct addition to food for human consumption; folic acid. Final rule." *Federal register*, 81(73): 22176–22183.

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).

Greene, Nicholas DE, and Andrew J Copp. 2014. "Neural tube defects." *Annual Review of Neuroscience*, 37: 221–242.

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.

Heckman, James J, Jora Stixrud, and Sergio Urzua. 2006. "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior." *Journal of Labor Economics*, 24(3): 411–482.

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.

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 Arynchyna-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.

Lewis, Christine J, Nancy T Crane, Dennis B Wilson, and Elizabeth A Yetley. 1999. "Estimated folate intakes: data updated to reflect food fortification, increased bioavailability, and dietary supplement use." *The American Journal of Clinical Nutrition*, 70(2): 198–207.

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.

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).

Murnane, Richard J, John B Willett, and Frank Levy. 1995. "The growing importance of cognitive skills in wage determination." *The Review of Economics and Statistics*, 77(2): 251.

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.

Rambachan, Ashesh, and Jonathan Roth. 2023. "A more credible approach to parallel trends." *Review of Economic Studies*, 90(5): 2555–2591.

Ray, Joel G, Gita Singh, and Robert F Burrows. 2004. "Evidence for suboptimal use of periconceptional folic acid supplements globally." *BJOG: An International Journal of Obstetrics & Gynaecology*, 111(5): 399–408.

Redpath, Ben, Vijaya Kancherla, and Godfrey P Oakley. 2018. "Availability of corn masa flour and tortillas fortified with folic acid in Atlanta after national regulations allowing voluntary fortification." *JAMA*, 320(15): 1600–1601.

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."

Yetley, Elizabeth A, Christine M Pfeiffer, Karen W Phinney, Zia Fazili, David A Lacher, Regan L Bailey, Sheena Blackmore, Jay L Bock, Lawrence C Brody, Ralph Carmel, et al. 2011. "Biomarkers of folate status in NHANES: a roundtable summary." *The American Journal of Clinical Nutrition*, 94(1): 303S–312S.

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.
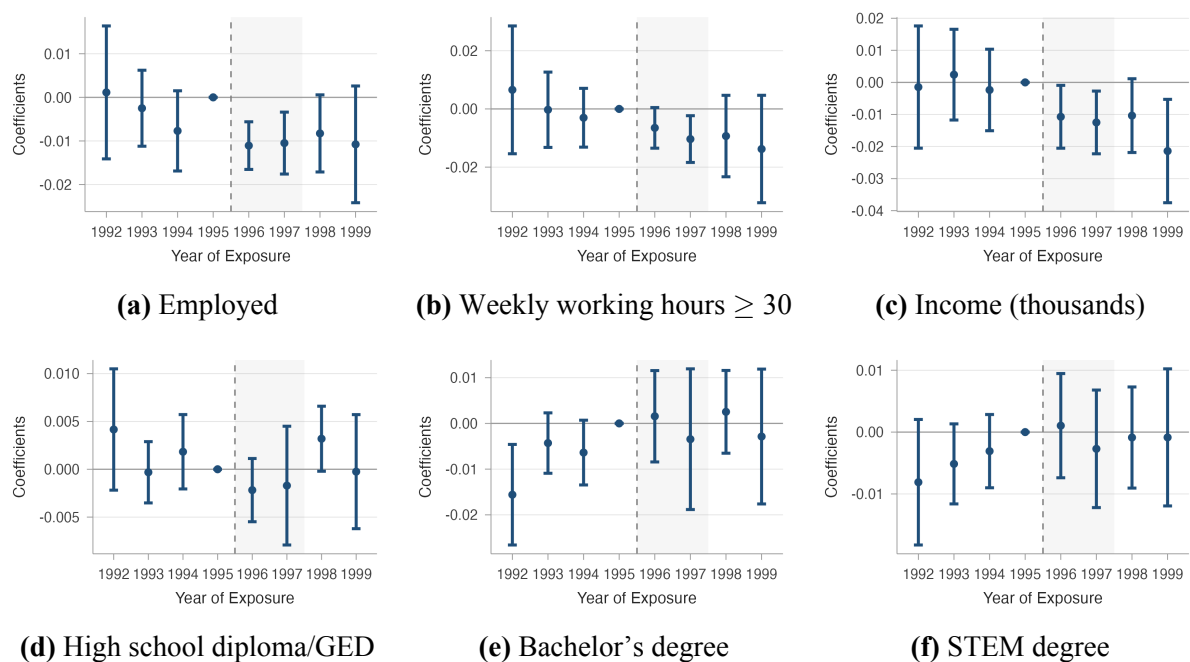
Appendix

A Tables and figures

Table A1: Fortification exposure and pre-1989 state/county characteristics

	High CNSA$_s$		
	State-level	County level	
	(1)	(2)	(3)
Demographic features			
Share of black (%), 1988	-0.002	-0.002	0.000
	(0.015)	(0.002)	(0.002)
Share of female (%), 1988	0.015	0.037*	0.008
	(0.204)	(0.021)	(0.019)
Share of under 5 (%), 1988	0.260	0.048*	-0.030
	(0.217)	(0.025)	(0.020)
Share of over 65 (%), 1988	-0.006	-0.011	-0.016
	(0.081)	(0.012)	(0.012)
Birth rate (%), 1988	-0.022	0.012	0.014**
	(0.057)	(0.008)	(0.006)
Death rate (%), 1988	0.162	0.008	-0.013
	(0.272)	(0.020)	(0.020)
Log population, 1988	0.093	-0.031	0.005
	(0.124)	(0.026)	(0.021)
Economic conditions			
Transfer income p.p. (million $), 1988	-0.307	-0.007	0.129
	(0.666)	(0.147)	(0.091)
Income p.p. (million $), 1985	107.878	23.606	3.413
	(113.606)	(17.774)	(18.994)
Federal funds p.p. (million $), 1986	-234.162	9.907	8.859
	(162.300)	(9.340)	(9.560)
Unemployment rate (%), 1986	-0.056	-0.005	-0.005
	(0.066)	(0.011)	(0.006)
Agriculture			
Value of produces sold per farm (million $), 1987	-4.560	-0.596*	-0.028
	(2.789)	(0.322)	(0.129)
Average farm size (million acres), 1987	152.637	-13.204	-0.314
	(121.444)	(8.099)	(4.841)
State FEs			Y
Observations	49	2,850	2,850
R^2	0.321	0.051	0.283
Dep. var. mean	0.294	0.224	0.224

Notes: This table presents results from regressing high CNSA—an indicator for states in the 4th quartile of baseline CNSA rates—on pre-1989 state or county characteristics. Regressions are weighted by 1988 population. Data sources: County Intercensal Estimates (demographic shares and population), County Databook 1988 (birth rate, death rate, farm value, and farm size), BEA Regional Economic Information System (transfers), and Bureau of Labor Statistics (unemployment). Standard errors in county-level regressions are clustered at the state level. ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.
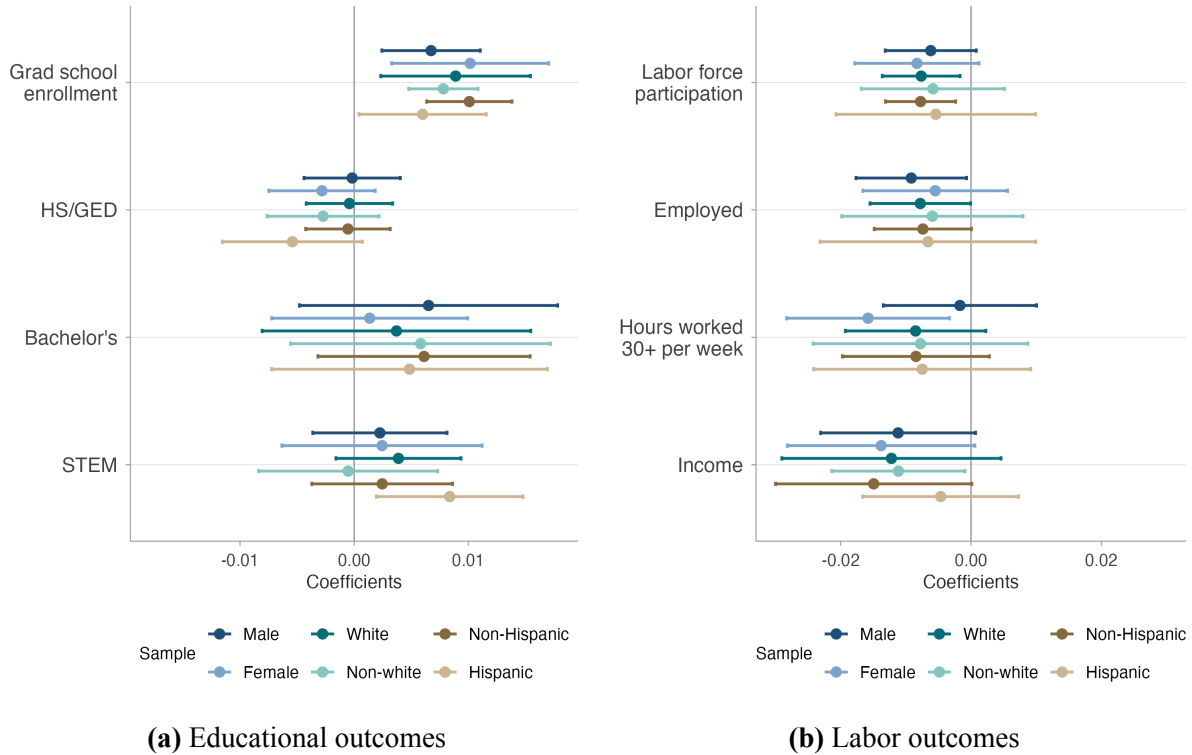
Figure A1: Long-run effects of folic acid fortification on other young adult outcomes, event-study plots

(a) Employed

(b) Weekly working hours $\geq$ 30

(c) Income (thousands)

(d) High school diploma/GED

(e) Bachelor's degree

(f) STEM degree

Notes: These figures plot event-study estimates with 95% confidence intervals. Standard errors are clustered by state of birth. Exposure year groups cohorts into four-quarter bins; for example, 1996 includes cohorts born in 1996Q4–1997Q3. Controls include state-of-birth-by-survey-year and quarter-by-year-of-birth fixed effects, along with gender, race, ethnicity, Medicaid eligibility, exposure to mental-health parity laws, and welfare reforms. The shaded region denotes partially exposed cohorts (exposed before January 1, 1998). Regressions are weighted by IPUMS person weights.

Figure A2: Heterogeneity by gender, race, and ethnicity

(a) Educational outcomes

(b) Labor outcomes

Notes: This table presents cohort-DiD estimates and their 95% confidence interval by gender, eace, and ethnicity. I control for state-of-birth-by-survey-year and quarter-by-year-of-birth fixed effects, along with gender, race, ethnicity, Medicaid eligibility, exposure to mental-health parity laws, and welfare reforms. Regressions and dependent-variable means are weighted by IPUMS person weights.

Table A2: Effects of 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)
High CNSA × Exposed cohorts	-2.328	0.001	0.000	0.002
	(2.072)	(0.001)	(0.011)	(0.001)
Observations	64,754	64,754	64,754	64,754
R^2	0.860	0.794	0.789	0.739
Dep. var. mean	3323.419	0.074	38.883	0.113
Est./Dep. var. mean	-0.070%	1.073%	0.000%	1.594%
State-of-birth FEs	✓	✓	✓	✓
Quarter-year-of-birth FEs	✓	✓	✓	✓
Control variables	✓	✓	✓	✓

Notes: This table presents standard DiD estimates with standard errors clustered at the state level. Natality records are aggregated to county-by-quarter-of-birth cells and merged with state-level exposure measures. High CNSA indicates states in the 4th quartile of baseline CNSA rates. Exposed cohorts are those whose first trimester ended after the March 1996 authorization of folic acid fortification. Controls include state-of-birth and quarter-by-year-of-birth fixed effects, along with time-varying covariates for Medicaid eligibility, mental-health parity laws, and welfare reforms. Regressions are weighted by cell birth counts. ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

Table A3: Effects of folic acid fortification on maternal characteristics

	Non-white	Hispanic	Age ≤ 22 29	Education < college	Unmarried	Inadequate prenatal care
	(1)	(2)	(3)	(4)	(5)	(6)
High CNSA × Exposed cohorts	0.005	-0.001	0.005**	0.003	0.007	0.010
	(0.004)	(0.005)	(0.002)	(0.004)	(0.007)	(0.011)
Observations	64,754	64,754	64,754	64,674	64,754	64,754
R^2	0.986	0.992	0.934	0.953	0.918	0.819
Dep. var. mean	0.211	0.179	0.272	0.555	0.325	0.240
Est./Dep. var. mean	2.585%	-0.770%	1.910%	0.480%	2.099%	4.001%
State-of-birth FEs	✓	✓	✓	✓	✓	✓
Quarter-year-of-birth FEs	✓	✓	✓	✓	✓	✓
Control variables	✓	✓	✓	✓	✓	✓

Notes: This table presents standard DiD estimates with standard errors clustered at the state level. Natality records are aggregated to county-by-quarter-of-birth cells and merged with state-level exposure measures. High CNSA indicates states in the 4th quartile of baseline CNSA rates. Exposed cohorts are those whose first trimester ended after the March 1996 authorization of folic acid fortification. Controls include state-of-birth and quarter-by-year-of-birth fixed effects, along with time-varying covariates for Medicaid eligibility, mental-health parity laws, and welfare reforms. Regressions are weighted by cell birth counts. ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

Table A4: Controlling for place-of-residence-by-survey-survey FEs

	Grad school enrollment		Labor force participation	
	(1)	(2)	(3)	(4)
High CNSA × Exposed cohorts	0.009***	0.008***	-0.007**	-0.007**
	(0.003)	(0.003)	(0.003)	(0.003)
Observations	1,065,355	1,065,355	1,065,355	1,065,355
R^2	0.013	0.059	0.013	0.052
Dep. var. mean	0.059	0.059	0.843	0.843
Est./Dep. var. mean	14.541%	14.327%	-0.865%	-0.862%
State-of-birth×Survey-year FEs	✓	✓	✓	✓
Quarter-year-of-birth FEs	✓	✓	✓	✓
Control variables	✓	✓	✓	✓
State-of-residence-by-survey-year FEs	✓		✓	
PUMA-of-residence-by-survey-year FEs		✓		✓

Notes: This table presents cohort-DiD estimates with standard errors clustered at the state of birth. High CNSA is an indicator for states in the 4th quartile of baseline CNSA rates. Exposed cohorts are those whose first trimester ended after the March 1996 authorization of folic acid fortification. Controls include state-of-birth-by-survey-year and quarter-by-year-of-birth fixed effects, gender, race, ethnicity, Medicaid eligibility, exposure to mental-health parity laws, and welfare reforms, as well as place-of-residence-by-survey-year fixed effects. Regressions and dependent-variable means are weighted by IPUMS person weights. ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.
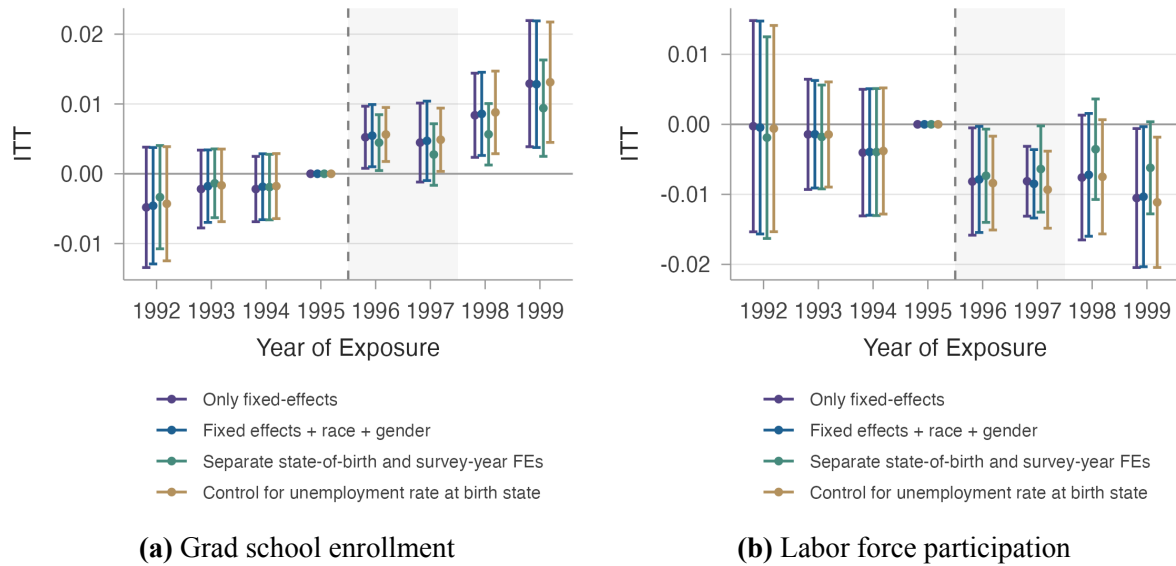
Table A5: Long-run effects on conditional unemployment and occupational quality indices of young adults

	Unemp.	Socio-economic Index, Hauser and Warren	Occ. prestige score, Siegel	Occ. prestige score, Nakao and Treas	Occ. income score	Occ. earnings score, 1950 basis	Occ. education score, 1990 basis
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
High CNSA × Exposed cohorts	0.001	12.775	1.180	1.254	0.024	-12.020	-10.456
	(0.002)	(13.284)	(1.626)	(1.647)	(0.062)	(10.082)	(9.948)
Observations	876,876	826,217	826,217	826,217	826,217	826,217	826,217
R^2	0.016	0.055	0.060	0.054	0.035	0.011	0.005
Dep. var. mean	0.062	3750.854	410.225	443.103	27.882	646.357	774.111
Est./Dep. var. mean	0.884%	0.341%	0.288%	0.283%	0.087%	-1.860%	-1.351%
State-of-birth×Survey-year FEs	✓	✓	✓	✓	✓	✓	✓
Quarter-year-of-birth FEs	✓	✓	✓	✓	✓	✓	✓
Control variables	✓	✓	✓	✓	✓	✓	✓

Notes: This table presents cohort-DiD estimates with standard errors clustered at the state of birth. High CNSA is an indicator for states in the 4th quartile of baseline CNSA rates. Exposed cohorts are those whose first trimester ended after the March 1996 authorization of folic acid fortification. Controls include state-of-birth-by-survey-year and quarter-by-year-of-birth fixed effects, along with gender, race, ethnicity, Medicaid eligibility, exposure to mental-health parity laws, and welfare reforms. Regressions and dependent-variable means are weighted by IPUMS person weights. In Column (1), unemployment is conditional on labor force participation. In Columns (2)-(7), occupational quality indices are conditional on being employed. $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.1$.
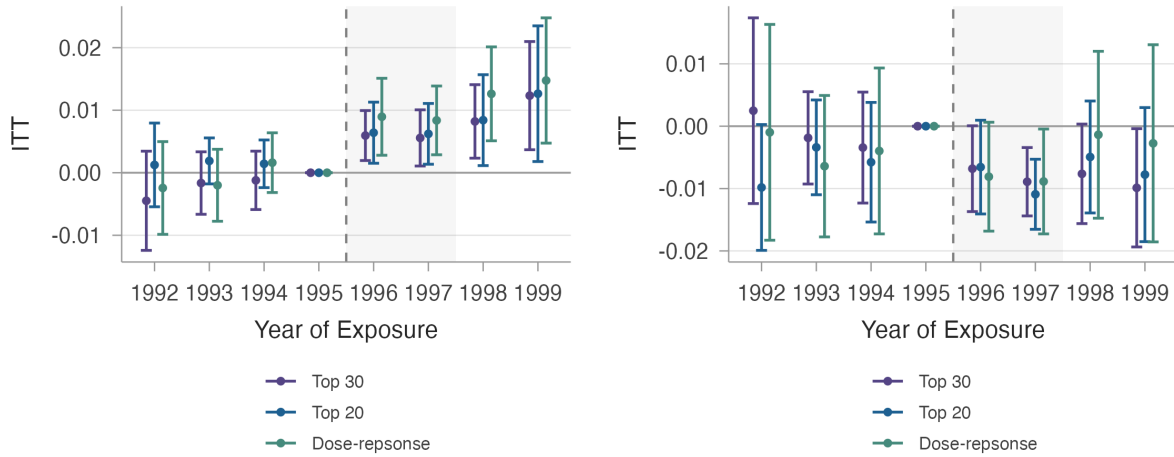
B Online appendix

Figure B1: Robustness to alternative model specifications

(a) Grad school enrollment   **(b)** Labor force participation

Notes: These figures plot event-study estimates with 95% confidence intervals. Standard errors are clustered by state of birth. Exposure year groups cohorts into four-quarter bins; for example, 1996 includes cohorts born in 1996Q4–1997Q3. The shaded region denotes partially exposed cohorts (exposed before January 1, 1998). Regressions are weighted by IPUMS person weights. From left to right: (i) parsimonious model with only state-of-birth-by-survey-year and quarter-by-year-of-birth fixed effects; (ii) preferred specification without confounding policy controls; (iii) preferred specification with fixed effects entered separately; and (iv) preferred specification with an added Bartik-style state-by-year unemployment rate.

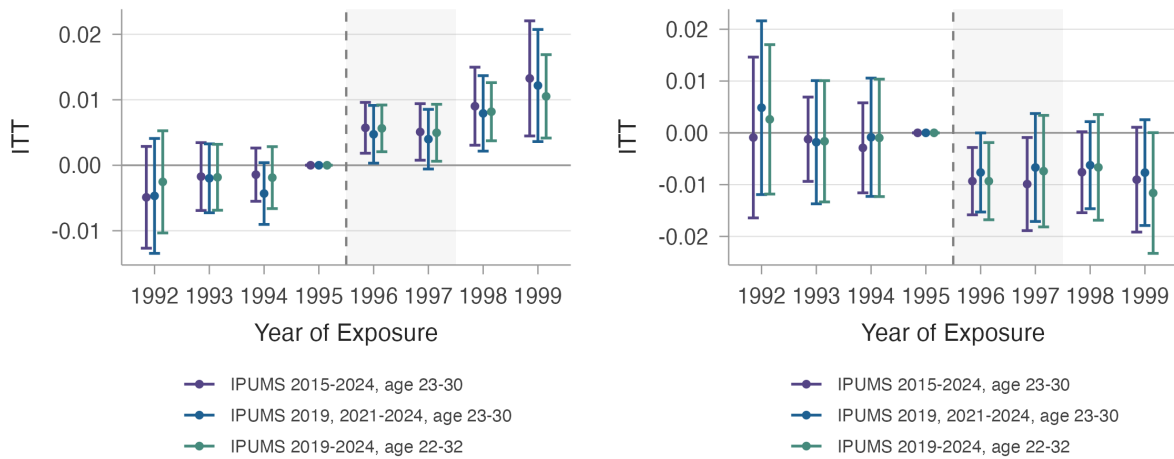**Figure B2:** Robustness to alternative exposure measures

(a) Grad school enrollment

(b) Labor force participation

Notes: These figures plot event-study estimates with 95% confidence intervals. Standard errors are clustered by state of birth. Exposure year groups cohorts into four-quarter bins; for example, 1996 includes cohorts born in 1996Q4–1997Q3. The shaded region denotes partially exposed cohorts (exposed before January 1, 1998). Regressions are weighted by IPUMS person weights. From left to right: (i) top 30% threshold for high CNSA; (ii) top 20% threshold for high CNSA; and (iii) dose-response model.

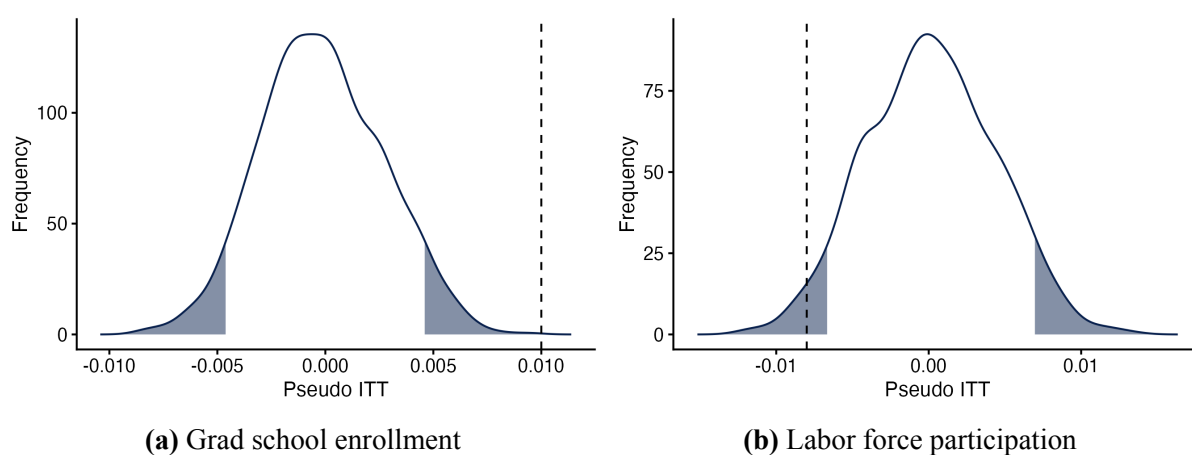**Figure B3:** Robustness to alternative samples

(a) Grad school enrollment
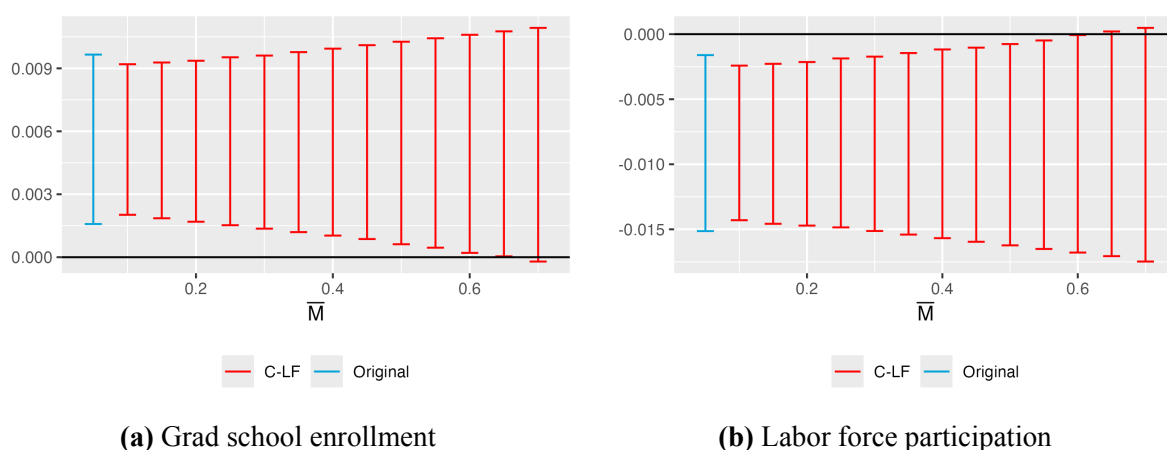
(b) Labor force participation

Notes: These figures plot event-study estimates with 95% confidence intervals. Standard errors are clustered by state of birth. Exposure year groups cohorts into four-quarter bins; for example, 1996 includes cohorts born in 1996Q4–1997Q3. The shaded region denotes partially exposed cohorts (exposed before January 1, 1998). Regressions are weighted by IPUMS person weights. From left to right: (i) PUMS 2015–2024; (ii) PUMS 2019–2024 excluding 2020; and (iii) PUMS 2019–2024 with a wider age range of 22 to 32.

Figure B4: Randomization test

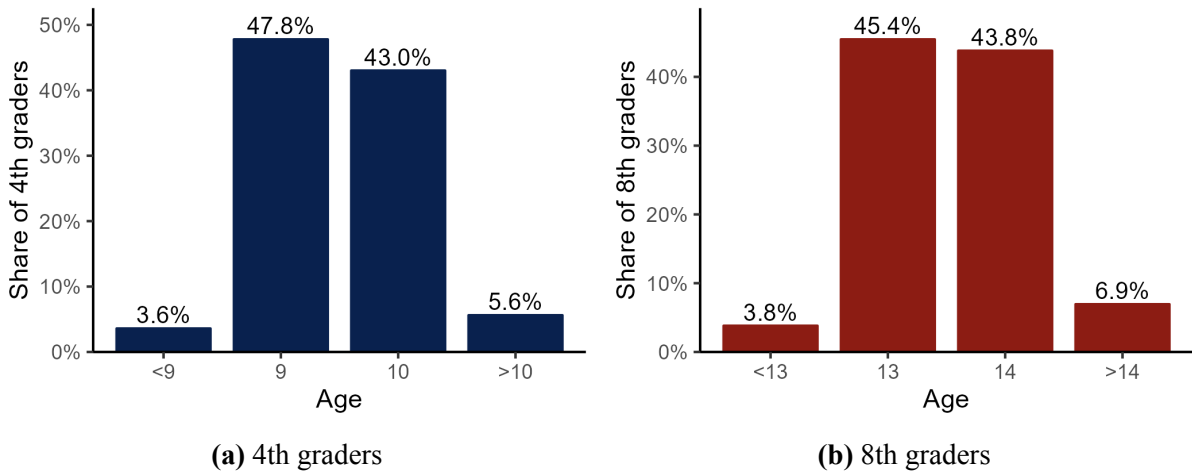**(a)** Grad school enrollment

(b) Labor force participation

Notes: These figures compare actual estimates (vertical dashed lines) with pseudo effects from 1,000 random re-assignments of fortification exposure, preserving its empirical distribution across states. Regressions are weighted by IPUMS person weights; percentile calculations are weighted by birth counts. Shaded areas represent the 5th and 95th percentiles of the simulated null distribution.

Figure B5: Sensitivity to hypothesized violation of parallel trend assumption

(a) Grad school enrollment

(b) Labor force participation

Notes: These figures present Rambachan and Roth (2023) sensitivity tests for the first exposed cohort at the 90% confidence level.

Figure B6: Age distribution of 4th and 8th graders

(a) 4th graders   **(b)** 8th graders

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

Table B1: Long-run effects of folic acid fortification on test scores of 4th and 8th graders, smaller assigned year of birth

	Math			Reading		
	Average score (1)	% ≥ basic (2)	% ≥ proficient (3)	Average score (4)	% ≥ basic (5)	% ≥ proficient
P(High CNSA) × Exposed cohorts	4.477***	3.648***	4.599***	-0.811	1.164	-0.987
	(1.083)	(1.058)	(0.857)	(1.078)	(1.477)	(0.616)
Observations	408	408	408	408	408	408
R^2	0.995	0.945	0.939	0.995	0.951	0.945
Dep. var. mean	259.991	75.502	34.361	241.256	69.351	31.305
Est./Dep. var. mean	1.722%	4.832%	13.385%	-0.336%	1.678%	-3.151%
State-of-school-by-year-of-birth FEs	✓	✓	✓	✓	✓	✓
Grade FEs	✓	✓	✓	✓	✓	✓
Control variables	✓	✓	✓	✓	✓	✓

Notes: This table presents cohort-DiD estimates with standard errors clustered at state of school. P(High CNSA) is fortification exposure adjusted for migration. Exposed cohorts are those with likely year of birth after 1996, determined by assigning age 9 to grade 4 and age 13 to grade 8. Controls include state-of-school-by-(likely)-year-of-birth, grade fixed effects, Medicaid eligibility, exposure to mental-health parity laws, and welfare reforms. The unit of observation is state-by-year-by-grade cells. Regressions are weighted by student counts from ACS PUMS. ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.