IRONING OUT DEFICIENCIES: EVIDENCE FROM THE UNITED STATES ON THE ECONOMIC EFFECTS OF IRON DEFICIENCY[1]

GREGORY NIEMESH
VANDERBILT UNIVERSITY

October 2012

Abstract: Iron deficiency reduces productive capacity in adults and impairs cognitive development in children. In 1943, the United States government issued War Food Order No. 1, which required the fortification of bread with iron to reduce iron deficiency in the working age population during World War II. This universal fortification of grain products increased per capita consumption of iron by 16 percent. I use the exogenous timing of the federal law and cross-place variation in dietary iron consumption before the order to measure the economic impact of the fortification program. Areas with lower levels of iron consumption prior to the mandate experienced greater increases in income and school enrollment between 1940 and 1950. A long-term follow up suggests adults in 1970 with more exposure to fortification during childhood earned higher wages and were less likely to live in poverty. Exposure to fortification did not influence completed years of schooling, suggesting that the increase in earnings was not driven by the quantity of schooling. *JEL* Codes: I12, I15, J24, N32, O10.

Contact Information:
VU Station B #351819
2301 Vanderbilt Place
Nashville, TN 37235-1819
703-909-6336
gregory.t.niemesh@vanderbilt.edu

[1] Research on this project was generously supported by the John E. Rovensky Fellowship, the Vanderbilt University Summer Research Awards Program in Arts and Science, Graduate School Dissertation Enhancement Program, Noel Dissertation Fellowship and Kirk Dornbush Summer Research Grant. I am grateful to William J. Collins, Jeremy Atack, Kathryn Anderson, Douglas C. Heimburger, Hoyt Bleakley, Martha Bailey, Alan Barreca and to discussants and seminar participants at UC-Davis, University of Michigan, Colgate University, the 2012 Stanford Institute for Theoretical Economics, the 2011 Cliometrics Conference, Population Association of America Meetings, the NBER Summer Institute, and the Economic History Association Annual Meetings.

I. Introduction

Micronutrient deficiencies plague the developing world. The World Health Organization estimates that over a quarter of the world's population suffers from iron deficiency, which leads to impaired cognitive development in children and reduced work capacity in adults (McLean et al. 2008). Renewed interest in combating micronutrient deficiencies in developing countries stems from the potentially large impact of health interventions on productivity and quality of life.[2] For instance, the Copenhagen Consensus of 2008 recommends iron and iodine fortification as a highly cost-effective development intervention (Lomborg 2009). Although fortification programs diffused rapidly after the first implementation in the early 1940s, significant potential remains for further gains. As of 2009, 63 countries had implemented flour fortification programs, but 72 percent of world refined flour production remained unfortified (Horton, Mannar and Wesley 2008).

Surprisingly few studies directly evaluate the effects of national-level fortification programs. Of those that do, data limitations and experimental design issues limit their usefulness for policy. Layrisse et al. (1996) find reductions in anemia and deficiency rates from a 1993 Venezuelan program, but no control group was used as a comparison and no economic outcomes were included in the analysis. Imhoff-Kunsch et al. (2007) use expenditure data to measure the *potential* for improvements in health based on reported consumption of fortified foods, but do not directly evaluate health or economic outcomes in response to the intervention. In practice, estimates of benefit-cost ratios for iron fortification typically proceeded by applying productivity estimates from supplementation field trials to prevalence measures from health surveys (Horton and Ross 2003).

This paper uses a sweeping change in federal policy in the United States in the 1940s to estimate both the short-term and long-term effects of fortification on labor market outcomes and human capital development. The discovery of vitamins and minerals during the early 20th century intensified the public health profession's interest in the nutritional status of Americans. A number of diet surveys and blood serum case studies during the 1930s showed a widespread prevalence of deficiencies in iron.[3] Low iron consumption was found in all socioeconomic classes, but the prevalence varied across geographic areas. For example, one study found that the proportion of the

[2] Until the 1990s the World Bank's view had been to rely on the secular increase in incomes to reduce nutritional deficiencies. According to the World Bank (1981), "Malnutrition is largely a reflection of poverty: people do not have enough income for food. Given the slow income growth that is likely for the poorest people in the foreseeable future, large numbers will remain malnourished for decades to come…The most effective long-term policies are those that raise the incomes of the poor." Subsequent research showed that rising incomes at the lower end of the distribution, whether through economic development or income redistribution, do not necessarily lead to immediate decreases in malnutrition (Behrman and Deolalikar 1987).

[3] Kruse et al.'s *Inadequate Diets and Nutritional Deficiencies in the United States* (1943) summarizes the results from a number of blood sample case studies. Stiebeling et al. (1941) summarizes the diet diaries from over 6,800 households surveyed from across the country in 1936.

population considered iron deficient ranged from 47 and 74 percent of white and African-American children in a rural Tennessee county to less than 5 percent of adult male aircraft manufacturing workers in Southern California (Kruse et al. 1943; Borsook, Alpert and Keighley 1943). Taking the case studies as a whole, the United States in the 1930s had rates of iron deficiency similar to those currently found in Turkey or Brazil (McLean et al. 2008). Concerns with worker health and production during World War II finally led to a national fortification program in 1943, but to my knowledge no formal evaluation has tested whether the program led to productivity gains.

In addition to the literature on health and micronutrient fortification in developing countries, this investigation ties to two other branches of research in economics. First, economic historians have linked improvements in nutrition to gains in income and health over the last three centuries (Fogel 1994; Floud, Fogel, Harris and Hong 2011; Steckel 1995). This literature has mainly focused on calorie and protein malnutrition, to the exclusion of the hidden hunger of micronutrient deficiencies. Unfortunately, the evolving and complex interaction of dietary trends, mortality, and income tends to obscure clear causal interpretations of the cotrending relationships in this literature. Second, applied microeconomists have linked health and productivity outcomes in individual-level datasets. A key theme of this literature is that isolating the causal impact of health is difficult but essential (Strauss and Thomas 1998; Almond and Currie 2011). In this paper I follow a strand of this literature that uses targeted public health interventions to estimate the impact of health insults on economic outcomes (Bleakley 2007; Feyrer, Politi and Weil 2010; Cutler et al. 2010; Field et al. 2009).

The paper's central empirical questions are whether places with relatively low iron consumption levels before the program's implementation experienced relatively large gains in labor market and schooling outcomes after the program's implementation, and if so, whether this pattern can be given a causal interpretation. The identification strategy relies on three main elements. First, as shown in diet surveys from the 1930s, there were significant pre-existing differences in iron consumption levels and the prevalence of deficiency across localities. These differences are only weakly correlated with pre-intervention income. Second, the timing of the federal mandate was determined by wartime concerns and technological constraints in the production of micronutrients. In this sense, the timing was exogenous. Finally, iron consumption has a non-linear effect on health. Therefore, a program that increases iron consumption across the entire population is likely to have disproportionate effects on the health of those who were previously iron deficient (Hass and Brownlee 2001).

Evaluating this particular program entails a number of data challenges. The ideal dataset would observe pre- and post-fortification nutrition, health outcomes, and economic outcomes in longitudinal micro-level data. But this ideal dataset does not exist. My approach combines the necessary pieces from a number of different sources. The "Study of Consumer Purchases in the United States, 1935-1936" provides detailed diet records and location information for households (ICPSR, USDOL 2009). I then use the USDA National Nutrient Database (USDA 2009) to convert the diets into the associated nutritional intakes. Labor market and schooling outcomes come from the 1910 through 1950 decennial census microdata (IPUMS, Ruggles et al. 2010). The datasets can be linked at the level of "state economic area" (SEA), essentially a small group of contiguous counties with similar economic and social characteristics circa mid-century.[4]

I find that after the iron fortification mandate in 1943, wages and school enrollment in areas with low iron intake did increase relative to other areas between 1940 and 1950. The regression results are generally robust to the inclusion of area fixed effects, regional trends, demographic characteristics, World War II military spending, and Depression-era unemployment. One standard deviation less iron consumption before the program is associated with a 2 percent relative increase in male wages from 1940 to 1950, and a 1 to 1.5 percentage point relative increase in school enrollment. These estimates are economically significant, accounting for 4 percent and 25 percent of the gains in real income and school enrollment over the decade in areas that started below the median level of iron consumption. I estimate a cost-benefit ratio of at least 14:1, which is within the range for those estimated in developing countries (Horton and Ross 2003).

I use the 1970 census microdata to undertake a separate analysis that suggests that iron deficiency had a lasting long-term impact on human capital formation and wages. The cohort analysis measures differences in childhood exposure to fortification by combining differences in years of potential exposure (based on year of birth) with geographic differences in pre-existing rates of iron deficiency (based on place of birth). Cohorts with more exposure to fortification had higher earnings and were less likely to be considered living in poverty by the census. Moving from no exposure to a full 19 years of exposure implies a 3.6 percent increase in earnings as an adult and a decrease in the likelihood of living in poverty by 0.48 percentage points. Increased quantity of schooling does not drive the increase in adult incomes.[5]

---

[4] The State Economic Area (SEA) is a concept used by the Bureau of the Census. An SEA consists of either a single county or a group of contiguous counties in the same state with similar economic characteristics shortly before the 1950 census. See Donald J. Bogue's "State Economic Areas" (1951) for a full description of the procedure used to group counties.

[5] Assessing the impact with randomized trials would prove difficult. The cost of tracking infants into adulthood and the ethical concerns about withholding treatment over an extended period time would both prove prohibitive. An historical accident provides the necessary variation in exposure during childhood for the current analysis.

II. Iron Deficiency and the Fortification of Flour and Bread

II.A. Health Effects of Iron Deficiency

Iron deficiency is the most common nutritional deficiency worldwide and is caused by low dietary intake, blood loss, growth, pregnancy, and impaired absorption. Iron has two main functions in the body: to transport oxygen throughout the body in the bloodstream, and to process oxygen in the cells of muscles and tissue.[6] A lack of iron causes reduced work capacity through a diminished ability to move oxygen throughout the body and a reduction in the tissue cell's ability to process oxygen. The reduction in oxygen manifests as reduced aerobic capacity, endurance, energetic efficiency, voluntary activity and work productivity (Hass and Brownlee 2001).[7] A lack of iron also affects productivity by reducing cognitive ability and skill acquisition.

Iron deficiency in infants and children causes developmental delays and behavioral disturbances, including decreased motor activity, social interaction, and attention (Beard and Connor 2003). Studies that follow the same children over time have found that iron deficiency can have long-lasting effects on neural and behavioral development of children even if the deficiency is reversed during infancy (Lozoff et al. 2006). Between the ages of 12 and 18, adolescents are at higher risk because of increased iron requirements. The risk subsides by the end of puberty for males, but menstruation keeps the risk high for women throughout the childbearing years. In treatment-control studies on subjects with iron deficiency or anemia, cognitive ability and work capacity in adolescents treated with iron-therapy improved relative to the placebo group (Groner et al. 1986; Sheshadri and Gopaldas 1989; Seomantri et al. 1985). Poor health during childhood can lead to reduced educational investment. Bobonis et al. (2006) find an economically large impact of iron supplementation on pre-school participation in a developing country context with a 69 percent baseline rate of iron deficiency.

II.B. The Fortification Movement and Federal Mandate

Before the intervention in the early 1940s, iron in the U.S. food supply was gradually declining as consumers reduced grain consumption and increased sugar and fat consumption (Gerrior, Bente and Hiza 2004). While acknowledging that the diets of many Americans were deficient in micronutrients in the early 20th century, the medical profession and regulatory authorities

[6] Daily iron requirements vary significantly by age and sex. The recommended daily allowance (RDA) for adult men is 8 mg per day, whereas the RDA for non-pregnant women of childbearing age is 18 mg per day. No differences in requirements exist between the genders during childhood, but requirements increase during periods of growth. The RDA from childhood to puberty ranges from 7 to 11 mg per day (Institute of Medicine 2001).

[7] For example, experiments have shown that Sri Lankan tea pickers are more productive when not suffering from a deficiency of iron (Hass and Brownlee 2001).

were initially steadfast in their opposition to the addition of any foreign substances to food products (Wilder and Williams 1944). The Food and Drug Administration (FDA) and the American Medical Association (AMA) reversed views in the 1930s during the debates over whether to allow vitamin D fortified milk as a tool to prevent rickets. In an important step for proponents of iron fortification, the AMA backed proposals to enrich bread and flour with iron and thiamin in 1939 (Bing 1939).

In May 1941, the FDA enacted regulations specifying the labeling of "enriched" wheat flour. No standard for bread was promulgated at the time. This did not require manufacturers to fortify their products, only that to use the label "enriched flour" the product must contain between 6 and 24 milligrams of iron and niacin and 1.66 to 2.5 milligrams of thiamin in each pound of flour (Federal Register 1941). These levels represent a *doubling to tripling of the micronutrient content of unenriched products.* Two years later, the FDA increased the minimums and maximums for enriched flour (Federal Register 1943). During this period, the National Research Council (NRC) promoted enrichment on a voluntary basis for bakers and millers. Anecdotal evidence from the NRC archives suggests 75 to 80 percent of flour and bread was voluntarily enriched by 1942 (Wilder and Williams 1944). Most parts of the U.S. had very high participation rates, but the South lagged behind, enriching only 20 percent of the flour consumed.

The first federal *requirement* to fortify bread came in War Food Order No. 1 in 1943, which mandated fortification at the "enriched" levels.[8] The mandate had a large, abrupt, and long-lasting impact on iron consumption in the United States. Figure I shows a 16 percent increase in the iron content of the U.S. food supply in the early 1940s, which is directly linked to the fortification of flour and bread.[9] By 1950 fortification provided 22 percent of all iron in the food supply, adding 2.7 milligrams daily. This increase by itself is 34 percent of the recommended daily allowance for men and 15 percent for women. The long secular decline in average iron consumption would have continued unabated into the 1950s in the absence of the intervention.

III. The Theoretical Impact of Fortification on Labor Market and Schooling Outcomes

This section briefly discusses a theoretical framework to clarify our intuition and interpret the empirical results. At the heart of this paper is health's role as an input into productivity and human capital accumulation. As such, health enters the production function for academic skill, wage equations, or labor supply choice as an input. I begin by defining how health is produced by the following production function:

[8] Flour fortification continued on a voluntary basis.
[9] The USDA constructed this series by using the disappearance method – production plus imports minus farm use and exports (Gerrior, Bente and Hiza 2004).

(1) $H = H_P(C, M, E, \eta)$

where C is consumption, M is medical inputs, E is the health environment, and η is innate healthiness with all inputs entering positively. Consumption is broadly defined and may include exercise, healthy eating, smoking, and alcohol consumption. I model micronutrient fortification as a positive shock to the health environment (E) instead of a reduction in prices of medical inputs or a direct shock to health.[10] The effect of an improvement in the health environment is then translated through an appropriate production function into an economic outcome.

Experiments in the lab often estimate a parameter that approaches the marginal effect of health in the production function for productivity by directly measuring physical work capacity (Rowland et al 1988; Zhu and Hass 1998; Perkkio et al. 1985). More commonly, a field study provides supplementation for a short period of time and then observes productivity in jobs where output can be directly measured (Edgerton et al. 1979; Ohira et al. 1979; Gardner et al. 1975). When these estimates are then used to measure cost-effectiveness of a policy, potential behavioral responses are not taken into account, but they should be.[11] People have the choice to adjust inputs along all margins in response to treatment while facing countervailing income and substitution effects. Behavioral responses can either attenuate or strengthen the effect beyond that of the direct structural effect of the production function relationship. In any case, policymakers usually are not ultimately interested in the structural effect per se; rather, the impact of a health intervention on economic outcomes after taking into account all behavioral adjustments is most useful for informing policy. The purpose of the model is to show how the results are filtered through a model of household choice. Thus, the estimated effect of iron fortification on school enrollment and income should be interpreted while keeping in mind that adjustments may be made along other margins. I will first discuss the model for childhood human capital acquisition, followed by the model for adult labor market outcomes.

III.A. Theoretical Impact on Children

The discussion that follows for childhood human capital acquisition closely follows that found in Glewwe and Miguel (2008). This one-period model illustrates many of the issues that arise in estimation even though it is quite simple. A child's skill is created using the following production function:

[10] A direct shock to the health stock abstracts away from adjustments to medical inputs and child consumption. I choose not to model the intervention as a reduction in the price of medical inputs because I do not see any evidence of families adjusting consumption of bread in response to enrichment.
[11] Thomas et al. (2006) account for a number of behavioral adjustments in response to a iron supplementation randomized control trial in Indonesia.

(2) $T = T_P(H, I, S, Q, \alpha)$

Skill increases with health (H), parental investment in the child's skill (I), years of schooling attained before entry to the labor market (S), school level inputs that parents do not control (Q), and innate ability (α). We can now see how the structural effect of fortification on skill would be produced. An improvement in the health environment (E) in equation (1) would directly increase health, which then enters equation (2) to increase skill.

The model requires a utility function to allow the actors to make adjustments in response to fortification. Parents make all decisions concerning investments in the child's human capital. The parents' utility includes their own consumption C, child health H, and the final skills of the child T:

(3) $U = U(C, H, T)$

Parents do not value their child's school enrollment directly, but only through its influence on skills.

In practice, I observe only years of schooling and enrollment status in the data. To gauge the effects on the demand for years of schooling, we must first determine the impact on demand for child skills. While it is clear this will be positive, the effect on demand for years of schooling is ambiguous because of the various behavioral responses available to parents. The total effect of a change in the health environment on the demand for academic skills can be decomposed into the following parts, where a subscript D denotes a demand function:

(4) $$\frac{\partial T_D}{\partial E} = \frac{\partial T_P}{\partial H}\frac{\partial H_P}{\partial E} + \frac{\partial T_P}{\partial H}\frac{\partial H_P}{\partial M}\frac{\partial M_D}{\partial E} + \frac{\partial T_P}{\partial I}\frac{\partial I_D}{\partial E} + \frac{\partial T_P}{\partial S}\frac{\partial S_D}{\partial E}$$

The first term in equation (4) is the direct effect of a change in the health environment on the demand for child skills. A positive shock to the health environment will directly increase the health status of the child through the health production function, and in turn the demand for skills. The parents experience a positive income effect from the shifting out of the production possibilities of child health and child skills. The parents can now choose among a wider range of choices among all three inputs in the utility function. The remaining terms arise from the behavioral response of the parents taking the opportunity to adjust medical inputs, parental investment in child skills, years of schooling, and parental consumption.

The second term illustrates how marginal adjustments to medical inputs affect demand for child skills. How parents change expenditures on medical inputs in response to a change in the health environment is ambiguous. For example, iron fortification can make iron supplements obsolete (i.e., $\frac{\partial M_D}{\partial E} < 0$). Alternatively, reduced iron deficiency might make some medical treatments more productive, reversing the sign. In an iodine supplementation RCT in Tanzania, Adhvaryu et al. (2012) finds that supplementation induces parents to increase other health investments in the child. Children that received supplementation *in utero* were more likely to receive vaccines and were breastfed longer.

The final two terms arise from adjustments to parental investments in child skills and years of schooling. Again, the sign of these two terms is ambiguous. We do not know how changes in the marginal utility of the three inputs and the marginal product of educational inputs and years of schooling respond to a change in the health environment. Income effects cause the parent to reallocate resources away from academic skills production and towards consumption. Price effects, on the other hand, work in the opposite direction. A health improvement raises the marginal productivity of investment in skills and years of schooling, reducing the shadow price of child skills. The rise in demand for skills can only be satisfied by increasing the demand for parental investment and years of schooling.

An examination of the demand function for years of schooling sheds light on the potential issues that arise during estimation. The demand for years of schooling can be expressed as:

(5) $S = S_D(W; r, p_C, p_M, p_I; E, Q; \alpha, \eta, \sigma, \tau)$

where W is parental wealth, r is the interest rate, p_X is price of input X, E is the health environment, Q is school level resources such as quality, P is parental schooling, α is innate ability in production of academic skills, η is innate healthiness, σ is a parental preference parameter for child academic skills, and τ is a parental preference parameter for child health. The estimation of $\frac{\partial S_D}{\partial E}$ potentially suffers from omitted variable bias if any of the parameters in (5) are correlated with E and unobserved by the econometrician.

III.B. Theoretical Impact on Adults

The relationship between adult productivity and health is complicated by feedback effects and simultaneity issues. Simultaneity of health, labor supply, and wages drive the important differences between the child schooling model and the adult labor market model. The following discussion closely follows the theory outlined in Strauss and Thomas (1998). Health is again, produced by the following production function:

(6) $H = H_P(M, L; A, E, \eta)$

Health is increasing in health inputs (M) and decreasing in labor supply (L), as work may be taxing on health. Health is likely to vary with socio-demographic characteristics such as age or gender, as well as parental background in terms of either parental health or parental education, all of which are collected in vector A. The health environment (E) mirrors that in equation (3). Wages follow a function that increases in health, and depends on observable socio-demographic characteristics (A) and educational attainment (S) and unobservable characteristics (α):

(7) $w = w(H; A, S, \alpha)$

Labor supply depends on health and wages, the price for consumption, observable characteristics, and unobservable characteristics (ξ):

(8) $L = L(H(L), w(H), p_C, S, A, \xi)$

If the health environment for adults is improved through fortification, then the consumption set expands just as in the child model. Adults can adjust their consumption, use of medical inputs, and labor supply in response to the improvement in the health environment.

The direct effect of improved health on wage (income) is positive. However, the total effect on wage (income) will typically be different from the direct effect because of adjustments to medical inputs and labor supply. Moreover, the labor supply response is ambiguous because of the typical income and substitution effect of a change in wage, but also because health enters the functions for labor supply and wages, and labor supply enters the health production function. Whether the income or price effects dominate the labor supply choice is difficult to intuit because of the simultaneous determination of health, wages, and labor supply.

IV. Identifying the Economic Impact of Iron Deficiency

The implementation of the U.S. food fortification program may provide plausibly exogenous variation in health improvements during the 1940s. The empirical strategy relies on three key elements: pre-existing differences in iron consumption and prevalence of deficiency, exogenous timing of the federal mandate, and the non-linear effect of iron consumption on health. Assigning a causal interpretation to the partial correlations between pre-program iron consumption and subsequent outcomes requires the absence of unobserved shocks and trends to outcomes that are correlated with pre-program iron consumption. Such shocks and trends are impossible to rule out completely, but further investigation suggests that the identifying assumption is tenable and that the regression estimates are likely to reflect a causal relationship.

First, diet surveys and blood sample case studies demonstrated the widespread but uneven prevalence of iron deficiency across the country in the 1930s. There was considerable variation across places, even within regions. A key concern is whether pre-program iron consumption is highly correlated with income, which could indicate a severe endogeneity problem. Figure II demonstrates that a strong relationship does not exist between iron consumption and income during the pre-program period at the SEA level.[12] A regression that controls for differences in black proportion of the population, home ownership, farm status, and the local Gini coefficient also shows a weak relationship between income and iron consumption. In sum, current local economic characteristics are not strong predictors of pre-program iron consumption, and in any case, regressions below will include controls for income in 1936, local fixed effects, regional trends, and more.[13]

The second piece of the identification strategy comes from the timing of the federal mandate. Its passage was external to what was going on in the low iron consumption areas. Fortification was mandated at the federal level in response to wartime concerns, reducing the scope for states, counties, and individual consumers to select into or out of treatment. Moreover, technological constraints made earlier implementation infeasible. The mandate clearly, quickly, and significantly increased per capita iron consumption by 16 percent (Gerrior, Bente and Hiza 2004).

Finally, iron consumption has a non-linear effect on health. Above a certain threshold additional iron consumption provides no health improvement; the excess simply gets filtered out of the body. Furthermore, severity of deficiency matters. Gains in health per unit of additional iron are proportional to the extent of the deficiency (Hass and Brownlee 2001). A non-linear effect implies that those with low pre-program iron consumption likely experienced larger improvements in health from fortification, even if fortification raised everyone's intake of iron. This non-linearity at the individual level translates into differential changes in deficiency rates at the level of geographic areas. Figure III shows that state economic areas with relatively high levels of iron deficiency before the intervention experienced relatively large declines in deficiency after the intervention.

The federal mandate, thus, provides a quasi-experiment in which the "treatment effect" varies across areas based on pre-intervention iron consumption. With census microdata, I estimate equations at the individual level of the following general form,

$$Y_{its} = \beta \cdot (IRON_s \times POST_t) + \delta_t + \delta_s + (\tilde{\delta}_r \times t) + X_{it}\theta_1 + X_{st}\theta_2 + \varepsilon_{its} \qquad (9)$$

[12] I use data from the "Study of Consumer Purchases" on diet, income, and demographic observables to explore the variation in iron consumption across areas. These data are discussed in more detail later in the paper.

[13] A significant portion of geographic variation in iron consumption remains unexplained, given that income and demographic characteristics are not good predictors. Contemporaneous price variation explains only half the variation. A full explanation of the variation in iron consumption can be thought of as an answer to "Why do people eat what they eat?", a question outside the scope of this paper. However, the inability of contemporaneous economic variables, especially income, to explain variation in diets is not a novel finding (Behrman and Deolalikar 1987; Logan and Rhode 2010). Relative prices from the past predict food consumption 40 years after the fact due to the process of taste formation (Logan and Rhode 2010). Research on the psychology of food and tastes provides further evidence for the importance of learning and the persistence of diets (Capaldi 1996). In sum, a number of hypotheses exist in economics and psychology that explain dietary choices based on factors other than current income and prices.

where Y signifies a labor market or schooling outcome, and δ_t and δ_s signify a set of year dummies and state economic area (SEA) indicators. Some specifications include a geographic-area-specific linear time trend $(\tilde{\delta}_r \times t)$ at the level *r*, where *r* denotes census divisions or SEAs depending on the span of time observed.[14] All regressions include a vector of individual-level controls denoted by X_{it}, with some specifications also including a vector of area-specific controls denoted by X_{st}.

The coefficient of interest is β. The variable $IRON_s$ denotes the pre-intervention average iron consumption in area *s*. Each individual is assigned the average pre-program iron consumption level for the SEA of residence. The variable $POST_t$ denotes an indicator equal to one if year t is after the intervention date of 1943. Interacting the two gives the variable of interest.[15] The hypothesis is that areas with low iron consumption before the intervention experienced larger health benefits from fortification and therefore larger gains in labor market and schooling outcomes. *If the hypothesis is correct, then estimates of* β *should be negative.*

Potential threats to a causal interpretation of estimates of β include a trend in or unobserved shock to the outcome that is correlated with pre-program iron consumption and is not absorbed by the control variables (X_{it} and X_{st}) or place-specific trends. I explore potential confounding factors further in the respective estimation section for each outcome.

V. Diet Data and Pre-existing Differences in Iron Consumption

V.A. Pre-intervention Iron Intake and the Study of Consumer Purchases

I use the wealth of household diet information contained in the "Study of Consumer Purchases in the United States, 1935-1936" to calculate household iron consumption and deficiency, as well as average iron consumption for states and state economic areas. The Data Appendix provides an in-depth discussion of the process used to construct the iron consumption measure. The food schedule portion of the survey provides a detailed account of the types, quantities, and cost of all foods consumed over seven days for a sample of 6,800 households from across the United States. The survey contains a surprising amount of detail on the food purchase and consumption patterns of the respondents: over 681 individual food items are recorded, as well as the number of meals

---

[14] When income is the outcome of interest, the place-specific trends cannot be specified at the SEA level because the census started inquiring about income in 1940. This restricts analysis to a two-period comparison (1940 and 1950). However, β can still be indentified in regressions that include census-division trends (there are 9 census divisions). The census has collected information on school attendance over a longer time span, which allows separate identification of SEA-specific trends and β.

[15] There is no coefficient estimate for $IRON_s$ because the regressions include area fixed effects.

provided for each member of the household. The survey included families in 51 cities (population of 8,000 and up), 140 villages (population of 500 to 3,200), and 66 farm counties across 31 states.

Each household diet is converted to iron intake by employing the USDA National Nutrition Database (USDA 2009). I calculate the average daily per person iron consumption for each household using the number of meals provided by the home. A daily measure of iron intake simplifies comparisons to the recommended daily allowances published by the USDA. Summary statistics are reported in table I. Average daily iron consumption for the sample is 10.2 mg, with a median of 10.4 mg. Note that the recommended daily allowance for men of working age is 8 mg and for women is 18 mg. About 68 percent of all households in the sample consume less than a household specific RDA determined by the age and gender mix of the household. Because the effects of iron deficiency are non-linear, I also report the proportion of the sample that consumes less than 75 percent and 50 percent of the household specific RDA. A surprisingly high proportion of the sample consume less than these lower cutoffs, 40 percent and 12 percent respectively. Assuming that insufficient intake translates to deficiency, then the United States of 1936 experienced similar rates of iron deficiency as Turkey or Brazil today.

Substantial variation exists across households in the daily consumption of iron. Figure IV plots the pre-intervention distribution of household per capita daily consumption of iron in the 1936 diets. No similar survey was conducted after the iron fortification program in the 1940s, and so I cannot construct a figure by applying the same procedure to later diets. Instead, to see whether the program plausibly affected iron consumption throughout the distribution, I construct an estimated iron distribution by applying the "enriched" iron levels to bread and flour consumption in the 1936 diets. Figure IV plots this estimate of the post-intervention distribution for comparison with the original 1936 distribution. Taking diets as given, fortification strongly shifts the distribution to the right, including significant gains for those who were originally at the lower left tail of the distribution. Average consumption of iron increases by 3.8 mg from 10.2 mg to 14 mg, and the proportion of households in the sample predicted to consume less than the recommended daily allowance declines from 68 percent to 23 percent.[16] In sum, because store-bought bread and flour were such common elements in Americans' diets, it is highly likely that the fortification program significantly boosted iron consumption throughout the distribution.

A later USDA dietary consumption survey from 1955 provides further evidence of the broad impact fortification had on the American diet.[17] By at least 1948, nearly all of the white bread

[16] The increase in iron consumption here is larger than that using the total U.S. food supply estimates because grain consumption declined by 20 percent from 1936 to 1950. Without this decline in grain consumption the two estimates would be similar.

[17] I rely on published reports from the 1955 survey as the microdata was subsequently destroyed.

purchased by the American public was enriched (USDA 1961). In 1955, survey respondents reported consuming enriched bread products across all regions, income groups, and urbanizations. Table II shows that substantial quantities of enriched grain products were consumed in all regions of the country despite concerns about compliance in the South. All areas of the United States were able to, and did, purchase enriched bread.

The fortification program reached those most in need of treatment. Increases in iron consumption occurred along the entire income distribution. In fact, the lower third of the income distribution experienced larger absolute and percentage increases in consumption between the 1936 and 1955 surveys. The counterfactual is stark. Average iron consumption would have been 20 percent less in 1955 without the enrichment program. By 1955, 90 percent of households met the RDA for iron and 98 percent of households consumed at least two-thirds of the RDA (USDA 1961).

To facilitate merging with census microdata and to capture the geographic variation in iron consumption before fortification, I calculated the mean daily iron consumption over households within each SEA. The required conversions to construct household iron consumption leave the sample with 3,545 observations across 82 SEAs in 30 states. The mean SEA consumes 10.5 mg of iron per person daily, with a standard deviation of 1.8 mg. In just under half of the SEAs the average household consumes less than 10 mg per day. The prevalence of deficiency varies significantly across SEAs and within regions. Figure V maps the variation across states in the proportion of the sample that consumes less than the household specific recommended daily allowance. Significant variation exists within and across census divisions.

V.B. Pre-existing differences in iron consumption are distributed quasi-randomly

The identification strategy relies on the assumption that pre-existing geographic differences in iron consumption across SEAs are not correlated with SEA heterogeneity in omitted characteristics that induce changes in the outcome. I conduct a direct, albeit partial, test of the identifying assumption by individually regressing SEA average iron consumption on several pre-program SEA characteristics. The assumption is supported if these characteristics do not predict iron consumption, and the exercise suggests specific controls if characteristics do have predictive power.[18] Point estimates provided in table III suggest that iron consumption is uncorrelated with several economic, labor market, agricultural, and demographic characteristics. For example, per capita war spending, New Deal spending, World War II mobilization rates, retail sales and manufacturing

[18] New Deal spending, retail sales, weather and migration data were compiled for Fishback, Kantor and Wallis (2003) and Fishback, Horrace and Kantor (2005, 2006). Copies of the data sets can be obtained at the following website: http://www.u.arizona.edu/~fishback/Published_Research_Datasets.html. All other data is from Haines (2010).

output are essentially unrelated to the pre-existing differences in iron consumption. Exceptions include the unemployment rate, fraction of the population native born, net migration, and growth in median home values over the 1930s. The results, therefore, suggest adding specific controls to the regressions in the following empirical sections. The empirical strategy provides plausibly exogenous variation in health improvements during the 1940s when specifications include the additional controls.

V.C. Outcome Data

All individual-level outcome data and demographic controls come from the Integrated Public Use Microdata Series (IPUMS, Ruggles et al. 2010), a project that harmonizes decennial census microdata. The basic specification uses census data from 1940 and 1950 as these years bracket the iron fortification program. Income data are limited to 1940 and 1950 as these are the only census years that provide both income and SEA identifiers. For the school enrollment regressions, additional data from 1910 through 1950 are used to control for gradually evolving SEA specific unobserved characteristics. Panel B of table I contains summary statistics for outcome variables, and the Data Appendix provides a more detailed discussion.

VI. Fortification's Effects on Contemporaneous Adult Labor Market Outcomes

In this section, I estimate the changes between 1940 and 1950 in income and labor supply that are associated with increases in iron consumption.[19] Table IV presents point estimates for β, the coefficient on $(IRON_s \times POST_t)$ from equation (9).[20] Each entry is from a separate regression. Because areas with low iron consumption experience larger relative gains, the coefficient is expected to be negative if the hypothesis is correct. Standard errors are clustered at the SEA-by-year level to allow for a common shock to income at the local level.[21]

[19] The full sample includes wage and salary workers aged 18 to 60 with positive income. I exclude observations without educational attainment information or that are recorded as full-time or part-time students from the income and weeks regressions. Because the 1940 census only inquired about wage and salary income, I exclude observations that list the main class of worker status as self-employed.

[20] I choose to use iron consumption to measure the area specific intensity of treatment for two reasons: it facilitates the calculation of a reduced-form effect of the program as a whole because we know how much consumption increased following the mandate, and biomedical researchers have recently turned to continuous measures of hemoglobin and serum ferritin instead of cutoffs. Response to treatment occurs even if the patient remains below the anemic cutoff, and functional decrements continue after falling below the anemic cutoff (Horton, Alderman, and Rivera 2009). Appendix table A.I reports point estimates for β from equation (9) using alternative measures of treatment intensity. Differences in the estimated effect compared to table IV are negligible.

[21] Standard errors are clustered at the SEA by year level according to the procedure developed by Liang and Zeger (1986). Correlation of unobserved shocks to individuals within the same SEA in the same year is the main concern. Serial correlation does not pose a serious problem as the time periods in the panel are separated by ten years

VI.A. Income Gains for Men[22]

Row (A) provides results from the base specification. The regression includes a census-division time trend, as well as individual-level indicators for industry, occupation, veteran status, marital status, race, four educational attainment categories, and an age quartic interacted with education category.[23] The point estimate suggests that iron fortification led to statistically and economically significant relative gains in income over the 1940s for men in areas with lower iron consumption. The result is robust to regional convergence in wages over the decade.

For a public health program that was relatively inexpensive at 0.50 dollars per capita annually (Wilder and Williams 1944), the economic impact on wage income alone is impressive. The base estimate for men suggests a one-standard deviation difference (1.8mg) in SEA average iron consumption was associated with a 2.6 percent difference in income growth.

Robustness Checks

Potential threats to a causal interpretation of β remain in the form of unobserved SEA-specific shocks to income that are correlated with iron consumption. For example, heterogeneity in local labor market conditions due to wartime spending and mean reversion from the depths of the Great Depression (perhaps) could be correlated with iron consumption in the 1930s. Similarly, a temporary negative shock might simultaneously cause low iron consumption and low income in 1936. As the temporary shock dissipates we would expect income gains correlated with low iron consumption even if fortification had no effect. To reduce the scope for such omitted variable bias, row (B) table IV includes the 1937 SEA-level unemployment rate, 1936 SEA average income, per capita World War II spending, total New Deal expenditures, net migration over the 1940s, and median house price growth, all interacted with $POST_t$. The results are little changed, suggesting that conditional on the control variables, geographic differences in pre-program iron consumption are uncorrelated with omitted heterogeneity causing income gains.[24] Using the estimate from row (B)

(Bertrand, Duflo and Mullainathan 2004). In the full sample, the number of clusters = 164. Results from regressions that aggregate to the SEA level using the procedure developed by Donald and Lang (2007) are consistent with those of the microdata regressions.

[22] I split the analysis by sex based on the radically different labor market incentives facing men and women during the 1940s. I focus the analysis on men in the main text because the sample sizes for working women are small. In general, unmarried women behave as if fortification had positive effects on productivity, however, married women do not. The appendix contains a more in depth analysis of the labor market impacts for women.

[23] The four educational attainment categories are less than high school, high school, some college, and college.

[24] Point estimates are similar when limiting the sample to native-born or foreign-born men. Cross-state migration does not explain the results. Regressions limiting the sample to native-born men residing in their state of birth provide identical point estimates to those in row (B). Additionally, the wage distribution compressed during the 1940s (Goldin and Margo 1992). Point estimates from adding the SEA Gini coefficient for male wage and salary

with the full set of controls, a one-milligram difference in average iron consumption is associated with a 1.03 percent difference in wage and salary income growth. Increasing SEA average iron consumption by 2 mg translates into a 2 percent increase in income between 1939 and 1949.[25] For perspective, this would account for 4 percent of the total income growth over the decade in the areas below the median of iron consumption.

A number of public health programs were conducted in the southern states simultaneously with the bread enrichment program. Hookworm and malaria eradication efforts continued into the 1940s in parts of the South. Moreover, some southern states allowed for voluntary enrichment of bread and flour with B vitamins starting in 1938. Surveys, however, indicated southern bakers and millers did not participate in the voluntary programs (Wilder and Williams 1944). Row (C) drops the southern states from the sample to limit identification of the effects of iron fortification to variation from states in the Northeast, Midwest, and West census regions. Dropping the South has little effect on the point estimate.

The choice of industry and occupation may be endogenous to a change in health caused by increased iron consumption. For example, a worker might upgrade to a higher paying occupation or industry because of increased endurance or work capacity. The coefficient estimates from regressions without controls for occupation and industry are essentially unchanged from before, as reported in row (D). Thus, the relative gains in income do not appear to be caused by occupational upgrading. A regression using the IPUMS occscore variable as the dependent variable gives a similar interpretation (results unreported).

Important differences appear between age groups. Rows (E) through (G) split the male sample into 10-year age categories. In general, younger men appear to have experienced a larger increase in income from iron fortification. A one standard deviation difference in iron consumption at the SEA level implies a 4.7 percent difference in income growth for men under the age of 28 and a 1.4 percent difference for men between the ages of 28 and 37. The impact of iron fortification fades at older ages., although estimates are noisy.

American diets underwent substantial changes during the 1940s in response to rationing and a large demand for food on the part of the U.S. military and allies. These changes, however, were not long lasting. Diets returned to their pre-war patterns shortly after rationing was discontinued in 1946. While fluctuations in the consumption of non-enrichment nutrients briefly improved overall

---

income suggests that the income gains correlated with low iron consumption are not explained by a compression of wages during the 1940s. However, this measure captures within-SEA compression of the wage distribution, but not wage compression across SEAs.

[25] The fortification program increased per capita daily iron consumption by 2 milligrams (Gerrior, Bente and Hiza 2004)

nutritional status, they do not appear to drive the empirical results I find in this section. The appendix provides evidence from the medical literature, the diet survey data, and regressions that include other micronutrients as explanatory variables.

Besides iron, enriched bread contained added amounts of niacin and thiamin, increasing per capita daily consumption of both vitamins during the early 1940s. An attempt to tease out the impact of each micronutrient individually was inconclusive.[26] Niacin, iron, and thiamin consumption is highly correlated at the individual level ($\rho = 0.83$) and SEA level ($\rho = 0.85$). Moreover, niacin or thiamin deficiency essentially implies iron deficiency. Iron inadequacy is much more prevalent than niacin and thiamin inadequacy in the sample, 68 percent versus 25 and 12 percent. Interpreting the reduced-form estimates of the program impact as coming solely from niacin would imply an unreasonable individual effect. Consequently I focus on iron deficiency, but at the very least the results can be interpreted as the total effect from reductions in deficiencies of all three micronutrients.

VI.B. Labor Supply of Men

The above regressions clearly point to relative gains in income over the 1940s correlated with low pre-program iron consumption for young male workers. Iron fortification may have promoted these gains through a number of potential channels: labor supply could change at the intensive or extensive margins, or productivity per unit of time could rise. As discussed in section III in relation to the theoretical model, all three channels may have worked simultaneously but not necessarily in accordance with each other. I attempt to shed light on these issues by conducting separate regression analyses for labor force participation, weeks worked, and hours worked.

Columns (2) and (3) of table IV offer a direct assessment of the labor supply channels. Labor force participation by men in 1940 was already quite high and had little room for improvement. Thus it is no surprise that only small changes on the extensive margin of work were associated with pre-intervention iron consumption. The results are generally consistent with small relative *decreases* in male labor force participation rates in areas with low iron consumption. To get a sense of the economic significance, a one-standard deviation difference in iron consumption implies a 0.28 percentage point difference in labor force participation. In 1940 and 1950, the average labor force

---

[26] Less than one percent of observations that are thiamin deficient and less than 5 percent of observations that are niacin deficient are not also deficient in iron. High collinearity notwithstanding, I regress income and school enrollment on niacin, thiamin, and iron consumption individually and combined. When included separately, the reduced form results are all similar due to the high levels of correlation between the measures. When included together, the results are inconclusive. For the specification including all three measures, the point estimates and standard errors are -0.014 (0.010) for iron, -0.003 (0.004) for niacin, and 0.005 (0.038) for thiamin from the income regression.

participation rate for men aged 18 to 27 was 87 percent and 81 percent respectively. Adjustments along the extensive margin do not seem prevalent and are unlikely to drive the results in column (1) of table IV.

The census inquired about labor force participation in 1920 and 1930, allowing me to include an SEA specific linear trend in an estimation of equation (9). Using the full set of data for 1920 to 1950 give similar estimates to those in table IV for β with and without including SEA specific linear trends. The additional data also allows for a replication of the two-period analysis using years prior to the intervention. As expected, replicating the regressions using data from 1920-1930 and 1930-1940 provide no statistically or economically significant estimates for β.

Changes at the intensive margin are explored using the weekly hours and weeks worked variables. Conditional on working positive hours, changes in weekly hours are not correlated with pre-intervention iron consumption (results unreported). Column (3) of table IV reports point estimates from regressions with a continuous measure of weeks worked as the dependent variable. The results are broadly consistent with relative increases in weeks worked correlated with low pre-intervention iron consumption, although the estimates are noisy. The point estimate for young male workers is again the largest, corresponding to an increase of roughly one-quarter of a week of work or 0.6 percent of the 1940 mean. Evaluated at the mean of weeks worked in 1940 (43) for the full sample and for a one-milligram increase in iron consumption, gains on the intensive margin of labor supply account for one-third of the total increase in income. In general, it appears that men responded to reductions in iron deficiency by adjusting labor supply along the intensive margin. However, changes in hours and weeks worked do not explain the full effect of iron fortification on income.

VI.C. Interpretation of Men's Labor Market Results

As an external check to the validity, I compare the results to those found in the development and medical literatures, which report average treatment effects on the treated for patients *pre-identified as anemic or iron deficient*. My results pertain to an aggregate level effect on the total population, not solely anemic patients.[27] I convert the average aggregate result to a parameter similar to an "average treatment effect on the treated" or "an intent to treat effect" for an iron deficient individual.

[27] Evidence suggests that patients with subclinical iron deficiency receive benefits from iron supplementation (Horton and Ross 2003). The point estimates in table IV include gains to sufferers of iron deficiency at stages less severe than anemia. They also include adjustments along the intensive margin of work and other health inputs.

The fortification program differentially affected individuals, with larger benefits accruing to those with lower initial levels of iron. I assume that only those individuals consuming less than 75 percent of their RDA experienced gains in health from the program, and thus gains in income. Concentrating the full reduced-form impact of the program onto this portion of the population gives the average income gain to the iron deficient individual. Using the results from row (B) of table IV, a difference of 1 mg in pre-intervention iron consumption implies a differential 1 percent gain in income. On average the program increased iron consumption by 2 mg per day (Gerrior, Bente and Hiza 2004). Therefore, the full reduced-form effect on income of the program was 2 percent. Dividing by the proportion of the sample that consumed less than 75 percent of the RDA suggests that the program increased incomes by 5 percent at the individual level. Applying this procedure to labor supply suggests that the program increased weeks worked by 1.6 percent at the individual level. The remainder, 3.4 percent, can be interpreted as the productivity effect (wage/hr). A major contribution of this paper is the result that increased labor supply makes up a large portion of the total increase in income associated with the iron fortification program. Estimated benefit-cost ratios that rely solely on the productivity impacts will underestimate the true benefits of an iron fortification campaign.

The result for the individual productivity effect is well within the range of values found in field experiments in the developing country context. Thomas et al. (2008) find a 30 percent increase in productivity for pre-identified anemic self-employed Indonesian males in a randomized study of iron supplementation. Rubber tappers in Indonesia were found to have increased productivity by 10-15 percent (Basta et al. 1979). Chinese textile workers increased production efficiency by 5 percent after supplementation (Li et al. 1994), and Sri Lankan tea pickers increased the amount of tea picked by 1.2 percent (Edgerton et al. 1979).

VII. Iron Fortification's Contemporaneous Effects on School Enrollment

In this section, I estimate the gains in school enrollment associated with increases in iron consumption. I start by focusing on the 1940-50 period, estimating specifications similar to those discussed above. Then, I extend the sample to the early 20th century to allow estimation with SEA-specific time trends.

School enrollment is measured as a binary indicator equal to one if the child attended school for at least one day during the census reference period. Table V presents point estimates for β, the coefficient on $(IRON_s \times POST_t)$. Each entry is from a separate estimation of equation (9), with the full sample limited to children of ages 8 through 17. *Again, because areas with lower iron* *consumption before fortification are hypothesized to have experienced larger improvements in health after fortification, the coefficient is expected to be negative if the hypothesis is correct.* Standard errors are clustered at the SEA-by-year level. Regressions control for race, sex, and race and sex interacted with $POST_t$, and age dummies, at the individual level. SEA average income in 1936 interacted with $POST_t$ along with year dummies and state economic area indicators are also included.

Column (1) of table V reports point estimates for β from the base specification of equation (9) for several demographic subsamples spanning 1940 to 1950. Results for the full sample of children aged 8-17 are in row (A) and are consistent with the hypothesis that fortification led to greater schooling. A 1 mg difference in iron consumption is associated with a 0.48 percentage point differential change in school enrollment rates.

As in the discussion above, potential threats to a causal interpretation of β remain in the form of unobserved area specific shocks to or trends in enrollment correlated with iron consumption. For example, regional convergence in school enrollment rates could confound the estimate to the extent that low enrollment areas also tended to be low iron areas. Column (2) presents results from a specification that includes a census-division-specific time trends. Identification comes from variation in enrollment gains within census divisions that is correlated with pre-intervention iron consumption. This specification has the benefit of controlling for any division-specific unobservable shocks to enrollment that might be correlated with iron consumption. The point estimates decrease by just over one-half for the full sample, but we still see an increase in school enrollment that is consistent with a positive impact from enriched bread and flour.

The wealth of data contained in the IPUMS allows me to extend the sample to include the 1910 through 1950 censuses and, therefore, SEA-specific time trends.[28] Identification of β now comes from deviations of enrollment during the 1940s from pre-existing SEA trends. In general, the point estimates in column (3) are larger in magnitude than those without controlling for a trend. The impact of iron fortification on school enrollment is economically significant; a one standard deviation difference in iron consumption implies a difference in school enrollment of 1.5 percentage points.[29]

Using the five decades of census data, I explore the timing of the relationship between school enrollment and iron consumption. Figure VI plots the estimated coefficient on $IRON_s$ from regressions using each year of census data as a separate sample. The correlation between iron consumption and school enrollment is stable and positive during the decades prior to the enrichment program. Only during the 1940s does the relationship make a sharp decline. The timing of relative

[28] State economic area is not included as a geographic identifier in later censuses.
[29] Iron drives the impact on school enrollment from enriched bread, not niacin or thiamin. For the specification including all three measures, the point estimates and standard errors are -0.7 (0.3) for iron, 0.08 (0.2) for niacin, and 0.016 (0.016) for thiamin.

gains in school enrollment for low iron consumption SEAs coincides with the federal fortification mandate. As a further test, I estimate the two-period specification of equation (9) separately for the 1910-1920, 1920-1930, and 1930-1940 samples. As expected, changes in school enrollment are not correlated with 1936 iron consumption in any of the decades.[30]

It is also possible that differential changes in parental income and education were correlated with the measure of iron consumption and therefore could confound interpretation of β. The time trends should control for this to some extent, but ideally, individual-level controls for parental education and income could be included. Unfortunately, the sampling procedures for the 1950 census instructed that detailed sample-line questions were to be asked of a single member of the household. School enrollment and income variables are never recorded together within the same respondent household. Therefore, I use average SEA measures of income and education as an alternative strategy. I calculate average real wage and salary income for males and the proportion of the population that has completed high school, some college, or college for each SEA in 1940 and 1950 using individuals between the ages of 25 and 50. As seen in row (B) of table V, the inclusion of parental controls reduces the magnitude of the point estimates. However, the results are still consistent with iron fortification having a positive impact on school enrollment rates. Moreover, parental income is a potentially endogenous control, and one could argue that it should be excluded. As such, row (B) may be interpreted as decomposing the full effect of iron fortification on enrollment into its "direct" effect and the portion from fortification's effect on parental income.

The effects of fortification do not seem to be concentrated in one single demographic group, although there are some important differences. The theory suggests that groups closer to the margin of school enrollment experience larger effects from iron fortification. Rows (C) through (H) of table V report point estimates for β from regressions using distinct demographic subsamples. The percentage point increase for 13-17 year olds is roughly twice that of the 8-12 year olds. School enrollment of the younger age group was already quite high in 1940 at 96 percent, whereas it was only 82 percent for the older age group.

The estimated effect for nonwhites is over twice that of whites; however, controlling for SEA time trends reduces the magnitude of the relationship. The rapid convergence of black and white enrollment rates in the South *before* 1940 might cause the reduction in the nonwhite point estimate moving from column (1) to (3) (Margo 1990).

Similar to the income results in section V, the point estimates suggest males experienced larger improvements in enrollment from iron fortification than did females. The coefficient for males

[30] Estimates by sample year: 1910-20: β = 0.0003(0.003); 1920-30: β = -0.0009(0.003); 1930-40: β = 0.0016(0.002).

is about twice that of females; once again, including SEA time trends diminishes the difference to the point where it is not statistically different from zero at common confidence levels. Overall, the results from rows (C) through (H) suggest a slightly larger effect for demographic subgroups that on average are closer to the margin of attending school.

As a final robustness check, column (4) reports point estimates from regressions that do not include the South census region. This does not considerably alter the point estimates. Identification does not seem to be coming solely from the South, and public health programs targeting the South therefore cannot be driving the main results.

VIII. Long-Term Effects on Children

The impact of iron deficiency during infancy and early childhood might extend to long-term effects manifested during adulthood. I follow up on children that potentially benefitted from the iron fortification mandate by looking at their corresponding adult outcomes using the 1970 decennial census microdata. Economic outcomes I examine include income, years of schooling, and poverty status as an adult. The cross cohort comparison comes from older cohorts having less time to gain during childhood from the fortification program. Similarly, children born in states with high pre-existing iron consumption also had less scope for improvements.

This type of cohort analysis poses a problem in how each observation as an adult is linked to the corresponding exposure to the fortification program as a child. Using 1970 state of residence introduces unnecessarily large measurement error and selection bias into the exposure variable. For this reason, I use the state of birth as the geographic unit in calculating exposure to fortification.

For this analysis, the variable of interest is the interaction of two variables that taken together measure the potential for gains in health from the fortification program. Variation in health gains comes across states of birth based on the average iron consumption in state *s* in 1936 – ($IRON_s$). Cross-cohort variation comes from the number of childhood years exposed to the iron fortification campaign – (EXP_{ik}). Childhood years are defined as time under the age of 19, as most children will have finished their educational choices by this age. The mandate came into force in 1943, thus adults born in 1924 and before received no exposure, with EXP_{ik} increasing linearly until equal to 19 for cohorts born in 1943 and after. I estimate equations of the following form at the individual level.

$$Y_{isk} = \beta \cdot (EXP_{ik} \times IRON_s) + \delta_s + \delta_k + X_{isk}\theta + \varepsilon_{isk} \quad (10)$$

State of birth and cohort fixed effects are included in the regression. Demographic controls include binary indicators for each *age* x *nonwhite* x *female* cell, state of birth interacted with nonwhite, female, and *nonwhite* x *female*. Identification of β comes from within cohort differences in iron consumption by state of birth and from across cohort differences in exposure to the program within states of birth. The hypothesis is that those born in states with high iron consumption had less to gain from fortification and those born earlier had less time during childhood to actually benefit from the intervention. Again, β is expected to be negative if the hypothesis is correct. The identifying assumption is that no unobserved shocks to the outcome were correlated with iron consumption and cohort exposure.

Table VI presents the results from the estimation of equation (10). Adults born in states with lower average pre-intervention iron consumption and with more exposure to fortification had higher income and were less likely to live in poverty than adults with less exposure. Moving from 0 to a full 19 years of exposure at a one standard deviation difference in iron consumption implies a 3.6 percent increase in total income as an adult, a 0.027 year increase in years of schooling, and a decrease in the likelihood of living in poverty by 0.48 percentage points.[31] Results suggest that quantity of schooling is not the causal channel. The point estimates in the schooling regressions are economically small and imprecisely estimated. Moreover, when years-of-schooling is added to the income regressions the point estimates for β do not change.

Mean-reversion poses a potential threat to a causal interpretation of β. Older cohorts might have been hit by a temporary shock that simultaneously caused lower productivity and lower iron consumption. Even without an effect of iron status on wages, younger cohorts would experience income gains as the temporary shock dissipated. I attempt to control for this possibility by including the natural log of average wage and salary income by state in 1940 interacted with age cohort.[32] At the same time, parental income gains were correlated with iron consumption over this period as argued in section VI. This control will account for fortification's impact through the parental income channel as well as mean reversion. Evidence of mean reversion exists for all three outcomes, but point estimates from income and poverty status regressions remain between one-half to two-thirds of the magnitudes without controlling for mean reversion.

The estimates in the first six columns might underestimate the effect of *in utero* and childhood iron deficiency by "culling" the weak babies and children. Potential bias enters the estimates through selective mortality, however I do not find any evidence supporting its presence. Columns (7) and (8) test for selective mortality by estimating equation (10) using as dependent variables the fraction female and the log of cohort size. A female fetus is more likely to survive an adverse health event while *in utero* than a male fetus, leading to a higher fraction female in the

[31] Results from estimations using 1980 microdata are similar to those in table VI, suggesting that the long-term impacts are persistent over the life-cycle. I find no evidence of an impact on labor force participation, employment status, intensive labor supply, occupational income score, college attendance, and receipt of a high school diploma.
[32] From author's calculations using the 1940 census microdata provided by IPUMS.

presence of selective mortality (Trivers and Willard 1973). I find no evidence that my estimates suffer from bias driven by selective mortality.

Differences across demographic groups in the impact of exposure to the enriched bread program appear in table VI. As in section VI, the impact on income seems to be concentrated in men. Nonwhites experience larger effects than whites. However, the point estimates are not significantly different from each other in the statistical sense.

VIII.A. Interpretation of Estimated Long-term Impacts

To put the magnitude of the results into context, they first need to be adjusted to an individual level. Moving from a cohort with zero exposure to a full 19 years of exposure to the enriched bread program is associated with a 2 percent increase in income for the full sample and a 3.4 percent increase in income for men.[33] I take these values as the full reduced-form estimate of the long-term impact of the fortification program. Dividing by the proportion of the sample that consumed less than 75 percent of the RDA suggests an individual level effect of 5 percent for the full sample and 8.5 percent for the men.

Recent years have seen a burgeoning literature on the long-term consequences of early childhood insults and remediation.[34] The foundational work on "developmental origins" in the economics literature uses extreme events such as pandemics and famines combined with short and sharp critical periods in child development (Almond 2006; Schulz 2010; Meng and Qian 2009). While these extreme events are useful for identifying a clear causal effect, the tradeoff is the utility of results for policy options on the margins currently under consideration. My work follows the literature by examining a milder health insult not defined by sharp distinctions between critical periods for potential impacts.

The "fetal origins hypothesis" has received much attention and has generally found large estimated impacts when using clearly delineated treatment and control groups. Field et al. (2009) use a large pre-natal iodine supplementation program in Tanzania to identify the impact of *in utero* iodine deficiency on cognition and human capital. Almond (2006) finds that birth cohorts exposed to the average maternal infection rate during the 1918 Spanish Flu Pandemic decreased schooling by 2.2 percent and decreased annual income by 9 percent. My results suggest that 19 years of iron deficiency decreases income by an amount similar to an *in utero* influenza infection.

Infections have received the most attention in an early childhood context. The estimated effects from these studies tend to be large (Almond and Currie 2011). Bleakley (2007) estimated that

[33] Because of the evidence that mean reversion is present, I use point estimates from column (2) of table V.
[34] For a review of a large portion of this literature see Almond and Currie (2011).

an entire childhood spent with a hookworm infection in the early 20th century United States was associated with a 40 percent decline in adult earnings. Barreca (2010) finds that *in utero* and postnatal exposure to malaria leads to a 13 percent decline in adult income, although the results are noisy. In an analysis of childhood nutrition on the other hand, Almond, Hoynes and Shanzenbach (2010) find mixed evidence of long-term effects on economic outcomes from the role-out of the Food Stamp Program.

Again, my estimates of the income gains for moving from 19 to 0 years of iron deficiency are within the range found for those in the early childhood context. Evidence from this literature also suggests that the causal mechanism works through the quality of schooling, not the quantity. Case and Paxson (2009) find that reductions in infectious disease mortality in a child's state of birth during the mid 20th century were associated with improved cognitive scores at older ages. Moreover, Bleakley (2007) finds that hookworm eradication induced gains in adult income were likely caused by improved quality of schooling. An application of the envelope theorem provides theoretical justification. Parents optimize quantity of schooling prior to the health improvement. Therefore, the major gains to income accrue from the inframarginal increases in quality, not the marginal increase in quantity of schooling (Bleakley 2010).

IX. Discussion and Conclusion

This study contributes to the literature on the effects of micronutrients, and health more generally, on economic activity by evaluating the iron fortification mandate of 1943 in the United States. The program significantly and quickly increased the iron content of the American diet, which is potentially significant because iron affects cognitive development, energy levels, and other aspects of health and because a sizable fraction of the American population was found to be iron deficient. The consumption of fortified grain products was extremely widespread in the United States, and therefore so was the increase in iron consumption associated with the fortification program. Even so, the biological effects of iron are such that health gains would have accrued primarily to those who were relatively undernourished. I combined information on household diets and information on labor market and schooling outcomes to show that places that had relatively low iron consumption before the program experienced relatively large improvements in outcomes after the program's implementation. This relationship is robust to controlling for a variety of observable individual and local characteristics and adjustment for trends that could otherwise bias the results.

I find that gains in income were concentrated in younger males, with results consistent with a hypothesis that the causal mechanism works primarily through an increased wage rate, as opposed to simply more work. Nonetheless, adjustments on the intensive margin of labor supply explain about one-third of the increase in income. This finding is important because many studies in the medical literature focus solely on the directly observable changes in productivity, omitting scope for adjustments on other margins that a model of behavioral choice predicts could occur in response to a health improvement.

To examine the cumulative and long-term effects, I analyze the adult outcomes of children exposed to the bread fortification program. Cohorts with more exposure to fortification had higher earnings and were less likely to be considered living in poverty by the census. In some demographic subgroups, however, I cannot rule out that the results are driven by mean reversion in outcomes or the impact of the fortification program working through a parental income channel.

The economic benefits from iron fortification are large relative to the costs required to implement the program. In 1943, enrichment cost 35 to 50 cents per person annually (Wilder and Williams 1944). The results suggest that annual income increased by $7 per capita, which represents only part of the total benefits of fortification. The national fortification mandate thus had a benefit-cost ratio of at least 14:1, which is within the range for those found in developing countries (Horton and Ross 2003).

Given that iron deficiency is still common, especially in developing nations, the policy implications of these findings are significant. In this regard, two additional and inter-related aspects of American economic history merit attention. First, implementing an effective fortification program in the U.S. in the mid-20th century was comparatively simple. Americans consumed market-purchased bread and flour in large quantities, ensuring wide coverage and treatment. Moreover, milling was highly centralized in the U.S., which facilitated enforcement of the mandate. Finding a similar widely-consumed, market-purchased staple food in many low-income countries is less straightforward (Imhoff-Kunsch et al. 2007), especially where subsistence farming and local milling are the norm.

Second, it is notable that rising income in the U.S in the early 20th century was associated with *declining* iron consumption, as Americans switched almost entirely to the consumption of white flour and bread, and as diets increased in reliance of sugars, fats, and oils. Technology, commercialization, and rising incomes led the American public to consume a *less* nutritious diet. In developing countries, even poor households use incrementally higher incomes to purchase food based on taste, not nutrition (Behrman and Deolalikar 1987). However, the structural changes that accompany economic development, such as concentration of food production and distribution networks, are likely to lower the costs and increase the coverage of food fortification programs.

The results of this paper leave open several questions for future research. Clearly, the "missing link" in this paper pertains to the direct health benefits of food fortification. It is reasonable to assume that such benefits exist based on the medical literature, and the results of this paper are certainly consistent with significant health benefits, but it would still be useful to document and measure the patterns of health improvement directly. I am continuing this line of research in papers that estimate reductions in maternal and infant mortality, as well as the late-life health effects of early childhood iron deficiency.

APPENDIX

Alternative Measures of Treatment Intensity

The results in table IV are not sensitive to alternative measures of treatment intensity. I re-estimate equation (9) with alternative measures of iron status. Table A.I reports point estimates for β from equation (9) for eight different measures of iron status, all measured at the level of the state economic area. For purposes of comparison, row (A) reproduces the base results from table IV for the full sample of men and for those between the ages of 18 and 27. Rows (B) and (C) separate the sample into HI, MEDIUM, and LOW bins, and an indicator for above median iron consumption. One common measure used in the nutrition literature is a binary indicator for iron deficiency. Row (D) uses the proportion of households that consume less iron than an age and gender specific RDA. While the full RDA might be the optimal level, the marginal benefit of additional iron is small at initial values near the RDA. Rows (E) through (G) reduce the cutoff for deficiency to consumption less than 75 percent, 50 percent, and 25 percent of the RDA. Row (H) uses the natural log of SEA average daily iron consumption. The results provide a similar interpretation across all measures of iron status. Gains in wage and salary income are correlated with lower iron status in all rows. To make comparison of the magnitudes easier, column (3) reports the difference in income associated with a one standard deviation difference in the corresponding measure of iron status. All measures vary between 0.8 and 2.1 percent. Results for the schooling analysis are similar (results unreported).

Iron Fortification's Effect on Labor Market Outcomes for Women

Medical surveys show that women are much more likely to be iron deficient then men, both in the United States (Brotanek et al. 2007) and in developing countries (Horton, Alderman and Rivera 2009). Women were also more likely than men to suffer from iron deficiency during the 1940s as well (Kruse et al. 1943), suggesting that women may have experienced larger declines in deficiency. If so, one might expect a larger estimated effect on labor market outcomes for women than for men. However, this does not appear to be the case.

No clear conclusions about the effects of iron fortification on women can be made on the basis of table A.II. In the full women's sample, the estimate of β has the opposite sign of what is expected and is marginally significant. Splitting the sample into married and unmarried women produces point estimates with opposite signs, and unfortunately leaves small sample sizes due to the 1950 sampling procedure. Married women in low iron areas saw a relative decline in their wage income over the 1940s, whereas unmarried women experienced a small and imprecisely estimated relative gain in income. Even after controlling for weeks worked, as in row (C) married women in high iron areas experience a relative increase in wage and salary income over the period.

The result for married women is puzzling, although married women during this period were facing different human capital investment and labor market incentives than men. Recall that the impact on potential wage/income is filtered through a behavioral household bargaining model to get the *observed* wage impacts. The institutional arrangements that shaped married women's labor force participation were complicated and evolving during the 1940s. So much so, that it is plausible that women were unable to – or preferred not to – translate improved health into labor market earnings.

Controls in the income regressions include occupation, industry, education category, and number of children. Thus, the explanation must be that either prices on observed characteristics or unobserved characteristics and their prices are changing differentially in a fashion correlated with pre-program iron consumption. In addition, changes in fertility, marital status, or farm residence are not correlated with iron consumption. Ideally, I would want to control for the income and labor market outcome of the spouse, but the 1950 sample does not allow for this possibility.

Other Changes in Nutrition and Diets During the 1940s.

The U.S. food supply changed considerably during the 1940s. Rationing and mobilization for World War II made their market on the mid-century diet. However, these changes were fleeting. Diets returned to their pre-war patterns shortly after rationing was discontinued in 1946. The main text argues that income and school enrollment effects stem from bread enrichment and increases in the consumption of iron. Availability of many nutrients spiked during the war years as figure A.I makes clear. Changes in consumption of other nutrients could potentially be correlated with consumption of the enrichment micronutrients confounding the estimates. While overall nutritional status did improve as a result, this section shows that increased nutrient intakes other than iron did not drive the observed gains in income and school enrollment associated with fortification.

The effect on the aggregate food supply can be easily seen from the graphs of the nutrient content of the U.S. food supply in figure A.I. Availability of each nutrient is indexed to the 1940 level. From the 1940 base to the 1945 peak, consumption of calcium surged 19 percent, vitamin A by 15 percent, phosphorus by 14 percent, protein and zinc by 12 percent, potassium by 10 percent, and magnesium by 9 percent. While not sustained, the spike in consumption of these nutrients during the treatment period could potentially cause biases in estimation.

Table A.III reports the estimated coefficient on ($IRON_s$ X $POST_t$) from separate regressions individually controlling for SEA average consumption of other nutrients. Row (1) reprints the baseline results from table IV and table V. In all regressions, the coefficient on the variable of interest is similar in magnitude to those of the baseline in row (1). The sole exception being zinc in the school enrollment regression. Moreover, the estimated coefficients on the control nutrients are never statistically or economically significant (estimates unreported). I take this as strong evidence that the estimated results are driven by fortification and not by other the micronutrients that experienced increased consumption during the 1940s.

Construction of Diets and Iron Consumption

Data on iron consumption comes from the "Study of Consumer Purchases, 1935-1936," (henceforth, SCP). The food schedule of the SCP provides a detailed account of the diet of each household by recording the consumption of over 600 individual types of food items over the seven days preceding the interview date. The full sample included 61,000 households, of which 6100 were digitized by the ICPSR. The committee chose to include communities of varying sizes across all regions of the United States. In total, the survey included families in 51 cities, 140 villages, and 66 farm counties across 31 states. The country was split into six regions, with interviews conducted in each region of one large city (252,000 - 302,000), two or three middle-sized cities (30,000-72,000) and four to nine small cities (8,000-19,000). In addition, Chicago and New York were included to cover metropolitan areas of more than 1,000,000 in population. Families from two or more groups of villages (500-3,200) were surveyed from each region, as well as two or more groups of farm counties.

The sample of households completing the food schedule is limited to households that include at least two members, married for at least one year, and with no more than ten boarders. Non-white households were surveyed only in New York City, Columbus, OH, and the South region. To be included, non-farm families were required to have at least one wage earner or be employed in a clerical, professional, or business occupation, and an income of at least $500 per year for the largest cities, and $250 for smaller areas. There were no upper limits on income. Families that received "direct relief" were excluded from the food schedule. Farm families were required to be full-time farmers. Of this sample, 3,545 observations have the information required to construct the iron measures. Some observations in the SCP list a quantity for a food item, but not the units in which the quantity is measured. I assign the most frequently reported unit of measurement for those observations.

Each variable, or food item, in the diet sample is assigned one or more serial numbers from the USDA National Nutrient Database corresponding to the appropriate food product. For food items that encompass a number of product serial numbers in the USDA database, I average the nutrient content across the serial numbers. For example, variable V1424 in the SCP records the quantity of "Beef, Steak, Round", to which I assign two serial numbers, weighted equally, from the USDA database: 13874 Beef, round, bottom round and 13877 Beef, round, eye of round. A full listing of SCP food items and assigned USDA National Nutrient Database serial numbers can be found on my website at www.gregoryniemesh.net. To construct the counterfactual diets under enrichment, I assign micronutrient contents to the fortified food products based on the amount of iron contained in their enriched forms mandated by law.

Summing across all food items gives the total amount of iron consumed by the household in the previous week. Daily iron consumption per person is constructed by taking weekly household iron consumption dividing by the number of meals provided by the home, and multiplying by 3 meals per day. This daily iron consumption per person measure is used throughout my analysis.

The food schedule does not ask about meals provided outside the home, such as meals purchased in restaurants or provided by schools. To the extent that meals provided outside the home were dissimilar in average iron content to meals in the home, my constructed daily iron consumption measure will contain measurement error from meals provided by restaurants, schools, and work places.

Deficiency Measures

Nutrient deficiency is determined at the household level using a household specific recommended daily allowance (RDA). RDAs are published by the Institute of Medicine in a series of Dietary Reference Intake publications. I construct household specific RDAs based on the age and gender composition of each household in the SCP and the RDAs published by the Institute of Medicine. A household is defined as deficient if the consumption of a particular micronutrient does not meet or exceed the household specific RDA.

Weights

The sampling procedure used to conduct diet interviews in the SCP makes the survey somewhat unrepresentative of the population within a state economic area. In particular, the diet survey under-samples lower-income households compared to the census sample, and at rates that vary across region. The receipt of relief funds removes a household from the SCP sample, thus differentially removing low-income households from the North compared to the South. As iron consumption is positively correlated with income, there is a concern that the sampling procedure artificially lowers iron consumption in the South relative to other regions. To address the problem, I weight the SCP sample to better represent the observable characteristics of the 1940 census IPUMS sample within the geographic level under consideration. In a combined SCP and census sample, I use real income of the household head, household size, and farm status to predict the probability of an observation to be in the 1940 IPUMS sample in a logistic regression. I then weight observations in the SCP sample by the predicted probability of being found in the 1940 IPUMS sample. In this fashion I attempt to weight the SCP sample to be more representative of the actual population. I construct weights for state economic areas and states. All summary statistics and regression results use the above weighting scheme.

Outcome Data

All individual level outcome data and demographic controls come from digitized census microdata for the years 1910 through 1950 for school enrollment regressions, 1940 through 1950 for income regressions, and 1970 for the cohort analysis. All data is provided by the Integrated Public Use Microdata Series (IPUMS, Ruggles et al. 2010) a project that harmonizes decennial census microdata. To be included in the sample for the income regressions, observations need to list their main occupation as wage and salary, have a positive income, not be enrolled in school, and be over 17 years old. Top-coded values are multiplied by 1.4. The top-coding cutoffs change from $5,001 in 1940 to $10,000 in 1950.

The school enrollment analysis includes observations between 8 and 17 years of age. The census recorded a child as attending school if the child was enrolled in a "regular school system" for the years 1940 and 1950. Attendance at any school, college, or educational institution would be recorded as enrolled in school during 1910 and 1920. In 1930, attendance at night school was explicitly added to the definition. The reference period for school attendance changed over census years. In 1910, the period included the 7.5 months before April 15th. The period changed to the 4 months prior to January 1st in 1920. For 1930, 1940, and 1950, the census day remained April 1st, but the reference period was 6 months, 1 month, and 2 months respectively.

For the long-term cohort analysis, I use 1970 decennial census microdata. Native-born observations born in states with diet data in the SCP are included. The sample is limited to observations between the ages of 20 and 60, corresponding to cohorts born between 1910 and 1950. The sample for income regressions is further limited to individuals not in school and with positive personal total income.

Controls

World War II spending data comes from the 1947 County Databook (Haines 2010). I calculate the per capita total war spending on contracts and facilities for the state economic area. Total war spending includes all spending from 1940 through 1945 on major war supply contracts for combat equipment and other, and also major war facilities projects, both industrial and military. Dividing by the total population in 1940 gives per capita measures.

State economic area unemployment is constructed from the 1937 Census of Unemployment (Haines 2010). In the main analysis I use the total unemployed divided by 1940 total population (TOTUNEMP/TOTPOP40). However, the results are similar using total male unemployed as well.

New Deal spending, retail sales, weather and migration data were compiled for Fishback, Kantor and Wallis (2003) and Fishback, Horrace and Kantor (2005, 2006). Copies of the data sets can be obtained at the following website:
http://www.u.arizona.edu/~fishback/Published_Research_Datasets.html.

References

Acemoglu, Daron, David Autor and David Lyle, "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury," *Journal of Political Economy*, 112 (2004), 497-551.

Adhvaryu, Achyuta and Anant Nyshadham, "Endowments and Investment within the Household: Evidence from Iodine Supplementation in Tanzania," Working Papers 998, Economic Growth Center, Yale University (2012).

Almond, Douglas, "Is the 1918 Influenza Pandemic Over? Long-Term Effects of *In Utero* Influenza Exposure in the Post-1940 U.S. Population," *Journal of Political Economy,* 114 (2006), 672-712.

Almond, Douglas and Janet Currie, Chapter 15 - Human Capital Development before Age Five, In: Orley Ashenfelter and David Card, Editor(s), *Handbook of Labor Economics*, Elsevier, 2011, Volume 4, Part B, 1315-1486.

Almond, Douglas, Hilary Hoynes and Diane Whitmore Shanzenbach, "Childhood Exposure to the Food Stamp Program: Long-run Health and Economic Outcomes," Working Paper. 2010.

Barreca, Alan, "The Long-Term Economic Impact of In Utero and Postnatal Exposure to Malaria," *Journal of Human Resources,* 45 (2006), 865-892.

Basta, Samir, Soekirman, Darwin Karyadi and Nevin Scrimshaw, "Iron deficiency anemia and the productivity of adult males in Indonesia," *The American Journal of Clinical Nutrition*, 32 (1979), 916-925.

Beard, John and James Connor, "Iron status and neural functioning," *Annual Review of Nutrition,* 23 (2003), 41-58.

Behrman, Jere and Anil Deolalikar, "Will Developing Country Nutrition Improve with Income? A Case Study for Rural South India," *Journal of Political Economy,* 95 (1987), 492-507.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan, "How Much Should We Trust Differences-in-Differences Estimates?," *Quarterly Journal of Economics,* 119 (2004), 249-275.

Bing, Franklin, "Report of the annual meeting of the council of foods," *Journal of the American Medical Association,* 113 (1939), 681.

Bleakley, Hoyt, "Disease and development: Evidence form hookworm eradication in the American South," *Quarterly Journal of Economics*, 122 (2007), 73-117.

Bleakley, Hoyt, "Health, Human Capital, and Development," *Annual Review of Economics*, 2 (2010), 283-310.

Bobonis, Gustavo, Edward Miguel and Charu Puri-Sharma, "Anemia and School Participation," *The Journal of Human Resources*, 41 (2006), 692-721.

Bogue, Donald, *State Economic Areas,* (Washington, DC: U.S. Bureau of the Census, 1951).

Borsook, Henry, Elmer Alpert and Geoffrey Keighley, "Nutritional Status of Aircraft Workers in Southern California," *Milbank Memorial Fund Quarterly*, 21 (1943), 115-157.

Brotanek, Jane, Jacqueline Gosz, Michael Weitzman, and Glenn Flores, "Iron Deficiency in Early Childhood in the United States: Risk Factors and Racial/Ethnic Disparities," *Pediatrics,* 120 (2007), 568-575.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrapped-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, 90 (2008), 414-427.

Capaldi, Elizabeth, ed., *Why We Eat What We Eat: The Psychology of Eating*, (Washington, DC: American Pyschological Association, 1996).

Case, Anne and Christina Paxson, "Early Life Health and Cognitive Function in Old Age," *American Economic Review Papers and Proceedings*, 99 (2009), 104–109.

Cutler, David, Winnie Fung, Michael Kremer, Monica Singhal, and Tom Vogl, "Early-life

exposure and adult outcomes: Evidence from Malaria Eradication in India," *American Economic Journal: Applied Economics,* 2 (2010), 72-94.
Donald, Stephen and Kevin Lang, "Inference with difference-in-differences and other panel data," *Review of Economics and Statistics*, 89 (2007), 221-223.
Edgerton, V. R., G. W. Gardner, Y. Ohira, K. A. Gunawardena and B. Senewiratne, "Iron-deficiency anaemia and its effect on worker productivity and activity patterns," *British Medical Journal*, (1979), 1546-1549.
Feyrer, James, Dimitra Politi and David Weil, "The economic effects of micronutrient deficiency: Evidence from salt iodization in the United State," Working Paper, (2010).
Field, Erica, Omar Robles, and Maximo Torero, "Iodine Deficiency and Schooling Attainment in Tanzania," *American Economic Journal: Applied Economics*, 1 (2009), 140–69.
Fishback, Price V., Shawn Kantor and John Wallis, "Can the New Deal's Three R's Be Rehabilitated? A Program-by-Program, County-by-County Analysis," *Explorations in Economic History*, 40 (2003), 278-307.
Fishback, Price V., William C. Horrace and Shawn Kantor, "Did New Deal Grant Programs Stimulate Local Economies? A Study of Federal Grants and Retail Sales During the Great Depression," *Journal of Economic History*, 65 (2005), 36-71.
Fishback, Price V., William C. Horrace and Shawn Kantor, "The Impact of New Deal Expenditures on Mobility During the Great Depression," *Explorations in Economic History*, 43 (2006), 179-222.
Floud, Roderick, Robert Fogel, Bernard Harris and Sok Chul Hong, *The Changing Body: Health, Nutrition, and Human Development in the Western World since 1700*, NBER Books, National Bureau of Economic Research, Inc. (2011).
Fogel, Robert, "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy," *American Economic Review*, 84 (1994), 369-395.
"Food and Drug Administration: Wheat Flour and Related Products (Definitions and Standards of Identity)," Federal Register, 6 (May 27, 1941), 2574.
"Food and Drug Administration: Wheat Flour and Related Products (Amendments to Definitions and Standards of Identity)," Federal Register, 8 (July 3, 1943), 9115.
Gardner, Gerald, Reggie Edgerton, James Barnard and Edmund Bernauer, "Cardiorespiratory, hematological and physical performance responses of anemic subjects to iron treatment," *The American Journal of Clinical Nutrition*, 28 (1975), 982-988.
Gerrior, S., L. Bente and H. Hiza, *Nutrient Content of the U.S. Food Supply 1909-2000. Home Economics Research Report No. 56.* (Washington, DC: U.S. Department of Agriculture, Center for Nutrition Policy and Promotion, 2004).
Glewwe, Paul and Edward Miguel, "The Impact of Child Health and Nutrition on Education in Less Developed Countries," in *Handbook of Development Economics*, Elsevier (2008).
Goldin, Claudia and Robert Margo, "The Great Compression: The Wage Structure in the United States at Mid-Century," *The Quarterly Journal of Economics*, 107 (1992), 1-34.
Groner, Judith, Evan Charney, Neil Holtzman and David Mellits, "A randomized trial of oral iron on tests of short-term memory and attention span in young pregnant women," *Journal of Adolescent Health Care*, 7 (1986), 44-48.
Haines, Michael R., and Inter-university Consortium for Political and Social Research. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002 [Computer file]. ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010. doi:10.3886/ICPSR02896.
Hass, Jere and Thomas Brownlee, "Iron deficiency and reduced work capacity: A critical review of the research to determine a causal relationship," *Journal of Nutrition,* 131 (2001), 676S-688S.
Horton, S., H. Alderman and J Rivera D., "Hunger and Malnutrition," in B. Lomborg (ed.) *Global*

Crises, Global Solutions, 2nd ed. (New York, NY: Cambridge University Press, 2009).
Horton, Susan, Venkatesh Mannar and Annie Wesley, "Best Practice Paper: Food Fortification with Iron and Iodine," Copenhagen Consensus Center Working Paper, October (2008).
Horton, S. and J. Ross, "The economics of iron deficiency," *Food Policy,* 28 (2003), 51-57.
Imhoff-Kunsch, Beth, Rafael Flores, Omar Dary, and Reynaldo Martorell, "Wheat Flour Fortification Is Unlikely to Benefit the Neediest in Guatemala," *Journal of Nutrition*, 137 (2007), 1017-1022.
Institute of Medicine. Food and Nutrition Board, *Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium and Zinc*, (Washington, DC: National Academy Press, 2001).
Kruse, H.D., Otto Bessey, Norman Jolliffe, James McLester, Frederick Tisdall, and Russell Wilder, *Inadequate Diets and Nutrional Deficiencies in the United States: Their Prevalence and Significance,* (Washington, DC: National Academy of Sciences, 1943. National Research Council Bulletin 109).
Layrisse, Miguel, Jose Felix Chaves, Hernan Mendez-Castellano, Virgilio Bosch, Eleonora Tropper, Betsi Bastardo, and Eglis Gonzalez, "Early response to the effect of iron fortification in the Venezuelan population," *The American Journal of Clinical Nutrition*, 64 (1996), 903-907.
Li, Ruowei, Xuecun Chen, Huaicheng Yan, Paul Deurenberg, Lars Garby and Joseph Hautvast, "Functional consequences of iron supplementation in iron-deficient female cottonmill workers in Beijing," *The American Journal of Clinical Nutrition*, 59 (1994), 908-913.
Liang, Kung-Yee and Scott Zeger, "Longitudinal data analysis using generalized linear models," *Biometrika*, 73 (1986), 13-22.
Logan, Trevon and Paul Rhode, "Moveable Feasts: A New Approach to Endogenizing Tastes, *Working Paper*, 2010.
Lomborg, Bjorn, ed., *Global Crises, Global Solutions*, 2nd ed. (New York, NY: Cambridge University Press, 2009).
Lozoff, Betsy, James Connor, Barbara Felt, Michael Georgieff, John Beard and Timothy Schallert, "Long-lasting neural and behavioral effects of iron deficiency in infancy," *Nutrition Reviews,* 64 (2006), S34-S91.
Margo, Robert, *Race and Schooling in the South, 1880-1950,* (Chicago, IL: University of Chicago Press, 1990).
McLean, Erin, Bruno de Benoist, Ines Egil and Mary Cogswell, *Worldwide prevalence of anemia 1993-2005: WHO Global Database on Anemia*. World Health Organization, (2008).
Meng, Xin and Nancy Qian, "The Long Term Consequences of Famine on Survivors: Evidence from a Unique Natural Experiment using China's Great Famine," *NBER Working Paper Series, No. 14917, (2009).*
Ohira, Y., V. R. Edgerton, G. W. Gardner, B. Senewiratne, R. J. Barnard and D. R. Simpson, "Work Capacity, Heart Rate and Blood Lactate Responses to Iron Treatment," *British Journal of Haematology*, 41 (1979), 365-372.
Perkkio, Mikko, Lennart Jansson, Scott Henderson, Canio Refino, George Brooks and Peter Dallman, "Work performance in the iron-deficient rat: improved endurance with exercise training," *American Journal of Physiology*, 249 (1985), E306-E311.
Rowland, Thomas, Molly Desiroth, Gerald Green and John Kelleher, "The Effect of Iron Therapy on the Exercise Capacity of Nonanemic Iron-Deficient Adolescent Runners," *American Journal of Diseases of Children*, 142 (1988), 165-169.
Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander, *Integrated Public Use Microdata Series: Version 5.0*. (Minneapolis, MN: Minnesota Population Center, 2010).
Schulz, Laura, "The Dutch Hunger Winter and the Developmental Origins of Health and Disease," *PNAS: Proceedings of the National Academy of Sciences*, 107 (2010), 16757-16758.

Seomantri, AG, Ernesto Politt and Insun Kim, "Iron deficiency anemia and educational Achievement," *The American Journal of Clinical Nutrition*, 42 (1985), 1221-1228.
Seshadri, Subadra and Tara Gopaldas, "Impact of iron supplementation on cognitive functions in preschool and school-aged children: the Indian experience," *The American Journal of Clinical Nutrition*, 50 (1989), 675-684.
Steckel, Richard, "Stature and the Standard of Living," *Journal of Economic Literature,* 33 (1995), 1903-1940.
Stiebeling, Harriet, Esther Phipard, Day Monroe, Sadye Adelson, and Faith Clark, *Family Food Consumption and Dietary Levels: Five Regions,* United States Department of Agriculture, Bureau of Home Economics, Miscellaneous Publication No. 452, 1941.
Strauss, John and Duncan Thomas, "Health, Nutrition, and Economic Development," *Journal of Economic Literature*, 36 (1998), 766-817.
Thomas, Duncan, Elizabeth Frankenberg, Jed Friedman, Jean-Pierre Habicht, Mohammed Hakimi, Nicholas Ingwersen, Jaswadi, Nathan Jones, Christopher McKelvey, Gretel Pelto, Bondan Sikoki, Teresa Seeman, James P. Smith, Cecep Sumantri, Wayan Suriastini, Wayan Suriastini, and Siswanto Wilopo, "Causal Effect of Health on Labor Market Outcomes: Experimental Evidence," California Center for Population Research-UCLA Working Paper, CCPR-070-06 (2006).
Trivers, Robert and Dan Willard, "Natural Selection of Parental Ability to Vary the Sex Ratio of Offspring," *Science,* 179 (1973), 90-92.
United States Department of Agriculture, "Dietary Evaluation of Food Used in Households in the United States," *U.S. Dept. Agr. Household Food Consumption Survey 1955*, Report 16, 1961.
United States Department of Agriculture, Nutrient Data Laboratory Home Page, "USDA National Nutrient Database for Standard Reference, release 22," This is an electronic document. Date of publication: November 30, 2009.
United States Department of Labor. Bureau of Labor Statistics. Cost of Living Division, United States Department of Agriculture. Bureau of Home Economics. Economics Division , United States National Resources Committee. Consumption Research Staff. Industrial Section, United States Central Statistical Board, and United States Works Progress Administration. Study of Consumer Purchases in the United States, 1935-1936 [Computer file]. ICPSR08908-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009-06-29.doi:10.3886/ICPSR08908.v3.
Wilder, Russel and Robert Williams, "Enrichment of Flour and Bread: A History of the Movement," *Bulletin of National Research Council*, no. 110, November, (1944).
World Bank, *World Development Report of 1980,* (Washington: World Bank, 1981).
Zhu, Y Isabel and Jere Hass, "Response of serum transferring receptor to iron supplementation in iron-depleted, nonanemic women," *The American Journal of Clinical Nutrition*, 67 (1998), 271-275.

Table I: Summary statistics

		Full sample	Areas with iron consumption Below median	Above median
Panel A: Iron Consumption				
Household level				
Iron consumption in mg	Mean	10.2		
	Median	10.4		
	St. dev	(4.39)		
% of households consuming less than RDA		68		
% less than 75% of RDA		40		
% less than 50%		12		
Observations		3,545		
SEA level				
Iron consumption in mg		10.5	9.1	11.8
		(1.79)	(0.75)	(1.5)
Iron consumption in mg after fortification		14	12.4	15.1
		(2.9)	(1.61)	(2.8)
% of households deficient		65	78	52
		(20)	(11)	(19)
% of households deficient after fortification		20	28	11
		(16)	(15)	(10)
Observations		82	41	41
Panel B: Outcomes				
Income variables				
Income in 1940		904	815	993
		(203)	(213)	(150)
Income in 1950 (1940$)		1,522	1,416	1,629
		(259)	(274)	(194)
Men's income in 1940		1,005	917	1,095
		(232)	(245)	(180)
Women's income in 1940		605	539	671
		(146)	(139)	(123)
School enrollment variables (in percent)				
1940 enrollment		88.8	87.4	90.3
		(4.5)	(5.4)	(2.8)
1950 enrollment		91.9	91.3	92.4
		(3.7)	(4.3)	(3.0)

NOTES: Iron consumption is the daily average for a household. Means are over SEA averages of variables with standard deviations displayed below in parentheses. Income data includes all observations with positive income, wage and salary employment as first occupation, and over 17 years old. Schooling variables include children between the ages of 8 and 17 unless otherwise noted. SOURCE: Iron consumption and deficiency come from author's calculations using "Study of Consumer Purchases, 1935-1936". See Data Appendix for more detail. Income and school enrollment data provided by census microdata (IPUMS). Recommended Daily Allowances constructed from Institute of Medicine (2001).

Table II: Enriched grain products were consumed by the entire population

Panel A: Consumption of enriched grain product by region in 1955 (weekly lbs per capita)

	All (Non-South)	Northeast	North Central	West	South
Grain Products (flour equivalent)	2.44	2.21	2.59	2.60	3.69
Enriched	1.91	1.72	2.05	2.01	2.50

Panel B: Iron consumption increased for all income groups

	1936	1955	Δ in mg	% Δ
All	11.8	17	5.2	44%
Lowest third in income	10.2	16.4	6.2	61%
Middle third in income	11.8	17	5.2	44%
Highest third in income	14	17.6	3.6	26%

SOURCE: USDA (1961)

Table III: State Economic Area predictors of average iron consumption

	Economic and labor market conditions			Migration	Retail Sales	
	War spending p.c.	Unemployment rate	Average income (1936)	Net migration (1930-40)	Retail sales p.c. (1939)	Retail sales growth (1929-39)
Point estimate	-0.21	60.2***	0.052	0.037*	0.095	-0.36
S.e.	(0.13)	(19.2)	(0.047)	(0.020)	(0.084)	(1.81)
R^2	0.03	0.09	0.05	0.03	0.016	0.001

	Fraction of population (1940)			Manufacturing (1940)		
	Native born	Black	Urban	Average wage	Output p.c.	Value added per worker
Point estimate	-5.63*	-0.011	-0.0004	0.001	0.0003	0.0004
S.e.	(3.16)	(0.011)	(0.0009)	(0.006)	(0.0009)	(0.002)
R^2	0.038	0.012	0.003	0.0001	0.001	0.001

	Housing (1940)					
	% Homeowner	Median home value ($1000)	% Δmed. home value (1930-40)	% w/ electricity	% with radio	% with refrigerator
Point estimate	0.008	-0.125	-2.96*	0.003	0.001	0.001
S.e.	(0.018)	(0.193)	(1.49)	(0.012)	(0.014)	(0.014)
R^2	0.002	0.005	0.047	0.001	0.0001	0.0001

	Agriculture (1940)			Weather (1930s)		
	Crop value per acre	Farm value per acre	Fraction tenant	Dustbowl county	Average temperature	Average rainfall
Point estimate	0.02	0.003	-0.99	-1.0	0.003	-0.11
S.e.	(0.014)	(0.002)	(1.08)	(2.5)	(0.031)	(0.17)
R^2	0.024	0.027	0.01	0.002	0.001	0.005

	New Deal spending				WWII Mobilization
	Total grants p.c.	Relief grants p.c.	Public work grants p.c.	Loans p.c.	State mobilization rate
Point estimate	0.001	0.013*	0.004*	0.20	4.65
S.e.	(0.003)	(0.008)	(0.0023)	(0.25)	(4.74)
R^2	0.007	0.074	0.012	0.01	0.01

NOTES: Each entry is a point estimate from a separate regression of SEA average daily iron consumption on each regressor individually. Heteroskedasticity-robust standard errors are reported in parentheses. There are 82 observations in each regression. SOURCES: New Deal spending, retail sales, weather and migration data were compiled for Fishback, Kantor and Wallis (2003) and Fishback, Horrace and Kantor (2005, 2006). Copies of the data sets can be obtained at the following website: http://www.u.arizona.edu/~fishback/Published_Research_Datasets.html. Income is from the "Study of Consumer Purchases." World War II mobilization rates come from Acemoglu, Autor and Lyle (2004). All other data is from Haines (2010).

Table IV: Results for contemporaneous adult male labor market outcomes

	(1) Log wage and salary income *(percent)*	(2) Labor force participation *(p.p.)*	(3) Conditional weeks *(wks per year)*
(A) All men	-1.45	0.18	-0.07
	(0.49)	(0.07)	(0.07)
(B) Add controls for unemployment, per capita war spending, 1936 income, net migration, new deal expenditures, and growth in median house price.	-1.03	0.14	-0.14
	(0.43)	(0.05)	(0.07)
(C) Drop South census region	-0.92	0.10	-0.04
	(0.46)	(0.05)	(0.07)
(D) Drop industry and occupation	-1.05	n.a.	n.a.
	(0.57)		
Age groups			
(E) Under 28	-2.61	0.54	-0.25
	(1.14)	(0.23)	(0.16)
(F) Aged 28-37	-0.79	-0.14	0.06
	(0.60)	(0.17)	(0.11)
(G) Aged 38-48	-0.47	-0.04	-0.23
	(0.65)	(0.01)	(0.11)
Industry and occupation controls	Yes	No	No
Census division time trend	Yes	Yes	Yes

NOTES: Each point estimate comes from a separate estimation of equation (9) and gives the difference in outcome corresponding to a 1 mg difference in iron consumption. Heteroskedasticity-robust standard errors have been corrected for correlation at the (SEA x year) level and are reported in parentheses. The full sample includes all men aged between 18 and 65 who are not full-time or part-time students. Column (1) report results from log income regressions that exclude primary job self-employed workers and those with non-positive wage and salary income. Column (2) reports results from regressions of a binary indicator of labor force participation. The dependent variable in column (3) is weeks worked and includes all workers indicating positive weeks. Demographic controls include veteran and marital status, race, educational attainment (<HS, HS, SC, C), and an educational category specific quartic in age. Row (B) includes SEA 1937 unemployment, per capita war spending, SEA average income in 1936, net migration over the 1940s, total New Deal expenditures, and growth in median house price. All interacted with $POST_t$. SOURCES: Individual outcomes and controls come from IPUMS. Unemployment, war spending, and house prices are from Haines (2010). New Deal spending and migration data were compiled for Fishback, Kantor and Wallis (2003) and Fishback, Horrace and Kantor (2005, 2006). SEA average iron consumption and income are calculated by the author from the "Study of Consumer Purchases."

Table V: Results for contemporaneous school enrollment

	(1) 1940-1950 Base	(2) 1940-1950 Census Division Trend	(3) 1910-1950 SEA Trend	(4) 1910-1950 Drop South Census Region
(A) Ages 8-17	-0.48	-0.22	-0.82	-0.62
	(0.19)	(0.16)	(0.26)	(0.20)
(B) Ages 8-17 with controls for parental income and education	-0.26	-0.29	-0.66	-0.56
	(0.13)	(0.16)	(0.22)	(0.18)
Demographic Subgroups				
(C) Ages 8-12	-0.40	-0.13	-0.40	-0.40
	(0.12)	(0.13)	(0.11)	(0.10)
(D) Ages 13-17	-0.48	-0.38	-1.12	-0.76
	(0.28)	(0.27)	(0.50)	(0.43)
(E) White	-0.37	-0.20	-0.69	-0.61
	(0.18)	(0.17)	(0.26)	(0.22)
(F) Nonwhite	-1.60	-0.17	-1.52	-1.07
	(0.45)	(0.68)	(0.45)	(0.47)
(G) Male	-0.66	-0.40	-0.94	-0.68
	(0.20)	(0.18)	(0.27)	(0.20)
(H) Female	-0.29	-0.04	-0.69	-0.59
	(0.21)	(0.22)	(0.29)	(0.26)

NOTES: Point estimates are for β from a linear probability regression of school enrollment as in equation (9). Each entry is the percentage point difference in school enrollment rates implied by a one milligram difference in pre-intervention state economic area average iron consumption. Standard errors clustered at the state by year level are reported in parentheses. The full sample includes children between the ages of 8 and 17. Demographic controls include sex, race, age dummies, and sex and race interacted with $POST_t$. All regressions include year and state economic area dummies, and 1936 SEA average income interacted with $POST_t$. Parental income and education is proxied by averages at the SEA level for men aged 25 to 50 for income and for men and women aged 25 to 50 for education. Parental education is entered as the proportion of the sample in the SEA with completed education at the high school, some college, and college levels (HS, SC, C). SOURCE: All individual level data comes from the 1910-1950 census microdata (IPUMS). Average iron consumption and 1936 income by SEA is calculated by the author from the "Study of Consumer Purchases, 1935-1936."

Table VI: Long-term follow up of children exposed to iron fortification program

Estimated effects from a one-standard deviation difference in iron consumption for a full 19 years of exposure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Controls for mean reversion	No	Yes	No	Yes	No	Yes	No	No	No
Dependent Variable	Log total income, 1969		Years of schooling		Poverty (=1)		P(Female)	Log cohort size	P(Black)
Years of exposure X average iron consumption	-3.6**	-2.0	-0.027	0.01	0.48*	0.22	0.02	0.005	0.005
	(1.6)	(1.3)	(0.06)	(0.06)	(0.27)	(0.21)	(0.21)	(0.060)	(0.006)
Control for years of schooling	-3.4***	-2.2***							
	(1.2)	(0.9)							
Subsamples									
Male	-5.7***	-3.4***	-0.055	0.00	0.40	0.21			
	(2.1)	(1.1)	(0.07)	(0.05)	(0.26)	(0.20)			
Female	-1.5	0.3	-0.005	0.03	0.56*	0.23			
	(2.0)	(2.2)	(0.07)	(0.07)	(0.33)	(0.24)			
White	-3.5**	-2.3*	-0.006	0.01	0.39	0.19			
	(1.4)	(1.2)	(0.06)	(0.06)	(0.24)	(0.19)			
Nonwhite	-5.3*	-1.8	-0.32**	-0.12	1.77**	0.81			
	(3.1)	(1.6)	(0.15)	(0.07)	(0.86)	(0.65)			

NOTES: Standard errors clustered on state of birth are reported in parentheses. To adjust for the small number of groups, critical values are constructed using the wild-bootstrap procedure described in Cameron, Gelbach, and Miller (2008). ***, **, and * indicate significance at the 1%, 5%, and 10% levels. Each entry is from a separate regression of equation (10). Reported point estimates are from a full 19 years of exposure at a one-standard deviation difference in iron consumption. Controls include interactions of state of birth with nonwhite and female, and nonwhite x female, and for each age x nonwhite x female cell. State-average iron consumption is matched to individuals based on their state of birth. Mean reversion is controlled for by the interaction of 1940 log state average income with cohort. The full sample consists of males and females of all races between the ages of 22 and 60. SOURCES: All individual level data comes from the 1970 census microdata (IPUMS). Average iron consumption by state is calculated by the author from the “Study of Consumer Purchases, 1935-1936.” The log of state mean wage and salary income is from author's calculations using 1940 census microdata provided by IPUMS.

Table A.I Alternative measures of treatment intensity

		Dependent variable - Log wage and salary income		
		(1)	(2)	(3)
Measure of Treatment Intensity		All Men	Under 28	Diff in income associated with 1 st. dev. diff in measure
(A)	mg daily iron consumption	-1.0**	-2.2**	1.8%
		(0.4)	(1.1)	
(B)	HI	-2.6	-7.2*	
		(2.2)	(4.3)	
	MED	-0.6	-1.7	
		(1.7)	(3.0)	
(C)	ABOVE MEDIAN	-4.4**	-6.7*	
		(1.7)	(3.7)	
(D)	Proportion of households consuming less than RDA	4.3	20.1**	0.9%
		(3.9)	(8.3)	
(E)	Proportion of households consuming less than 75 percent of RDA	10.1***	17.1**	1.8%
		(3.6)	(6.9)	
(F)	Proportion of households consuming less than 50 percent of RDA	12.5	34.4***	1.2%
		(7.6)	(12.2)	
(G)	Proportion of households consuming less than 25 percent of RDA	38.0***	31.0**	2.1%
		(6.1)	(13.9)	
(H)	Log daily iron consumption	-11.0**	-25.8**	1.8%
		(4.7)	(12.1)	

NOTES: See table IV for details on the specification. Point estimates in row (A) repeat base results from table IV. Row (B) uses indicators for SEA iron consumption in the top third of the distribution (HI), middle third (MED) with the bottom third omitted. An indicator for above median iron consumption is used in row (C). Deficiency measures in rows (D) through (G) are the proportion of households in an SEA that consume a level of iron that is less than a given percentage of a household specific RDA based on the household's age and gender mix. Note that point estimates are expected to be positive for the deficiency measures if the hypothesis is correct. Column (3) reports the difference in log wage and salary income for the full sample of men associated with a one standard deviation difference in the measure of iron status used in the corresponding row. SOURCE: See table IV.

Table A.II: Results for contemporaneous adult labor market outcomes for women

	(1) Log Wage and Salary Income	(2) Labor Force Participation	(3) Conditional Weeks
(A) All women	1.14	0.68	0.14
	(0.52)	(0.14)	(0.15)
Married women			
(B) No control for weeks worked	2.91	0.45	0.37
	(1.1)	(0.15)	(0.25)
(C) Control for weeks worked	1.56	n.a	n.a.
	(0.80)		
(D) Unmarried women	-0.55	0.94	-0.22
	(0.89)	(0.32)	(0.19)
Industry and occupation controls	Yes	No	No
Census division time trend	Yes	Yes	Yes

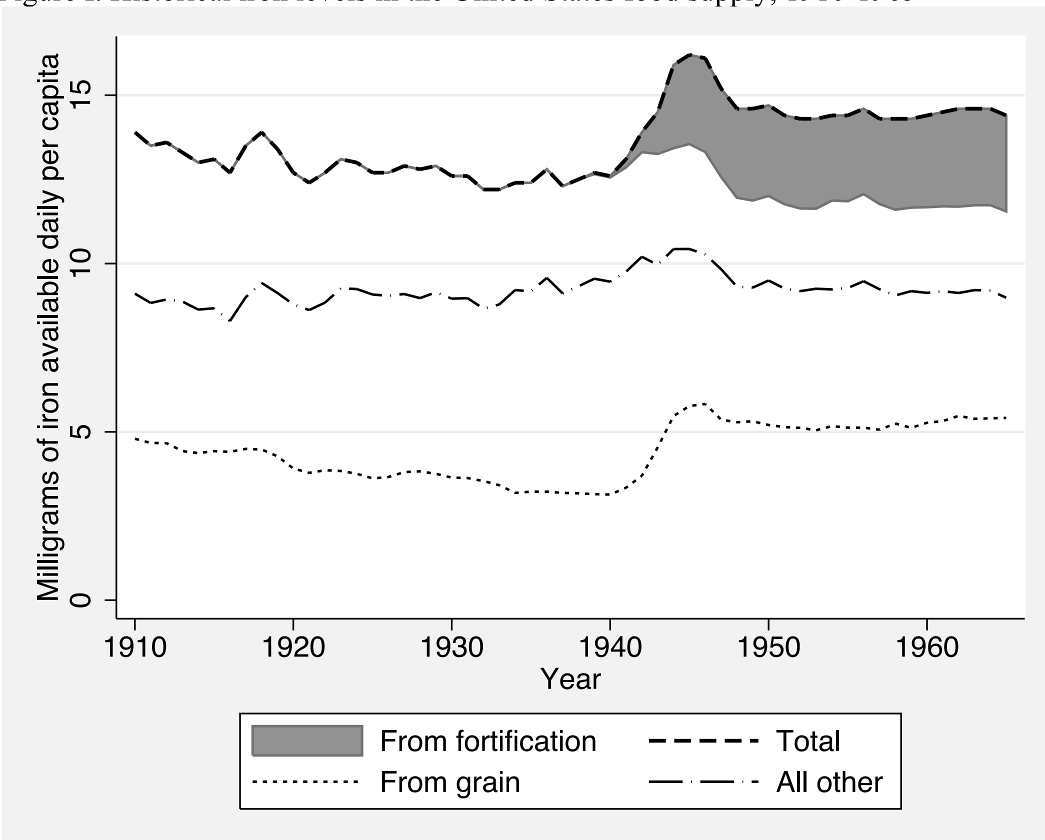
NOTES: Each point estimate comes from a separate estimation of equation (9) and gives the difference in outcome corresponding to a 1 mg difference in iron consumption. Heteroskedasticity-robust standard errors have been corrected for correlation at the (state x year) level and are reported in parentheses. The full sample includes all women aged between 18 and 65 who are not full-time or part-time students. Column (1) report results from log income regressions that exclude primary job self-employed workers and those with non-positive wage and salary income. Column (2) reports results from regressions of a binary indicator of labor force participation. The dependent variable in column (3) is weeks worked and includes all workers indicating positive weeks. Demographic controls include veteran and marital status, race, educational attainment (<HS, HS, SC, C), and an educational category specific quartic in age. Female regressions include the number of own children in the household. All regressions include as controls the SEA 1937 unemployment rate, per capita war spending, and SEA average income in 1936 interacted with $POST_t$. SOURCE: Individual outcomes and controls come from IPUMS. Unemployment and war spending are from Haines (2010). SEA average iron consumption and income are calculated by the author from the "Study of Consumer Purchases."

Table A.III: Robustness checks for other nutrients

	Men Income	Ages 8-17 School Enrollment
Baseline results	-1.7	-0.8
	(0.5)	(0.3)
Protein	-1.8	-0.8
	(0.5)	(0.3)
Vitamin D	-1.8	-0.7
	(0.4)	(0.3)
Vitamin A	-1.6	-0.30
	(0.4)	(0.25)
Calcium	-1.9	-0.8
	(0.4)	(0.3)
Magnesium	-1.4	-1.0
	(0.5)	(0.4)
Zinc	-2.3	-0.03
	(0.6)	(0.30)
Phosphorus	-1.9	-0.8
	(0.5)	(0.3)
Potassium	-1.4	-0.7
	(0.5)	(0.4)
Census Division Trend	Yes	No
SEA Trend	No	Yes

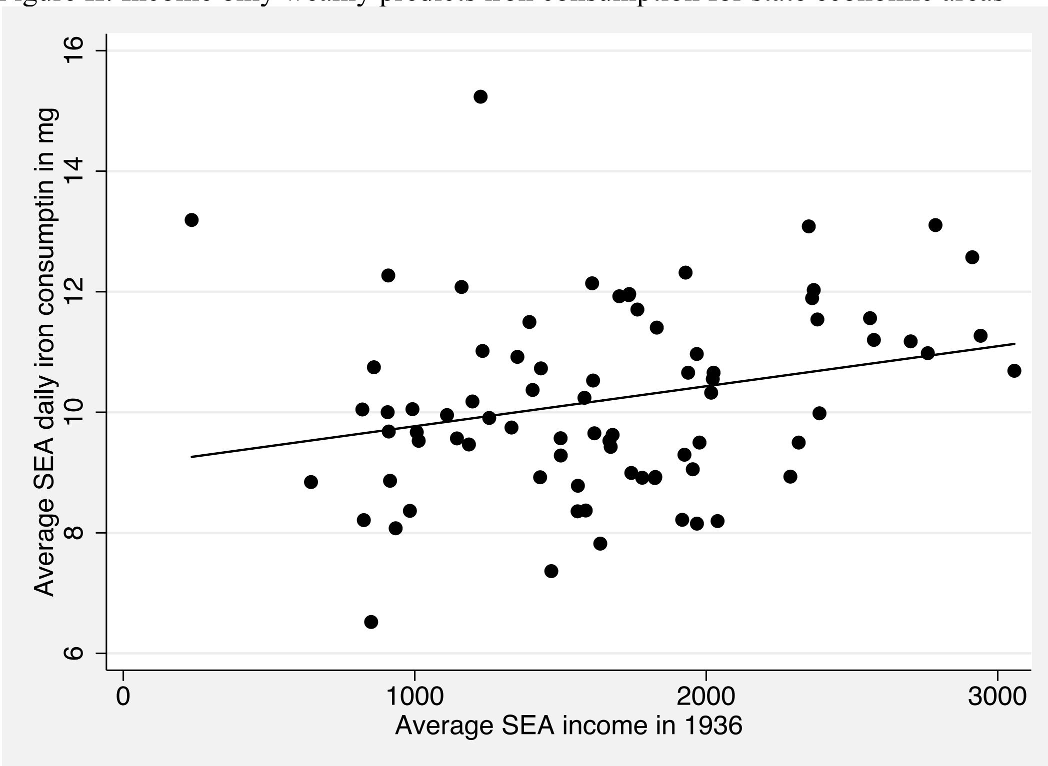
NOTES: Each entry is the point estimate of β on ($IRON_s$x $POST_t$) from an OLS estimation of equation (9). Rows list the additional nutrient interacted with $POST_t$ and included as an independent variable in the regression. See documentation for table IV and table V in the main text for further details and sources.

Figure I. Historical iron levels in the United States food supply, 1910-1965



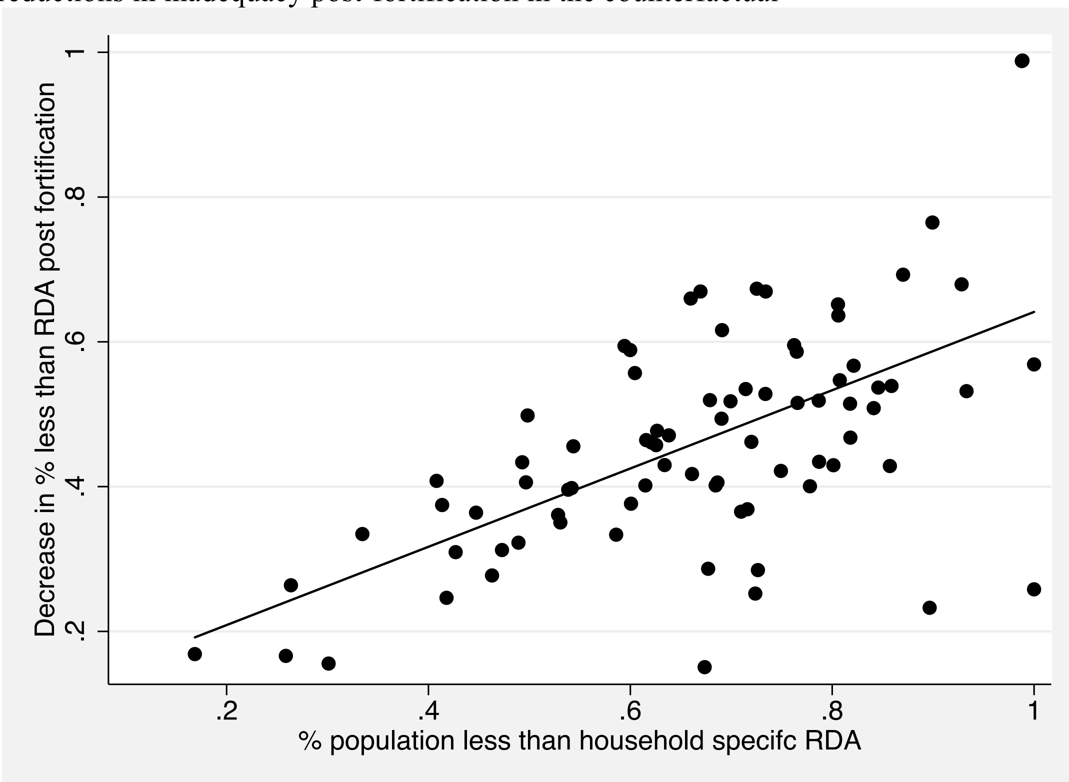
SOURCE: Gerrior, Bente and Hiza (2004)

Figure II: Income only weakly predicts iron consumption for state economic areas



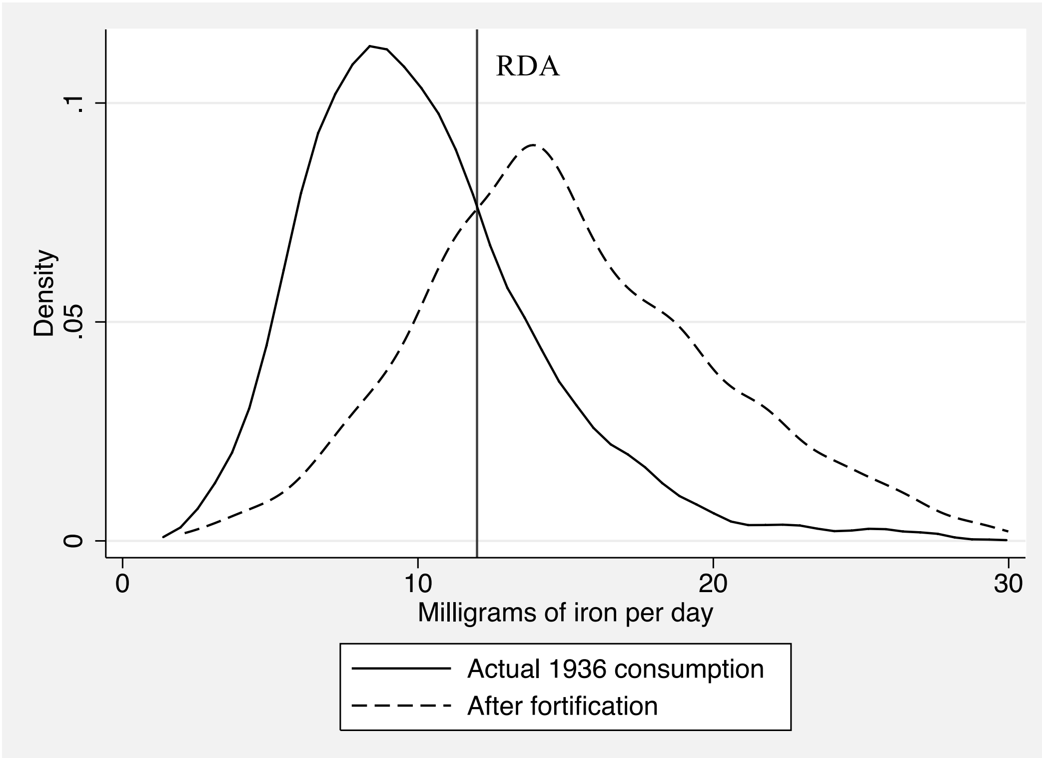
NOTES: The point estimate from the regression line is 0.00052 [s.e. = 0.00047, R^2 = 0.03]. A one-standard deviation difference in income is associated with a 0.3 milligram increase in daily iron consumption.
SOURCE: Author's calculations using "Study of Consumer Purchases, 1935-1936" and USDA (2009). See data appendix for more details on variable construction.

Figure III: State economic areas with higher rates of inadequate diets in 1936 experienced larger reductions in inadequacy post-fortification in the counterfactual



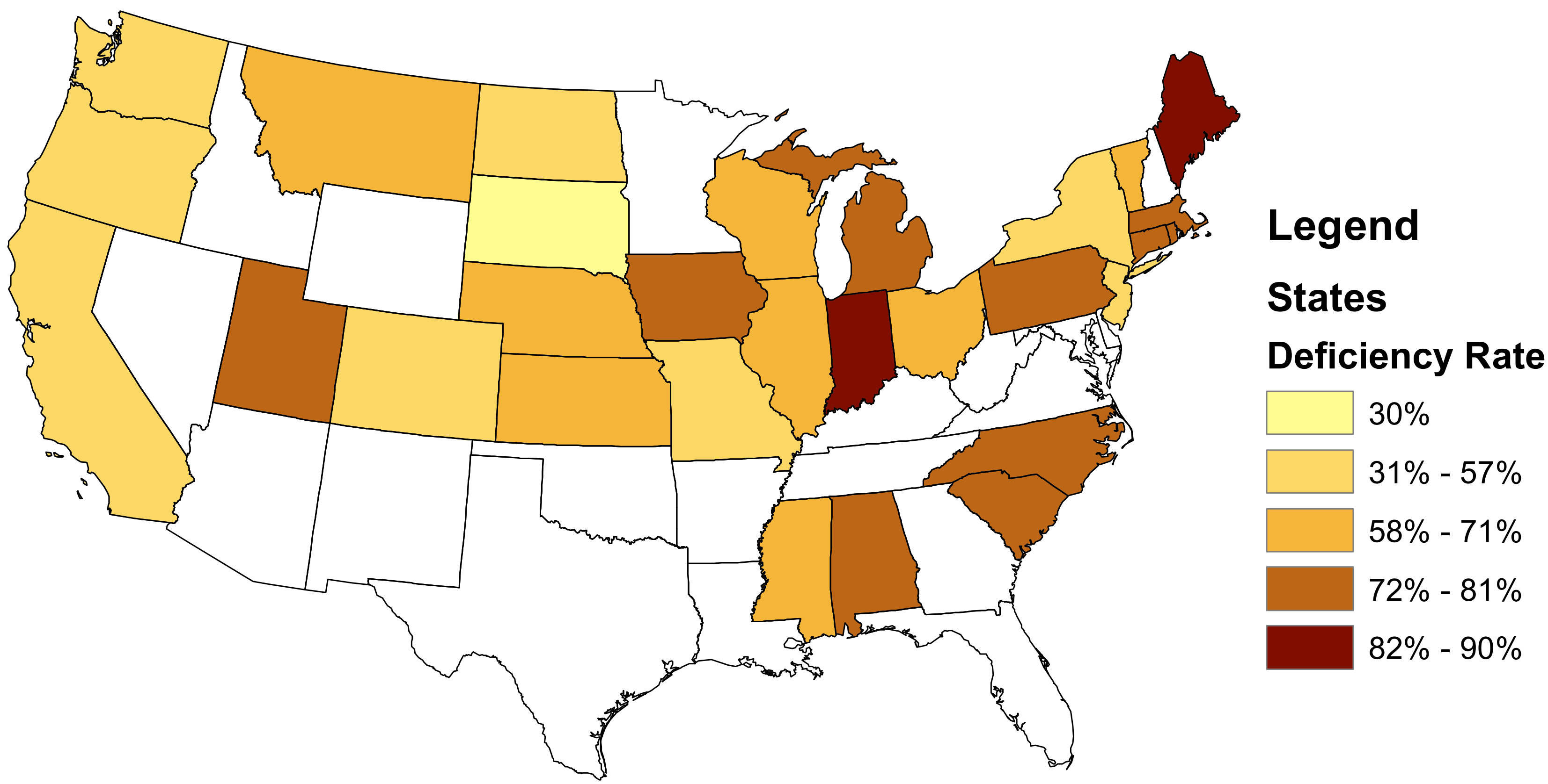
SOURCE: Author's calculations using "Study of Consumer Purchases, 1935-1936" and USDA (2009). See data appendix for more details on variable construction.

Figure IV: Frequency of household iron consumption in 1936 and counterfactual distribution after fortification



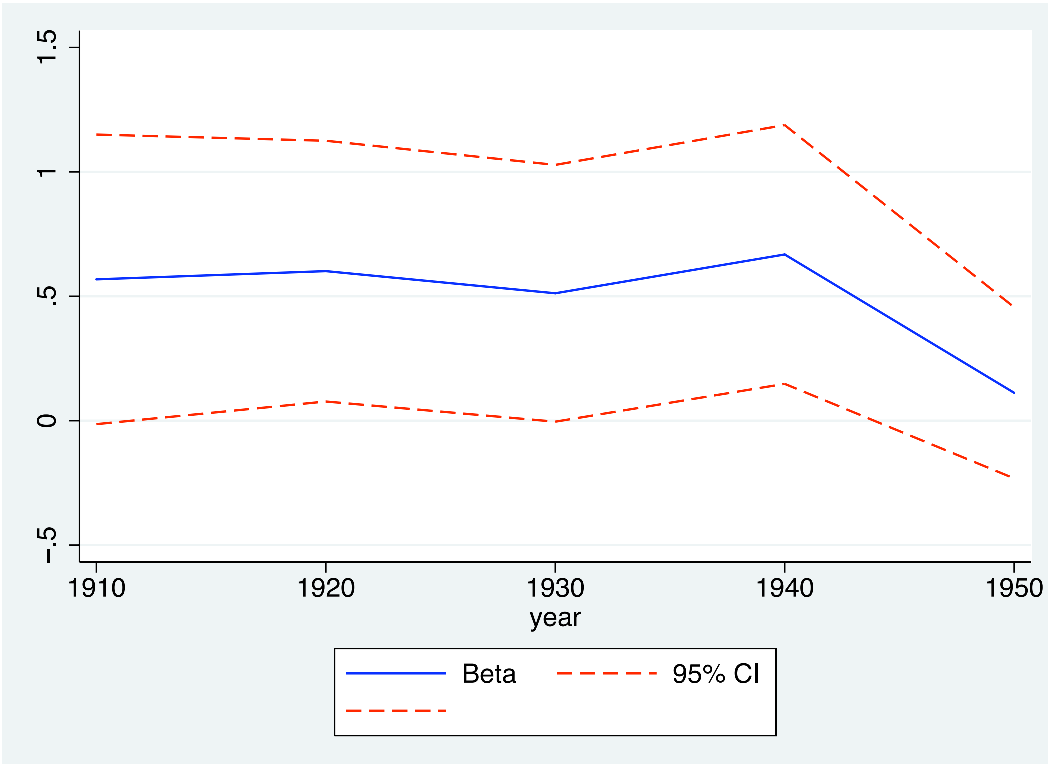
NOTES: The vertical line represents the average of household specific recommended daily allowances. For clarity, the RDA for adult men is 8 mg and 18 mg for adult women.
SOURCE: Author's calculations using "Study of Consumer Purchases, 1935-1936" and USDA (2009).

Figure V: Percent of state population consuming less than the recommended daily allowance in 1936.



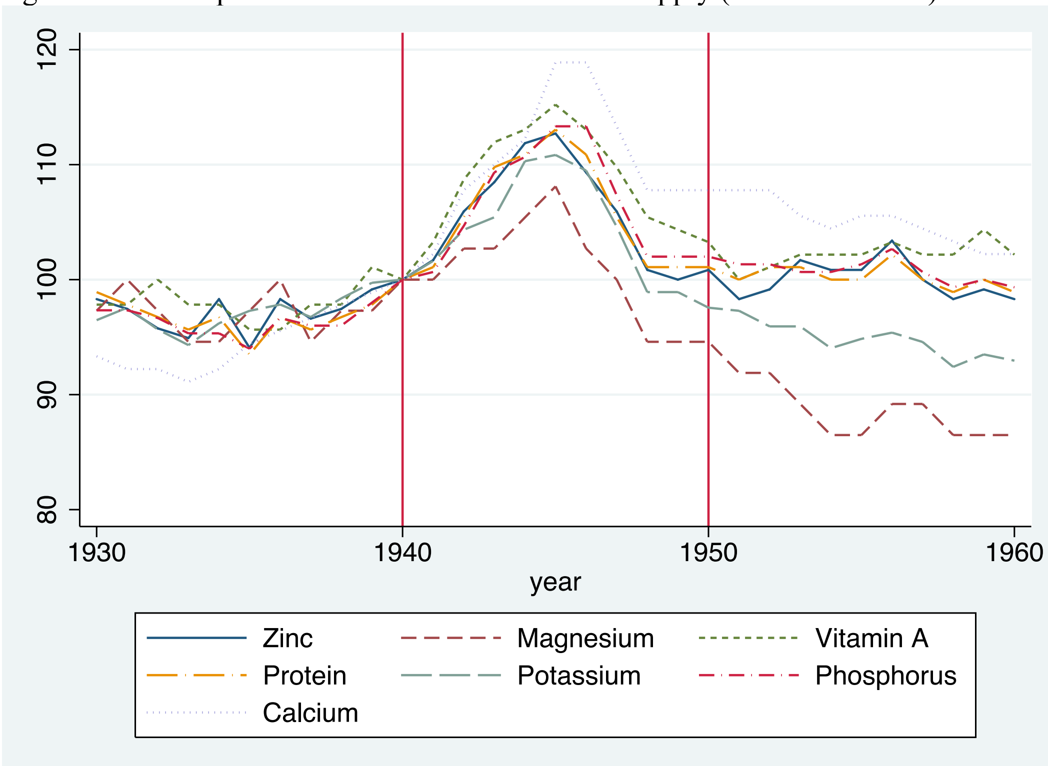
NOTES: The “Study of Consumer Purchases” does not contain diet data for the non-shaded states.
SOURCE: Author’s calculations using “Study of Consumer Purchases, 1935-1936” and USDA (2009)

Figure VI: Timing of school enrollment gains in low iron consumption SEAs coincides with enrichment program.



NOTES: The x-axis plots census year. The y-axis is the estimated coefficient (and standard errors) from regressing a school enrollment indicator on 1936 iron consumption averaged to the level of state economic area. Additional controls include dummies for age, race, sex, and race interacted with sex. The sample includes all children aged 8 to 17 in the state economic areas for which diet information was collected in 1936. Standard errors are clustered by state economic area.

Figure A.I: Per capita nutrient content of U.S. food supply (indexed to 1940)



DATA SOURCE: USDA, Nutrient Content of U.S. Food Supply: http://65.216.150.146/