THE ECONOMIC AND HEALTH BENEFITS OF IRON FORTIFICATION IN THE UNITED STATES

By

Gregory Thomas Niemesh

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

August, 2012

Nashville, Tennessee

Approved:

Professor William J. Collins

Professor Jeremy Atack

Professor Kathryn Anderson

Professor Douglas C. Heimburger, MD

To Lindsey, whose love and support means everything to me

and

To my parents, Ron and Debby, for instilling in me the desire to question and explore

ACKNOWLEDGEMENTS

I wish to thank my committee members Kathryn Anderson, Jeremy Atack, William Collins, and Douglas Heimburger who were more than generous with their support and expertise. A special thanks to William Collins, my advisor, to whom I am forever indebted for his countless hours of reading, reflection, encouragement, and most of all, for providing me the model of how to thrive in the profession. I can only hope to be as charitable, considerate, and diligent as Bill.

This dissertation was generously supported by the John E. Rovensky Fellowship, and grants from the Vanderbilt University Summer Research Awards Program in Arts and Science, the Graduate School Dissertation Enhancement Program, the Noel Dissertation Fellowship and the Kirk Dornbush Summer Research Grant.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

I. INTRODUCTION

Micronutrient deficiencies plague the developing world. For example, the World Health Organization estimates that over a quarter of the world's population suffer 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.[1] For instance, the Copenhagen Consensus of 2008 lists iron and iodine fortification as the third most cost-effective development intervention (Lomborg 2009). As of 2009, 63 countries had implemented flour fortification programs, but 72 percent of all flour produced remains unfortified (Horton, Mannar and Wesley 2008).

Surprisingly few studies directly evaluate the effects of national-level fortification programs. Moreover, data limitations and experimental design problems limit their usefulness for policy purposes. Layrisse *et al.* (1996), for example, 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

[1] 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).

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 apply productivity estimates from supplementation field trials to prevalence measures from health surveys. Willingness-to-pay or quality-adjusted life year parameters are applied to the same prevalence measures to value health improvements (Horton and Ross 2003). Estimated benefit-cost ratios that rely solely on the productivity impacts will underestimate the true benefits of an iron fortification campaign. This dissertation will include the potentially large adjustment in labor supply when calculating an income-based benefit-cost ratio.

This dissertation 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 health, labor market outcomes and human capital investment. 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.[2] 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 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

[2] 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.

(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. Chapter II provides the historical background on fortification as a policy option in the United States and describes the trend of iron consumption and deficiency in the population over the 20th century.

In addition to the literature on health and micronutrient fortification in developing countries, this investigation ties to two other branches of literature 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 focused on calorie and protein malnutrition, and neglected 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 co-trending 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 dissertation 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 2008; Cutler *et al*. 2010). Chapter III discusses the identification strategy in detail, and provides a basic theoretical model of how iron fortification affects observed labor market and schooling outcomes. The investigation's central empirical questions are whether places with relatively low-iron consumption

levels before the program's implementation experienced relatively large gains in health, labor market, and schooling outcomes after the program's implementation, and if so, whether this pattern can be given a causal interpretation. The analysis's 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 income. Second, the timing of the federal mandate was determined by wartime concerns and technological constraints and, in this sense, was exogenous. Finally, iron consumption has a non-linear effect on health. Therefore, a program that increases iron consumption across the 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 runs into 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, essentially a small group of contiguous counties with similar economic and social characteristics circa mid-century.[3]

[3] 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

Chapter IV and chapter V estimate the short-term gains during the 1940s in labor market, education, and health outcomes that can be associated with the iron fortification program. The basic empirical question is whether the program increased productivity, labor supply, and income for adults, and whether school enrollment increased for children. I find that after the iron fortification mandate in 1943, wages and school enrollment in areas with low-iron intake increased relative to other areas. 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. Interpreting the estimates causally, the full-reduced form effect of the program was to increase male wages by 2.8 percent on average from 1940 to 1950, and increase school enrollment by 1 to 1.5 percentage points. 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).

The fortification program was meant to directly improve the iron status and health of the adult population. While the large increase in iron consumption and wide reach of the program make improvements in health plausible, a lack of health data from the period makes a direct test challenging. The data requirements severely limit the type of health outcomes available. The empirical strategy requires pre- and post-intervention outcome measurements for at least some level of geographic detail. To my knowledge, county-level mortality rates are the sole measure available. Using county-level total and infant

---

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.

mortality rates from 1931 to 1950, I find no evidence that fortification had any effect on overall mortality. Estimates are consistent with the program reducing infant mortality and stillbirths by 2 percent, although identification assumptions may not be satisfied for this particular specification.

In chapter VI, I use the 1970 census microdata to undertake a separate analysis that suggests that iron deficiency may have 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). Results suggest that 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 to fortification 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 (9 percent of the sample lived in poverty in 1970). Increased quantity of schooling does not drive the increases in adult income.[4]

Repeating the exercise using late-life measures of chronic diseases and mortality from the Health and Retirement Study provides evidence for the "developmental origins" hypothesis.[5] More exposure to the iron fortification campaign during the first five years of life is associated with a lower incidence of a number of chronic diseases: high blood pressure, diabetes, low HDL cholesterol, and disabilities limiting daily life activities. I

[4] 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.

[5] The "developmental origins" hypothesis suggests that early-life environmental conditions, during both *in utero* development and childhood, have an important and long-lasting effect on adult health.

find evidence of an independent biological causal impact of iron deficiency during childhood. Fortification induced changes in income, wealth, education, and parental income do not seem to play large roles as the causal mechanism.

CHAPTER II

II. FORTIFICATION, IRON DEFICIENCY, AND THE HISTORY OF BREAD ENRICHMENT IN THE UNITED STATES

Fortification as a Policy Option

In the early decades of the 20th century, developments in nutritional science and social science converged to show that nutritional deficiencies were indeed a significant problem in the United States, while also providing the tools for a cure. Chemists and nutritionists made a rapid succession of advances in the 1920s and 1930s by isolating specific chemical substances required by the human body for survival. They named these substances "vitamins" and "minerals," together called micronutrients. Discoveries were made by inducing nutritional deficiencies in laboratory rats by withholding the micronutrient, the symptoms of which could be linked to commonly occurring diseases: anemia to a deficiency of iron in 1932, beriberi to thiamin in 1934, and pellagra to niacin in 1937 (Health, Strauss and Castle 1932; Williams 1935; Koehn and Elvehjem 1937). The cure for these nutritional deficiencies was now as a simple as consuming a large dose of the nutrient.

Prevention of nutritional deficiencies at the population level followed a different path. Food fortification became the favored strategy of public health officials during the mid-20th century. For instance, milk fortified with vitamin D milk was introduced in the mid-1930s to reduce childhood rickets. In 1927, the Lever Brothers in England began producing margarine with vitamin A, which was widely distributed in the United States

by the mid-1930s.[6] A voluntary flour enrichment program in a number of the Southern states of the United States added niacin and thiamin to wheat and corn flours in the late 1930s in response to the high prevalence of pellagra and beriberi in that region (Park *et al.* 2004).

The rapid decline in deaths from rickets, pellagra, and beriberi testify to the effectiveness of food fortification as a public policy. Indeed, looking back on the 20th century, the U.S. Centers for Disease Control (CDC) lists the elimination of these nutritional disorders through better diets and fortification among the most important accomplishments of the American public health community (Semba 2007). Nonetheless, it is important to recognize that policymakers have several options to weigh when it comes to addressing nutritional deficiencies.

In general, there are three possible approaches to reducing micronutrient deficiency: fortification, supplementation, and food-based approaches. Fortification is the artificial addition of micronutrients to a commonly consumed staple food. Nutritionists differentiate between "fortification" and "enrichment." Fortification adds nutrients that are absent from the natural food, or adds nutrients to a level greater than what occurs naturally. Vitamin D fortified milk is an example, as milk does not naturally contain the nutrient. Enrichment, on the other hand, adds a nutrient to a processed food until equivalent to the naturally occurring level found in the unprocessed product. Iron-enriched bread during the 1940s, the focus of this study, is an example of enrichment. Iron was added to white-wheat flour until the nutrient content was equivalent to that found in whole-wheat flour. This distinction between "enrichment" and "fortification" is

[6] Until the 1940s, however, food fortification was implemented on a voluntary basis.

irrelevant for the purpose of estimating the gains from health improvements. I use the terms interchangeably throughout the remainder of this manuscript.

With regular consumption of a fortified food, an individual will reduce the risk of deficiency. An important advantage of fortification programs is the ability to reach a large population and play a preventative role. The cost of fortification is low relative to the other options. Iron fortification of flour and bread costs roughly $0.10 per person annually (Horton and Ross 2003). Moreover, a public health authority does not need to screen the population and single out at-risk individuals for treatment; the entire population that consumes the fortified food receives treatment.

To the extent some people must change their eating habits to achieve the recommended daily allowance (RDA) through fortified foods, however, it might be difficult to reduce their micronutrient deficiencies. This disadvantage applies to other food-based programs as well, in which an individual is given information about a proper diet to encourage more nutritious eating habits. For this reason, supplementation has been the preferred choice for therapeutic programs. Supplementation involves providing a patient with a dose of the micronutrient. This is simple once an individual has been identified as deficient in a particular nutrient. Daily or frequent doses of the supplement are needed to rectify the deficiency, which requires a stable distribution network and compliant patient. In contrast, once a fortification program is in place, its success in reducing nutritional deficiencies requires only that people consume a staple food—not that they change their diets, buy special supplements, or comply with a doctor's instructions.

A debate continues on the cost-effectiveness of supplementation versus food-fortification for a continuing prevention program (Horton and Ross 2003). Studies of the actual costs of implemented programs have contradictory results dependent on place and time (Howson et al. 1998). In both historical and contemporary contexts, the effectiveness of a fortification program depends on the centralization and commercialization of the food supply. Highly concentrated production or distribution markets require less monitoring and regulation to implement a fortification intervention.

For example, the salt industry in the U.S. during the 1920s was highly concentrated with 36 percent of production by companies in Michigan, 18 percent in New York, and 14 percent in Ohio. In 1924, the Michigan legislature passed legislation requiring the addition of iodine to all salt destined for the Michigan retail market. Producers found it more profitable to produce *only* iodized salt and distribute it to the entire country, rather than offer differentiated products (Feyrer, Politi and Weil 2008). Likewise, part of the success of flour fortification in the United States can be attributed to the high concentration of the flour industry by the 1940s. A small number of commercial flourmills and large regional commercial bakeries made monitoring the mandate relatively easy. Evidence also points to some economies of scale in the fortification process. Large bakeries and flourmills were more likely to adopt fortification voluntarily before the federal mandate (Wilder and Williams 1944).

Prevalence of Iron Deficiency in the United States Before World War II

The economic history of the United States does not include episodes of widespread famine, but large segments of the population have suffered serious bouts of micronutrient deficiency, even in the 20th century. For instance, pellagra, which is caused by a lack of niacin, claimed approximately 5,000 lives annually during the 1920s in the U.S. (Park et al. 2000). Rickets, which is caused by vitamin D deficiency in childhood, led to approximately 400 deaths annually in the 1920s (Weick 1967), which suggests that rickets was widespread because it was rarely fatal. Physical examinations conducted for the World War I draft provided some of the first nationwide data on the population's health, and the results were illuminating (Love and Davenport 1920). About 0.5 percent of draftees suffered from simple goiter and cretinism, both caused by severe and prolonged iodine deficiency. In the Upper Midwest, rates of goiter reached as high as 3 percent.

While not as deadly or symptomatic as pellagra and rickets, iron deficiency and anemia were likely even more prevalent. During the 1930s and early 1940s, diet surveys and case-studies that measured levels of iron in blood serum were conducted throughout the country. The percent of the 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 in Southern California (Kruse et al. 1943).[7] The studies vary in the sex, race, and income of the populations under observation. None followed a sampling procedure. Kruse et al. (1943) summarizes the results of various case-studies as follows,

[7] Health professionals of the period did not differentiate iron deficiency and anemia.

> "All the evidence from numerous surveys over the past ten years to the present among persons of all ages in many localities is without exception in complete agreement that inadequate diets were widespread in the nation."

Taking the case studies as a whole, the United States in the 1930s had rates of iron deficiency between 20 and 30 percent, similar to those currently found in Turkey or Brazil (McLean et al. 2008).

Dietary intake data provide an alternative to biological assay methods, and I rely heavily on dietary information to characterize nutritional intake in this dissertation. In particular, 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 remarkable 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 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.

I converted the diet of each household into the amount of iron provided using the USDA National Nutrition Database (USDA 2009). Using the number of meals provided by the home, I calculated the average daily per person iron consumption for each

Table 1: Summary statistics of iron consumption for diet data

			Areas with iron consumption	
		Full sample	Below-median	Above-median
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

Notes: Iron consumption is the daily average for a household. Means are over SEA averages of variables with standard-deviations displayed below in parentheses.
Sources: Iron consumption and deficiency come from author's calculations using "Study of Consumer Purchases, 1935-1936". See Data Appendix for more detail. Recommended Daily Allowances constructed from Institute of Medicine (2001).

household. A daily measure of iron intake simplifies comparisons to the recommended daily allowances published by the Institute of Medicine.[8] Summary statistics are reported in

[8] As part of its work related to nutrition, the same group at National Research Council that promoted bread enrichment, also developed the original RDAs in 1943. These standards have become refined overtime thanks to advancements in food intake assay methods. I use the most recent published RDAs throughout the analysis, as they represent the most scientifically advanced estimates of the body's requirements.

Table 1. 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.

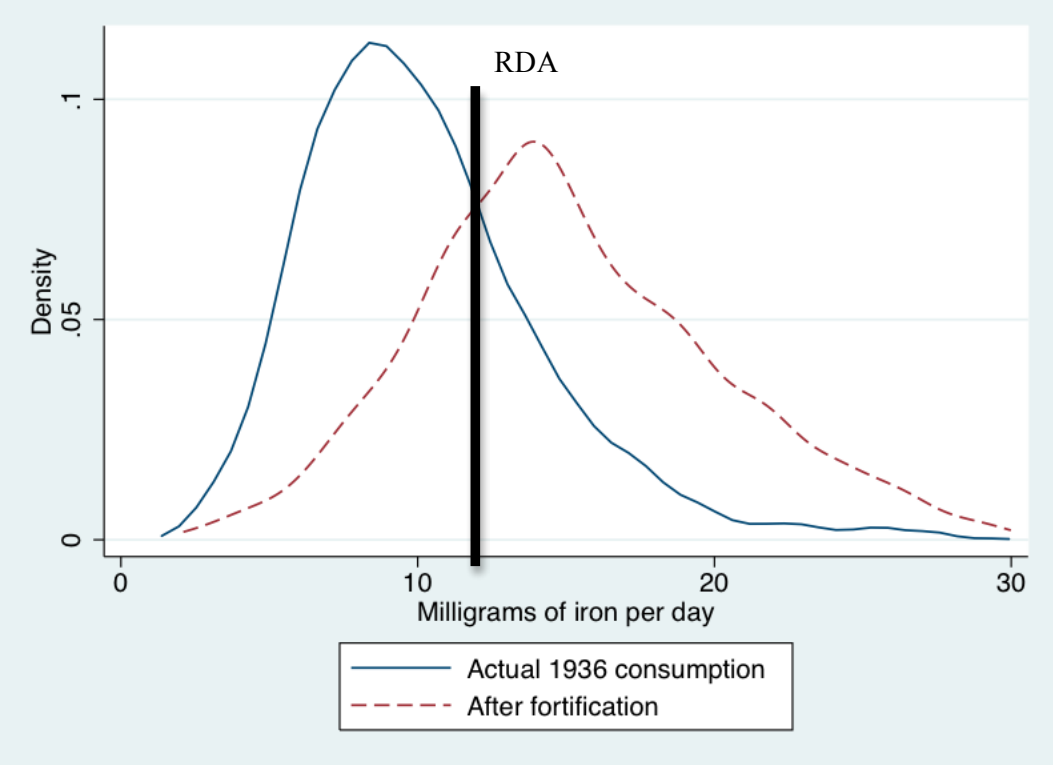

Figure 1: Frequency of household iron consumption in 1936 and counterfactual distribution after fortification

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.
Sources: Author's calculations using "Study of Consumer Purchases, 1935-1936" and USDA (2009).

There was substantial variation across households in the daily consumption of iron. Figure 1 plots the pre-intervention distribution of household per capita daily consumption of iron in the 1936 diets. Unfortunately, no similar survey was conducted soon after the iron fortification programs 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 1 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 increases by 3.8 mg from 10.2 mg to 14.0 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.[9] In sum, because store-bought bread and flour were such common elements in American diets, it is highly likely that the fortification program significantly boosted iron consumption throughout the distribution.

To capture the geographic variation in iron consumption before fortification and facilitate merger with census microdata, I calculated the mean daily iron consumption over households within each state economic area (SEA). An SEA consists of a group of contiguous counties that were economically similar in 1950. The number of counties in an SEA can range from a single county for a large urban area to 8 counties that are agriculturally similar. The required conversions to construct household iron consumption

[9] 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.

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 2 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, and this variation will prove useful in identifying the effects of iron fortification.

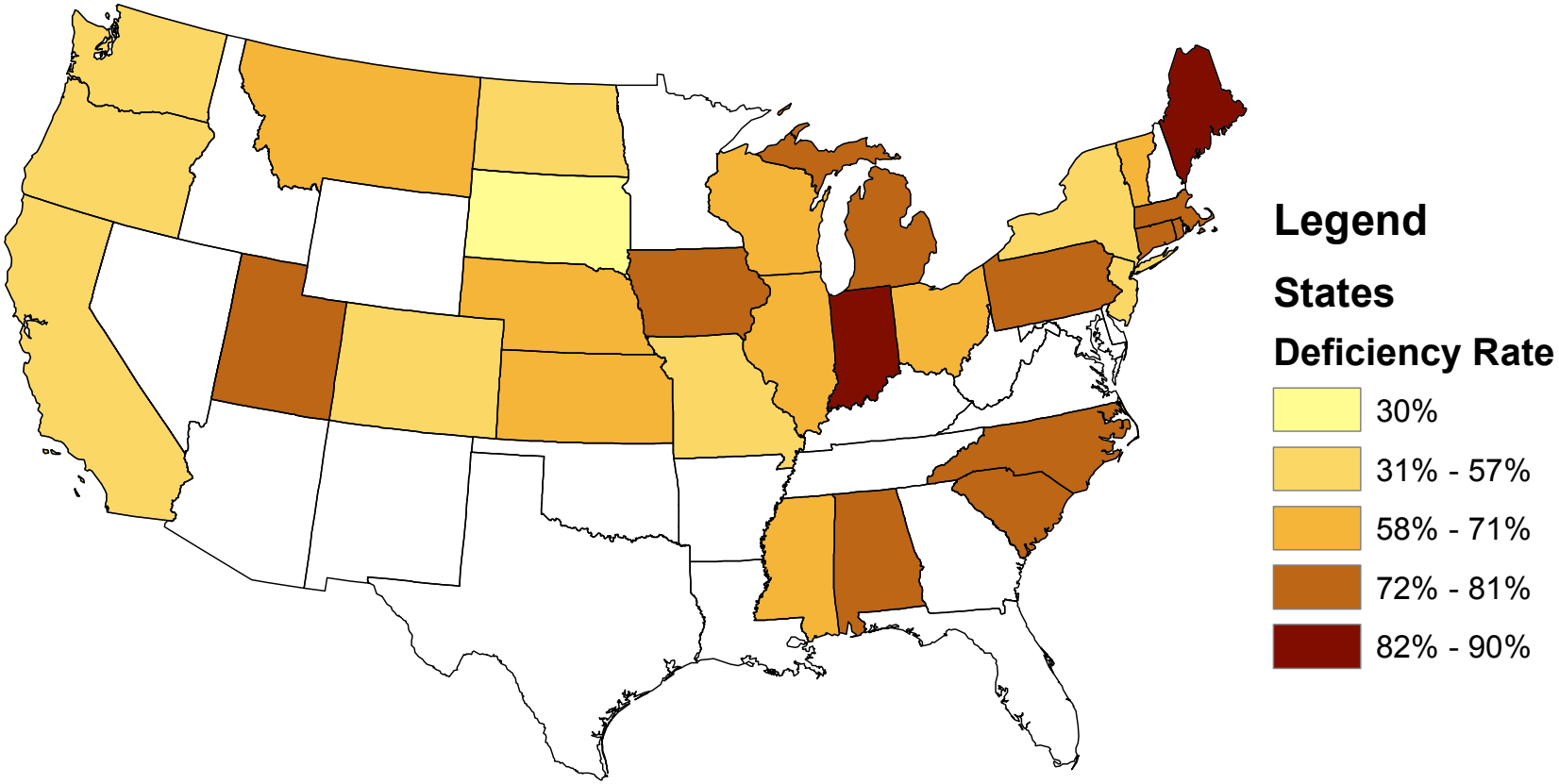

Figure 2: Percent of 1936 state population consuming less than the recommended daily allowance

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).

History of Bread Enrichment in the United States

"The improvement of the quality of bread was basic to any nutrition program that might be undertaken for national defense."

Surgeon General Parran
National Nutrition Conference for Defense (1941)

Before 1943, average iron availability 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 were initially steadfast in their opposition to the addition of any foreign substances to food products (Wilder and Williams 1944). Commercially produced bread had always been susceptible to the addition of cheap fillers such as sawdust to reduce the cost of production (Bobrow-Strain 2012). Public interest in securing the food supply from these practices led to action on the part of public officials. Food regulation was one of the major campaigns of the public health profession during the first decades of the 20th century. The Pure Food and Drug Act of 1907 had only recently been enacted into law, allowing the Food and Drug Administration (FDA) to regulate the labeling and contents of food.

The views of the FDA and the American Medical Association (AMA) became more nuanced during the debates over whether to allow vitamin D-fortified milk in the 1930s. They acknowledged a distinction between health-augmenting additives that did not alter the concept or ideal of the particular food, such as micronutrients, and additives that fundamentally changed the product. A consumer does not view as "bread" a loaf that contains sawdust. The AMA's Council on Foods and Nutrition first discussed fortified

grain products in their March 1936 meeting, and issued the following announcement in the *Journal of the American Medical Association* (1936, 107(1), 39),

> "While the Committee has shown by its actions that it is not opposed to the addition of vitamin concentrates to foods under all circumstances, yet it desires that such fortification shall not be made unless (a) it has been demonstrated that there is a need for such fortification and (b) unless suitable experimental evidence shows that the fortified food suitably serves the purpose desired."

The announcement was the first step to define *when* fortification was appropriate to use as a tool to fight nutritional disorders. The most powerful health lobby now supported the cause of fortification, at least in cases where the evidence of need was clear. In an important step for proponents of fortification, the Council on Foods and Nutrition backed proposals to enrich bread and flour with iron and thiamin in 1939 (Bing 1939).

The movement towards an enrichment program was accelerated by events overseas during the summer of 1940. In reaction to the Battle of Britain, the British government decided to fortify all bread and flour with thiamin, iron, and calcium in July 1940 (Moran and Drummond 1940).[10] Shortly thereafter, the Subcommittee on Medical Nutrition of the National Research Council (NRC) recommended that the U.S. Army and Navy require suppliers to fortify all purchases of flour with iron and thiamin.[11]

The most important event of the summer of 1940 was perhaps the scheduled public hearings by the FDA to determine a standard for white flour and bread in September. Then, in May 1941, the FDA enacted regulations stipulating the labeling of

[10] The British government was never able to actually accomplish the policy. Supply shortages and production constraints imposed by the war effort made thiamine production a low priority. The bread of expeditionary forces was fortified, and the home front saw the increased use of "wheat meal loaf" made from under-milled flour. There were also worries about the possibility of war profiteering on the part of thiamin and bread producers. Canada also decided to use long-extraction wheat flour, with the product marketed as "Canada Approved" starting in January 1942.

[11] The military would actually implement the recommendation in 1943, but not because the soldier's diet was thought to be inadequate. Officials believed that the military's opinion carried weight in the decisions of civilians towards enriched bread.

"enriched" wheat flour. Manufacturers were not required to fortify, but to use the label "enriched flour" the product had to 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.* No standard for bread was promulgated at the time.

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 in the voluntary enrichment program, 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. The mandate had a large, abrupt, and long-lasting impact on iron consumption in the United States. Figure 3 shows the iron availability in the total food supply of the United States over the first half of the 20th century. [12] The vertical axis measures the per capita milligrams of iron available daily. As noted above, average iron consumption in the American diet gradually declined over the first decades of the century. Substitution of sugars, fats, and oils for grains, as well as a gradual decline in calories, reduced average iron consumption.

It is interesting to find the deterioration along this one dimension of nutrition in the face of a long secular improvement in overall, or net, nutrition. The anthropometric

[12] The USDA constructed this series by using the disappearance method – production plus imports minus farm use and exports (Gerrior, Bente and Hiza 2004).

evidence suggests a clear 50-year increase height for cohorts born starting in at least 1900. Average male height increased by 3 percent (5 cm) between the 1900 and 1950 birth cohorts (Costa and Steckel 1995). The list of potential causes for the improvement is long: the germ theory of disease and rise of public health investments, purified water supplies, sewage treatment and disposal, vaccinations, improved hygiene, and economic growth to provide better diets for expectant mothers and growing children (Steckel 1995).

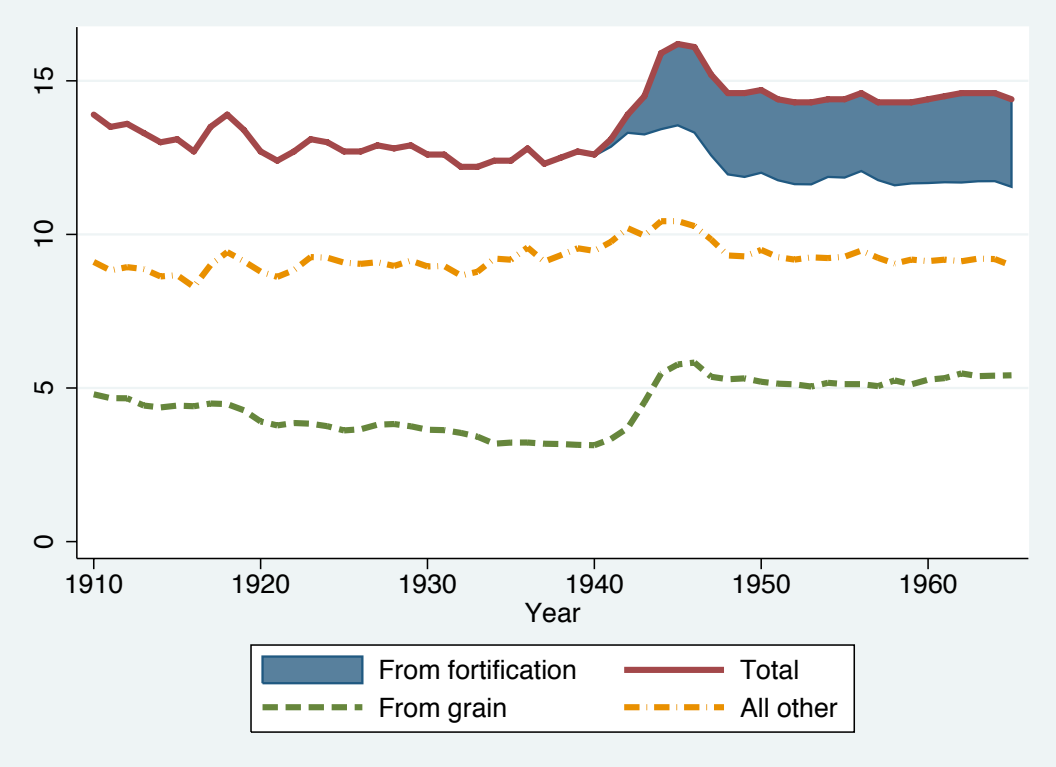

Figure 3: Historical iron levels in the United States food supply, 1910-1965

Sources: Gerrior, Bente and Hiza (2004)

Coinciding with the timing of the enriched bread campaign, the large and abrupt increase in iron consumption is striking in the face of its long secular decline. Between 1940 and 1946, we see a 16 percent increase in the iron content of the U.S. food supply,

which is directly linked to the fortification of flour and bread given by the shaded area of figure 1. By 1950, fortification provided 22 percent of all iron in the food supply, adding 2.7 milligrams daily. This increase by itself equates to 34 percent of the recommended daily allowance for men and 15 percent for women. The lower edge of the shaded area in Figure 3 provides the counterfactual average diet in the absence of the program. The long-term decline in iron consumption would have continued unabated.

Legal opinion in the 1940s was that without the emergency powers granted during wartime, the federal government had no authority to require enrichment. It was understood by all involved that War Food Order No. 1 would expire at the end of the war. This finally occurred on October 18, 1946, with the end of rationing in the United States. The proponents of enrichment were prepared for this day, as they had long been promoting regulation at the state level. As early as 1942, individual states began to pass their own enrichment legislation, often going farther than the federal order. South Carolina was the first and mandated the enrichment of both bread and flour, and later cornmeal. Louisiana and Alabama followed closely behind by passing legislation in 1943. In general, enthusiasm for government enrichment mandates was greater in the Southern states, which had higher levels of deficiency as well as less enrichment on a voluntary basis. By 1950, 26 states had passed regulations for enrichment..

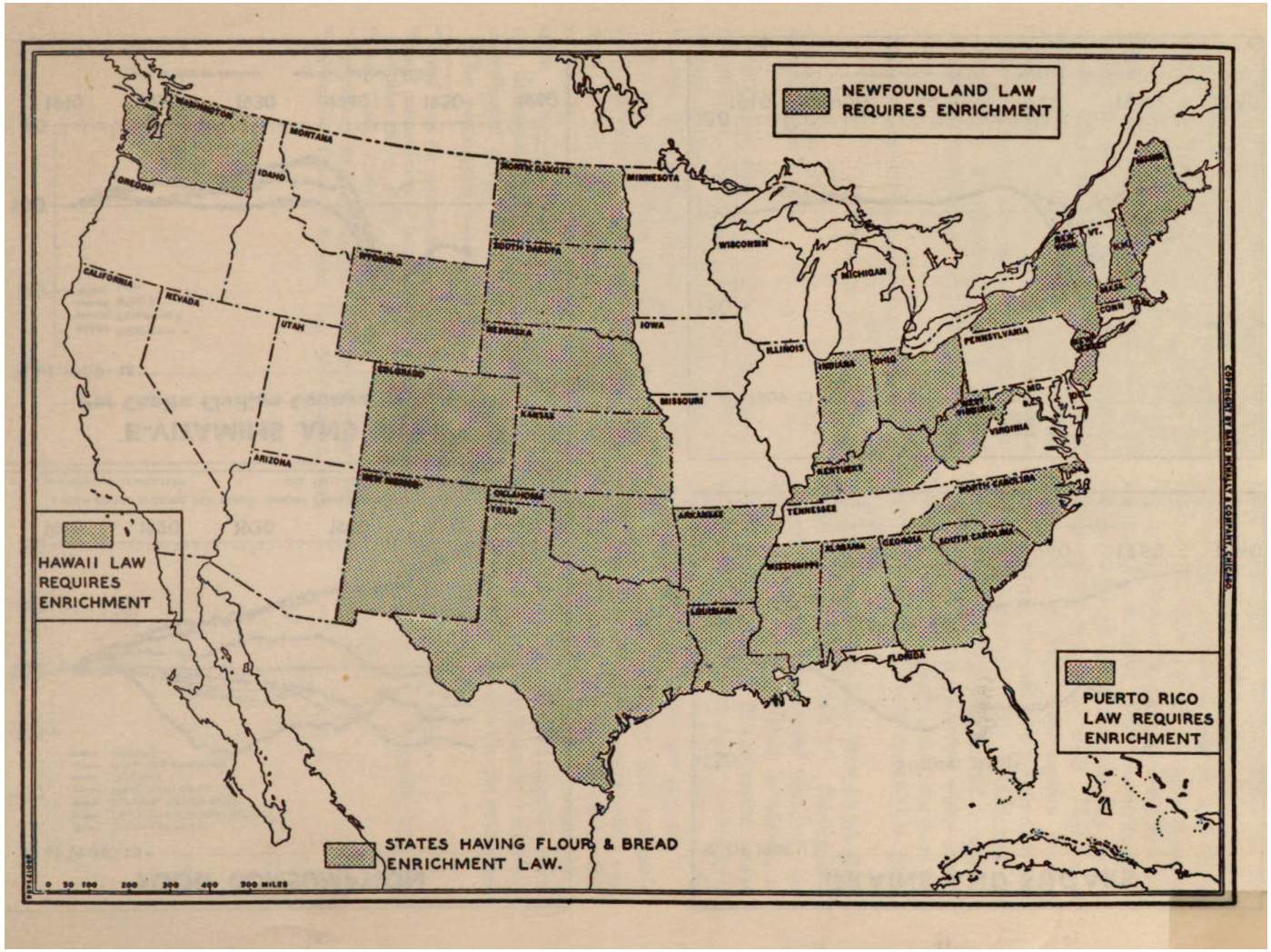

Figure 4: Extent of state enrichment legislation in 1950

Source: Committee on Cereals (1958).

Iron Deficiency after Bread Enrichment

Despite the sweeping nature of the program and the government's extraordinary effort to manipulate the contents of the food supply, information on the effectiveness of the program is scant and anecdotal. A study by Jeans et al. (1952) is typical of the type of evidence gathered during the period. Jeans and his associates calculated the contribution of enrichment to the diets of 400 low-income pregnant women in Iowa. Iron intake for these women would have been 8 to 18 percent lower if cereal foods were unenriched. Their conclusion exemplifies the views of the health community.

> "The fact that the bread is enriched has certainly aided the majority of women reported in this study from having even greater deficiencies in iron, thiamine, and riboflavin. The usefulness of the enrichment program is obvious." (Jeans *et al*. 1952, 33).

It appears that fewer cases of diseases linked directly to nutritional deficiency were addressed in physicians' clinics, and the rapid reductions in related deaths were clearly evident in the death registration data.[13] Like Jeans and his co-authors, much of the profession viewed the effectiveness of fortification as obvious and a foregone conclusion—nutritional scientists and public health professionals had revealed and then alleviated the scourge of micronutrient malnutrition. While plausible, the lack of careful statistical analyses, let alone comparisons between treatment and control groups, makes it difficult to know just how strong the program's effects were. Surely fortification in the 1940s helped alleviate some deficiencies, but to what extent and with what effects? Later

[13] Dr. Jolliffe, of the New York University College of Medicine, could testify as follows at a public hearing held by the War Food Administration, January 21, 1943: "I attribute to bread enrichment a marked and unmistakable decrease in florid beri beri and florid pellagra in my wards at Bellevue Hospital. In 1938-39 little bread was enriched; in 1942-43 seventy-five percent or more has been enriched in New York City. This has been accompanied by a decrease of three-fourths in our cases of florid beri beri and of two-thirds in florid pellagra," (Wilder and Williams 1944).

and large-scale National Health and Nutrition Examination Surveys of the 1960s and 1970s showed that the "hidden hunger" lingered in large pockets of the population.

Dietary intake surveys can help fill the statistical gap in the years immediately after World War II in the absence of national health examination surveys. The 1955 USDA dietary consumption survey provides further evidence of the broad impact fortification had on the American diet.[14] By at least 1948, nearly all of the white bread 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 2 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).

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

Table 2: 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

Source: 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, Table 8.

Panel B: Iron consumption increased for all income groups

	1936				1955			
	All	Lowest Income Third	Middle Income Third	Highest Income Third	All	Lowest Income Third	Middle Income Third	Highest Income Third
Iron (mg)	11.8	10.2	11.8	14	17	16.4	17	17.6

	Change in average iron consumption, 1936-1955			
	All	Lowest Income	Middle Income	Highest Income
Increase in mg	5.2	6.2	5.2	3.6
% Change	44%	61%	44%	26%

Source: 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, Table 19.

Prevalence data does not become available until the 1970s with the introduction of the National Health and Nutrition Examination Survey (NHANES). The results from these surveys are consistent with an important role for enrichment. The proportion of males with anemia had declined to 2.6 percent in 1976 and had fallen to below 1 percent by 1988 (Dallman, Yip, and Johnson 1984; CDC 2002). For infants, children, and women of childbearing age, the NHANES surveys exposed a lingering prevalence of deficiency and anemia: infants and toddlers 10 percent, white women 9 percent, black women 17 percent and Hispanic women 21 percent (Brotanek et al. 2008, Looker et al. 1997).

The public health community reassessed the adequacy of the enrichment program in light of the new evidence, introducing new nutritional guidelines and programs. Fortification of infant formula took on an urgency that had been previously been absent; a federal mandate was issued as part of the Infant Formula Act of 1980 (Anderson *et al.* 1982). New guidelines for supplementation during pregnancy were published. By the early 1990s, iron supplements were taken by 72 percent of women in the third trimester, and by 69 percent of those lactating (Cogswell et al. 2003). The WIC nutritional program played an increasingly important role in providing fortified formula for infants and fortified cereals for low-income children. Food producers rode the same wave of parental concern by introducing highly fortified breakfast cereals. Together, the measures caused a rapid increase in per capita daily iron availability in the food supply of the United States, which can be seen in **Error! Reference source not found.**. The size of the increase post-1970s is even larger than that of the 1943 bread and flour enrichment program. However, diminishing marginal returns in health set in fairly quickly with additional iron consumption. It is unclear whether the large post-1970s upsurge in the average is driven by increases in consumption at the lower end of the distribution for those most in need, or by increases for people already satiated. The most recent NHANES survey suggests that significant pockets of deficiency remain even today. In 2000, the prevalence was 10 percent for white women, 19 percent for black – non-Hispanic, and 22 percent for Hispanic women (CDC 2002). Further fortification of staple foods is unlikely to reach these at risk groups, and alternative strategies to reduce iron deficiency are needed.

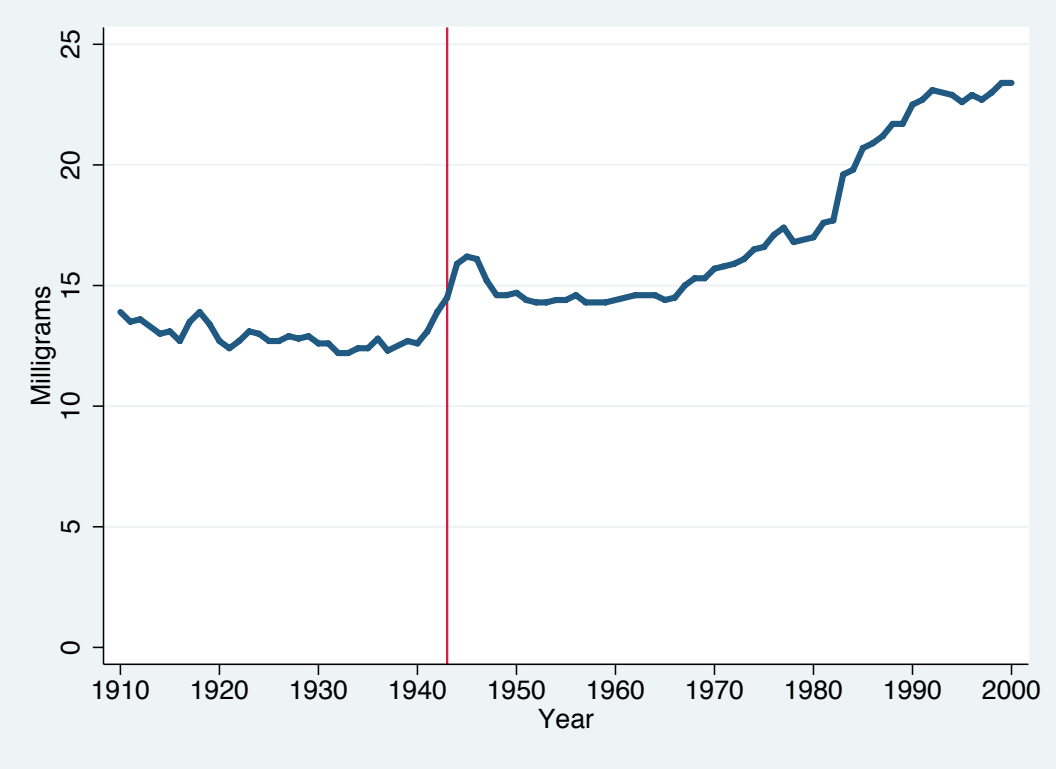

Figure 5: Per capita daily iron availability in the U.S. food supply, 1910-2000

Source: Data from Gerrior, Bente, and Hiza (2004).

CHAPTER III

III. THEORETICAL AND EMPIRICAL FRAMEWORK

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.[15] A lack of iron causes reduced work capacity. The mechanism works through a reduced 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).[16] 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

[15] 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).

[16] 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).

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).

Iron balance is a complicated mechanism that depends on a wide variety of dietary and bodily factors. The amount of iron in the body is determined by intake, loss, and storage. The body regulates its balance of iron mainly through adjusting the absorption rate. When iron stores are high or sufficient, the body absorbs less of the iron consumed. The absorption rate is determined by a number of factors: amount of iron in the body, rate of red blood cell production, amount and kind of iron in the diet, and the presence of substances that can inhibit or promote absorption. It varies anywhere from 1 percent to more than 50 percent of dietary iron consumed. The iron in meat is more easily absorbed than that in plants. The presence of vitamin C in the meal enhances absorption, while polyphenols in vegetables, tannins in tea, and calcium tend to inhibit iron absorption (Institute of Medicine 2001).

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 the analysis 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:

(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.[17] 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. 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

[17] 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.

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.

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 the model is quite simple. A child's skill is created using the following production function:

(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) will directly increase health, which then enters equation (2) to increase skill.

The model needs 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 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, such as antimalarial medication, reversing the sign.

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.

*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.

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. As mentioned above, 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 6 plots iron consumption against income at the SEA level. There is no strong relationship in the pre-program period. 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. I directly test this identifying assumption later in this section.

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.[18] 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 manuscript. However, the inability of contemporaneous economic variables, especially income, to fully 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

[18] R-squared is 0.51 from a regression of SEA average iron consumption in 1936 on the prices of beef, pork, poultry, seafood, eggs, milk, vegetables, beans, fruit, wheat flour, cornmeal, and bread. Prices are calculated as the quantity-weighted average of prices within a food group in a state economic area.

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.

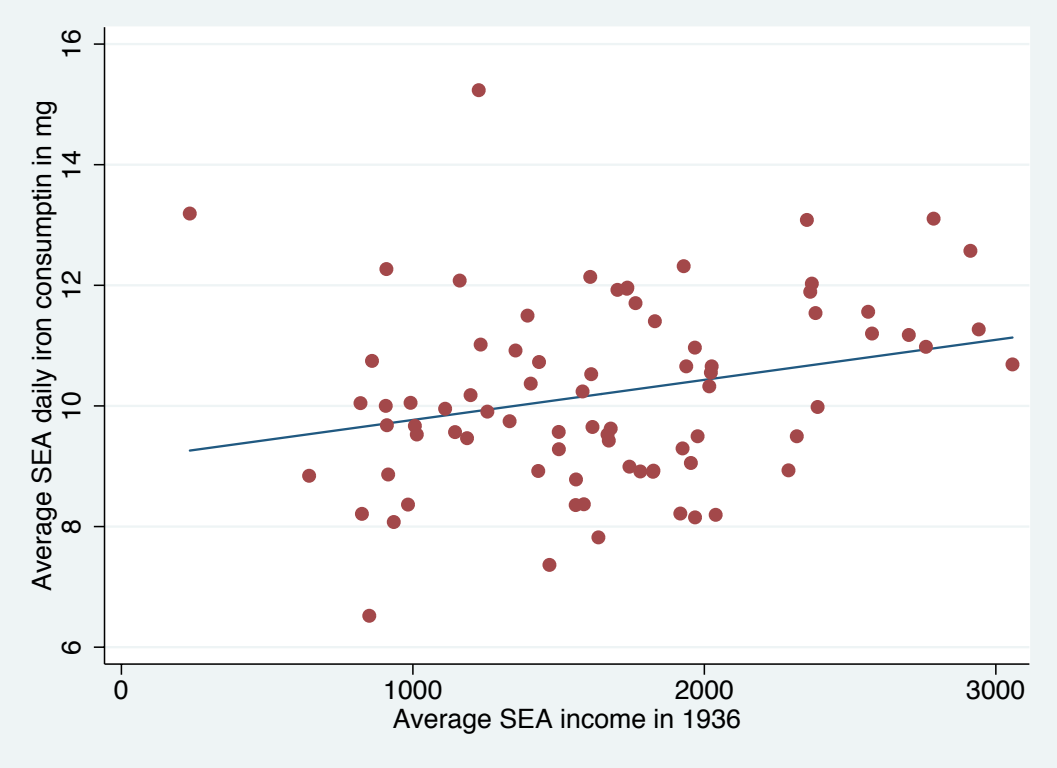

Figure 6: Income weakly predicts iron consumption for state economic areas

Notes: The point estimate from the simple regression 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.
Sources: Author's calculations using "Study of Consumer Purchases, 1935-1936" and USDA (2009). See data appendix for more details on variable construction.

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 experienced larger improvements in health from fortification, even if fortification raised everyone's intake of iron. The non-linearity of the effect of iron extends from the individual level to the geographic area. Figure 7 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,

$$(8) \qquad Y_{its} = \beta \cdot (IRON_s \times POST_t) + \delta_t + \delta_s + (\tilde{\delta}_r \times t) + X_{it}\theta_1 + X_{st}\theta_2 + \varepsilon_{its}$$

where Y signifies a health, 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.[19] All regressions include a

[19] When income is the outcome variable 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.

vector of individual level controls denoted by X_{it}, with some specifications also including a vector of area specific controls denoted by X_{st}.

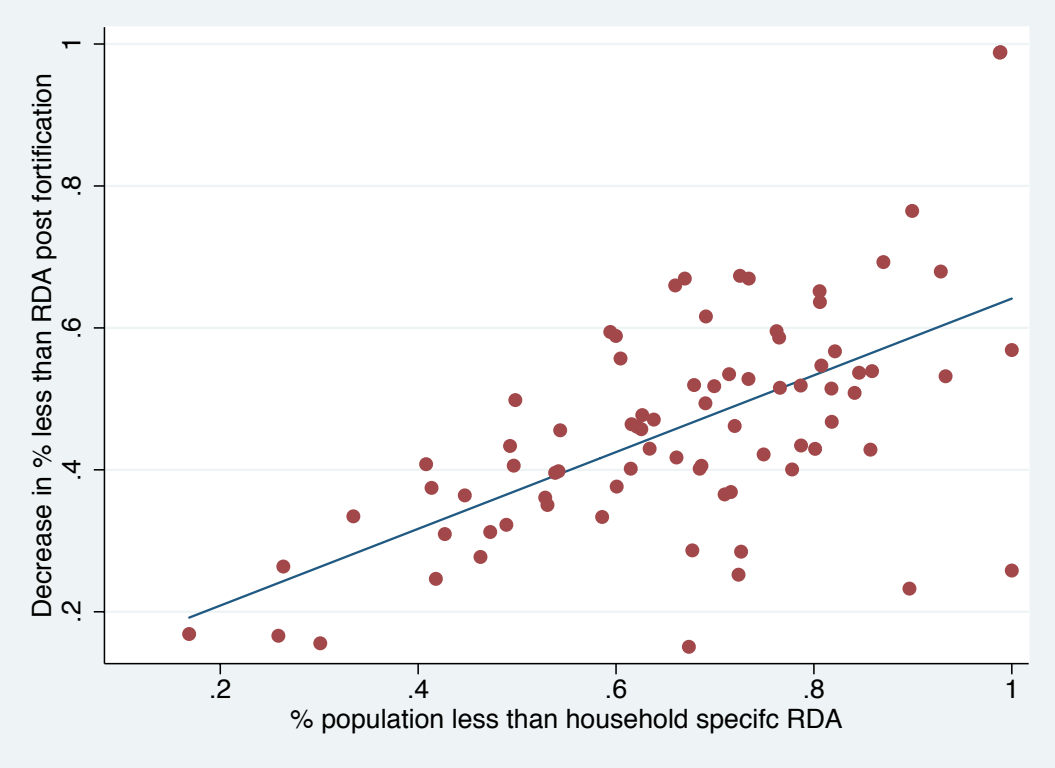

Figure 7: State economic areas with higher rates of inadequate diets in 1936 experienced larger reductions in inadequacy post-fortification in the counterfactual

Notes: RDA stands for recommended daily allowance and is specific to each household. The RDA is based on the age and gender mix of the household.
Source: Author's calculations using "Study of Consumer Purchases, 1935-1936" and USDA (2009). See data appendix for more details on variable construction.

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 pre-program average iron consumption level in his SEA. 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.[20] The hypothesis is that areas with low-iron consumption before fortification had 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.

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 test of the identifying assumption by individually regressing SEA average iron consumption on several pre-program SEA characteristics. The assumption is validated if these characteristics do not predict iron consumption, and the exercise suggests specific controls if characteristics do have predictive power.[21] Point estimates provided in Table 3 suggest that iron consumption is uncorrelated with several economic, labor market, agricultural, and demographic characteristics. For example, per capita war spending, New Deal spending, retail sales and manufacturing 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

[20] There is no coefficient estimate for $IRON_s$ because the regressions include area fixed effects.
[21] 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).

over the 1930s. Results suggest adding specific controls to the regressions in the following empirical sections.

Table 3: Predictors of average iron consumption in state economic areas

	Economics and labor market conditions			Migration	Retail Sales	
	War spending p.c.	Unemployment rate (1937)	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)					
	% Own home	Median home value ($1000)	Growth med. home value (1930-40)	% with 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 precipitation
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			
	Total grants p.c.	Relief grants p.c.	Public work grants p.c.	Loans p.c.
Point estimate	0.001	0.013*	0.004*	0.20
S.e.	(0.003)	(0.008)	(0.0023)	(0.25)
R^2	0.007	0.074	0.012	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." All other data is from Haines (2010).

CHAPTER IV

IV. CONTEMPORANEOUS EFFECTS ON ADULTS

Income and Labor Market Outcomes for Men

In this chapter, I estimate the changes between 1940 and 1950 in wage and salary income, labor force participation, and weeks worked that are associated with increases in iron consumption. 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. The full sample includes wage and salary workers aged 18 to 60 with positive income. I exclude from the income and weeks regressions observations without educational attainment information or that are recorded as full-time or part-time students. 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. Table 4 contains summary statistics for outcome variables, and the Data Appendix provides a more detailed discussion.

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).

Table 4: Labor market summary statistics for men

		Areas with iron consumption	
	Full sample	Below-median	Above-median
Income in 1940	1,005	917	1,095
	(232)	(245)	(180)
in 1950 (1940$)	1,522	1,495	1,550
	(259)	(236)	(281)
Labor force participation in 1940	0.86	0.86	0.86
	(0.03)	(0.03)	(0.03)
in 1950	0.87	0.87	0.87
	(0.03)	(0.03)	(0.02)
Weeks worked in 1940	43.6	44.0	43.2
	(2.4)	(2.2)	(2.6)
in 1950	44.5	44.8	44.2
	(1.8)	(1.8)	(1.8)
Hours worked in 1940	47.4	47.4	47.3
	(3.8)	(3.6)	(3.9)
in 1950	47.1	47.1	47.1
	(3.5)	(3.2)	(3.7)
Observations	82	41	41

Notes: Averages over SEAs are reported with standard deviation below in parentheses. Income data includes all observations with positive income, wage and salary employment as first occupation, and over 17 years old.
Source: Decennial census microdata provided by IPUMS.

Table 5 presents point estimates for β, the coefficient on $(IRON_s \times POST_t)$, as described in equation (8).[22] Each entry is from a separate regression for labor market outcomes. The coefficient is expected to be negative if the hypothesis is correct, as areas with low-iron consumption should experience relatively large gains. Standard errors are clustered at the SEA by year level according to 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 (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.

Income Gains for Men[23] 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.[24] 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 3.1 percent difference in income growth.

[23] I split the analysis by sex based on the radically different labor market incentives facing men and women during the 1940s.
[24] The four educational attainment categories are less than high school, high school, some college, and college.

Table 5: Results for contemporaneous adult male labor market outcomes

	(1) Log wage and salary income	(2) Labor force participation	(3) Conditional weeks
(A) All men	-1.72	0.54	-0.058
	(0.46)	(0.18)	(0.068)
(B) Add controls for 1937 unemployment, per capita war spending, and 1936 income	-1.22	0.52	-0.19
	(0.45)	(0.16)	(0.07)
(C) Drop South census region	-1.28	0.37	-0.15
	(0.49)	(0.16)	(0.07)
(D) Drop industry and occupation	-1.31	n.a.	n.a.
	(0.577)		
Age groups			
(E) Under 28	-3.25	1.46	-0.32
	(1.11)	(0.35)	(0.16)
(F) Aged 28-37	-1.4	-0.13	-0.22
	(0.69)	(0.16)	(0.11)
(G) Aged 38-48	0.77	0.30	-0.14
	(0.75)	(0.22)	(0.08)
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 (8) 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 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, and SEA average income in 1936. All interacted with $POST_t$.
Sources: 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."

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) includes the 1937 SEA-level unemployment rate, 1936 SEA average income, and per capita World War II spending, 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.[25]

Using the estimate from row (B) with the full set of controls, a one-milligram difference in average iron consumption is associated with a 1.22 percent difference in wage and salary income growth. The fortification program increased per capita daily iron consumption by 2 milligrams (Gerrior, Bente and Hiza 2004). Increasing SEA average iron consumption by 2 mg translates into a 2.4 percent increase in income between 1939 and 1949. 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 (Bleakley 2007; Barreca 2010).

[25] The addition of controls for new deal spending, net migration, and growth in home values does not substantially change the point estimates (results unreported). Moreover, the wage distribution compressed during the 1940s (Goldin and Margo 1992). Point estimates from adding the SEA Gini coefficient for male wage and salary 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. 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).

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 in 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 or industrial shifts. 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 (1.8 mg) at the SEA level implies a 5.9 percent difference in income growth for men under the age of 28 and a 4.3 percent difference for men between the ages of 28 and 37. The impact of iron fortification fades at older ages.

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 (Gerrior, Bente and Hiza 2004). While fluctuations in the consumption of non-enrichment nutrients briefly improved overall nutritional status, they do not appear to drive the empirical results I find in this section. Appendix II provides evidence from the medical literature, diet survey data, and regressions including 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 (ρ = 0.83) and SEA level (ρ = 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 based on the estimated prevalence of niacin deficiency during the early 1940s (Kruse *et al.* 1943). 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.

Alternative Measures of Treatment Intensity: The results in Table 5 are not sensitive to alternative measures of treatment intensity. Table 6 reports point estimates for β from equation (8) for six different measures of iron status, all measured at the level

[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.

Table 6: 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 iron measure
(A)	mg daily iron consumption	-0.0122 (0.00449)	-0.0324 (0.0111)	2.2%
(B)	Proportion of households consuming less than RDA	0.0596 (0.0404)	0.277 (0.0845)	1.2%
(C)	Proportion of households consuming less than 75 percent of RDA	0.0910 (0.0363)	0.142 (0.0757)	1.6%
(D)	Proportion of households consuming less than 50 percent of RDA	0.136 (0.0737)	0.354 (0.130)	1.3%
(E)	Proportion of households consuming less than 25 percent of RDA	0.391 (0.0612)	0.258 (0.155)	2.2%
(F)	Log daily iron consumption	-0.127 (0.0497)	-0.356 (0.122)	2.0%

Notes: See Table 5 for details on the specification. Point estimates in row (A) repeat base results Table 5. Deficiency measures in rows (B) through (E) 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.
Sources: All individual level data comes from the 1940 and 1950 census microdata provided by IPUMS. SEA-specific 1937 unemployment rates, average income in 1936, and WWII contract spending comes from Haines (2010). Average iron consumption by SEA is calculated by the author from the "Study of Consumer Purchases, 1935-1936." Recommended Daily Allowances (RDAs) for iron are published by the Institute of Medicine (2001).

of the sate economic area. For purposes of comparison, row (A) reproduces the base results from Table 5 for the full sample of men and for those between the ages of 18 and 27. One common measure used in the nutrition literature is a binary indicator for iron deficiency. Row (B) 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 (C) through (E) reduce the cutoff for deficiency to consumption less than 75 percent, 50 percent, and 25 percent of the RDA. Row (F) 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 1.3 and 2.2 percent.

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 chapter 3 in relation to the theoretical model, all three channels may have worked simultaneously, but not necessarily in accordance with one another. 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 5 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.[27] 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. Men under the age of 28 are the only exception,

[27] Column (2) adds men attending school back into the sample.

with a relatively large point estimate. To get a sense of the effect's economic significance, a one standard-deviation difference in iron consumption implies a 2.6 percentage point difference in labor force participation. In 1940 and 1950, the average labor force participation rate for men aged 18 to 27 was 87 percent and 81 percent respectively.

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 (8). Using the full set of data from 1920 to 1950 give similar estimates for β to those in Table 5 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.

One concern is that the income results could be driven by these differential changes in participation rates. That is, low-wage workers might be dropping out of the workforce in low-iron areas or high-wage workers dropping out in high-iron areas. However, it appears that much of the difference can be explained by larger increases in young men attending college in the low-iron areas. Likely, these college-going men would have been relatively high-income earners, biasing the income result towards zero.

Adjustments along the extensive margin do not seem prevalent and are unlikely to drive the income results. 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). Again, the one exception is for young men of ages 18 to 27, who worked 46

hours per week on average in 1940 in areas with below-median iron consumption. For this age group, a one standard-deviation difference in pre-intervention iron consumption implies a difference in weekly hours worked of around three-fourths of an hour, or 1.6 percent of mean hours worked (44) in 1940.

Column (3) of Table 5 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. The point estimate for young male workers is again the largest, corresponding to an increase of roughly one-half week of work or 1.2 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.

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 commonly report estimates of the average treatment effect 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.[28] I convert the average

[28] Evidence suggests that patients with subclinical iron deficiency receive benefits from iron supplementation (Horton and Ross 2003). The point estimates in Table 5 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.

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.

The fortification program affected individuals differentially, 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 5, a difference of 1 mg in pre-intervention iron consumption implies a differential 1.2 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 2.4 percent. Dividing by the proportion of the sample that consumed less than 75 percent of the RDA suggests that the program increased incomes by 6.1 percent at the individual level. Applying this procedure to labor supply suggests that the program increased weeks worked by 2.2 percent at the individual level. The remainder, 3.9 percent, can be interpreted as the productivity effect (wage/hr). A major contribution of this study 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 to 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).

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. Table 7 reports summary statistics for women's labor market outcomes.

No clear conclusions about the effects of iron fortification on women can be made on the basis of Table 8. In the full women's sample, the estimate of β in the income regression 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 in the income regressions, 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.[29]

Table 7: Labor market summary statistics for women

		Areas with iron consumption	
	Full sample	Below-median	Above-median
Income in 1940	605	539	671
	(146)	(139)	(123)
in 1950 (1940$)	868	837	900
	(156)	(156)	(152)
Labor force participation in 1940	0.26	0.27	0.24
	(0.07)	(0.07)	(0.07)
in 1950	0.28	0.28	0.27
	(0.05)	(0.05)	(0.05)
Weeks worked in 1940	39.0	39.4	38.5
	(3.4)	(3.0)	(3.7)
in 1950	35.9	36.2	35.7
	(3.7)	(3.7)	(3.7)
Hours worked in 1940	41.5	41.3	41.7
	(2.4)	(2.4)	(2.5)
in 1950	38.1	38.1	38.2
	(1.8)	(1.9)	(1.8)
Observations	82	41	41

Notes: Averages over mean SEA values are reported with standard deviation below in parentheses. Income data includes all observations with positive income, wage and salary employment as first occupation, and over 17 years old.
Source: Decennial census microdata provided by IPUMS.

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

[29] Labor force participation rates for both married and unmarried women increased more in high-iron consumption areas. Similar to the men, the majority of this increase can be explained by a larger increase in the school enrollment rate for women over the age of 18 in the low-iron areas.

Table 8: Results for contemporaneous adult female labor market outcomes

	(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 (8) 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 $POST_t$.
Sources: 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."

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.[30] Moreover, this period is the start of the baby boom and important changes in family structure and fertility choices. Perhaps health benefits from iron fortification caused larger changes in expected, but not yet materialized, children in the low-iron areas, with an accompanying impact on labor market decisions.

Finally, the prevalence of iron deficiency in women might not have declined as much as that of men in response to the enriched bread program. NHANES II suggested that in the late 1970s iron deficiency remained widespread in women at 11 percent, but not in men at under 1 percent (Looker *et al*. 1997).

Contemporaneous Adult Health Outcomes

This study hypothesizes gains in economic outcomes from iron fortification in the United States derived from improvements in nutrition and health. The most immediate improvement in health would have been increased iron stores in the body. A dataset that contains biological assay measures of iron deficiency from both before and after the intervention would be necessary to credibly assess the effectiveness of the program on this primary goal. Unfortunately, no large national surveys collected blood samples in the years immediately following the enriched bread campaign. The first large-scale health survey to provide plausible prevalence estimates did not occur until the mid-1970s

[30] The 1950 census asked a detailed set of questions of only one person in the household, called the sample-line person. Within a household, the husband's wage is never reported if the wife's wage is given, and vice-versa.

with the NHANES I survey. As such, we must rely on anecdotal and suggestive evidence that the rate of iron deficiency declined in response to the fortification campaign.

That iron enriched bread and flour improved the health of the U.S. population seems plausible. Multiple historical sources suggest that iron deficiency was widespread before the intervention, that nearly the entire population consumed white bread in large quantities, and that iron deficiency had declined a few decades later. Blood sample case studies from before 1943 point to a prevalence rate of anemia of 20 to 30% (Kruse *et al*. 1943). The actual rates of deficiency were likely higher because we know from modern surveys that anemia rates underestimate the full extent of iron deficiency (Yip 1994). The prevalence of anemia had fallen to 5 to 15% by the early 1970s (Meyers *et al*. 1983). The 1943 bread enrichment campaign, therefore, had an upper bound of the reduction in deficiency of 25 percentage points. Note, however, that iron status is a continuous measure, and a person's health would improve even if the clinical cutoff for deficiency were not crossed.

The remainder of this section examines the subsequent health impact of improvements in iron status associated with fortification. One observable, but extreme, measure of health is mortality. Iron deficiency is rarely the primary cause of death, but evidence is growing on its importance as an underlying cause. Iron deficiency is often a contributory factor to other illnesses, exacerbating morbidity and mortality. Higher mortality rates are almost always associated with anemia (Nissenson and Goodnough 2003). The nutritional deficiency is not just a symptom of the primary illness. Low levels of hemoglobin are associated with heightened risks of mortality in sufferers from

renal failure, congestive heart failure, and ischemic heart disease (Collins *et al*. 2001; Kushang *et al*. 2009; Carson *et al*. 1996).

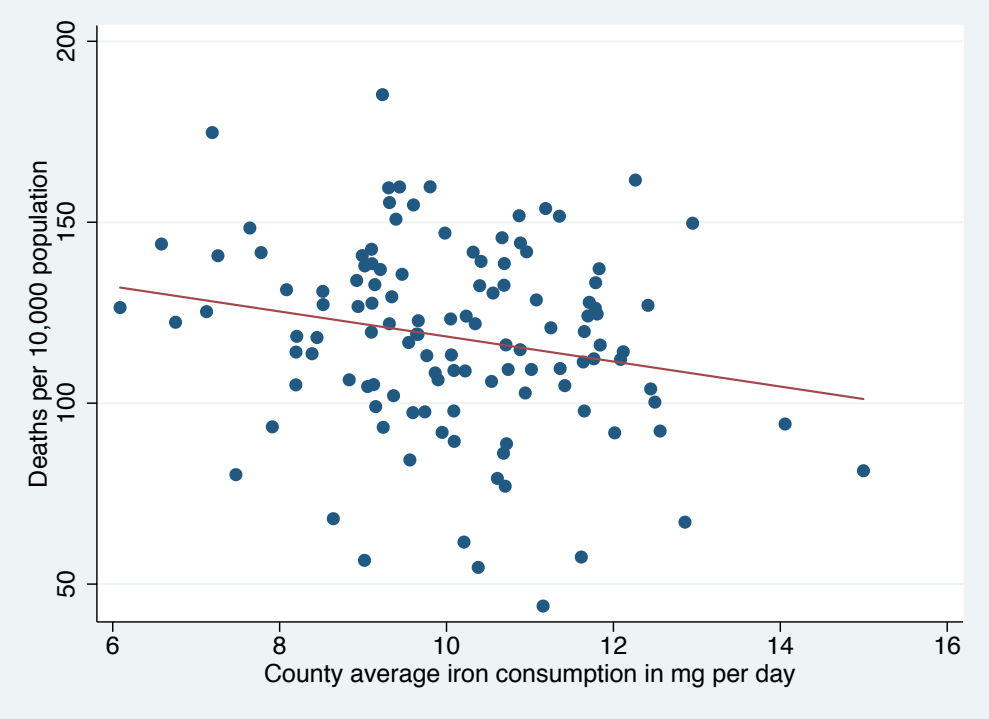

Figure 8: High-iron consumption correlated with low mortality rates in the cross-section

Notes: Each dot represents the mortality rate and average iron consumption of a state economic area in 1936.
Sources: Mortality data is from the published volumes of the *Vital Statistics of the United States* and provided by Price Fishback. Iron consumption is from the author's calculations using the "Study of Consumer Purchases in the United States, 1935-1936," and the *USDA National Nutrient Database.*

The enriched bread campaign may have improved crude mortality rates in counties with low-iron consumption relative to high consumption counties. Figure 8 illustrates that higher levels of iron consumption are associated with lower death rates, at least in the cross-section for 1936. While the figure is suggestive that a correlational relationship exists, we must take care to make a causal interpretation. Unobservable

time-invariant county characteristics or unobservable gradually evolving factors possibly drive the observed relationship. Therefore, I move on to regression analyses to control for a number of these omitted factors.

A number of advances in health care concurrent with the enrichment campaign complicate identification of the impact of iron fortification on mortality. First, penicillin was introduced in 1943 and rapidly diffused across the nation. Secondly, Congress passed the Hill-Burton Act in 1946, which provided federal grants in the form of matching money for the construction of hospitals. Hill-Burton led the way for the general expansion of access to health care during the 1940s.

Complications notwithstanding, I explore the possibly that iron fortification reduced mortality by improving the iron status in the population. The identification strategy roughly follows the framework outlined above in chapter III, but now with counties as the unit of analysis rather than SEAs. The published volumes of the *Vital Statistics of the United States* provide the number of total deaths occurring yearly in each county of the death registration area for the period 1930 to 1950.[31] County population estimates, used as the denominator to construct mortality rates, come from the decennial census counts and are interpolated for intercensal years using a constant growth rate. I estimate equations of the following general form:

$$(9) \quad \ln(M_{ct}) = \beta \cdot (IRON_c \times POST_t) + \delta_c + \delta_t + (\tilde{\delta}_c \times t) + X_{ct}\theta + \varepsilon_{ct}$$

The dependent variable is the natural log of the crude mortality rate for county c in year t. The variable of interest is the level of iron consumption in county c measured in 1936 interacted with $POST_t$, an indicator equal to 1 for observations in 1944 and after. County

[31] County death counts were kindly provided to me by Price Fishback.

fixed effects, a set of year indicators, and a county-specific linear time trend are included. The vector X contains a set of variables that vary by county and year that include the proportions of the county that are black, urban, and foreign born (Haines 2010). The sample is a balanced panel of 130 counties from 1930 to 1950. Standard errors are clustered by county because of the possibility for serial correlation in the error term.

Table 9 reports estimated coefficients from separate regressions of the log crude mortality rate. The fortification program was implemented in 1943 as a preventative strategy consisting of small daily increases in iron intake. Iron stores improve gradually as the body slowly absorbed the augmented dietary intake. The second row of results limits the impact of the program to begin in 1944.[32] Results are consistent with iron fortification causing a slight decline in crude mortality rates. The first column uses county average iron consumption as the measure of treatment intensity, and suggests that a 1 mg difference in pre-intervention iron consumption is associated with 0.9 percent decline in mortality rates.

The second column uses the proportion of a county's population that consumes less than the recommended daily allowance as the measure of treatment intensity. The results are roughly similar to those found in the first column. Moving from an entire population of a county consuming inadequate amounts of iron to none of the population is associated with a decline in the crude mortality rate of 6.8 percent. The remaining columns report results using consumption and deficiency rates for niacin and thiamin as the measure of treatment intensity. I find no evidence that fortification of these nutrients is associated with relative changes in mortality rates over the 1940s.

[32] Results using 1945 as the cutoff for the $POST_t$ indicator are similar.

Table 9: No evidence of a strong relationship between fortification and mortality rates

	Iron		Niacin		Thiamin	
	mg	% Def.	mg	% Def.	mg	% Def.
Break in year						
1943	0.91*	-6.8*	-0.01	-2.7	0.8	-10.3
	(0.485)	(4.1)	(0.22)	(5.1)	(1.4)	(6.6)
1944	0.88*	-10.7**	-0.07	-0.3	1.6	-7.1
	(0.517)	(4.27)	(0.22)	(4.8)	(1.0)	(6.6)

Notes: Each point estimate comes from a separate estimation of equation (9) and gives the difference in log crude mortality rate corresponding to a 1 mg difference in iron consumption, or moving from 0 to 100 percent of the population deficient. Heteroskedasticity-robust standard errors have been corrected for correlation at the county level and are reported in parentheses. All regressions include county specific linear time trends and include a vector of controls for percent black, percent urban, and percent foreign born. The first row of results has $POST_t$ equal to 1 for all observations in 1943 and after, whereas the second row allows the fortification program to have an effect starting in 1944.
Sources: Mortality data come from the *Vital Statistics of the United States*. Total, urban, black, and foreign born population come from Haines (2010). Average consumption of iron, niacin, and thiamin are calculated by the author from the "Study of Consumer Purchases."

For the counties in my sample, the unweighted mean mortality rate in 1943 was 107.2 deaths per 10,000 population.[33] Doubling the coefficient estimate gives a full reduced-form effect of the program of reducing mortality rates by 1.8 percent. Thus, taking the estimates at face value, the program appears to have reduced deaths by 2 per 10,000 population on average. The decline accounts for about 15 percent of the total decline in mortality rates between 1943 and 1950. The size of the estimated effect seems implausibly large in relation to what one might expect from medical advances from the same period. The results here must be qualified by the absence of controls for the introduction penicillin, expanded hospital access, and the successful treatment of tuberculosis and diarrhea that were all innovations during the mid-1940s.

[33] The mean mortality rate weighted by county population was 141.8 per 10,000. The two bracket the 109 per 10,000 crude mortality rate for the entire United States in 1943 (*Vital Statistics of the United States 1943).*

CHAPTER V

V. CONTEMPORANEOUS EFFECTS ON CHILDREN

School Enrollment and Attendance

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.

Table 10: School enrollment summary statistics

		Areas with iron consumption	
	Full sample	Below-median	Above-median
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: Sample includes children between the ages of 8 and 17 unless otherwise noted.
Sources: School enrollment data provided by census microdata (IPUMS).

All individual-level outcome data and demographic controls come from the Integrated Public Use Microdata Series (IPUMS, Ruggles *et al*. 2010). The basic specification uses census data from 1940 and 1950 as these years bracket the iron fortification program. Additional data from 1910 through 1950 are used to control for gradually evolving SEA-specific unobserved characteristics. Table 10 contains summary statistics for school enrollment variables.

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.[34] Table 11 presents point estimate for β, the coefficient on $(IRON_s \times POST_t)$. Each entry is from a separate estimation of equation (8), 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 11 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.

[34] See the data appendix for additional details on the school enrollment indicator.

Table 11: 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 (8). 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).
Sources: 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."

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.[35] 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.[36]

Using the five decades of census data, I explore the timing of the relationship between school enrollment and iron consumption. Figure 9 plots the estimated coefficient on $IRON_s$ from regressions using each year of census data as a separate

[35] State economic area is not included as a geographic identifier in later censuses.
[36] 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 -0.007 (0.003) for iron, 0.0008 (0.002) for niacin, and 0.016 (0.016) for thiamin.

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 and include zero. The timing of relative gains in school enrollment for low-iron consumption SEAs coincides with the federal fortification mandate.

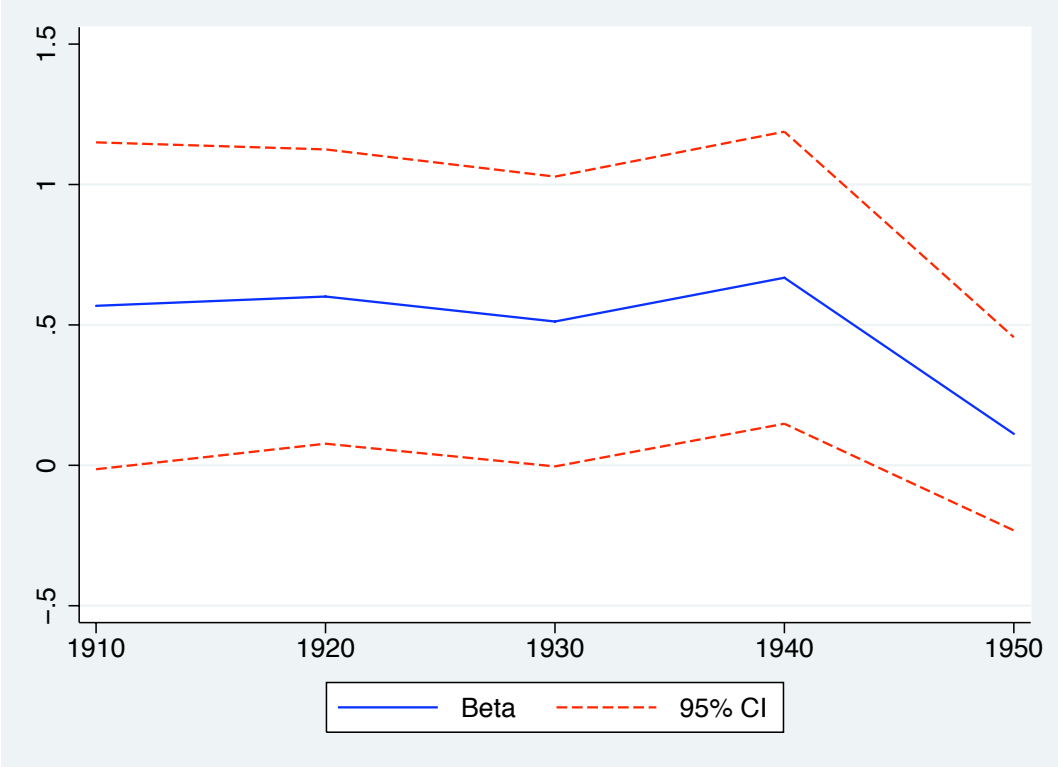

Figure 9: School enrollment gains in low-iron consumption areas coincide with enrichment program

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.
Sources: School enrollment data is from the IPUMS decennial census microdata. Average iron consumption and 1936 income by SEA is calculated by the author from the "Study of Consumer Purchases, 1935-1936."

As a further test, I estimate the two-period specification of equation (8) separately for the 1910-1920, 1920-1930, and 1930-1940 samples. In none of these decades are changes in school enrollment correlated with 1936 iron consumption.[37]

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. As an alternative strategy, I use average SEA measures of income and education. 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 11, 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 potentially endogenous, 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

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

that groups closer to the margin of school enrollment experience larger effects from iron fortification. Rows (C) through (H) of Table 11 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 chapter III, the point estimates suggest males experienced larger improvements in enrollment from iron fortification than did females. The coefficient for males 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.

Contemporaneous Infant Health Outcomes

Nutrition while *in utero* and infancy plays an important role in the growth and development of newborns. Maternal malnutrition, and iron deficiency in particular, can lead to stillbirths, slow intra-uterine growth, pre-term births, and low birth weight (Scholl and Hediger 1994, Steer 2000, Scholl 2005, Zeng *et al*. 2008). In this section, I test whether the fortification program had an observable effect on stillbirths and infant mortality through improved maternal and neonatal iron status.

The empirical strategy follows that outlined in chapter II. The analysis is at the county-level instead of the state economic area. Infant mortality and stillbirth data was digitized from the published volumes of *Vital Statistics of the United States*, and kindly provided by Price Fishback. The dataset includes births, infant deaths and stillbirths in a balanced panel of 130 counties from 1931 to 1950. Dependent variables include log mortality and stillbirth rates, calculated relative to the number of births in county c in year t. All specifications include a county-specific linear time trend to control for any gradually evolving trends in the dependent variables. A vector of controls for percent black, percent urban, and percent foreign born are included in each regression. Standard errors are clustered at the county level.

The timing of treatment is important because of the use of annual data and the year in which the trend break is specified to occur. I include specifications that allow the trend break to occur in different periods. The fortification program began in January 1943. By definition, treatment occurs while *in utero* for stillbirths. A trend break in 1943 implies fortification began in 1943, early enough to improve the iron status of expectant mothers.

Defining treatment is slightly more complicated for infant deaths. If the main impacts occur *in utero* then infants born before, and even into, 1943 do not receive treatment. However, if infant iron deficiency has an independent effect on mortality *and* fortified bread was consumed by infants, then a 1943 trend break is appropriate. Bread consumption by infants was minimal. Thus, the main mechanism for improved infant iron status would be through increased iron available in the mother's milk. While the proportion of mothers who breastfed their newborns was historically high in the 1940s, still only around 46% of mothers did so (Hendershot 1984). Iron fortified bread and flour would have little impact on the iron status of newborns outside of mother's milk.

Table 12: Impact of fortification on infant mortality and stillbirths

	Iron consumption		Niacin consumption		Thiamin consumption	
	1943	1944	1943	1944	1943	1944
Infant mortality	1.12	1.28	1.10*	0.839*	4.94	1.42
	(1.41)	(1.12)	(0.63)	(0.50)	(3.19)	(2.71)
Stillbirths	2.80*	2.73	2.12***	2.34**	1.14*	6.24
	(1.64)	(1.84)	(0.803)	(1.03)	(6.75)	(6.19)

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each point estimate comes from a separate estimation of equation (9) and gives the difference in log infant mortality or stillbirth rate corresponding to a 1 mg difference in iron consumption. Heteroskedasticity-robust standard errors have been corrected for correlation at the county level and are reported in parentheses. All regressions include county specific linear time trends and include a vector of controls for percent black, percent urban, and percent foreign born. The first column of results has $POST_t$ equal to 1 for all observations in 1943 and after, whereas the second column allows the fortification program to have an effect starting in 1944.

Sources: Mortality data come from the *Vital Statistics of the United States* provided by Price Fishback. Total, urban, black, and foreign born population come from Haines (2010). Average consumption of iron, niacin, and thiamin are calculated by the author from the "Study of Consumer Purchases."

Results are reported in Table 12. Rows correspond to dependent variables and columns to specifications. I separately examine the effect of each fortification nutrient: iron, niacin, and thiamin. Each point estimate comes from a separate estimation of equation (9) and gives the percent difference in the infant mortality or stillbirth rate that corresponds to a 1 mg difference in consumption of the particular micronutrient. The first column for each nutrient allows for the trend break in 1943, whereas the second column allows the fortification program to have an effect starting in 1944.

I do not find conclusive evidence in favor of fortification improving infant or fetal health. In general, the estimated coefficients across all nutrients are consistent with fortification reducing infant mortality and stillbirths. The results for iron and thiamin consumption are imprecisely estimated but large in magnitude. Counties that started with higher niacin consumption before the intervention, experienced relatively smaller gains in stillbirth rates after fortification. Estimates suggest that the additional 3 mg of niacin provided by the program to the average diet reduced stillbirth rates by just over 2 percent per one milligram difference in consumption (Gerrior, Bente and Hiza 2004).

The results are far from conclusive. Concerns remain that the identification assumptions are invalid. The empirical strategy assumes any deviation from trend not absorbed by common year effects can be ascribed to fortification. It appears that counties with relatively high-iron and niacin consumption before the program were hit by an unobserved shock to the mortality rate that caused positive deviations from trend. Mortality rates declined at a faster pace in the subsequent years as the shock dissipated. Figure 10 plots average rates over time for counties above and below-median consumption. Shocks to mortality and stillbirths in 1943 are implausibly large to

attribute to iron deficiency and fortification alone. The results must be qualified by the absence of controls for penicillin introduction, the Emergency Maternal and Infant Care program, new treatments for diarrhea, and expanded hospital access in general. Characteristics of parents who chose to have a child during the war years may also bias the results if changes did not occur commonly over counties.

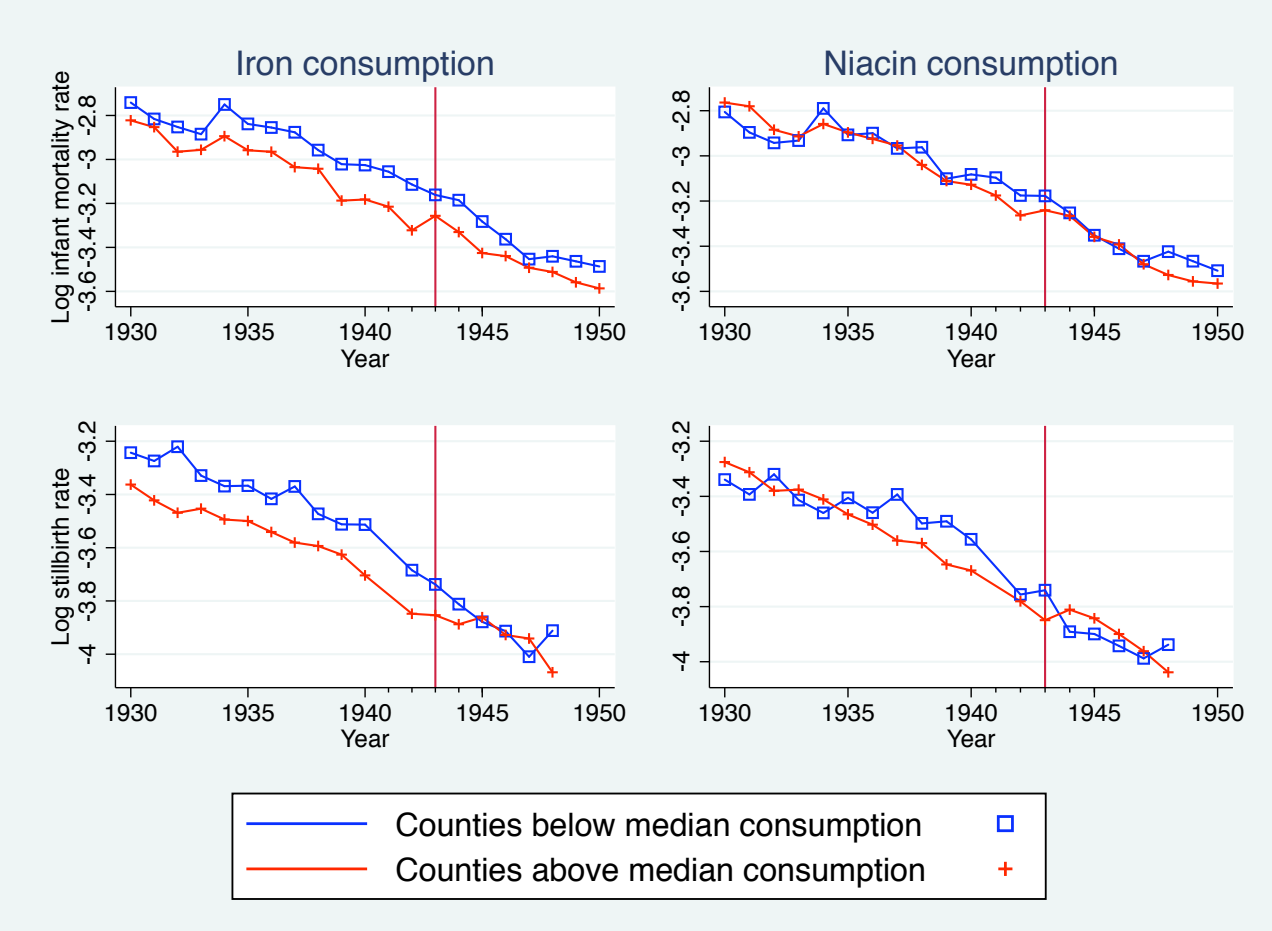

Figure 10: Presence of 1943 shocks to infant mortality correlated with iron consumption

Notes: Each point represents the average log mortality or stillbirth rate in a given year for the portion of the sample that is below or above-median iron or niacin consumption. The vertical line identifies 1943.
Sources: Mortality data come from the *Vital Statistics of the United States* provided by Price Fishback. Average consumption of iron and niacin are calculated by the author from the "Study of Consumer Purchases."

CHAPTER VI

VI. 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 who potentially benefitted from the iron fortification mandate by looking at their corresponding adult outcomes using the 1970 decennial census microdata and Health and Retirement Study data from 2008. I find that early-life iron deficiency leads to lower income as an adult and an increased incidence of chronic disease and mortality. 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.

Iron Deficiency and the "Developmental Origins" Hypothesis

Recent work in the biological sciences argues that poor early-life nutrition irreversibly alters the developmental path of a child, leading to the so-called "developmental origins" hypothesis. In the health fields, early adverse conditions have been linked to a heightened propensity for chronic disease, morbidity, and mortality during adulthood (Gluckman and Hanson 2004). Economists contributed to the literature by showing the consequences of deprivation on economic outcomes, and by emphasizing identification of causal effects (Almond and Currie 2011). That the impact of early-life health insults might extend into adulthood is intriguing. Reductions in

chronic disease would be, while not totally unexpected, a frequently neglected benefit of maternal and infant health programs.

Historical analyses are uniquely situated to explore these long-term impacts. Assessing the "developmental origins" hypothesis with randomized trials would prove difficult. The cost of tracking infants into adulthood and the ethical concerns about withholding treatment for an extended period time would both prove prohibitive. An historical accident provides the necessary variation in exposure during childhood for the current analysis.

The groundbreaking observational studies on "fetal origins" in the economics literature relied on negative historical events, such as famines and pandemics, with severe health impacts and clearly defined critical periods (Meng and Qian 2009, Schulz 2010, Almond 2006). While famines and pandemics are useful for identification, the tradeoff is in policy relevance. My work on nutritional deficiencies extends the literature by examining deprivations of a milder exposure that are informative in forming policy on the margins relevant today. Additionally, identification in the current analysis comes from the positive impact of a public health intervention that was actually implemented.

The literature on "developmental origins" linked to nutritional deprivations focuses almost solely on protein and calorie malnutrition and the use of famines as an identification strategy (Schulz 2010, Meng and Qian 2009). The exclusion of micronutrient deficiencies from these studies has more to do with the difficulty in isolating a single deprivation rather than a disinterest in the phenomena. A multitude of deprivations occur during a famine, making it impossible to disentangle the independent effects of deficiencies in macro- and micronutrients. Estimates from famine studies

provide the total impact of *all* deficiencies caused by the famine. The introduction of iron-enriched bread provides a useful quasi-experiment to isolate the causal role of micronutrient deficiency on later adult health and labor market outcomes. Iron deficiency was quite prevalent in the U.S. population during the first half of the 20th century, while calories and protein were consumed in relative abundance.

Empirical Strategy

Ecological cohort analyses that link observations of adult outcomes to location of birth or childhood require us to take a stand on which location's ecological measure to use. Ideally, location information for the entire childhood could be used in combination with a known function to aggregate exposures across ages. The ideal data is rarely available. As a solution, long-term impact analyses typically use either adult state of residence or state of birth. 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 assigning exposure to fortification.

For this analysis, we are interested is the interaction of two variables that taken together measure the potential for gains in health from the fortification program. Variation in 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 for the analysis of economic outcomes. Most children will have finished their educational choices by this age, and the biomedical literature finds evidence of delays in non-cognitive skill development for adolescents suffering from iron

deficiency (Beard and Connor 2003). The mandate came into force in 1943, thus adults born in 1924 and before received no childhood exposure, with EXP_{ik} increasing linearly until equal to 19 for cohorts born in 1943 and after. In the analysis of health outcomes I reduce potential exposure time to the first-five years of life.[38] The time path of exposure follows a similar explanation as described above.

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*. Additional control variables are included and discussed in more depth in the results sections of each analysis.

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. When a larger value of the dependent variable indicates a beneficial change, β is expected to be negative if the hypothesis is correct; β is expected to be positive in specifications for dependent variables that indicate poverty or the presence of a chronic health condition. The identifying assumption is that no unobserved shocks to the outcome were correlated with iron consumption and cohort exposure.

[38] While the true potential exposure time is unclear, 19 years seems appropriate in the analysis of economic outcomes. The biomedical literature finds that adolescent iron deficiency contributes to attention and behavioral problems that could lead to income differences in the future. The literature on "developmental origins" for health outcomes suggests that critical periods for most deprivations are concentrated in the first-five years of life. A shortened potential exposure index seems appropriate here.

Income and Human Capital Accumulation

Table 13 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.[39] 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.

[39] Results from estimations using 1980 microdata are similar to those in Table 13, suggesting that the long-term effects are persistent over the life-cycle.

Table 13: 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)
Controls for mean reversion	No	Yes	No	Yes	No	Yes
Dependent Variable	Log total income, 1969		Years of schooling		Poverty status (poverty =1)	
Years of exposure X average iron consumption	-3.6	-2.0	-0.027	0.01	0.48	0.22
	(1.6)	(1.3)	(0.06)	(0.06)	(0.27)	(0.21)
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. Each entry is the coefficient 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. Demographic controls include indicators for interactions of state of birth with nonwhite and female, and nonwhite x female. Indicators for each age x nonwhite x female cell are included as well. 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 mean total income with cohort. The full sample consists of all males and females of all races between the ages of 22 and 60 and born in one of the states with iron consumption information.

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.

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.[40] At the same time, parental income gains were correlated with iron consumption over this period as argued in chapter III. 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.

Differences across demographic groups in the impact of exposure to the enriched bread program appear in Table 13. As in chapter III, 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.

Table 14 reports estimated coefficients from specifications with a variety of other economic outcomes as the dependent variable. In general, I find no evidence of a long-term impact of fortification on income and socioeconomic indices, probability of welfare, labor force participation, employment status, or probability of finishing high school or college. Results are consistent with fortification causing small, marginally significant gains in labor supply on the intensive margin.

[40] From author's calculations using the 1940 census microdata provided by IPUMS.

Table 14: Long-term impact on other economic outcomes

	(1) Occupational Income Score	(2) Duncan Socioeconomic Index	(3) Positive welfare income (=1)	(4) Labor force participation (=1)	(5) Employed (=1)
ß	-0.1	0.12	0.01	(0.010)	-0.08
s.e.	(0.11)	(0.55)	(0.16)	(0.008)	(0.18)
Units	$100		p.p.	p.p.	p.p.
N	592,055	592,055	592,055	592,055	418,943

	(6) Weeks worked (intervalled)	(7) Hours worked (intervalled)	(8) High school graduate	(9) College graduate
ß	-0.069*	-0.0541*	-0.018	0.011
s.e.	(0.036)	(0.032)	(0.015)	(0.007)
Units	Interval	Interval	p.p.	p.p.
N	592,055	592,055	592,055	592,055

Notes: Standard errors clustered on state of birth are reported in parentheses. Each entry is the coefficient 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. Coefficients from weeks and hours worked equations come from an ordered probit specification. Demographic controls include indicators for interactions of state of birth with nonwhite and female, and nonwhite x female. Indicators for each age x nonwhite x female cell are included as well. State-average iron consumption is matched to individuals based on their state of birth. The full sample consists of all males and females of all races between the ages of 22 and 60 and born in one of the states with iron consumption information.
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.

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.[41] I take these values as the full reduced-form estimate of the

[41] Because of the evidence that mean reversion is present, I use point estimates from column (2) of Table 13.

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.

The "fetal origins hypothesis" has received much attention and has generally found large estimated impacts when using clearly delineated treatment and control groups. 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 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. 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) found 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. Quantity of schooling was already optimized prior to the health improvement. The major gains to income, therefore, accrue from the inframarginal increases in quality, not the marginal increase in quantity of schooling (Bleakley 2010).

Developmental Origins of Non-communicable Chronic Health Conditions

In this section, I explore the "developmental origins" effect of iron deficiency using adult health information from the Health and Retirement Study (HRS). This research is exploratory in nature, as the medical literature has been unable to isolate the effects of childhood nutritional deficiencies on late-life health. A hypothesis for a causal medical mechanism has yet to be developed.

Moreover, previous findings related to "fetal origins" of disease are somewhat non-intuitive. Calorie and protein deprivation while *in utero* seems to *increase* the likelihood of obesity during adulthood (Shulz 2010). One theory is that nutritional want causes a fetus to develop a "thrifty phenotype" metabolism to prepare for a lifetime of want (Barker). Weight gain then occurs when an adult with the "thrifty phenotype" finds themselves in an environment of calorie abundance. The point is that neither the medical literature, nor intuition, provide guidance on what adult impairments to attempt to link to childhood iron deficiency. This chapter follows a similar path to earlier "fetal origins" studies by examining all health indicators available.

Health and Retirement Study Data: The analysis of long-term health impacts from iron deficiency requires a dataset that measures health late in life. Chronic health conditions during early- to mid-adulthood are low probability events that require large datasets to identify the treatment effect of fortification. The datas*et al*so needs to include a range of cohorts with variation in exposure to the iron fortification program. The Health and Retirement Study meets these requirements.[42] The HRS is an ongoing longitudinal survey of adults age 50 and above conducted every two years starting in 1992. The original sample from 1992 included the cohorts born between 1931-1941 (HRS Cohort). In 1998, the HRS added respondents born between 1924-1930, and 1942-1947 (Children of the Depression and War Baby Cohorts). Finally, cohorts born between 1948 and 1953 (Early Boomers Cohort) were added to the sample in 2004. Geographic identifiers are removed from the public use HRS files. Because I link respondents to iron consumption in their state of birth, I use the restricted "Cross-Wave Geographic Information (State) [1992-2008]" data file.

Table 15 reports basic descriptive statistics. I use the full 2008 sample, as opposed to 2006, because health conditions become more prevalent as the cohorts age. The sample includes all observations born between 1924 and 1953 alive at the time of interview in 2008, and born in a state for which I have iron consumption data from the "Study of Consumer Purchases in the United States, 1935-1936". The 2006 interview included biological measures of health taken from a blood sample and physical

[42] The HRS is sponsored by the National Institute of Aging (Grant NIA U01AG009740) and is conducted by the University of Michigan. NHANES is also a potential dataset that contains the appropriate cohorts, measures health late in life, and contains state-of-birth information. While the total sample sizes are larger in NHANES, the HRS has larger samples for the cohorts used in the analysis.

examination. In addition to the main analysis, I estimate a series of duration models for which I use the entire panel of data from 1992 to 2008.

Health Measures in the HRS: The respondent is asked in each round whether they have ever been diagnosed by a physician for hypertension, diabetes, a psychiatric condition, chronic lung disease, heart disease, stroke, or cancer. The second set of health outcomes I examine are the presence of disabilities, whether the respondent has difficulty with any activities of daily living (ADLs), instrumental ADLs, and whether their health limits their ability to work. ADLs include the following activities: walking across a room, dressing, bathing, eating, getting in/out of bed, and using the toilet. Instrumental ADLs include: using a map, using a phone, managing money, taking medication, shopping for groceries, and preparing meals. The last set of health conditions I examine include BMI, height, and a self-reported health index. Research in the nutrition literature finds that early-life nutritional deprivations can "program" the body to be much more likely to be obese when later childhood and adult conditions are such that food is plentiful. Thus, early life iron deficiency may lead to an increased BMI.

From the 2006 interview-year, I use the biological health measures taken from a blood sample and physical examination. Outcomes include a continuous measure of HDL cholesterol, a binary indicator of low HDL cholesterol, and the Framingham Risk Score used to measure 10-year cardiovascular risk. The score is based on the respondent's age, sex, blood pressure, and cholesterol, and whether the respondent smokes tobacco, and suffers from diabetes. Low risk is considered having a 10 percent or less change of developing chronic heart disease in the next 10 years, intermediate risk 10-20 percent, and high risk 20 percent or more (NCEP 2002).

Table 15: Health and Retirement Study summary statistics

	Mean	Std. Dev.	Range	N
Panel A: Outcomes				
Health conditions				
Health limits work	0.30	0.46	{0,1}	9220
Any adla	0.31	0.90	{0,1}	10067
Any iadla	0.13	0.49	{0,1}	10066
Any psychiatric condition	0.22	0.54	{0,1}	9573
High blood pressure	0.59	0.49	{0,1}	10055
Cancer	0.16	0.36	{0,1}	10046
Diabetes	0.20	0.40	{0,1}	10059
Chronic lung disease	0.11	0.31	{0,1}	10058
Heart disease	0.25	0.44	{0,1}	10056
Stroke	0.087	0.28	{0,1}	10062
BMI	28.3	5.92	[10.9,74.4]	9952
Height	1.69	0.10	[1.22,2.32]	10071
Memory	0.046	0.21	{0,1}	10067
Self-reported health	2.82	1.09	[1,5]	10066
Framingham risk score	20	9.8	[1,30]	4289
kHDL	57	14	[21,133]	4289
Low HDL	0.08	0.27	[0,1]	4289
Health care utilization (past 2 years)				
Any hospital stay	0.27	0.44	{0,1}	10048
# hospital stays	0.48	1.30	[0,50]	10034
Nursing home stay	0.040	0.19	{0,1}	10058
# of spells spent in nursing home	0.053	0.64	[0,56]	10044
# nights spent in nursing care	8.48	71.9	[0,1614]	10039
Any homecare use	0.077	0.27	{0,1}	10007
Panel B: Controls				
Female	0.55	0.50	{0,1}	15232
Black	0.17	0.38	{0,1}	15232
Years of education	12.7	2.79	[0,17]	15181
High school drop out	0.19	0.39	{0,1}	15229
GED	0.043	0.20	{0,1}	15229
High school	0.33	0.47	{0,1}	15229
Some college	0.23	0.42	{0,1}	15229
College	0.21	0.41	{0,1}	15229
Alcoholic drinks per month	2.64	6.67	[0,105]	10049
Household net assets ($1,000)	555	2297	[-2m,9m]	10071
Household income ($1,000)	75	607	[0,6000]	10071
Ever smoked tobacco	0.59	0.49	{0,1}	10022
Smokes now	0.14	0.34	{0,1}	10021
Vigorous phys. act. (5 pt. scale)	4.00	1.34	[0,5]	10063

Source: Individual interview responses from the 2008 interview wave of the Health and Retirement Study. The Framingham risk score is calculated from the 2006 HRS physical examination data, which also provides kHDL and low HDL cholesterol.

These are self-reported measures of conditions, which may lead to reporting bias. Some individuals may not be aware they are suffering from one of the health conditions but the bias should not vary greatly with year and state of birth. Selective mortality becomes a concern if early-life childhood iron deficiency reduces the ability of the body to withstand chronic health conditions. The remaining sample of the low-exposure low-consumption cohort would be left with relatively healthy respondents who are less likely to suffer from chronic disease. However, this selective mortality should bias the results against finding any meaningful relationship between early childhood exposure to iron fortification and late-life health.

Health care utilization measures in the HRS: Directly examining utilization provides a test for the impact on health care spending. The HRS does not contain expenditure data, so utilization is an alternative to monetize the benefits from improved late-life health. The utilization measures include the number of hospital stays in the past 2 years, any hospital stay in the past 2 years, any nursing home stays in the past 2 years, the number of times/spells in nursing care, the number of nights spent in nursing care, and whether any homecare was used in the past 2 years.

Main Health Results: The estimation equation (10) is reprinted below for convenience.

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

The critical period now ends at age 5, instead of age 19, as this age seems to be the appropriate cutoff for the critical period of iron status on health. Adults born in 1937 and before received no exposure to the fortification campaign, with EXP_{ik} increasing linearly

until equal to 5 for cohorts born in 1943 and after. The dependent variables are health or health care outcomes and denoted by Y_{isk}.

Health Outcomes: Table 16 presents point estimates for β from equation (10). Each entry is from a separate regression and can be interpreted as the change in the dependent variable associated with a full 5 years of exposure to the fortification program at a two milligram difference in iron consumption – the average increase in iron intake caused by the program. Each row contains results for a dependent variable and each column a specification. Because areas with low-iron consumption experience larger gains in health, the coefficient is expected to be positive in specifications where the dependent variable is a binary indicator for a chronic health condition. The coefficient is expected to be negative for dependent variables where higher values imply better health. Standard errors are clustered by state of birth by estimating the asymptotic variance matrix (Liang and Zeger 1986).

I find that early-life iron deficiency tends to increase the likelihood of a number of adult health conditions. Column (1) reports estimates from the base specification of equation (10). The results are consistent with the fortification program reducing the incidence of diabetes (2.4 p.p.), high blood pressure (3.1 p.p.), and disabilities (Work limiting: 4 p.p.; ADL: 1.9 p.p.; IADL: 1.8 p.p.). I cannot reject the null hypothesis that iron deficiency during the first 5 years of life has no impact on chronic heart disease, memory loss, psychiatric conditions, or chronic lung disease.

Estimates from regressions using biological health measures present mixed evidence for the "developmental origins" hypothesis. I find no evidence of an impact on height, BMI, or the Framingham Risk Score. However, results are consistent with

fortification improving the HDL cholesterol levels of the population. Higher levels of HDL cholesterol are associated with lower risk of cardiovascular events, ostensibly through the removal of plaque from the arteries. It is surprising that I find an effect for high blood pressure and cholesterol, but no impact on the Framingham Risk Score even though it uses the two as components.

Table 16: Early-life iron deficiency increases incidence of chronic health conditions

	(1) Base	(2) Own Education	(3) Own Wealth	(4) Own Income	(5) Mean Reversion	(6) Full specification	(7) N
Self-reported health	0.0021	0.0013	0.0017	0.0016	0.0020	0.0019	10227
	(0.0025)	(0.0025)	(0.0026)	(0.0029)	(0.0025)	(0.0027)	
Chronic health conditions							
Diabetes	2.37**	2.07*	2.13*	2.06*	2.52**	2.38**	9673
	(1.11)	(1.18)	(1.17)	(1.17)	(1.06)	(1.01)	
Chronic heart disease	0.40	0.04	0.13	0.14	-0.11	-0.19	9671
	(0.84)	(0.81)	(0.83)	(0.86)	(0.69)	(0.70)	
Memory	0.92	0.93	0.96	1.03	1.11	1.16	9682
	(1.19)	(1.14)	(1.14)	(1.17)	(1.11)	(1.15)	
High blood pressure	3.09***	3.2***	3.17***	3.22***	3.13***	3.12***	9669
	(0.79)	(0.78)	(0.84)	(0.81)	(0.87)	(0.95)	
Psychiatric condition	-0.99	-1.29	-1.18	-1.09	-1.14	-1.11	9728
	(1.54)	(1.52)	(1.60)	(1.61)	(1.62)	(1.69)	
Stroke	1.32*	1.58**	1.61**	1.56**	1.09	1.00	9676
	(0.65)	(0.64)	(0.62)	(0.64)	(0.66)	(0.67)	
Chronic lung disease	-2.65	-2.2	-2.86	-3.05	0.47	0.03	9576
	(5.05)	(5.46)	(5.61)	(5.69)	(6.26)	(6.2)	
Disability							
Health limits work	4.03**	3.77**	3.83**	3.97**	3.88**	3.89**	8980
	(1.72)	(1.71)	(1.72)	(1.60)	(1.66)	(1.49)	
Any adl	1.93*	1.98**	2.04**	2.1**	1.65*	1.64*	9686
	(1.11)	(0.96)	(0.95)	(0.91)	(0.87)	(0.81)	
Any iadl	1.8**	2.04**	2.14***	2.31***	1.43**	1.55**	9685
	(0.74)	(0.75)	(0.71)	(0.75)	(0.56)	(0.59)	

Table 16 (continued) : Early-life iron deficiency increases …

	(1) Base	(2) Own Education	(3) Own Wealth	(4) Own Income	(5) Mean Reversion	(6) Full specification	(7) N
Biological measures							
Framingham risk score	-0.035	-0.039	-0.038	-0.043	-0.042	-0.045	4474
	(0.031)	(0.030)	(0.029)	(0.032)	(0.035)	(0.036)	
kHDL	-0.10**	-0.08*	-0.09*	-0.08*	-0.11**	-0.11**	2902
	(0.0458)	(0.0442)	(0.0434)	(0.0465)	(0.0437)	(0.0458)	
Low HDL	2.18**	1.84**	1.87**	1.74**	2.44**	2.27**	3777
	(0.96)	(0.85)	(0.88)	(0.84)	(0.94)	(0.89)	
Height	0.00017	9.66E-05	0.000149	0.000155	0.000887	0.000933	
	-0.00129	-0.00136	-0.00138	-0.0014	-0.00131	-0.00132	
BMI	0.0144	0.0668	0.0628	0.0675	-0.007	-0.0134	
	-0.108	-0.116	-0.12	-0.119	-0.117	-0.119	

Notes: Standard errors clustered on state of birth are reported in parentheses. Each entry is the coefficient from a separate regression of equation (10). Reported point estimates are from a full 5 years of exposure at a one standard-deviation difference in iron consumption. State-average iron consumption is matched to individuals based on their state of birth. Base results in column (1) include indicators for interactions of state of birth with nonwhite and female, year of birth with nonwhite and female, and nonwhite x female, the main effects, veteran status, the log of the infant mortality rate in the state and year of birth of the respondent, and parental education interacted with year of birth. Base results also include measures of health behavior for ever smoking tobacco, # of alcoholic drinks consumed per week, and amount of vigorous activity. Column (2) adds in indicators for educational attainment category (<HS, GED, HS, SC, C). Columns (3) and (4) add net wealth and log income measured in 2008 (2006 for the biological measure). Mean reversion is controlled for by the interaction of 1943 log state mean total income with year of birth. The full sample consists of all males and females of all races and born in one of the states with iron consumption information. Applying Bonferroni's adjustment because of 22 multiple comparisons requires a corrected significance level of $\alpha = 0.00227$ to recover an overall significance level of $\alpha = 0.05$ for the combined tests. Under this more conservative test, only the null hypothesis for high blood pressure is rejected.
Sources: All individual level data for the chronic health condition regressions comes from the 2008 interview wave of the Health and Retirement Study. Biological measure regressions use individual level data from the 2006 HRS interview wave. Average iron consumption by state is calculated by the author from the "Study of Consumer Purchases, 1935-1936." Log state average income in 1943 is published by the BEA. Infant mortality rates are from the published volumes of the *Vital Statistics of the United States.*

Causal Pathways: Does iron deficiency have its own direct biological impact on adult chronic disease or do other channels intermediate the relationship? The ensuing three columns of Table 16 explore the causal mechanisms.

A large literature in economics documents the positive correlation between health and own education, income and wealth, and parental income (Currie 2009). In previous chapters I show that the introduction of enriched bread leads to increases in school enrollment and parental income during the 1940s and to increased adult incomes for the children exposed to the program. The observed health improvements associated with fortification may not be directly caused by the program, but only indirectly through improvements in income and education. I include – separately and combined – measures of own income, wealth, and education. Alternatively, the time-path of these variables might be correlated with the variable of interest. The exercise functions as a robustness check to ensure the result is not entirely driven by income and education gains unrelated to the campaign.

The results in columns (2) through (3) are not substantively different from the base specification, suggesting that health gains from fortification were not working through increased income, wealth, or education. The test conducted above is not ideal. The respondent's reported income in the year prior to the interview is the only suitable variable available in the HRS. However, much of the sample is retired and living on investment and pension income. A measure of full lifetime income would be preferred. However, I do include controls for educational attainment and health behaviors, which potentially absorb a portion of the correlation between lifetime income and health.

Finally, I examine whether fortification improved later-life health by increasing parental income during the respondent's childhood. The HRS does not inquire about parental income, and the publicly available HRS data files do not contain parental occupation. The files do contain parent's education, which can account for some of the variation in income while in childhood. I include parent's education interacted with year of birth as a control in all specifications.

As a second best solution to account for fortification induced changes in income, and as a control for mean reversion, I include state average personal income for 1943 interacted with year of birth. This allows for a flexible time-path of health to shift with state income. Gains in parental income between 1924 and 1953 that are correlated with state average iron consumption are removed from the estimate of β. The addition of the mean reversion controls in column (5) attenuate the estimates somewhat, but we can draw the same conclusions as before. Column (6) reports estimates from the full specification. While gains in parental income from fortification, or alternatively reversion to the mean such as from the depths of the Great Depression, might explain a portion of the observed correlation between childhood iron deficiency and late-life health, there seems to be an important role for a direct effect of iron deficiency.

Health care utilization: Table 17 summarizes results from a series of estimations of equation (10) with dependent variables that measure health care utilization. The results provide unclear evidence for an effect of childhood iron deficiency on later-life health care utilization. Iron deficiency seems to have a long-term effect on the utilization of home care and nursing home services, but not on hospital services. Again, the effect of iron deficiency does not seem to be fully accounted for by income or wealth effects, or

by mean reversion and parental income. Although, the results are less clear than for health conditions.

Table 17: Health care utilization is weakly associated with childhood iron deficiency

	(1) Base	(2) Own Education	(3) Own Wealth	(4) Own Income	(5) Mean Reversion	(6) Full specification	(7) N
Any hospital stay	0.15	0.14	0.15	0.16	0.13	0.13	9665
	(0.12)	(0.11)	(0.11)	(0.12)	(0.11)	(0.12)	
# nights in hospital	0.38	0.39	0.41	0.43	0.32	0.31	9665
	(0.47)	(0.46)	(0.47)	(0.48)	(0.48)	(0.47)	
Any home care	0.14**	0.14**	0.14**	0.14**	0.16**	0.14**	9631
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	
Any nursing home stay	0.076	0.097*	0.098*	0.098*	0.085*	0.080	9674
	(0.052)	(0.050)	(0.050)	(0.054)	(0.048)	(0.053)	
Spells spent in nursing home	0.027	0.033	0.034	0.034	0.030*	0.030*	9662
	(0.018)	(0.020)	(0.020)	(0.021)	(0.017)	(0.017)	
# nights in nursing home	3.1**	3.7**	3.7**	3.5**	2.8**	2.4*	9658
	(1.4)	(1.5)	(1.5)	(1.6)	(1.2)	(1.4)	

Notes: Standard errors clustered on state of birth are reported in parentheses. Each entry is the coefficient from a separate regression of equation (10). Reported point estimates are from a full 5 years of exposure at a one standard-deviation difference in iron consumption. State-average iron consumption is matched to individuals based on their state of birth. Base results in column (1) include indicators for interactions of state of birth with nonwhite and female, year of birth with nonwhite and female, and nonwhite x female, the main effects, veteran status, the log of the infant mortality rate in the state and year of birth of the respondent, health behaviors, and parental education interacted with year of birth. Column (2) adds in indicators for educational attainment category (<HS, GED, HS, SC, C). Columns (3) and (4) add net wealth and log income measured in 2008 (2006 for the biological measure). Column (5) controls for mean by the interaction of 1943 log state mean total income with year of birth. The full sample consists of all males and females of all races and born in one of the states with iron consumption information.
Sources: All individual level data comes from the 2008 interview wave of the Health and Retirement Study. Average iron consumption by state is calculated by the author from the "Study of Consumer Purchases, 1935-1936." The log of state average income in 1943 is published by the Bureau of Economic Analysis. Infant mortality rates are from the published volumes of the *Vital Statistics of the United States.*

Selective mortality should be working against finding an impact on health care utilization. However, the bias might be larger than for health conditions. Home care and nursing home stays are often precipitated by an acute health event. If childhood iron deficiency makes a person less likely to survive an acute event brought on by the chronic ailment, then they might be less likely to use services like home care or nursing home care. The dead are in no need of care. The healthier sufferers – i.e. those that did not suffer from iron deficiency during childhood – would be more likely to use these types of services.

Impact of Childhood Iron Deficiency on Adult Mortality

In this section I directly assess the impact of the fortification program on mortality. Iron deficiency during childhood seems to increase the incidence of chronic disease and adult income. We might expect lower death rates and larger increases in life expectancy for adults born in states with low-iron consumption and with more exposure to the fortification program. The impact could be working through either a direct effect on health or by raising incomes, which is negatively correlated with mortality (Deaton 2002). I examine fortification's impact on mortality in a number of ways. First, I conduct a regression analysis of mortality rates between 1979 and 2004 using the aggregate vital statistics data supplied by the Centers for Disease Control. Secondly, I estimate a duration model using the individual level HRS data.

Mortality Regressions Using Death Registration Data: I use the yearly death count data from the Multiple Cause-of-Death Mortality Data from the National Vital Statistics System of the National Center for Health Statistics. The data is derived from

the underlying census of death certificates each year. For the years 1979 to 2004, state of birth was included in the public-use data sets. The number of deaths is summed for each *birth cohort x state of birth x age* cell. The population denominator in each cell is calculated using the IPUMS decennial census data. Intercensal population estimates are created by the average of two series: 1.) working forward from census years by subtracting deaths, and 2.) working backward from census years by adding deaths. I estimate a series of linear regressions of the following general form:

$$(11) \quad \ln(M_{ask}) = \alpha + \beta \cdot (IRON_s \times EXP_k) + \delta_a + \delta_s + \delta_k + \varepsilon_{ask}$$

The dependent variable is the log mortality rate for a *age x state of birth x birth cohort* cell. A series of binary indicators for age, state of birth, and birth cohort are included to absorb unobserved group effects. The variable of interest is average iron consumption in 1936 in state *s* interacted with the birth cohort exposure to the fortification program. Standard errors are robust to correlated errors within each state of birth by year of birth group. The concern is of serial correlation of the errors as we follow the time-path of age-specific mortality rates within a birth cohort from state *s*.

Table 18 reports point estimates of ß from separate specifications that vary the years of potential exposure. I find no evidence of an effect of fortification on adult mortality using this procedure. The dataset only includes 26 years of observations, limiting the overlap of observed ages across birth cohorts. The age of first observation in the dataset ranges from 29 years old for the cohort born in 1950 to 59 for the cohort born in 1920. For the core group of cohorts born around 1943, the age-specific mortality rates included here represent a period of relatively low mortality risk. In any case, I find no

evidence of selective mortality using the death registration data, including ages between 40 and 50.

Table 18: Panel regressions using death registration data

	Years of potential exposure			
	19 years	10 years	5 years	2 years
($IRON_s$ x EXP_k)	0.00228	0.00261	0.00450	0.0104
	(0.00159)	(0.00183)	(0.00326)	(0.00748)
Observations	15,280	15,280	15,280	15,280
R-squared	0.943	0.943	0.943	0.943

Notes: Each entry is from a separate regression of the log mortality rate in state of birth x year of birth x age cells. Point estimates can be interpreted as the percent difference in mortality rate for each year of exposure to the fortification program at a one-milligram difference in state average iron consumption. Standard errors are clustered by state of birth interacted with year of birth.
Sources: Annual deaths in each birth x state cohort between1979 and 2004 are calculated by the author using individual data from the Multiple Cause-of-Death Mortality Data from the National Vital Statistics System of the National Center for Health Statistics. Population denominators calculated using IPUMS decennial census data. Intercensal estimates are the average of two procedures: 1.) working backward from census years adding in deaths, and 2.) moving forward from census years subtracting deaths.

Duration Model of Mortality: We can think of length of life in terms of a survival model, which is used to model time to an event. In my case, I model the time until death measured in years since birth.[43] In the standard terminology, the event is called a failure and the time period under consideration is called "time at risk." Thus, a person will be at-risk starting at birth until failure at the time of death.

[43] The impact of childhood iron deficiency on the time to acquiring one of the chronic diseases is also of interest. However, the sampling structure of the HRS causes serious censoring issues, which combined with a competing risk duration model make estimation infeasible. Many respondents enter the sample already suffering from the chronic condition. Left- and right-censoring impede estimation as well.

Estimating survival time condenses to estimating the hazard function, which gives the instantaneous rate of failure at a given time period t. I use the Cox proportional hazards model because it allows for a flexible estimation that does not depend on the underlying base hazard rate. The Cox model imposes the feature that any variables used in the estimation shift the base hazard rate proportionally. The probability of failure of respondent *i* in time period t given survival to the period is:

$$(12) \qquad h(t,h_0 \mid \mathbf{x}_i) = \lim_{\Delta t \to 0} \frac{\Pr(t + \Delta t > T > t \mid T > t)}{\Delta t} = h_0(t)\exp(\mathbf{x}_i \boldsymbol{\beta}_x)$$

where T is the failure date, $h_0(t)$ is the baseline hazard rate that varies with t and is common to all observations. Explanatory variables are contained in $\mathbf{x}_j$ and enter the equation through the exponential function. The vector $\boldsymbol{\beta}_x$ provides the estimated coefficients that, once exponentiated, can be interpreted as the percent change in the instantaneous failure rate from a unit change in x. The variable of interest is, again, $(IRON_s \; x \; EXP_k)$, the interaction of state average iron consumption and the five-year exposure measure. The vector **x** contains a rich set of indicators for each *year of birth (yob), state of birth (sob), veteran status, nonwhite, sex, education category, yob x sex, sob x sex, sob x nonwhite* cell, and the log infant mortality rate in the state and year of birth of the respondent. Standard errors are clustered by state of birth.

The inclusion of year of birth and state of birth indicators implies that identification of the β of interest comes from within birth cohort variation in state of birth average iron consumption, and from across cohort variation in the number of the first five years of life spent exposed to the fortification program.

Table 19 reports coefficients from the semiparametric Cox proportional hazard model from estimation of equation (12). The elements of the table give the proportional

shift of the hazard function associated with a one-unit change in the explanatory variable measured as a percentage.[44] Results suggest that individuals born in states with low levels of iron consumption and with more time spent exposed to the fortification program experienced larger gains in life expectancy. The probability of death in a given period is 1.3 percent lower for each additional year of exposure to the program during the first-five years of life, and for each additional milligram of iron consumption difference.

The impact becomes even more substantial as the impact accumulates over the entire childhood. An adult that experienced a full five years of exposure at the two milligram average increase in iron consumption from the program would be 13 percent less likely do die in a given period relative to the baseline hazard rate. The other explanatory variables behave as expected. Higher educational attainment reduces mortality, whereas veteran status increased age-specific mortality rates. The total effect of gender indicates that males have a higher probability of failure than females.[45]

Parametric survival analysis techniques require the researcher to take a stand on the baseline hazard. Is it constant, monotonically increasing, or monotonically decreasing? When the researcher cares only about the effect of explanatory variables on the hazard function, not the baseline function itself, then the semiparametric Cox proportional hazard model becomes the technique of choice. It allows for a fully flexible estimation of the baseline hazard. The cost is in decreased efficiency.

[44] Coefficients have been exponentiated and multiplied by 100.
[45] Results are unreported because of the interactions of gender with state of birth, year of birth, and nonwhite.

Table 19: Fortification decreased mortality hazard rate

	(1) Cox model	(2) Weibull model
$(IRON_s \times EXP_k)$	-1.3**	-1.2*
	(0.6)	(0.7)
Log infant mortality rate	4.3	-14.8
	(23.0)	(27.5)
GED	-34.8***	-35.1***
	(9.1)	(9.0)
High school	-36.2***	-36.5***
	(4.3)	(4.4)
Some college	-38.7***	-39.1***
	(4.7)	(4.7)
College or more	-51.0***	-51.3***
	(6.0)	(6.1)
Veteran	14.2**	14.8***
	(5.2)	(5.2)
Constant		-7.8***
		(1.3)
ln_p		0.86***
		(0.03)
Number of subjects	14618	
No. of failures	3035	
Mean time to failure	80 years	

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on state of birth are reported in parentheses. Reported point estimates give the percent proportional shift of the hazard rate from one additional year of exposure at a one milligram difference in iron consumption. State-average iron consumption is matched to individuals based on their state of birth. Indicators are included for interactions of state of birth with nonwhite and female, year of birth with nonwhite and female, and nonwhite x female, the main effects, and veteran status. The log infant mortality rate in the state and year of birth of the respondent is included as a control variable. Educational attainment is indicated as one of five categories: GED, HS, SC, C, and less than HS omitted. Column (1) reports results from the semiparametric Cox proportional hazard model. The Cox model does not identify a constant. Column (2) reports results from the parametric duration model using a Weibull distribution. The constant provides the estimated scale parameter, and the estimate of ln p implies that the hazard function is monotonically increasing in age. The full sample consists of all males and females of all races and born in one of the states with iron consumption information.
Sources: All individual level data for the chronic health condition regressions comes from the full panel of the Health and Retirement Study form 1992 to 2008. Average iron consumption by state is calculated by the author from the "Study of Consumer Purchases, 1935-1936." Infant mortality rates are from the published volumes of the *Vital Statistics of the United States.*

Interval censoring enters the analysis because of the sample structure of the HRS. The main sample I use consists of respondents born between 1924 and 1953. The original HRS cohort included only individuals born between 1931 and 1941. The other years were not added until the 1998 and 2004 waves. The age at which respondents enter the period of observation varies with birth cohort. Ideally, I would want to begin observation of each cohort at age 50. But for the oldest cohorts – those born between 1924 and 1930 – entry age varies from 68 to 74. Censoring is correlated with the time exposed to the fortification program. The problem condenses to one of selective mortality during the ages where observations are censored.

Results from a parametric Weibull estimation are reported in column (2) of Table 19. The implications are similar to those of the Cox model. Moreover, the estimated parameter ln p implies that the baseline hazard is monotonically increasing in age.

Figure 11 and Figure 12 provide graphical evidence of the impact of exposure to the fortification program during the first-five years of life on the hazard rate of death and the probability of survival to a given age conditional on living to age 50. The hazard rate plots how the instantaneous probability of death evolves with age along the x-axis. The baseline hazard is estimated from equation (12) at the mean value of explanatory variables with exposure equal to zero. The full exposure hazard plots the proportional shift in the probability of death from a full 5-years of exposure at two milligram difference in iron consumption. The following figure plots the probability of survival to the age on the x-axis. The base and full exposure plots are calculated as described previously.

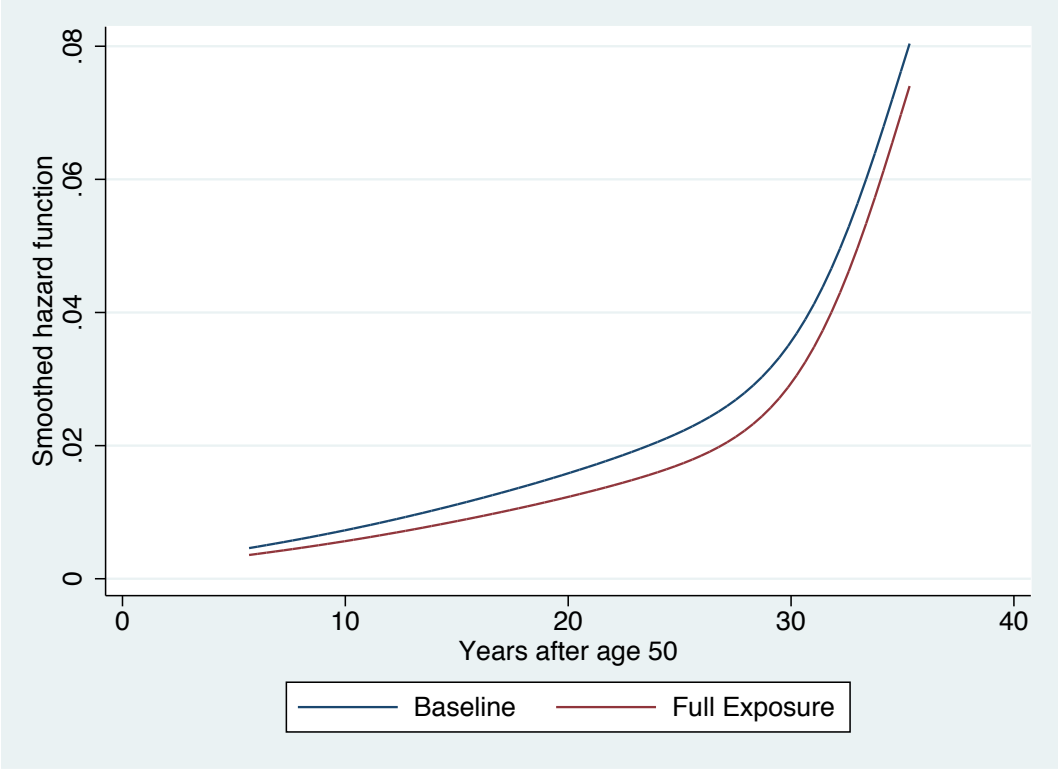

Figure 11: Impact of childhood exposure to fortification on the adult hazard of death.

Notes: Hazard rates from estimation Cox proportional hazard estimation of equation (12). Baseline hazard is estimated at the mean of explanatory variables with exposure set equal to zero. Full exposure hazard is equal to the baseline plus the effect of 5 years exposure to the fortification program at a 2 milligram difference in iron consumption. Source: See notes from Table 19.

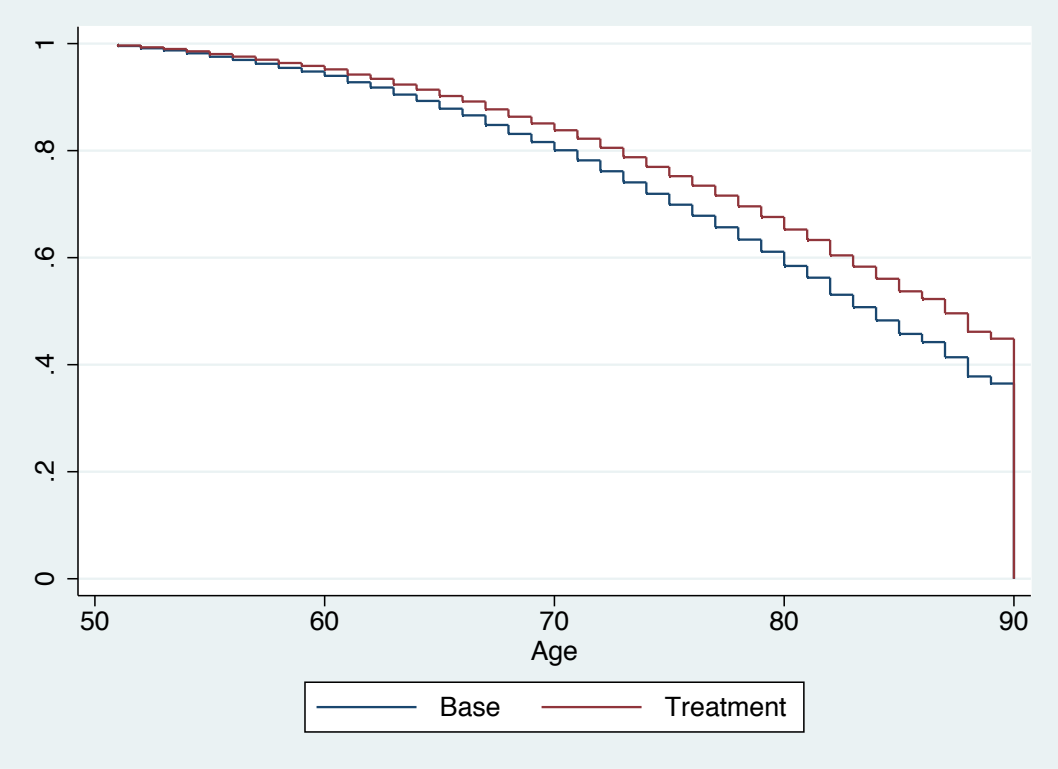

Figure 12: Probability of survival to a given age conditional on reaching age 50

Notes: Survival functions estimated using Cox proportional hazard model. Baseline survival is estimated at the mean of explanatory variables with exposure set equal to zero. Full exposure survival probability is equal to the baseline plus the effect of 5 years exposure to the fortification program at a 2 milligram difference in iron consumption.
Sources: See notes from Table 19

CHAPTER VII

VII. DISCUSSION

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 Americans' 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 in response to health improvements. Estimated benefit-cost ratios that rely solely on the productivity impacts will underestimate the true benefits of an iron fortification campaign.

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.

Similarly, the fortification program likely reduced the incidence of late-life chronic diseases in the cohorts with most to gain from the intervention. I find evidence that cohorts with more exposure to fortification were less likely to suffer as adults from diabetes, high blood pressure, disabilities, and low HDL cholesterol. In addition, these cohorts were less likely to utilize home care or nursing home care. Results from a duration model of mortality suggest that fortification reduced the hazard rate of death. I find mixed evidence of selective mortality during early and late adulthood, and thus cannot rule out this bias in the above estimates. However, selective mortality biases the coefficients away from finding a relationship.

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.

APPENDIX A

A. DATA APPENDIX

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 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 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.

*Economic 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.

Health and Retirement Study Data: The HRS is an ongoing longitudinal survey conducted every two years starting in 1992 of adults aged 50 and over. The original sample from 1992 included the cohorts born between 1931-1941 (HRS Cohort). In 1998, the HRS added the cohorts born between 1924-1930, and 1942-1947 (Children of the Depression and War Baby Cohorts). Finally, cohorts born between 1948 and 1953 (Early Boomers Cohort) were added to the sample. The restricted HRS data file contains state and year of birth.

Interview outcome variables come from the RAND HRS Data File (v.L). Biomarker and physical examination data are contained in the 2006 Biomarker Data File (v.1.0) and the 2006 HRS Core (Final) (v.2) file. Finger pricks provide blood samples for labs to measure total cholesterol, HDL cholesterol, and A1c. The respondent's blood pressure, pulse, and waist circumference were recorded during an in-home physical examination for a subset of the full HRS sample. Combined with age and gender information, the biomarkers can be used to construct the Framingham Risk Score, which is used to estimate an individual's 10-year cardiovascular risk. Individual-level demographic, income, and education controls are drawn from the RAND HRS Data File (v.L).

The sample includes all observations born between 1924 and 1953 alive at the time of interview in 2008, and born in a state for which I have iron consumption data from the "Study of Consumer Purchases in the United States, 1935-1936". The 2006

interview included biological measures of health taken from a blood sample and physical examination. In addition to the main analysis, I estimate a series of duration models for which I use the entire panel of data from 1992 to 2008.

APPENDIX B

B. 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 mark 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 13 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 13. 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 20 reports the estimated coefficient on ($IRON_s$ X $POST_t$) from regressions that control for SEA average consumption of other nutrients individually. Row (1)

reprints the baseline results from Table 5 and Table 11. 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.

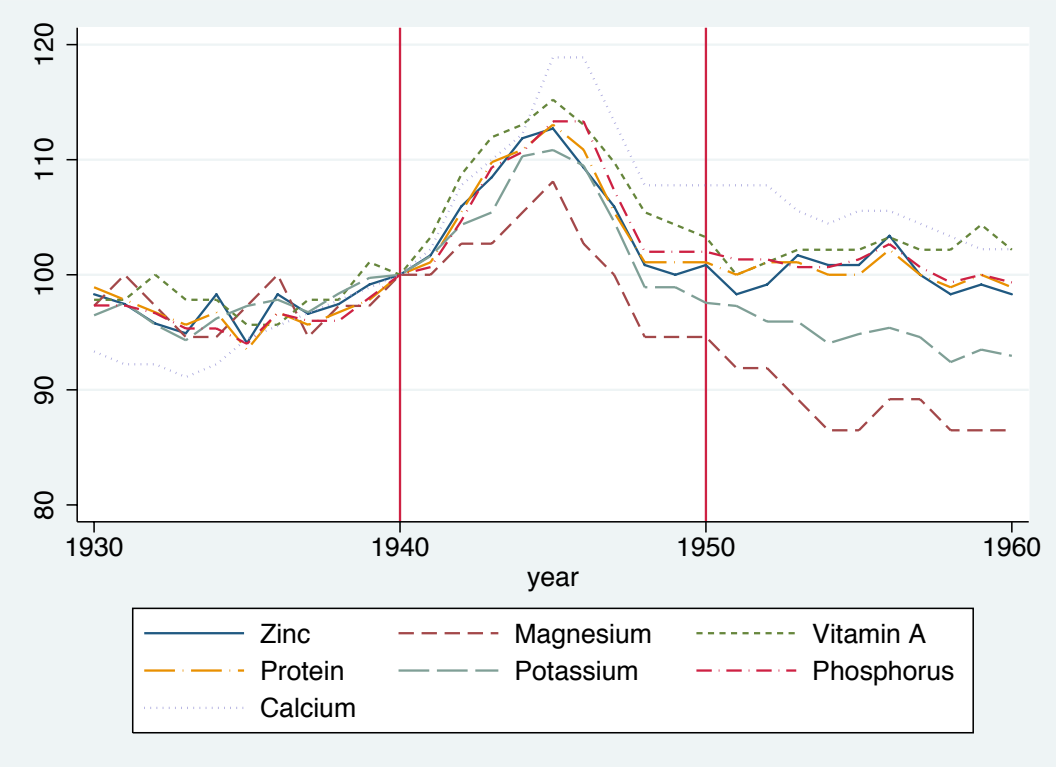

Figure 13: Per capita nutrient content of U.S. food supply (indexed to 1940)

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

Table 20: Robustness to changes in nutrients other than iron

	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 (8). Rows list the additional nutrient interacted with $POST_t$ and included as an independent variable in the regression. See documentation for Table 5 and Table 11 in the main text for further details and sources.

REFERENCES

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.

Barker, David JP, "Maternal nutrition, fetal nutrition, and disease in later life," *Nutrition*, 13 (1997), 807-813.

Barreca, Alan, "The Long-Term Economic Impact of In Utero and Postnatal Exposure to Malaria," *Journal of Human Resources,* 45(2010), 865-892.

Basta, Samir, Soekirman, Darwin Karyadi and Nevin Scrimshaw, "Iron deficiency anemia and the productivity of adult males in Indonesia," *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.

Bobrow-Strain, Aaron, *White Bread: A Social History of the Store-bought Loaf*, Beacon Press, (2012).

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.

Brotanek, Jane, Jacqueline Gosz, Michael Weitzman and Glenn Flores, "Secular trends in the prevalence of iron deficiency among us toddlers, 1976-2002," *Archives of Pediatrics & Adolescent Medicine,* 162 (2008), 374-381.

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

Carson, Jeffrey, Amy Duff, Roy Poses, Jesse Berlin, Richard Spence, Richard Trout, Helaine Noveck and Brian Strom, "Effect of anaemia and cardiovascular disease on surgical mortality and morbidity," *The Lancet*, 348 (1996), 1055-1060.

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

Centers for Disease Control and Prevention, "Iron deficiency--United States, 1999-2000," MMWR Morb Mortal Wkly Rep. 2002 Oct 11; 51(40), 897-9.

Cogswell, Mary, Laura Kettel-Khan and Usha Ramakrishnan, "Iron Supplement Use among Women in the United States: Science, Policy and Practice," *Journal of Nutrition*, 133 (2003), 1974S-1977S.

Collins, Allan, Suying Li, Wendy St. Peter, Jim Ebben, Tricia Roberts, Jennie Ma and Willard Manning, "Death, hospitalization, and economic associations among incident hemodialysis patients with hematocrit values of 36 to 39%," *Journal of the American Society of Nephrology*, 12 (2001), 2465-2473.

Committee on Cereals of the Food and Nutrition Board, *Cereal Enrichment in Perspective, 1958*, National Academy of Sciences – National Research Council: Washington, D.C. (1958).

Committee on Foods, "Annual Meeting of the Committee on Foods ," *JAMA*, 107 (1936), 39.

Costa, Dora and Richard Steckel, "Long-term Trends in Health, Welfare, and Economic Growth in the United States," in *Health and Welfare During Industrialization*. Eds.: Richard Steckel and Roderick Floud. Chicago: University of Chicago Press (1995).

Currie, Janet, "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development," *Journal of Economic Literature*, 47 (2009), 87–122.

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.

Dallman, Peter, Ray Yip and Clifford Johnson, "Prevalence and causes of anemia in the United States, 1976 to 1980," *The American Journal of Clinical Nutrition*, 39 (1984), 437-445.

Deaton, Angus, "Policy Implications Of The Gradient Of Health And Wealth," *Health Affairs,* 21 (2002), 13-30.

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.

Federal Security Agency, *Proceedings of the National Nutrition Conference for Defense*, May 26-28, 1941: U.S Government Printing Office, (1942).

Feyrer, James, Dimitra Politi and David Weil, "The economic effects of micronutrient deficiency: Evidence from salt iodization in the United State," Working Paper. (2008).

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).

Gluckman, Peter and Mark Hanson, "Living with the Past: Evolution, Development, and Patterns of Disease," *Science*, 305 (2004), 1733-1736.

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).

Heath, C.W., M.B. Strauss and W.B. Castle, "Quantitative aspects of iron deficiency in hypochromic anemia (the parenteral administration of iron)," *Journal of Clinical Investigation*, 11 (1932), 1293–1312.

Hendershot, Gerry, "Trends in Breast-Feeding," *Pediatrics*, 74 (1984), 591-602.

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.

Howson, Christopher, Abraham Horwitz and Eileen Kennedy, editor, *Prevention of Micronutrient Deficiencies: Tools for Policymakers and Public Health Workers, (Washington, DC:* National Academy Press, 1998).

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).

Jeans, P. C., M. B. Smith and Genevieve Stearns, "Dietary habits of pregnant women of low income in a rural state," *Journal of the American Dietetic Association*, 28 (1952), 27-34.

Koehn, C.J. and C.A. Elvehjem, "Further Studies on the Concentration of the Antipellagra Factor," *The Journal of Biological Chemistry*, 118 (1937), 693-699.

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).

Kushang, Patel, Luigi Ferruci, William Ershler, Dan Longo and Jack Guralnik, "Red Blood Cell Distribution Width and the Risk of Death in Middle-aged and Older Adults," *Archives of Internal Medicine*, 169 (2009), 515-523.

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," *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," *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).

Looker, Anne, Peter Dallman, Margaret Carroll, Elaine Gunter and Clifford Johnson, "Prevalence of Iron Deficiency in the United States," *The Journal of the American Medical Association*, 277 (1997), 973-997.

Love, Albert and Charles Davenport, *Defects Found in Drafted Men. Statistical Information Compiled from the Draft Records Showing the Physical Condition of the Men Registered and Examined in Pursuance of the Requirements of the Selective-Service Act*, (Washington, D.C.: Government Printing Office, 1920).

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).*

Meyers, Linda, Jean-Pierre Habicht, Clifford Johnson and Cavell Brownie, "Prevalences

of Anemia and Iron Deficiency Anemia in Black and White Women in the United States Estimated by Two Methods," *American Journal of Public Health*, 73 (1983), 1042-1049.

Moran, T. and J.C. Drummond, "Reinforced White Flour," *Nature*, 146 (1940), 117-118.

National Cholesesterol Education Program, "Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report," *Circulation*, 106 (2002), 3143-421.

Nissenson, Allen and Lawrence Goodnough, "Anemia: Not Just and Innocent Bystander?," *Archives of Internal Medicine*, 163 (2003), 1400-1404.

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.

Park, Youngmee, Christopher Sempos, Curtis Baron, John Vanderveen, and Elizabeth Yetley, "Effectiveness of Food Fortification in the United States: The Case of Pellagra," *American Journal of Public Health*, 90 (2000), 727-738.

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).

Scholl, Theresa, "Iron status during pregnancy: setting the stage for mother and infant," *The American Journal of Clinical Nutrition*, 81S (2005), 1218S-1222S.

Scholl, T O and M L Hediger, "Anemia and iron-deficiency anemia: compilation of data on pregnancy outcome," *The American Journal of Clinical Nutrition*, 59 (1994), 492S-500S.

Schulz, Laura, "The Dutch Hunger Winter and the Developmental Origins of Health and Disease," *PNAS: Proceedings of the National Academy of Sciences*, 107(39), 16757-58, (2010).

Semba, Richard, "The Impact of Improved Nutrition on Disease Prevention," in *Silent Victories: The History and Practice of Public Health in Twentieth-Century America*, John Ward and Christian Warren ed., (New York, NY: Oxford University Press, 2007).

Seomantri, AG, Ernesto Politt and Insun Kim, "Iron deficiency anemia and educational Achievement," *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," *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.

Steer, Philip, "Maternal hemoglobin concentration and birth weight," *The American Journal of Clinical Nutrition*, 71S (2000), 1285S-1287S.

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).

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.

Weick, Sister Mary Theodora, "A History of Rickets in the United States," *The American Journal of Clinical Nutrition,* 20 (1967), 1234-1241.

Wilder, Russell 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).

Yip, Ray, "Iron Deficiency: Contemporary Scientific Issues and International Programmatic Approaches," *Journal of Nutrition*, 124 (1994), 1479S-1490S.

Zeng Lingxia,Yue Cheng, Shaonong Dang, Hong Yan, Michael J Dibley, Suying Chang and Lingzhi Kong, "Impact of micronutrient supplementation during pregnancy on birth weight, duration of gestation, and perinatal mortality in rural western China: double blind cluster randomised controlled trial," *BMJ*, 337 (2008), 1-11.

Zhu, Y Isabel and Jere Hass, "Response of serum transferring receptor to iron supplementation in iron-depleted, nonanemic women," *American Journal of Clinical Nutrition*, 67 (1998), 271-275.