The Cognitive Effects of Micronutrient Deficiency: Evidence from Salt Iodization in the United States

James Feyrer*, Dimitra Politi†, and David N. Weil‡

February 27, 2015

Abstract

Iodine deficiency is the leading cause of preventable mental retardation in the world today. The condition, which was common in the developed world until the introduction of iodized salt in the 1920s, is connected to low iodine levels in the soil and water. We examine the impact of salt iodization on cognitive outcomes in the US by taking advantage of this natural geographic variation. Salt was iodized over a short period of time beginning in 1924. We use military data collected during WWI and WWII to compare outcomes of cohorts born before and after iodization, in localities that were naturally poor and rich in iodine. We find that for the one quarter of the population most deficient in iodine this intervention raised IQ by approximately one standard deviation. Our results can explain roughly one decade's worth of the upward trend in IQ in the US (the Flynn Effect). We also document a large increase in thyroid-related deaths following the countrywide adoption of iodized salt, which affected mostly older individuals in localities with high prevalence of iodine deficiency.

JEL classification: I12, I18, J24, N32

Keywords: Cognitive ability, Flynn Effect, human capital, productivity

*Department of Economics, Dartmouth College, and NBER. Email: james.feyrer@dartmouth.edu.

†School of Economics, The University of Edinburgh. Email: Dimitra.Politi@ed.ac.uk.

‡Department of Economics, Brown University, and NBER. Email: David_Weil@brown.edu. We would like to thank Hoyt Bleakley, Kenneth Chay, Andrew Clausen, Joseph Ferrie, seminar participants at Tel Aviv University, Boston College, Brown University, Northwestern University, and conference participants at the NBER's Cohort Studies Meeting, AEA Meetings and SIRE Young Researchers' Forum for helpful comments. Desislava Byanova, Federico Droller, Bryce Millett Steinberg, and Young Min Kim provided excellent research assistance.

1 Introduction

Deficiencies in the consumption of micronutrients, the vitamins and minerals required in trace amounts for proper metabolic functioning, are a major contributor to poor health in developing countries. For example, in 2005, some 667 thousand deaths of children under five (6.3% of all such deaths worldwide) were estimated to be attributable to vitamin A deficiency. The mortality burden of zinc deficiency was roughly two-thirds as large[1]. The World Health Organization estimates that nearly 50 million people suffer some degree of mental impairment due to iodine deficiency, and that this is the leading cause of preventable mental retardation in the world. Anemia, the most important cause of which is inadequate intake of iron, affects 1.6 billion people around the globe[2].

In developed countries, as recently as the first half of the twentieth century the toll of micronutrient deficiencies was also significant. This included rickets due to inadequate vitamin D in industrial cities; pellagra due to inadequate niacin consumption among populations dependent on maize; and iodine deficiency. A combination of dietary improvements and widespread food fortification has eliminated these problems.

Around the developing world, numerous programs of micronutrient supplementation are in place. Between 1999 and 2007, the fraction of children aged 6-59 months receiving full vitamin A supplementation in UNICEF-priority countries rose from 16% to 72%. Since 2004, 46 countries have adopted zinc supplementation as part of their child health policies. In China, the fraction of salt iodized rose from 30% to 96% between 1990 and 2000, and India banned the sale of non-iodized salt for human consumption in 2006[3]. In both Vietnam and the Philippines there are National Micronutrient Days, featuring education and distribution of supplements.

The 2008 Copenhagen Consensus ranked micronutrient supplementation for children

[1]By comparison, stunting, severe wasting, and intrauterine growth restriction, all primarily resulting from protein-energy malnutrition, were together responsible for 21% of deaths in this age group (Black, Allen, Bhutta, Caulfield, De Onis, Ezzati, Mathers and Rivera 2008).

[2]de Benoist, Mclean, Egli and Cogswell, eds (2008), World Health Organisation (2013).

[3]Mannar and Bohac (2009) and Micronutrient Initiative (2009).

(specifically zinc and vitamin A) as the number one most cost effective potential intervention for advancing global welfare. Specifically, the scholars on the Copenhagen panel estimated that the ratio of benefits (better health, fewer deaths, and increased future earnings) to costs of supplementation would be 17 to 1. Another package of micronutrient supplementation (iron and iodine) was estimated to be the third most cost effective potential intervention, with a benefit-cost ratio of 9 to 1[4].

Evaluating the effects of large-scale micronutrient interventions in practice on either health or labor productivity is difficult because of problems in measurement and identification. For example, micronutrient interventions often take place at the same time as many other changes in the health and economic environment. Further, as with many other health interventions (for example, bednets for mosquitoes and vaccination), take-up can be correlated with other characteristics that affect health outcomes. Evaluation is particularly difficult in the case of interventions, the benefits of which may not be observable for several decades. In a notable study that takes identification seriously, Field, Robles and Torero (2009) exploit delays and gaps in the implementation of the iodization program in Tanzania in order to identify a within-district (and within-household) effect on school grade attainment. They find that treatment of mothers with iodated oil resulted in a rise in schooling of 0.33 years among children, with a larger effect for girls than boys.

In this paper we study the long-run effects of iodine fortification in the United States[5]. The most important effects of iodine deficiency are on the cognitive development of fetuses in utero, and these effects are not reversible later on in life. Thus one would expect the benefits of iodization in terms of adult health to be apparent only many decades after the rollout of the intervention. Because iodine deficiency affects mental development, the elimination of

[4]Copenhagen Consensus Center (2008).

[5]In using historical data from a developed country to study what is now viewed as a development issue, we are following along the lines of Bleakley (2007), who examined the effects of hookworm eradication in the American South in early 20th century and finds significant effects on education and future incomes of those cohorts that benefited from the intervention. Similarly, Ferrie, Rolf and Troesken (2012) finds that increases in lead exposure through water pipes resulted in lower test scores among US Army recruits during World War II and Watson (2006) finds that improvements in sanitation of Indian reservations in the 1960's explain a big part of the convergence in infant mortality rates between Whites and Native Americans.

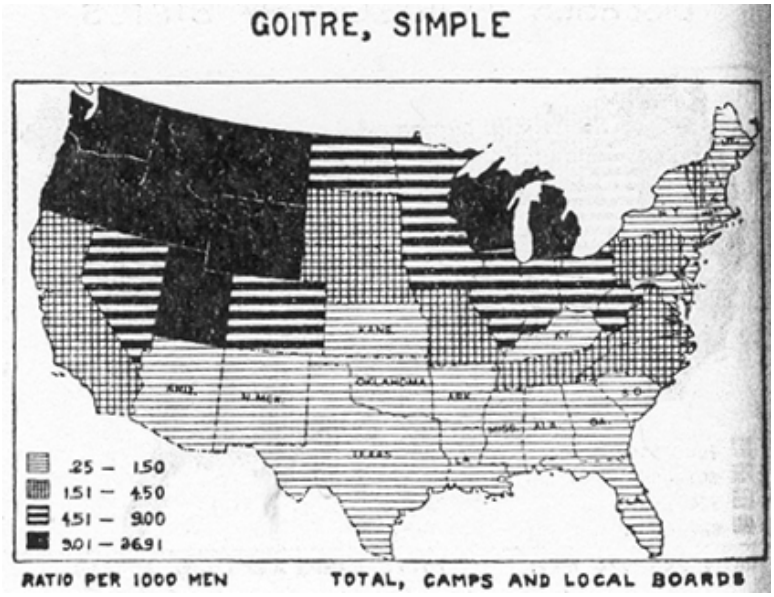
this deficiency is a candidate explanation for part of the Flynn Effect, the gradual rise in measured IQ over many decades that has been observed in developed countries. As we show below, the benefits of iodization were large, but there were also very significant health costs (several thousand premature deaths) associated with the program.

We identify our estimates by looking at the interaction of pre-existing geographic variation in iodine deficiency with a national program of salt iodization. Iodine deficiency is linked directly to geography through the food and water supply. In adults, the most noticeable symptom of iodine deficiency is goiter, the enlargement of the thyroid gland. Prior to salt iodization, endemic goiter was present in specific regions of the US and absent from others, depending on the iodine content of the soil and water. Figure 1 illustrates the geographic distribution of goiter in the US as measured among World War I recruits (we discuss the data further below). In 1924 iodized salt was introduced in the United States explicitly to reduce the goiter rate. This intervention rapidly reduced the incidence of iodine deficiency. Since there are large in utero effects of iodine deficiency, we expect to see a significant difference between those born before and after the introduction of iodized salt in locations with low levels of environmental iodine. Those living in high iodine regions provide a control group.

We exploit two unique data sources to look at the effects of iodine deficiency eradication on cognitive ability. After World War I, statistics from draft physicals were compiled by geographic location. From this source we know the incidence of goiter for 151 geographic regions before the introduction of iodized salt. This provides us with a measure of iodine deficiency prior to treatment.

Our outcome measure is provided by an extensive dataset of men who enlisted in the Army during World War II. The timing of the war generates a large sample of men born in the years 1920-1927, neatly covering the introduction of iodized salt. Upon enlistment, each recruit took the Army General Classification Test (AGCT), a forerunner to the AFQT. The Air Force was assigned draftees with significantly higher average test scores than the Ground Forces. We exploit this non-random assignment in our identification strategy.

Figure 1: Distribution of goiter during World War I



Source: Love and Davenport (1920)

The probability of assignment to the Air Force rises significantly in low iodine (i.e. high goiter) locations in the years after the introduction of iodized salt. In the lowest iodine regions, our estimates suggest a 3.8-10 percentage point increase in the probability of a man being assigned to the Air Forces after iodization. Using information about average scores of Air and Ground Force recruits we can infer a one standard deviation increase in average test scores in these regions. The average level of iodine deficiency in the US was significantly lower than in the highest regions, so the overall effect in the US was much more modest. The increase in cognitive ability due to salt iodization may have contributed to the rise in measured IQ that took place over the course of the twentieth century, the so-called "Flynn Effect".

In addition to the positive cognitive effects of iodizing salt we document a significant negative side effect. Chronic iodine deficiency, followed by a sudden large intake of iodine, can cause iodine-induced hyperthyroidism, which can be fatal. We show that there was a large rise in thyroid-related deaths in iodine deficient regions following the introduction of

iodized salt. We estimate that at least 10,000 deaths over the period 1925-1942 were the result of iodization. These deaths affected mostly older age groups, which had been deficient for a longer period of time.

The paper proceeds as follows: Section 2 provides some background on iodine deficiency disorders. Section 3 outlines the history of salt iodization in the US. Section 4 describes our data and provides some background on their collection. In Section 5 we present previously undocumented evidence on the spike of thyroid-related mortality following iodization. Section 6 explains our identification strategy and section 7 presents our results. In Section 8 we present three robustness checks which corroborate our main findings. Section 9 puts our results in the context of related research on the effects of iodization on cognitive ability. Section 10 concludes.

2 Iodine Deficiency Disorders

Recent work has shown that the quality of maternal health and nutrition during pregnancy has persistent effects on adult health outcomes. For instance, Almond (2006) shows that cohorts exposed to the Spanish Influenza of 1918, either in utero or during the first months of life, had worse health and socio-economic outcomes in their lifetime. Behrman and Rosenzweig (2004) find that differences in birth weight among identical twins are reflected in differences in school attainment and adult earnings[6].

Iodine is one of the "big three" micronutrients, deficiencies in which are a major source of ill health in developing countries (the other two are vitamin A and iron). Iodine deficiency, in particular, is the leading cause of preventable mental retardation in the world. WHO estimates that nearly 50 million people suffer some degree of mental impairment due to iodine deficiency[7]. 1.88 billion people -nearly one third of the world's population- are at risk, in the sense that their iodine intake is considered insufficient. 241 million school-age

[6]For an excellent review of the literature, see Almond and Currie (2011)

[7]source: WHO, http://www.who.int/features/qa/17/en/index.html.

children receive inadequate amounts of iodine in their diet (Andersson, Karumbunathan and Zimmermann 2012).

Iodine Deficiency Disorders (IDD) is a term used to describe a set of conditions, ranging from goiter to cretinism, that result from inadequate intake of iodine. Around 70-80% of the total iodine content in the human body is found in the thyroid gland (Fleischer, Forbes, Harriss, Krook and Kubota 1974). The thyroid gland uses iodine to produce thyroxin, a hormone that regulates metabolism. When there is too little iodine in the diet, the thyroid compensates for the shortage by enlarging in order to produce more thyroxin per unit of iodine. When dietary iodine is only slightly inadequate, the enlarged thyroid will be able to produce sufficient thyroxin for normal body functioning. This enlargement is known as euthyroid goiter. At lower levels of dietary iodine, the enlarged thyroid will produce inadequate thyroxin, a condition known as hypothyroid goiter, characterized by slow metabolism, lethargy, and weight gain.

Goiters due to iodine deficiency are often reversible after an increase in dietary iodine. However, in some people with this condition, increased iodine consumption results in the thyroid gland producing excessive quantities of thyroxin, resulting in hyperthyroidism. This is called iodine-induced thyrotoxicosis. Hyperthyroidism is characterized by an overly accelerated metabolism, with symptoms including rapid heartbeat, weight loss, temperature elevation, nervousness, and irritability[8,9]. Iodine-induced thyrotoxicosis is most likely to occur in individuals who have experienced long periods of iodine deficiency[10].

Iodine deficiency early in pregnancy causes irreversible, serious brain damage to the foetus. Severe iodine deficiency can result in cretinism, which is characterized by "profound mental deficiency, dwarfism, spastic dysplasia and limited hearing"(Scrimshaw 1998,

[8]Reference: "Goiter" in Health A to Z, www.healthatoz.com

[9]Goiter and hyperthyroidism can also result from Graves Disease, also called Basedow disease, an immune condition in which the thyroid is stimulated to produce excess thyroxin.

[10]Iodine-induced thyrotoxicosis is also called Jod-Basedow disease. "Jod" is German for iodine. The name indicates that iodine consumption is resulting in the symptoms of Basedow disease. Examining the rise in thyroid disease that followed the introduction of iodized bread in Tasmania, Connolly (1971) found that most patients with iodine-induced thyrotoxicosis had pre-existing goiter, and few had Graves disease.

p.364)[11]. In endemic areas cretinism can affect up to 15% of the population (de Benoist, Andersson, Egli, Takkouche and Allen, eds 2004). However, as Scrimshaw points out, "even in areas where cases of cretinism due to iodine deficiency in the mother are few, the linear growth of the infant, its intellectual capacity, and certain other of its neurological functions are permanently compromised to varying degrees" (Scrimshaw 1998, p.351). In other words, even if iodine deficiency does not result in cretinism, an iodine-deficient population will experience a continuum of cognitive impairment. Deficient populations typically see a leftward shift in the entire IQ distribution[12]. Bleichrodt and Born (1994), in a meta-analysis of 18 studies, estimate that the average IQ of iodine-deficient groups is 13.5 points lower than non-deficient groups. While widely cited, this result is based primarily on cross-sectional comparisons of observational data, which makes the identification of a causal effect questionable.

Historically, the geographical characteristics of an area determined the endemicity of iodine deficiency. Ocean water is rich in iodine, which is why endemic goiter is not observed in coastal areas. Iodine is transferred from the ocean to the soil by rain. Heavy rainfall can cause soil erosion, in which case the iodine-rich upper layers of soil are washed away. The last glacial period had the same effect: iodine-rich soil was substituted by iodine-poor soil from crystalline rocks (Koutras, Matovinovic and Vought 1980). This explains the prevalence of endemic goiter in regions that were marked by intense glaciation, such as Switzerland and the Great Lakes region of the USA. Iodine is taken up by plants when it is present in the soil, and can reach humans either through plants or animals which have eaten them. Iodine is also present in subsurface water in some locations. Finally, deposits of mineral salt (the remains of evaporated seawater) contain iodine, but this is lost when the salt is refined. The human body does not naturally store much iodine, so seasonal variations in iodine consumption may

[11] According to one interpretation, the word cretin comes from the French term for Christian. Cretinism was endemic in the French Alps, where the term was apparently invented for those who were too mentally limited to commit a sin and who, therefore, were good Christians.

[12] Scrimshaw (1998) provides a list of studies and experiments documenting the hindering effects of iodine deficiency in utero on mental development.

result in seasonal manifestations of iodine deficiency.

Ancient civilizations treated goiter with burnt sponge or seaweed, which are rich in iodine (Curtis and Fertman 1951, Langer 1960). Iodine was first isolated in 1811 by Courtois. Over a century of clinical research proved iodine's essential role in prophylaxis against iodine deficiency disorders. Doctors and public health officials have used different methods of iodine supplementation. Salt iodization has emerged as the cheapest and widest-reaching way to protect a population from iodine deficiency[13].

3 Iodine Deficiency and Salt Iodization in the United States

Micronutrient deficiencies in developed countries were only eliminated in the last century. For example, vitamin D was added to milk in the 1930s, and B vitamins (including niacin) were added to baked goods starting in the 1940s (Bishai and Nalubola 2002)[14]. Salt iodization was the first step in the systematic fortification of food for public health reasons. It was made possible by the nearly simultaneous discovery of a widespread health problem and of its underlying cause.

In the First World War draft, a little more than 2.5 million draftees were examined for various physical and mental shortcomings. A lengthy collection of countrywide data was compiled from these examinations. The statistics showed the geographic distribution of many diseases and defects across the United States (Love and Davenport 1920). Goiter was among the diseases that were measured, because an unexpectedly high number of soldiers had trouble wearing a uniform because of their enlarged thyroids.

[13] Alternatives have included the iodization of water supplies and bread, as well as the provision of iodine-enriched chocolates or milk to babies and schoolchildren, and injections of slow-releasing iodated oil. Iodization of water supplies proved wasteful since only a small proportion of water is used for drinking and cooking purposes. Bread iodization was used in the Netherlands as a wartime measure (Matovinovic and Ramalingaswami 1960).

[14] Vitamin D was used to fight rickets, and B vitamins were used to fight pellagra.

According to the draft examinations almost 12,000 men had goiter and a third of these were judged unfit for service, because the size of their neck was too big for the military tunic to be buttoned (Kelly and Snedden 1960, p.34). Most of them came from states in the Northwest (Washington, Oregon, Idaho, Montana) and the area around the Great Lakes. In Northern Michigan, for instance, more draftees were judged unfit for service "for large and toxic goiters than for any other medical disorder" (Markel 1987, p.221). On the other hand, goiter was rare in people coming from coastal areas.

The realization of the problem led to multiple surveys of goiter, which confirmed the geographical variation in the prevalence of the disease. By that time, there was medical and veterinary evidence showing that goiters could be reduced by adding iodine to the diet[15]. These observations prompted a public debate on whether and how to provide iodine prophylaxis to the American population. Some objections were raised as to the potential side-effects of universal measures of prophylaxis. It had been documented that large amounts of iodine could cause hyperthyroidism in some adults and thyrotoxicosis in others. Despite these concerns, the medical consensus was that small amounts of iodine in the diet were beneficial for the vast majority of iodine-deprived populations, and this was confirmed by experimental results in schoolchildren.

Public health authorities in Michigan, one of the worst-afflicted states, held a symposium on thyroid disease in 1922. The idea of salt iodization (which had been proposed by researchers in Switzerland, and was first implemented in that country in 1922) was introduced by David Murray Cowie, M.D. as a cheap and effective means of providing iodine to all population groups, regardless of social status. The Iodized Salt Committee was set up to investigate the matter further. The Committee, chaired by Cowie, produced reports on the low iodine content of drinking water in Michigan and the possibility of effective prevention of goiter through iodized salt. Its members agreed upon the launch of a statewide educational

[15]Experiments with school children confirmed that the size of goiters decreased after receiving iodine. The first such experiment took place in Akron, Ohio in 1917, under the direction of David Marine and O.P. Kimball. For details see Marine and Kimball (1921), and Carpenter (2005).

campaign on goiter and its prevention though iodized salt, sponsored by the Michigan State Medical Society. The campaign, launched in 1922, included lectures to physicians and the general public delivered across the state.

The Committee also contacted the state's salt manufacturers. The salt producers, although convinced about the public-service character of the project, had initial qualms about its economic feasibility and profitability; it would be financially impossible to separate the salt intended for the Michigan market and then add iodine to it. Instead, the Salt Producers Association decided to launch iodized salt nationwide; they saw the new product as an improved commodity for which there would be a much larger market -and corresponding profits- than that of Michigan. As a result, in May 1924, Michigan was the first state to introduce iodized salt[16]. The actions of the Michigan salt producers were important, because Michigan was the largest producer of salt for human consumption in the country[17].

The salt companies contributed to the educational campaign through aggressive advertising of the "new salt" throughout the country (Markel 1987, p. 224). Figures 2 and 3 shows two ads from this period. Figure 4 is a copy of a newspaper clipping from the era.

After Michigan introduced iodized salt, the market penetration of the new product was quite rapid. McClure (1934) reports that as of 1930, 89% of total salt sales in Michigan were iodized. Salt producers made the new product available nationwide, in order to make it financially feasible. Morton Salt Company, the largest producer in the country at the time,

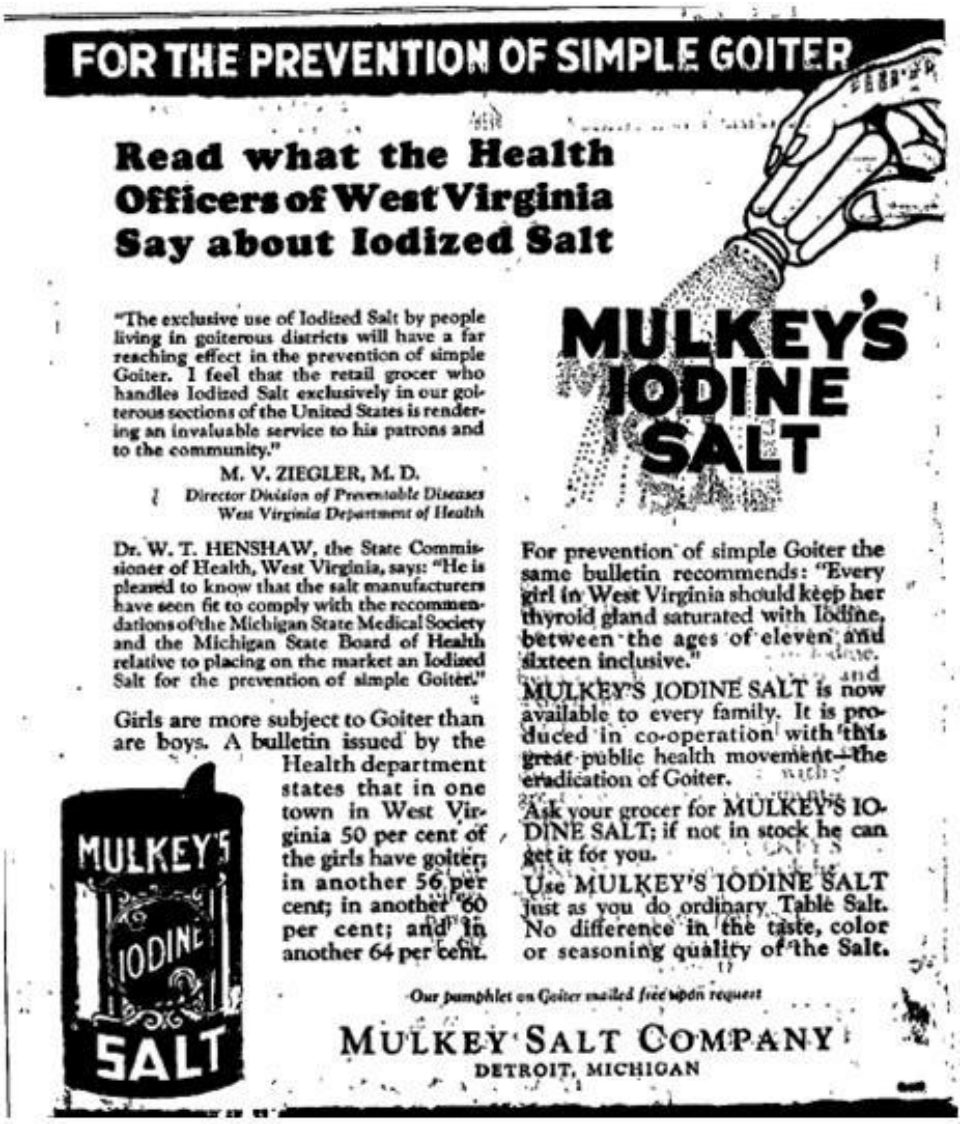
[16]Before that, in April 1923, public authorities in Rochester, New York, introduced iodine in the water supplies of one reservoir, in what is known as "the Rochester experiment" (Kohn 1975). Subsequent goiter surveys show an important decrease in incidence. However, it seems unlikely that this decrease was due to the iodization of the water supply, for the following reasons: first, only one of the reservoirs was iodized. Second, by that time, iodized salt was available and widely used in Rochester. Third, because of bigger awareness and improved medical monitoring, doctors were more likely to prescribe iodine supplements to anybody with a palpable goiter.

[17]Salt production takes three forms: evaporated, rock salt, and the production of liquid brine. In 1924 the quantities produced by these three methods were 2.22, 2.06, and 2.51 million short tons, respectively. Brine was used exclusively as a feedstock by the chemical industry. According to the Salt Institute (http://www.saltinstitute.org), as of today, virtually all food grade salt sold or used in the United States is produced by evaporation. This was the case in 1924 as well (personal communication from Richard Hanneman, President, Salt Institute, March 6, 2008). In 1924, Michigan was the largest producer of evaporated salt in the country, accounting for 36% of total salt production. The next largest producers were New York (18%) and Ohio (14%) (Katz 1927).

Figure 2: Morton's advertisement for iodized salt

Source: *"The Bee"*, Danville, Virginia, May 11th 1925

Figure 3: Mulkey's advertisement for iodized salt



FOR THE PREVENTION OF SIMPLE GOITER

Read what the Health Officers of West Virginia Say about Iodized Salt

"The exclusive use of Iodized Salt by people living in goiterous districts will have a far reaching effect in the prevention of simple Goiter. I feel that the retail grocer who handles Iodized Salt exclusively in our goiterous sections of the United States is rendering an invaluable service to his patrons and to the community."

M. V. ZIEGLER, M. D.
Director Division of Preventable Diseases
West Virginia Department of Health

Dr. W. T. HENSHAW, the State Commissioner of Health, West Virginia, says: "He is pleased to know that the salt manufacturers have seen fit to comply with the recommendations of the Michigan State Medical Society and the Michigan State Board of Health relative to placing on the market an Iodized Salt for the prevention of simple Goiter."

Girls are more subject to Goiter than are boys. A bulletin issued by the Health department states that in one town in West Virginia 50 per cent of the girls have goiter; in another 56 per cent; in another 60 per cent; and in another 64 per cent.

MULKEY'S IODINE SALT

For prevention of simple Goiter the same bulletin recommends: "Every girl in West Virginia should keep her thyroid gland saturated with Iodine, between the ages of eleven and sixteen inclusive."

MULKEY'S IODINE SALT is now available to every family. It is produced in co-operation with this great public health movement—the eradication of Goiter.

Ask your grocer for MULKEY'S IODINE SALT; if not in stock he can get it for you.

Use MULKEY'S IODINE SALT just as you do ordinary Table Salt. No difference in the taste, color or seasoning quality of the Salt.

Our pamphlet on Goiter mailed free upon request

MULKEY SALT COMPANY
DETROIT, MICHIGAN

Source: "Middlesboro Daily News", Kentucky, June 14th 1924

Figure 4: A newspaper article discussing iodized salt

Iodized Salt Now in General Use

New Way of Putting Iodine Into Food.

Do you ever pay particular attention in some of Decatur's restaurants when the waiter slips up and fills the salt shaker? If you have you probably have noticed that often the carton from which the salt is sifted into the shaker is labelled "iodized salt." And having seen it did you ever wonder why?

Restaurants use it because of an ever increasing demand for iodized salt in the food, and its use is a prevention of disease.

CAN DO NO HARM.

"I am glad to see it used," said Dr. Keister Monday morning when the matter was mentioned to him. "I do not know that I had noticed its use, but it can do no harm as I see it and its use may do much good. Iodine is recommended for use in cases of simple goiter and some splendid results have been reported in its use."

This section of the country, he explains is known for the great number of goiter cases, and a large per cent of the children in the schools have enlarged thyroids. Iodine is gone from the land here and therefore from the water. At the sea shore there rarely is a case of goiter developed. Mountain regions and lake regions and inland regions have many cases, it is believed because of this lack of iodine in the food and the water.

FROM SEA WEED.

Oysters and all kinds of sea food contain iodine, and the iodine itself is obtained from sea weed, and a Japanese sea weed in particular. As a rule the farther distant a section is from the sea the greater the amount of goiter.

There are three ways of administering it and by salt is the simplest. It merely puts into the food what it has failed to obtain from the land or water. Another way is to administer it medicinally and the third is by use in drinking water.

GIVEN TO CHILDREN.

In Akron, O., some very fine results have been obtained by giving it to school children. It was given over a period of ten years and the results watched. One half of the 10,000 children were given it twice a year. Among those to whom it was administered who were in a normal condition, not one developed goiter. Among those who did not take it 27.6 developed goiter in some form. Others who had the trouble but who took it showed that sixty per cent returned to normal condition.

Source: *"The Decatur Review"*, Illinois, December 1st 1925

began selling iodized salt on a nationwide basis in the fall of 1924[18]. At the same time, public awareness of the problem, especially in those areas that were worst afflicted, was gaining momentum. Articles in newspapers and magazines around the country advocated the use of the new salt for all cooking and eating purposes, making references to successful goiter prophylaxis in Switzerland[19]. Most state health authorities urged the public to use iodized salt. In advertisements of iodized salt, the new commodity bore the endorsement of state or national medical associations and educational booklets were provided by the salt companies upon demand. We have not been able to find precise numbers on the penetration of iodized salt (or other iodine dietary supplements) for the rest of the country, but all the evidence suggests that the new product became very popular very quickly, especially in the goiter-belt region, where it mattered the most. We have many newspaper sources that show the generalized availability and use of iodized salt from 1924 onwards. In addition, as we show in section 5, the rapid rise in thyroid-related mortality, not only in Michigan but throughout the country, further indicates that iodine supplementation diffused quite rapidly.

The focus of the salt iodization campaigns was goiter eradication. The link between iodine deficiency and mental function was still unknown. The decrease in mental retardation that resulted from iodization campaigns was a very positive unintended consequence. These cognitive benefits are the main motivation for modern iodization campaigns.

4 Data Sources

We wish to examine the impact of salt iodization in 1924. In order to do so, we need two pieces of information. First, who was likely to benefit from salt iodization? Second, how did

[18](Markel 1987). Collusion in the evaporated salt industry was widespread, and Morton acted as the price setter. Many companies literally made copies of Morton's price schedule, simply replacing their company letterheads for that of Morton (Fost 1970). Morton's decision to iodize salt in 1924 would thus likely have affected a large percentage of households, both directly and through Morton's influence on smaller companies.

[19]For example: Lima News (Lima, Ohio), on August 29, 1924, reports that iodized salt is now marketed "thru the regular grocery trade". In Appleton Post Crescent (Appleton, Wisconsin), on January 28, 1926, it is mentioned that "iodized salt is now sold by grocers everywhere, and families can use it instead of ordinary salt".

the outcomes of these individuals change after the iodization of salt in 1924?

As previously described, the presence or absence of endemic iodine deficiency depends on the iodine content of the soil and ground water. Individuals born in low iodine areas are much more likely to be iodine deficient than those born in high iodine areas. Iodizing salt will have a large effect on the former and no effect on the latter. The abrupt switch to iodized salt completes the natural experiment. Individuals born in high iodine areas should see no effect from salt iodization. Individuals born in low iodine areas, by contrast, should see effects whose size is related to the severity of iodine deficiency before the treatment. The latter group provides our treatment group and the former group our control.

To implement this strategy we need data on the prevalence of iodine deficiency before 1924, as well as data on outcomes of individuals born in these areas both before and after 1924. Our primary data sources take advantage of two previously unused surveys of prime age American males in the early part of the twentieth century.

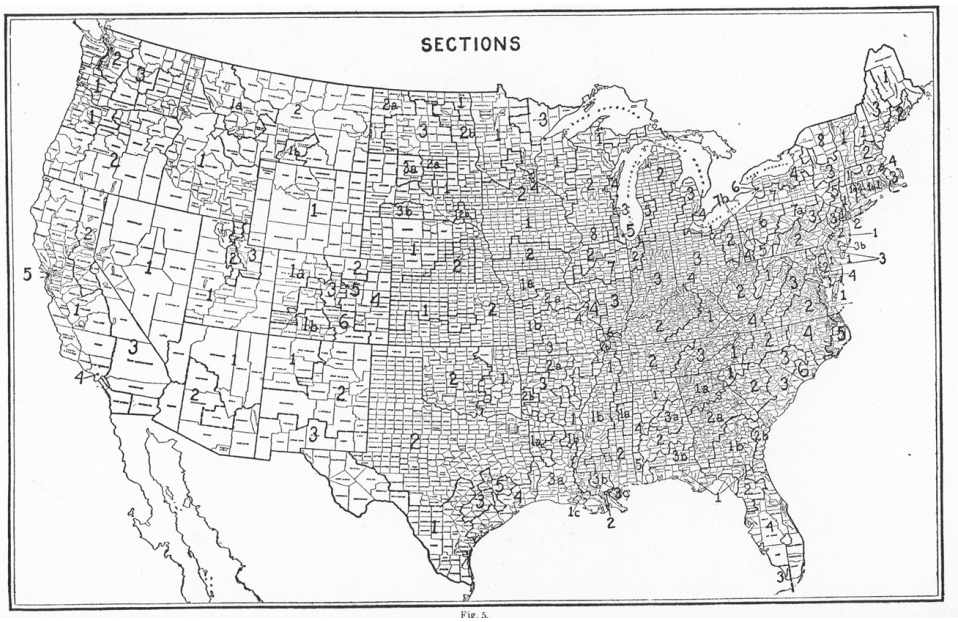
4.1 World War I: Defects Found in Drafted Men

For data on the prevalence of iodine deficiency before 1925 we use a volume entitled *Defects Found in Drafted Men* (henceforth referred to as *Defects*), published by the War Department in 1920 (Love and Davenport 1920)[20]. *Defects* summarizes the results of all the physical exams performed on draftees during World War I for both accepted and rejected men. Data on prevalence rates per 1,000 are recorded for 271 different medical conditions. The data are regional, organized by units called sections. All but the lowest population states are broken down into multiple sections. Illinois and New York, for example, are each broken down into 8 sections. Each section is defined as a collection of counties[21]. In total, *Defects* has data on 151 separate regions of the country. Figure 5 is a map of the US showing the locations of sections from *Defects*.

[20]Many thanks to Hoyt Bleakley for making us aware of this marvelous book.

[21]Since county borders in the US are relatively static, it is straightforward to map the *Defects* sections to present day US counties.

Figure 5: Sections of U.S. counties in World War I



source: Love and Davenport (1920)

The medical condition of interest for our study is simple goiter, which is a direct result of iodine deficiency. We use simple goiter as a proxy for underlying iodine deficiency. Simple goiter is relatively common in the data, with a population weighted average prevalence of about 5 cases per 1,000 and a median prevalence of 2.5 per 1,000. The prevalence rates range as high as 29.85 cases per 1,000. Figure 6 is a histogram of the distribution of goiter across sections in *Defects*. Though there are no sections with a zero rate of simple goiter, about one third of the sections have rates of less than 1 per 1,000.

Figure 7 shows the simple goiter rate at the state level. Rates range from a high of almost 27 per 1,000 in Idaho to a low of 0.25 per 1,000 in Florida. The fact that the data are at a finer level of aggregation than the state level is important because there is significant regional variation within the high goiter states. For example, in the five sections in Michigan the rates reported in *Defects* range from over 25 in the Upper Peninsula to less than 10 in Detroit and the surrounding areas.

The fact that spatial variation in goiter rates, as documented in *Defects*, was caused

Figure 6: Distribution of goiter rates across World War I sections

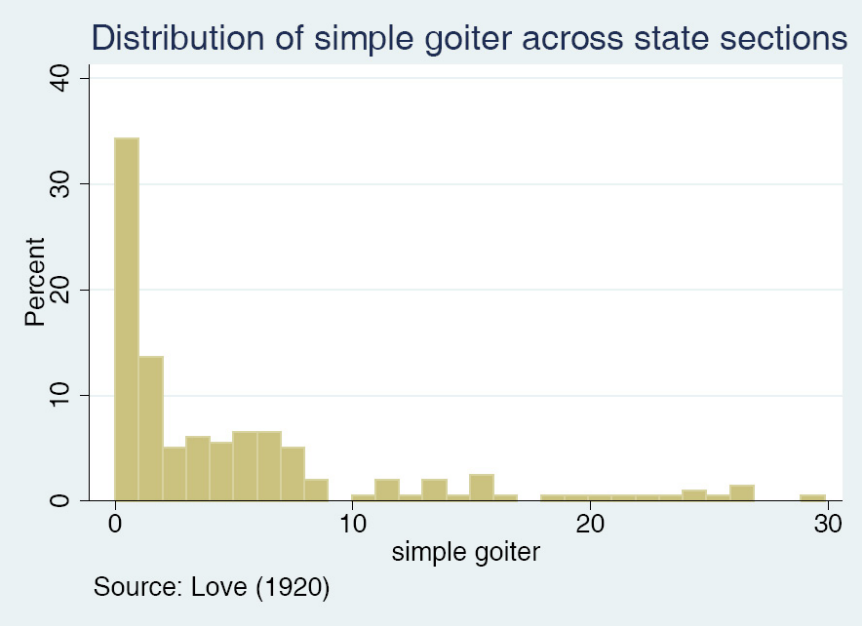


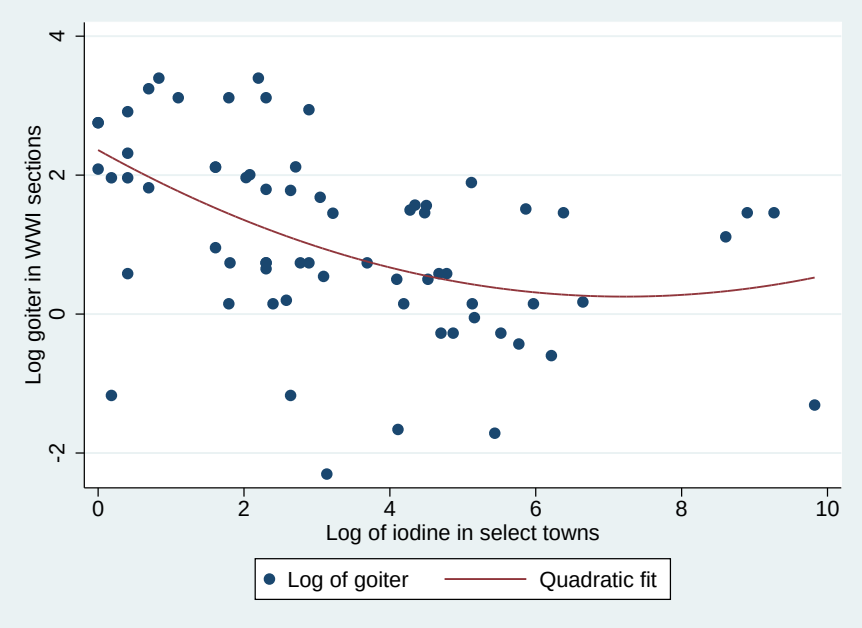
Figure 7: Simple goiter rates by state, as recorded during World War I draft

TABLE 18.—*Grand total for goiter, simple, with ratio per 1,000 men.*

State.	Number of cases.	Ratio per 1,000.	State.	Number of cases.	Ratio per 1,000.
Idaho	336	26.91	Kentucky	90	1.41
Oregon	421	26.31	District of Columbia	16	1.39
Washington	832	23.40	Kansas	48	1.25
Montana	576	21.00	Arizona	10	1.21
Utah	185	15.72	New York	308	1.19
Wyoming	102	15.37	Maryland	35	.94
Wisconsin	886	14.02	South Carolina	37	.94
Alaska	16	13.14	Connecticut	32	.89
Michigan	1,131	11.43	New Mexico	9	.88
North Dakota	156	8.73	Oklahoma	44	.72
Minnesota	578	8.04	New Hampshire	6	.70
West Virginia	307	7.89	Maine	13	.66
Illinois	1,397	7.79	Mississippi	24	.64
Iowa	458	6.68	Louisiana	32	.62
Indiana	464	6.49	Delaware	3	.59
Nevada	21	6.38	Alabama	29	.56
Ohio	798	5.59	Rhode Island	8	.55
Colorado	119	5.29	Georgia	33	.52
California	359	4.45	New Jersey	33	.43
Pennsylvania	829	4.10	Arkansas	17	.40
South Dakota	85	4.09	Massachusetts	29	.32
Missouri	342	3.99	Texas	36	.30
Virginia	188	3.38	Florida	6	.25
Nebraska	63	2.14	State not specified	186	1.96
Vermont	18	2.14			
Tennessee	120	1.96	Total	11,971	4.35
North Carolina	100	1.81			

source: Love and Davenport (1920)

Figure 8: Goiter and iodine content of water



source: McClendon (1924) and Love and Davenport (1920)

by variation in the underlying geological characteristics across locations is also supported by the fact that goiter rates correlate well with measures of the iodine content of water across localities. In a paper published in the *Journal of the American Medical Association* in 1924, J.F. McClendon and Joseph C. Hathaway provided measures of the iodine content of drinking water from 67 localities (some of them in the same county) across the US[22]. These measures came from lakes, springs, rivers and wells. Their paper includes US maps with the low-iodine areas being shaded, and other maps where the high-goiter areas are shaded (their data on goiter come from the *Defects* Book). The two shaded areas in the two maps largely overlap (McClendon 1924). Figure 8 is a scatterplot of the log of iodine in a certain locality against the log of simple goiter in the corresponding section from *Defects*.

Figure 8 shows that goiter prevalence is a matter of geography; goiter is higher where the iodine content of water is lower. The variability in goiter across low iodine areas arises, to some extent, from measurement issues. On the one hand, the goiter data come from com-

[22]Parts of Iodine per hundred billion parts of drinking water.

prehensive physical examinations of World War I recruits. The data on the iodine content of water, on the other hand, come from selected lakes, springs, rivers and wells. So it is entirely possible that, while the water source in one area was iodine-poor, neighboring areas (belonging to the same section) had access to an iodine-rich supply of water, driving the section goiter prevalence down. Another source of variability is distance from the ocean. Even if a town had a low content of iodine in its water supply, the population in the surrounding areas would have have been spared if iodine were plentiful in the air. In our estimation, we do not use the iodine content of water as our main measure of iodine deficiency because it is only available for a small number of counties. We do, however, use it as an alternative measure of iodine deficiency in a robustness check, and show that it supports our main results.

We expect that the effect of iodizing salt will be larger in areas where the prevalence of goiter in *Defects* is highest. Since the medical literature suggests that the largest impact of iodine deficiency is in utero, we would like to sample individuals born in high and low goiter areas both before and after 1924. Luckily for our purposes, the United States government performed a survey of this population during the draft for World War II.

4.2 World War II: Army Enlistment Data

The World War II enlistment data are from the National Archives and Records Administration (NARA). The data has its origin in punch cards produced during the enlistment process for the United States Army, including the Women's Army Auxiliary Corps. It includes both volunteers and draftees. After the war these punch cards were converted to microfilm. In 1994 NARA hired the census department to scan the microfilm into a collection of over 9 million enlistment records from 1938-1946. Though not complete, these records represent the majority of the enlistment into the Army during this period. Of the 1,586 rolls of microfilm, 1,374 (87%) were successfully scanned, leaving approximately 1.5 million punch cards unrecorded. The missing rolls are not sequential, and there are no indications that the records are missing in any systematic fashion. In addition to the missing rolls, several

hundred thousand individual records were unreadable. As far as we know, the only other study to date using this data is Ferrie et al. (2012) which studies the effect of lead exposure on the outcome of World War II enlistees.

Though the format of the punch cards changed somewhat over the course of the war, the coding for basic demographic information was consistent. The demographic fields are name, serial number, state and county of residence, place and date of enlistment, place and year of birth, race, education, and marital status. In addition, the particular branch of the Army that the enlistee entered is coded. One can also infer through the serial number whether the person was drafted. There are no records for individuals who entered the Navy or Marines.

The sample we have available to us is obviously very large and the timing of the draft is nearly perfect for our purposes. Limiting our sample to white men, we have data on over 300,000 from each birth year between 1921 and 1927, giving us extremely good coverage on both sides of the 1924 salt iodization date. Unfortunately, the data do not include the county of birth, only the state. We therefore limit our sample to individuals whose birth state is identical to their state of residence and we assume that the county of residence upon enlistment is the county of birth[23]. This eliminates 24.1% of the sample. Further limiting our attention to white males born between 1920 and 1928 and enlisting between 1940 and 1946 leaves us with 2.27 million records.

4.3 Test Scores and Assignment to the Air Forces

All enlistees were given the Army General Classification Test (AGCT), a predecessor to the AFQT that is currently given to enlistees. This test score would be an ideal outcome for our study, since the primary effect of iodine deficiency in utero is reduced cognitive ability. Unfortunately, the score only appears in our data for a brief time period too early in the war to be useful for our study[24]. We can, however, make some crude inferences about the test

[23]Classical measurement error in the county of birth will introduce attenuation bias in our estimates, which should therefore be seen as lower bounds.

[24]See Ferrie et al. (2012) for a more detailed description of AGCT scores in the NARA data.

Table 1: Percent of enlistees assigned to the Army Air Forces by year.

Year	Number of enlistees	Percentage in Air Force
1940	73373	13.7
1941	108808	32.6
1942	641862	18.9
1943	569361	03.8
1944	263872	10.8
1945	317258	13.9
1946	299995	19.2
Total	2274529	14.0

Source: National Archives and Records Administration (2002)

scores by examining which army branch the enlistees were assigned to.

Each test taker was assigned a grade of I, II, III, IV, or V, with class I being the highest score on the test and V the lowest. Jobs within the Army were deemed to require soldiers from different groups. For example, skilled positions like mechanics tended to get class I or II enlistees, while lower skill jobs, like cooks tended to get class IV or V enlistees. We do not know the particular job assignment of each recruit. However, we can identify enlistees who were assigned to the Army Air Force (AAF) versus those who were assigned to the Army Ground Forces (AGF)[25]. Roughly 14% of all enlistees were assigned to the Air Force over the course of the war, though this proportion varied from year to year. The year to year variation is described in Table 1[26].

Table 2 is a summary of our enlistment data. It gives the total sample size and the

[25]During World War II military aviation was part of the Army and not a separate branch. The Army Air Force was established in June of 1941 as a semi-autonomous group within the Army. Prior to this reorganization the aviation wing of the Army was known as the Army Air Corps. We will refer throughout to the Army Air Force even though our data span this renaming.

[26]The year to year variation is largely driven by changes in US participation in the war. In 1941, the US was largely engaged in the Air War and had less need for ground forces. By 1943 the need for ground forces rose considerably.

percentage of recruits going to the Air Force for each cohort and each enlistment semester. The number of enlisted men started increasing in 1942, and it peaked in the second semester of that year, as well as the first semester of 1943. People born after iodization enlisted in large numbers starting the first semester of 1943. The proportion of recruits going to the Air Forces was particularly low in 1943. This was a reflection of changes in war demands, as ground operations in Europe began in November 1942, so there was a higher need for Ground Forces enlistees.

Air Force enlistees were systematically different than enlistees in the Ground Forces. There is ample evidence that the Air Force enjoyed preferential assignment of inductees compared to other Army branches[27]. This is discussed in more detail below.

[27]See, for example, Palmer, Wiley and Keast (1948, p.21)

Table 2: Sample size and probability of joining the AAF by birth year and enlistment semester

semester of enlistment	Year of birth 20	21	22	23	24	25	26	27	28	Total	
40.01	19.05	12.90	6.67							13.43	% AAF
	42	62	30							134	Sample size
40.02	14.64	12.93	13.64							13.70	% AAF
	24580	28167	20492							73239	Sample size
41.01	19.13	15.78	20.58	35.94						19.39	% AAF
	19593	17651	15445	3052						55741	Sample size
41.02	37.70	68.03	60.74	56.87						46.52	% AAF
	33757	6628	6171	6511						53067	Sample size
42.01	16.45	33.07	45.40	39.76	30.34	21.69				25.35	% AAF
	73433	32665	11995	10164	2864	83				131204	Sample size
42.02	7.17	6.61	16.32	59.17	59.91	24.44				17.22	% AAF
	105570	182286	148683	44544	29485	90				510658	Sample size
43.01	5.60	8.65	4.26	1.67	1.38	1.60	8.51			2.49	% AAF
	19894	22683	59447	152767	141765	28716	94			425366	Sample size
43.02	7.00	8.44	8.08	5.69	4.30	9.04	31.25			7.53	% AAF
	10404	10382	10175	14951	27783	70236	64			143995	Sample size
44.01	3.74	5.66	6.40	5.23	4.39	18.11	26.33			14.32	% AAF
	14374	10830	9159	9563	10969	40914	29866			125675	Sample size
44.02	1.06	1.55	1.57	0.79	0.31	1.45	13.02			7.63	% AAF
	11368	9021	6879	8652	10938	15779	75560			138197	Sample size
45.01	0.91	1.15	0.87	0.26	0.13	0.21	19.70	7.24	10.00	11.26	% AAF
	4734	4255	4032	4967	6724	8227	56122	29750	50	118861	Sample size
45.02	28.10	25.59	22.52	19.43	14.94	10.63	11.51	12.18	37.18	15.54	% AAF
	8694	9598	9346	10773	13091	15767	40479	81021	9628	198397	Sample size
46.01	37.42	32.82	29.79	25.00	19.12	14.20	12.09	11.05	27.09	19.11	% AAF
	5320	6553	6669	7583	9229	11306	19160	78868	65417	210105	Sample size
46.02	40.69	39.51	38.92	36.70	30.73	21.53	16.78	10.43	18.82	19.25	% AAF
	1826	1987	2297	2455	3114	3604	5000	20148	49459	89890	Sample size
Total	14.23	12.49	15.60	15.97	10.50	9.53	16.17	10.89	24.58	14.00	% AAF
	333589	342768	310820	275982	255962	194722	226345	209787	124554	2274529	Sample size

Source: National Archives and Records Administration (2002)

4.4 The Battle for the Best Enlistees

The Army Air Force (AAF), which was still part of the Army during World War II, had a large proportion of jobs that required skilled recruits relative to the Army Ground Forces (AGF). Throughout the war, the AAF pushed to have a large proportion of the more highly skilled recruits assigned to them. In February 1942, the AAF successfully got the "75 percent rule" put into place. Under this rule, 75 percent of the men assigned to the AAF were to have scored above 100 (the median score) on the AGCT[28]. From this, we can infer that individuals assigned to the AAF during this period have, on average, higher test scores than those assigned to the AGF.

Unsurprisingly, the AGF was not pleased with this arrangement, and this rule was not in place for the entire war. Though lower skilled recruits could easily be used in the infantry, the AGF was concerned about having a supply of recruits who could become high-quality combat leaders. The AGF successfully lobbied the War Department to change the rule on August 1, 1942 so that the proportion of above average men received by the AAF was reduced to 55 percent[29].

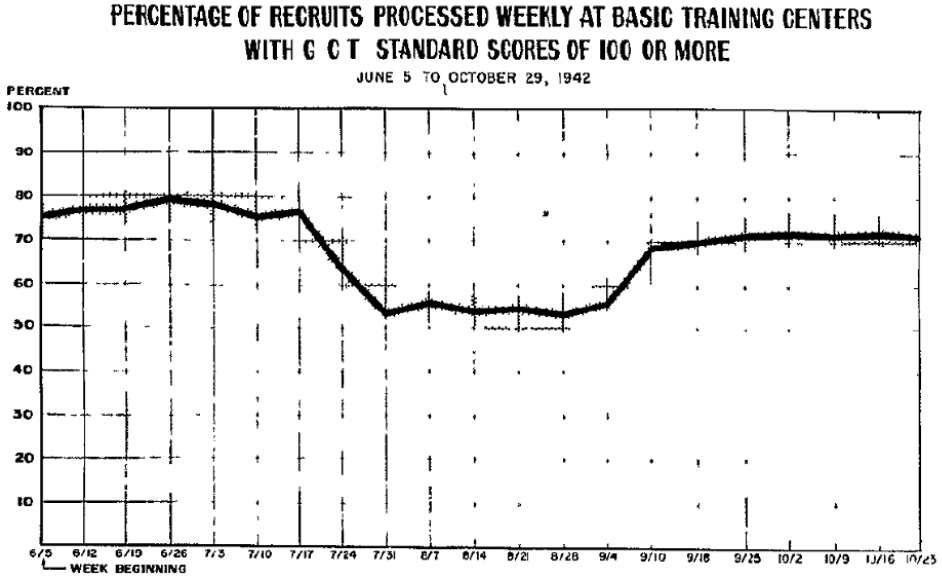
The AAF fought back against this change by using a second test, the mechanical aptitude (MA) test, as a screen for AAF recruits. At first, they simply requested that a higher proportion of men assigned scored above average on the MA test. This was later formalized. From December 1942 until June 1943, the AAF was supposed to be assigned 55 percent of their new recruits from the group with scores greater than the mean on both the AGCT and the MA tests. Combining the two tests was obviously more restrictive than just using one test. In fact only 37.5 percent of all recruits were above average on both tests[30]. This rule was allowed to expire, but it is clear that the AAF continued to get higher quality of recruits even after the rule formally expired. For example, for those inducted in 1943,

[28]US Air Force Historical Study #76, Classification and Assignment of Enlisted Men in the Air Arm 1917-1945, p. 44.

[29]ibid, p.46.

[30]ibid, p.56.

Figure 9: Percentage of Air Forces Recruits with above average AGCT scores



source: US Air Force Historical Study #76:
Classification and Assignment of Enlisted Men in the Air Arm, 1917-1945.

41.3% of soldiers assigned to AAF were class I or II. This percentage is higher than the one corresponding to Ground Combat Arms (29.7%) and Services (36.5%) (data come from Palmer et al. (1948)).

Figure 9 is a graph from the War Department showing the percentage of recruits assigned to the AAF with above average AGCT scores during the period of time that these rule changes were occurring. During the early part of the graph, the 75 percent rule was clearly in operation. At the end of July, the abolition of the 75 percent rule can be seen, with the AAF only getting 55 percent of recruits from the above average group. By September, the AAF had managed to return to the old proportions via the 55 percent mandate on both the AGCT and MA tests.

Additional evidence on the positive selection of men with high cognitive function into the AAF comes from an anomaly in the WWII enlistment data. As mentioned above, all new enlistees in the army took the AGCT test, but this test result was not generally recorded on the punch cards that are the primary source for the enlistment data. However, from March to

May 1943, AGCT scores were recorded in the fields marked "weight" for almost all recruits[31]. The fact that AGCT was coded in this field for some subset of the war is suggested in the documentation. Observing the actual distribution of values in the "weight" field confirms this is true for a subset of observations. Examining observations from enlistments through 1942, the weight field has a mean of 147 and a standard deviation of 28. For the period of March through May of 1943, the mean is roughly 106 and the standard deviation is about 80. Once we eliminate values which are below 20 and over 180, which clearly do not correspond to test scores (3.7% of observations), the mean becomes 98 and the standard deviation 21. These numbers are very close to the intended mean and standard deviation of the AGCT[32].

Figure 10 shows a histogram of AGCT scores for recruits entering the Air Forces alongside a histogram for all other recruits.[33] For the non Air Forces group there are 175,731 reported scores with a mean of 97.7 and a standard deviation of 21.1. This distribution matches the normal distribution of the AGCT (which was supposed to have a mean of 100 and standard deviation of 20). The second panel shows the histogram of recruits entering the AAF in this period. The mean (124.1) suggests that the AAF was, at this point, receiving substantially better recruits than the ground forces. However, as seen in Table 2, this was also a period in which very few recruits entered the Air Forces. This histogram is based on only 541 AAF recruits[34].

The preferential treatment of the AAF lasted until the end of 1943, when the "infantry crisis" broke out. The need for high-quality ground forces grew more acute in 1944 and

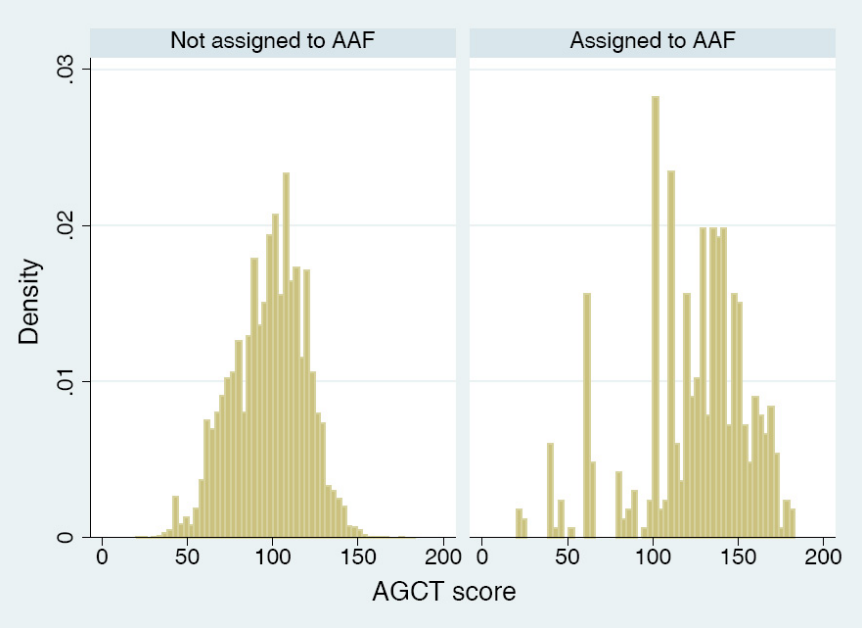
[31]Many thanks to Joe Ferrie for pointing this out to us.

[32]We must also note that correlation of the "weight" field with education is much stronger for this limited time period, at about 65%, compared with no correlation for the rest of the sample. This strong correlation, however, is only true for cohorts born up to 1925. It drops to 35% for the 1926 cohort, which only consists of 28 observations. There are no observations for the 1927 and 1928 cohorts.

[33]We only include values between 20 and 180, which are realistic test score values. The vast majority of observations (96.3% of the sample) are in this range. In addition, because the recording of AGCT scores in this field was for such a short period of time, there is some question as to whether all enlistment places coded this field the same way. For this reason we also drew the histograms using only data from enlistment places with over 500 recruits and where the mean of the weight field is between 80 and 120 within the enlistment place. This eliminated less than 11,000 observations and did not substantially change the distributions.

[34]Out of the 541 recruits who were assigned to the Air Forces over that period, none of them was born in 1926 or later, and only 37 individuals were born in 1925.

Figure 10: AGCT Scores for AAF versus all other enlistees



source: National Archives and Records Administration (2002)

lasted until the end of the War, while, at same time, air operations were not as important as in previous years. This meant that priorities between the AAF and the AGF shifted in favor of the latter, and the Army classification system was revised to allow for better-quality soldiers to join the Ground Forces[35]. With this classification procedure, as well as with transfers within the Army commands, the AGF had an influx of high-quality men, especially in the end of 1944 and afterwards, as opposed to the Air Forces. At the same time, the distribution of recruits among the Army commands changed, and most of the new inductees were assigned to the Ground Forces.

Through our enlistment records we know whether or not an individual was assigned to the AAF or the AGF. We know that individuals assigned to the AAF have higher test scores than the average enlistee for the initial phase of the war. We can also see the selection into

[35]The New classification procedure was based on "The Physical Profile System", which became operative in 1944, and classified recruits into three profiles, according to their ability to withstand strenuous combat conditions. 80% of men assigned to the AGF had to belong to the top profile, whereas only 10% of the AAF recruits came from the top group.

Table 3: Percent of high school graduates by birth year and branch.

	High School Graduation Rate		
Birth Year	Air Force	Ground Forces	Total
1920	76.1	42.9	47.6
1921	76.1	41.4	45.7
1922	70.3	39.8	44.6
1923	60.6	43.7	46.4
1924	52.2	39.5	40.8
1925	56.9	35.9	37.9
1926	51.0	36.8	39.1
1927	33.3	38.5	37.9
1928	26.4	40.7	37.2
Total	59.2	40.2	42.9

Source: National Archives and Records Administration (2002)

the AAF over the course of the war by looking at the proportion of high school graduates assigned to each branch. Table 3 gives the proportion of high school graduates in each branch by birth year. Figure 11 shows the proportion of high-school graduates by enlistment month, separately for the Air Forces and the rest of the Army.

Table 3 shows that the younger cohorts were less likely to have a high-school degree than the older cohorts. This is because, as the war progressed, the demand for manpower led to the drafting of ever younger men. It is also clear that there was positive selection into the Air Force, which became less pronounced over the course of the war. By the time the 1928 cohort enlisted, the Ground Forces were getting equally, if not better-qualified recruits than the Air Forces.

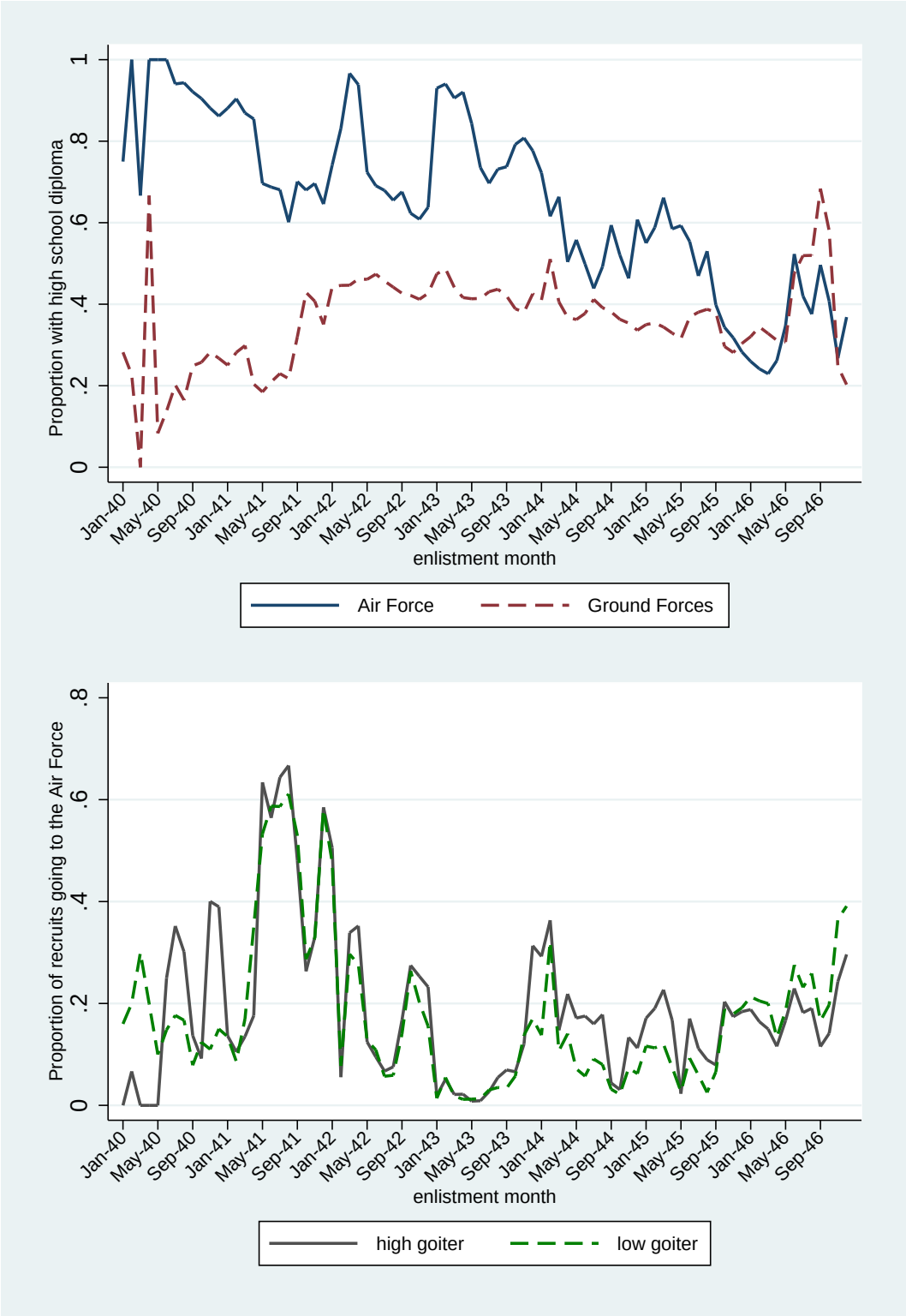
Figure 11 shows selectivity and demand for Air Force recruits over the course of the war. The lower figure plots demand for Air Force recruits, separately for high and low goiter sections. In particular, for each enlistment month we plot the percentage of total recruits from high and low goiter sections who were assigned to the Air Forces. Fluctuations in these curves are driven by war events that dictated specific Army needs. For example, the US

first engaged in air combat, driving up the demand for Air Force recruits in the beginning of the war. Right after ground operations got underway, in the end of 1942, there was higher demand for recruits in the Ground Forces, with a corresponding decrease in the numbers assigned to the Air Force for 1943. The upper plot in Figure 11 shows selectivity among Army branches, defined as the percentage of total recruits who had a high-school degree in the Air Force and the Ground Forces. There is a downward trend in the proportion of Air Force recruits who had high-school diplomas, which became steeper towards the end of 1943, when the infantry crisis broke out and priority was given to the Ground Forces. Figure 11 shows the spike in the proportion of high-school graduates in the Air Forces around February 1942, when the 75% rule was put into effect. Figure 11 also shows the temporary decrease in quality of AAF recruits in the second half of 1942, when the 75% rule was withdrawn. In late 1942 and early 1943, when the Mechanical Aptitude Test was put into use, the Air Force returned to the preferential-treatment status it enjoyed before. Positive selection into the Air Force ended after November 1945. After that, the Air Force no longer got more educated recruits (as measured by high school graduation rates) than the Ground Forces. About half of the 1927 and most of the 1928 birth cohorts enlisted after positive selection for the Air Force had ended.

It is clear that the Air Forces selected higher quality enlistees for the cohorts born from 1920 until 1926 (and, to some extent for those born in 1927). Luckily, this gives us good coverage of pre and post treatment enlistees. As we show below, the fact that there was *no* positive selection into the Air Force for the 1928 birth cohort is consistent with our empirical results.

We take advantage of selection into the Air Forces in two ways. First, if iodine deficiency affects cognitive ability we should expect a jump in the relative rate of assignment to the AAF for recruits born after 1924 in those counties where goiter rates were high in the *Defects* data. Second, by exploiting the normal distribution of the AGCT test, we can interpret our estimation results in terms of IQ gain for those regions where iodine deficiency was eradicated.

Figure 11: Air Force demand and selectivity in WWII



Source: National Archives and Records Administration (2002)

5 The rise in thyroid-related deaths

As discussed earlier, the treatment designed to help goiter sufferers sometimes ended up killing them. The doctors and public health officials who worked for the introduction of iodized salt were aware of the potential danger of iodine-induced hyperthyroidism, but viewed the danger as minimal. In Europe, the potential negative side-effects of iodine treatment had been discussed as early as the nineteenth century (see McClure (1934) and Carpenter (2005)) and made universal iodine prophylaxis a controversial public health measure. In this section, we document a large increase in deaths linked to the adoption of iodized salt, which we have not seen previously discussed in the literature.

Figure 12 shows the annual rate of deaths in the US over the period 1910-1960 due to exophthalmic goiter, which accounted for the overwhelming majority of deaths due to thyroid disease over this period[36]. There is an extremely large rise in the death rate at the time of iodization. Between 1921 and 1926 the death rate nearly doubled from 2.1 to 4 per 100,000. Deaths due to goiter remained elevated for at least a decade. There was also a large gender disparity in death rates. In 1926 death rates were over 6 times as high for women as for men[37,38]. The population of the United States in 1926 was 117 million, and so the rise of approximately two deaths per 100,000 people represented an extra 2,340 deaths in that year. Over the period 1925-1942 there appear to be at least 10,000 excess deaths that resulted from the introduction of iodized salt. We have found little discussion in the literature of what appears to be a high short-term price the country paid for long-run benefits resulting from this public health intervention[39].

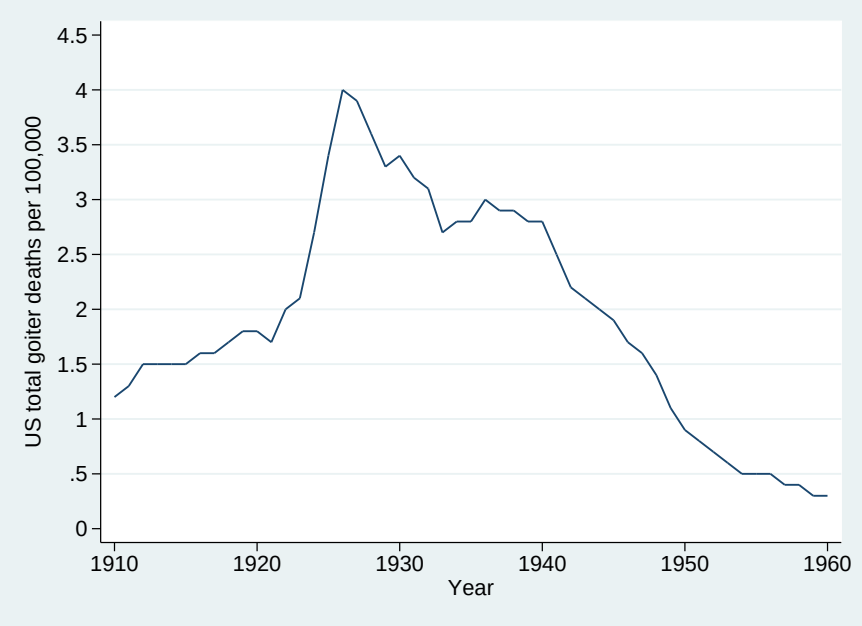
[36]Source of data: Grove and Hetzel (1968). Exophthalmic goiter is an enlargement of the thyroid accompanied by bulging of the eyes, which is sign of hyperthyroidism. While some medical dictionaries use the definition stated here, others define exophthalmic goiter as being synonymous with Graves disease, which is the dominant cause of the condition today. It is clear that the definition used in the text was being applied in the vital statistics data.

[37]1.1 per 100,000 for men, 7.0 per 100,000 for women.

[38]Field et al. (2009) finds differential effects between girls and boys in iodine deficient populations in Tanzania

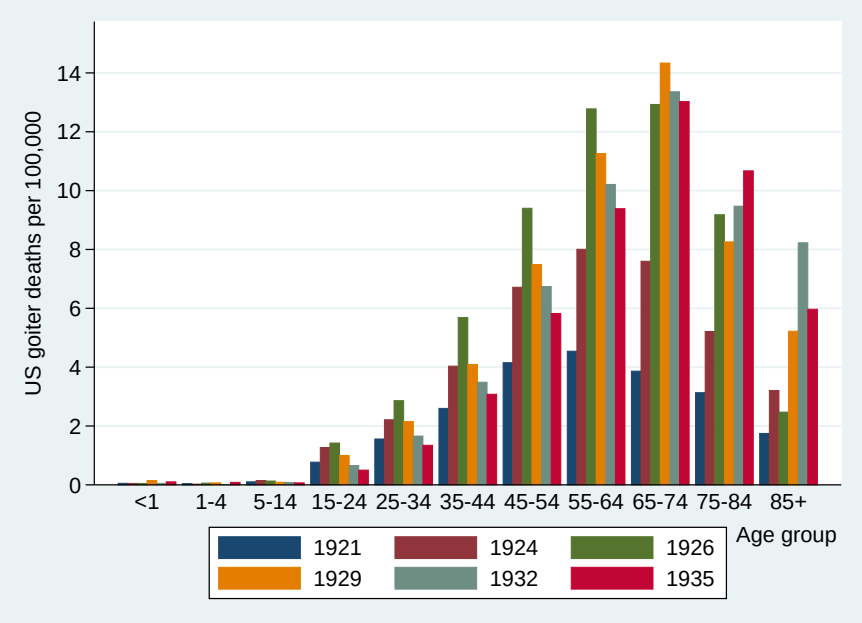
[39]An exception is McClure (1934) who notes that goiter deaths in Detroit spiked in the second year following iodization.

Figure 12: US deaths from exophthalmic goiter, 1910-1960



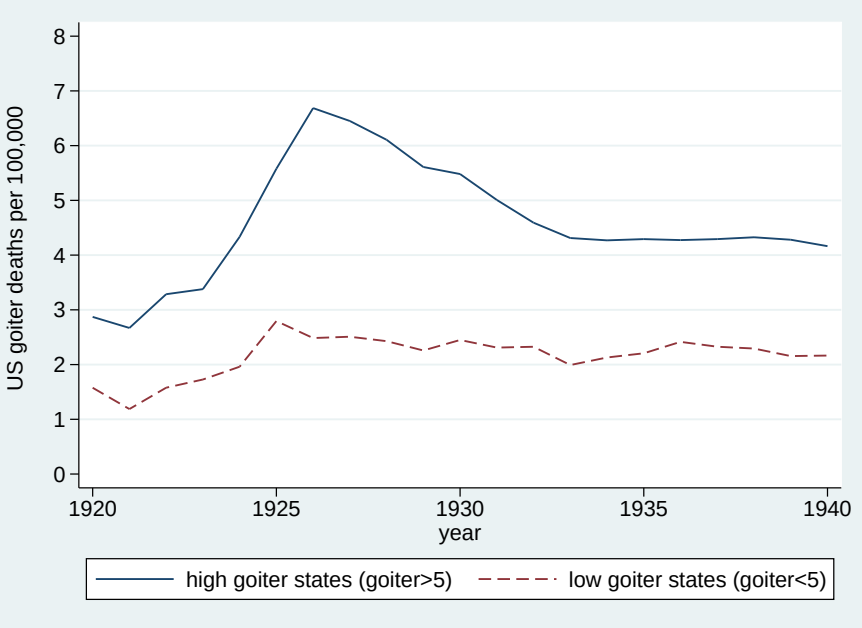
Source: Grove and Hetzel (1968)

Figure 13: US goiter deaths by age group, 1921-1935



Source: U.S. Department of Commerce (various years)

Figure 14: Deaths from exophthalmic goiter, by state group, 1920-1940



Source: U.S. Department of Commerce (various years)

Who was affected the most? The medical literature suggests that iodine-induced hyperthyroidism is most common among those with long-standing iodine deficiency. Consistent with the medical evidence, the rise in the death rate in the US was highest, and persisted the longest, among older age groups. As Figure 13 shows, deaths in the 25-34 age category less than doubled from 1921 to 1926, and had fallen below their 1921 level by 1935. In the 65-74 age category, deaths more than tripled between 1921 and 1926, and were still three times their 1921 level in 1935.

The link between iodine deficiency and the rise in deaths at the time of iodization is also apparent looking across states. Figure 14 plots goiter deaths over the period of salt iodization separately for high and low goiter states. We label a state as high-goiter if its goiter prevalence in *Defects* is higher than 5 cases per 1,000[40]. The trends between the two state groups are similar except for the period from 1925 to 1932. There is a sharp rise in

[40]High-goiter states are: Idaho, Oregon, Washington, Montana, Utah, Wyoming, Wisconsin, Michigan, North Dakota, Minnesota, West Virginia, Illinois, Iowa, Indiana, Nevada, Ohio, and Colorado.

deaths in high-goiter states in 1925 and 1926, which we do not observe for low-goiter states. As expected, large increases in mortality all took place in states that had high levels of goiter due to iodine deficiency. In addition, the rise in goiter deaths occurs sharply after 1924, which is when salt iodization began in Michigan. The evidence from mortality rates suggests, therefore, that iodization spread fast around the country after its introduction in Michigan in 1924.

6 Identification

Our identification strategy relies on two important sources of variation. The first is the pre-existing extent of underlying iodine deficiency. In regions where goiter was low among men drafted in WWI, we expect little change after the introduction of iodized salt. In regions where goiter was high we expect to see cognitive improvements after iodization, leading to a significant increase in the proportion of men being assigned to the Army Air Force. Our second source of variation arises from the timing of the intervention. The introduction and rapid spread of iodized salt provides for a clean distinction between the treated and the untreated. We expect to see a sharp difference between cohorts born before and after 1924. High goiter areas therefore provide the treatment group and low goiter areas the control group. Years before 1924 are pre-treatment and years after 1924 are post-treatment.

Our baseline regression specification is the following Linear Probability Model:

$$y_{iste} = \alpha + \sum_{t \neq 1924} \beta_t [goiter_s \cdot I(t = \text{birth year})_t] + controls_{ste} + \epsilon_{iste}$$

where y_{iste} is an indicator variable equal to 1 if an individual i born in year t in locality s and joining the Army at time t was assigned to the Air Force, and equal to 0 otherwise.

The goiter rate interacted with a set of birth year dummies provides the main coefficients of interest. We choose 1924 as the excluded category, so that the coefficients of interest correspond to percentage changes relative to the 1924 cohort. Iodized salt first became available in May 1924, and take-up, while rapid, was not instantaneous. Field et al. (2009)

provide evidence that the first trimester of gestation is the most relevant period for the cognitive impact of iodine deficiency. Thus the vast majority of those born in 1924 would not have been treated at all. The pattern of coefficients on the interaction of goiter and birth year dummies will show how the relationship between the geologically determined level of iodine deficiency (as measured by goiter rates in WWI enlistees) and cognitive ability (as measured by the likelihood of entering the Air Force) changes over time. The iodization of salt in 1924 should make these coefficients significantly larger (or less negative) for the later cohorts. Note, however, that we do not impose any particular pattern in the coefficients of interest; we let the data tell the story of whether and how the probability of enlisting in the Air Force changed for cohorts born in affected areas, and we check whether that matches the introduction of iodized salt.

In all the regressions we include a number of additional controls. We include a full set of *Defects* section fixed effects, as well as a set of state-specific trends. The section fixed effects control for pre-existing, time-invariant characteristics of sections which might be correlated both with the probability of enlisting in the Air Force and the section goiter rate in WWI. The state-specific trends control for gradual changes at the state level, which might affect each cohort's probability of enlisting in the Air Force. We also include a full set of birth year fixed effects interacted with enlistment year indicator variables, to control for the changing demands of the war and the need for Air Force recruits. The interaction of birth year and enlistment year fixed effects will also control for any systematic relationship between birth cohort, enlistment period and the propensity to enter the Air Force which is shared across all members of the cohort. This implicitly provides a control for enlistment age. In another specification, we include full sets of both birth year fixed effects and enlistment month fixed effects. The proportion of enlistees being assigned to the Air Force varies dramatically over the course of the war, so such controls are necessary.

In our final sample we include all white males born between 1920 and 1928, who joined the Army between January 1940 and December 1946, who reside in their state of birth, and

who were between 17 and 26 years old when they enlisted[41]. Since our identification relies on variation within a given birth - enlistment year cohort, we only use birth year - enlistment year cells with more than 100 observations[42]. Finally, because our theory relies on the Air Force getting preferential treatment, for most of the analysis we focus on the enlistment period up to November 1945, which is when positive selection for the Air Force comes to an end.

We estimate a Linear Probability Model of assignment in the Air Force. We use two-way clustering for standard errors, at the *Defects* section level and the birth year level. We also estimate a Logit Model, which gives similar results[43]. To corroborate our findings, we run a few robustness checks: first, we check that our estimates are not confounded by changes in demand for Air Force recruits. The enlistment year - birth year interactions should control for changing demand conditions, but we also show that our results hold for different parts of the distribution of demand. Second, we use an alternative measure of underlying iodine deficiency, the iodine content of drinking water, which is available for a few locations. The estimates from this check are consistent with the results using goiter as our measure of iodine deficiency. Finally, we do a falsification exercise, using all the other diseases that were measured at the section level in *Defects*, to show that our pattern of coefficients is not generally replicated with diseases that were not affected by iodization[44,45].

[41]We drop recruits from Alaska, because there was no variation in their enlistment branch (only one individual is registered in the Air Force throughout our enlistment period).

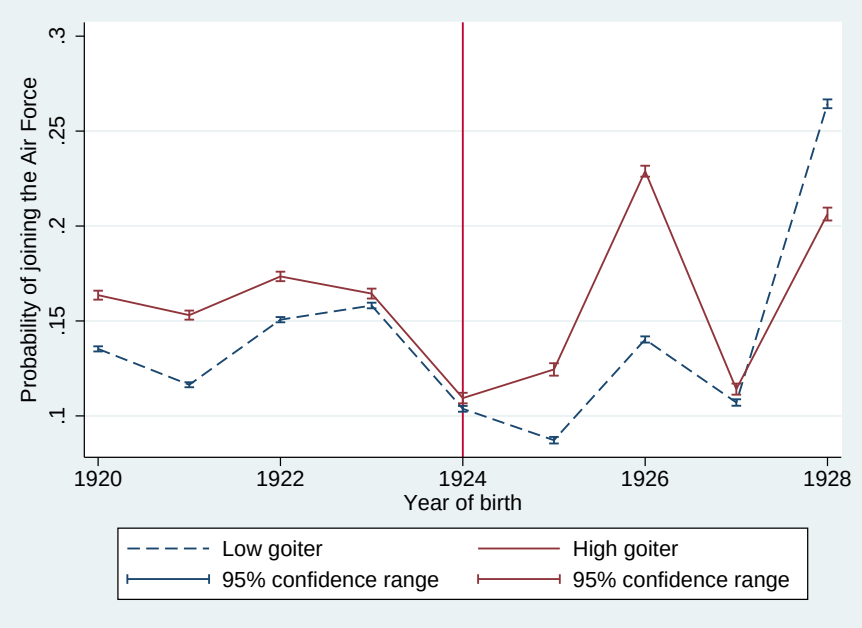
[42]This means that enlistment for the 1923 birth cohort starts in January 1941, for the 1924 cohort it starts in January 1942, and for the 1927 cohort it starts in January 1945.

[43]The Logit results are available upon request.

[44]We have also checked whether our estimates are robust to aggregating everything at the *state* of birth, rather than the *section* of birth, and get similar results.

[45]We ran similar specifications with the same datasets using education as an outcome variable, but found no effect of iodization on education. We also tried using Census Data to identify an effect of iodization on education, using the goiter level at the *state* of birth, rather than the *section* of birth. We have not found any effect of iodization on education levels. While surprising, this finding is consistent with the fact that in our data we find a very weak correlation between AGCT scores and education. Specifically, the correlation between AGCT scores and graduation from high school is 14%, while the correlation between AGCT scores and having one year of college or more is just 9%.

Figure 15: Probability of joining the Air Forces



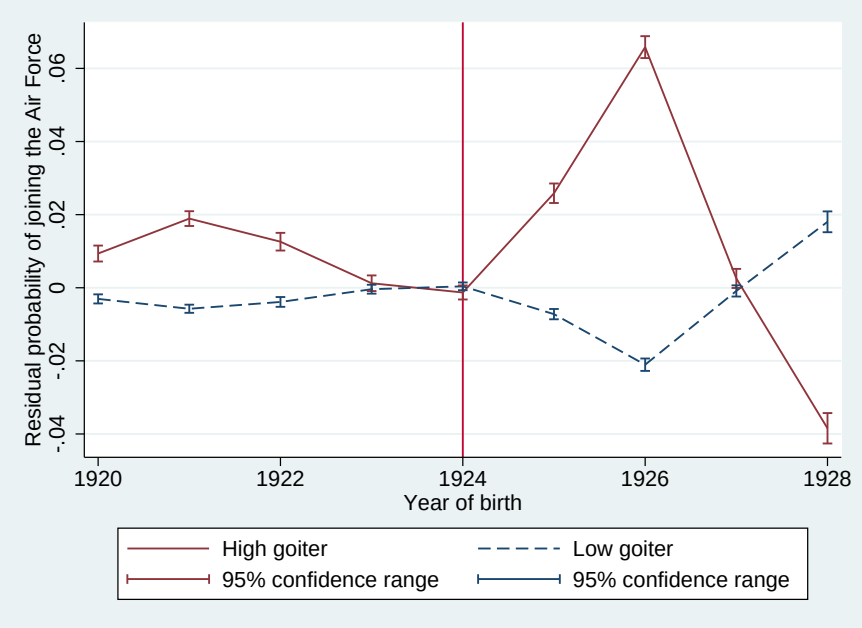
source: National Archives and Records Administration (2002) and Love and Davenport (1920)

7 Results

Figure 15 is a first graphical preview of our results. We plot the probability of joining the Air Force for each cohort of recruits, by high-goiter and low-goiter group, according to the goiter level in their section of birth. The 95% confidence band around the means is very small, because our sample is big. The high goiter group contains *Defects* sections which are at the top 25% of the distribution with a cutoff of 5.4 goiter cases per 1000. Figure 15 shows a jump in the probability of joining the Air Forces for the 1925 cohort coming from a high-goiter area, relative to the same cohort coming from a low-goiter area. The jump is even more pronounced for the 1926 cohort. Note that half of the 1927 cohort and all of the 1928 cohort only enlisted after the Air Force stopped receiving preferential treatment among Army commands.

Figure 16 is a regression-adjusted graphical presentation of our results. Figure 16 plots

Figure 16: Residual probability of joining the Air Forces



source: National Archives and Records Administration (2002) and Love and Davenport (1920)

the average residuals for each cohort - goiter group combination, after regressing an Air Force indicator variable on enlistment month fixed effects for each birth cohort separately. The 95% confidence band is, again, very narrow, since each point in the figure is an average, and our sample size is big. Figure 16 accounts for changing demand conditions over the course of the war, and it shows that those cohorts born after iodization in a previously high-goiter area, had a sudden increase in their probability of joining the Air Force (as long as they enlisted when there was positive selection into the Air Force), as opposed to those coming from low-goiter areas. Figure 16 also shows evidence of *negative* selection into the Air Force for the 1928 cohort, who enlisted during the final phase of the war.

Table 4 displays estimates from a linear probability model of assignment to the Air Force. The dependent variable is an indicator switching on if the individual entered the Army Air Force (AAF). The regressors of interest are the interactions of birth year dummy variables with a continuous measure of goiter in the section where the recruit was born. The excluded

Table 4: Baseline results: Probability of enlisting in the Air Force

	(1) All enlistment months	(2) All enlistment months	(3) Positive selection months	(4) Positive selection months
Goiter X Born in 1920	-0.318	-0.312	-0.0229	0.00668
	(0.225)	(0.233)	(0.184)	(0.173)
Goiter X Born in 1921	-0.156	-0.106	0.0491	0.113
	(0.156)	(0.162)	(0.137)	(0.124)
Goiter X Born in 1922	-0.134	-0.122	-0.0352	-0.0187
	(0.0972)	(0.128)	(0.0707)	(0.0641)
Goiter X Born in 1923	-0.167***	-0.197***	-0.135**	-0.169*
	(0.0520)	(0.0685)	(0.0533)	(0.0934)
Goiter X Born in 1925	0.509***	0.468***	0.516***	0.494***
	(0.168)	(0.127)	(0.158)	(0.124)
Goiter X Born in 1926	0.819**	0.771**	0.726***	0.693***
	(0.323)	(0.324)	(0.259)	(0.243)
Goiter X Born in 1927	0.462**	0.394**	0.491**	0.441**
	(0.233)	(0.198)	(0.209)	(0.181)
Goiter X Born in 1928	-0.0501	-0.152	-0.370	-0.518
	(0.415)	(0.440)	(0.399)	(0.337)
Constant	149.9***	-146.1	77.03***	107.2
	(10.80)	(2,437)	(16.81)	(2,516)
Observations	2,274,529	2,274,529	1,935,444	1,935,444
R-squared	0.131	0.134	0.153	0.167
Birth X Enlist. year FE	YES	NO	YES	NO
Section FE	YES	YES	YES	YES
State trends	YES	YES	YES	YES
Birth year FE	NO	YES	NO	YES
Enlist. month FE	NO	YES	NO	YES

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; Coefficients correspond to changes in percentage points. Disturbances are clustered at the section and birth year levels. Columns (1) and (2) include enlistment months January 1940 to December 1946. Columns (3) and (4) include enlistment months January 1940 to November 1945.

Table 5: Probability of enlisting in the Air Force for high-goiter sections

	(1) top 20th percentile	(2) top 25th percentile	(3) top 30th percentile
High goiter X Born in 1920	1.602	-0.0967	0.905
	(2.249)	(2.787)	(2.781)
High goiter X Born in 1921	2.249	1.432	2.784
	(1.680)	(2.066)	(2.076)
High goiter X Born in 1922	0.0927	0.185	0.467
	(1.268)	(1.606)	(1.619)
High goiter X Born in 1923	-0.930*	-1.462**	-0.908
	(0.496)	(0.680)	(0.661)
High goiter X Born in 1925	3.796***	5.580***	5.267***
	(1.223)	(1.802)	(1.582)
High goiter X Born in 1926	6.886***	10.00***	8.581***
	(1.528)	(2.421)	(2.093)
High goiter X Born in 1927	4.421***	6.698***	3.507*
	(1.654)	(2.151)	(2.071)
High goiter X Born in 1928	-4.181	-7.360	-7.568
	(2.904)	(5.431)	(5.276)
Constant	59.24***	67.71***	54.88***
	(12.35)	(14.97)	(15.95)
Observations	1,935,444	1,935,444	1,935,444
R-squared	0.153	0.154	0.154
Birth X Enlist. year FE	YES	YES	YES
Section FE	YES	YES	YES
State trends	YES	YES	YES

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; Coefficients correspond to changes in percentage points in the probability of enlisting in the Air Force for enlistment months January 1940 to November 1945. Disturbances are clustered at the section and birth year levels. Column (1) treats the top 20% of the population-weighted goiter distribution as high goiter. Column (2) treats the top 25% of the population-weighted goiter distribution as high goiter. Column (3) treats the top 30% of the population-weighted goiter distribution as high goiter.

year is 1924, which is when iodized salt was introduced. The coefficients correspond to changes in percentage points in the probability of joining the Air Force. Columns (1) and (2) of table 4 use all enlistment months in the estimation. In columns (3) and (4) we focus on the positive selection period, which ends in November 1945. Our theory relies on the Air Force getting preferential treatment in the cognitive ability of its enlistees, so we focus on this enlistment period in all other specifications. We include section fixed effects and state-specific trends in our regressions, and we cluster standard errors both at the section and birth year levels. Columns (1) and (3) include interactions of birth year and enlistment year fixed effects. Columns (2) and (4) include birth year fixed effects and enlistment month fixed effects as controls.

The pattern of coefficients suggests a clear break before and after 1924, which is consistent with the timing of iodization. Cohorts born after 1924 in high goiter sections are more likely to join the Air Force compared to other recruits. The 1923 cohort, which was the last cohort born before iodization in high goiter areas, was less likely to be assigned to the Air Force compared to the cohorts that would have enlisted at the same time. The positive effect disappears for the 1928 cohort. This is consistent with the fact that the 1928 cohort enlisted when positive selection into the Air Force was largely over, as can be seen in Figure 11. It also reflects negative selection into joining the Army for this cohort, since individuals born in 1928 who are included in our sample would have been only 17 or 18 year old at the time of enlistment.

The effects that we estimate are large in high goiter areas. The coefficients for 1925, 1926, and 1927 are between 0.4 and 0.8 percentage points. The highest goiter areas in our sample have levels of roughly 30 per 1000 cases. Multiplying the coefficients by 30 indicates that the *highest* goiter areas saw a 12-24 percentage point increase in the likelihood of joining the AAF if they were born after 1924.

In Table 5 we run a specification similar to column (3) of Table 4, but instead of using a continuous measure of goiter, we separate the sample into high and low goiter sections,

and include interactions of birth year dummy variables with a high goiter indicator as our regressors of interest. We define high goiter to correspond to the top 20% (column (1)), the top 25% (column (2)) and the top 30% (column (3)) of the population-weighted goiter distribution[46].

The results of Table 5 echo the earlier results, but they are easier to interpret. Depending on how we define "high goiter" and which cohort we are looking at, individuals born post-iodization in high goiter areas see a 3.8-10 percentage point increase in the probability that they enter the Air Forces compared to earlier cohorts. Given that the average rate of assignment to the Air Force is roughly 14% for the entire sample, this represents a large effect.

In summary, the pattern of coefficients from the linear probability model suggests a big, sudden increase in the probability of assignment to the Air Force for later cohorts born in high goiter areas. The timing of the change corresponds to the introduction of iodized salt, and is consistent with cognitive gains in utero for the treated populations.

In the following section we present three robustness checks which corroborate our baseline results.

8 Robustness checks

8.1 Robustness check: Changes in demand for Air Force recruits

In our main specification, the birth and enlistment cohort control variables account for changes in the overall demand for Air Force recruits over time, which, as shown in Figure 11, were dictated by war circumstances. In this section, we check directly if our main results are affected by overall demand for manpower in the Air Force, by cutting the sample across different demand periods.

[46]The high goiter cutoff for each percentile of the distribution is 6.65 goiter cases per 1000 for the top 20th percentile, 5.4 goiter cases per 1000 for the top 25th percentile, and 5.21 for the top 30th percentile.

We define demand as the proportion of total recruits enlisting in a particular month who were assigned to the Air Force. Table 6 shows the results of this exercise. We run a linear probability model similar to Column (3) of Table 4. Column (1) of Table 6 looks at enlistment months when demand was above the median, whereas column (2) looks at enlistment months when demand was below the median for the entire sample[47]. Columns (3) and (4) omit periods of very high and very low demand. In particular, column (3) keeps enlistment months within the 10th-90th percentiles of the distribution of demand, whereas column (4) keeps enlistment months within a smaller range, corresponding to the 25th-75th percentiles[48].

As Table 6 shows, our main results are largely unaffected when we cut the sample across different parts of the distribution of demand for Air Force recruits. The estimates for post-iodization cohorts in column (2), which looks at periods of below-median demand, are smaller than the estimates of column (1), which focuses on periods of above-median demand, but the pattern of coefficients is very similar to our baseline specifications. We lose statistical significance for the post-1925 cohorts, which is partly a reflection of lower point estimates. We also note that keeping only those enlistment months that fall below the median means that the sample size for post-1925 cohorts decreases. So the loss in statistical significance is also due to decreased sample size and corresponding loss in statistical power. The smaller point estimates in column (2) are consistent with lower (but still active) preferential treatment of the Air Force when the demand for Ground Forces was bigger. The estimates in columns (3) and (4) are consistent with our earlier results, showing that our results are not driven by periods of exceptionally high (or low) demand for the Air Force. We note that the fact that coefficients for the 1927 cohort are not always robust to changes in demand is not surprising, given that only half of the 1927 cohort enlisted when there was positive selection into the Air Force in the first place.

[47]The median demand for Air Force recruits across the sample is 10.6%.

[48]The 10th percentile of the distribution of demand is 1.5%. The 90th percentile is 22.6%, the 25th percentile is 5.4% and the 75th percentile is 17.8%.

Table 6: Robustness check: Probability of enlisting in the Air Force in different demand periods

	(1) Above median demand for AAF	(2) Below median demand for AAF	(3) 10th-90th percentile demand for AAF	(4) 25th-75th percentile demand for AAF
Goiter X Born in 1920	-0.0290	0.284	-0.0140	-0.0999
	(0.227)	(0.269)	(0.182)	(0.140)
Goiter X Born in 1921	0.170	0.237	0.0227	-0.102
	(0.173)	(0.199)	(0.123)	(0.0768)
Goiter X Born in 1922	0.00252	0.161	1.33e-05	-0.186***
	(0.108)	(0.107)	(0.0551)	(0.0653)
Goiter X Born in 1923	-0.0987	-0.0339	-0.0928*	-0.0903
	(0.0831)	(0.0701)	(0.0527)	(0.0661)
Goiter X Born in 1925	0.841***	0.221***	0.466***	0.385***
	(0.273)	(0.0731)	(0.161)	(0.130)
Goiter X Born in 1926	0.935***	0.472	0.628**	0.458*
	(0.344)	(0.327)	(0.266)	(0.241)
Goiter X Born in 1927	0.736**	0.116	0.392**	0.141
	(0.308)	(0.198)	(0.176)	(0.161)
Goiter X Born in 1928	0.0709	-0.566	-0.613	-0.855*
	(0.431)	(0.423)	(0.426)	(0.504)
Constant	176.1***	91.02	16.60***	-11.89
	(25.65)	(21,243)	(5.164)	(9.083)
Observations	990,063	945,381	1,429,928	902,809
R-squared	0.135	0.091	0.139	0.126
Birth X Enlist. year FE	YES	YES	YES	YES
Section FE	YES	YES	YES	YES
State trends	YES	YES	YES	YES

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; Coefficients correspond to percentage point changes in the probability of enlisting in the Air Force. Disturbances are clustered at the section and birth year levels. Demand for Air Force recruits is defined as the proportion of all recruits who enlist in the Air Force in a given enlistment month, for the enlistment period January 1940 to November 1945. Column (1) (column (2)) includes enlistment months above (below) the median of the population-weighted distribution of demand for Air Force recruits. Column (3) (column (4)) includes enlistment months within the 10th-90th (25th-75th) percentiles of the population-weighted distribution of demand for Air Force recruits.

8.2 Robustness check: Iodine content of drinking water

The statistics compiled during the WWI draft and summarized in *Defects* provide an impressive, rare snapshot of the health status of young American adults in early 20th century. The data are comprehensive, detailed, consistently and systematically collected, and they cover all US states. For the purposes of our study, we are lucky that goiter was one of the recorded diseases. To our knowledge, there is no higher quality dataset on iodine deficiency across the US prior to the introduction of iodized salt.

That said, goiter is not the only measure of iodine deficiency. In this section we use the iodine content of drinking water in a given location as another measure of underlying iodine deficiency. Our data come from McClendon (1924), and they were collected from lakes, springs, rivers and wells across 67 localities in the US (see discussion in section 4.1). These data are not as comprehensive as the goiter data, and they suffer from measurement issues, since the iodine content of a single source of drinking water might not be representative of a whole county's access to iodine. They can serve, however, as a robustness check of our main results. Figure 8 shows that, as expected, the iodine content data correlate negatively with the goiter data from *Defects*.

We match the localities for which we have iodine content data to the county in which they belong. Some counties had more than one measures of iodine content. In those cases we calculate the average for the county. We end up with data on the iodine content of drinking water in 59 counties across 30 states. We then merge the iodine content data with our WWII enlistment data, and run a linear probability model similar to the one in column (3) of Table 4, using interactions of birth year dummy variables with the logarithm of iodine content in a given county. Our sample size is much smaller since we only have data from 59 counties; we lose almost 90% of observations. Instead of section fixed effects, we include county fixed effects. Similarly to column (3) of Table 4, we include birth year fixed effects interacted with enlistment year fixed effects, and we cluster standard errors both at the county and the birth year levels. We include state-specific trends, but they are only identified for the 14 states

Table 7: Robustness check: Probability of enlisting in the Air Force using the log of iodine content of drinking water in 59 counties

	(1) Air Force	(2) Air Force
log iodine X Born in 1920	-0.964	0.581
	(0.831)	(0.538)
log iodine X Born in 1921	-0.771	0.178
	(0.718)	(0.574)
log iodine X Born in 1922	-0.398	-0.0141
	(0.585)	(0.576)
log iodine X Born in 1923	-0.378	-0.315
	(0.234)	(0.251)
log iodine X Born in 1925	-1.454	-1.373*
	(0.978)	(0.725)
log iodine X Born in 1926	-1.830	-2.720
	(1.426)	(2.193)
log iodine X Born in 1927	-1.048*	-1.776
	(0.563)	(1.276)
log iodine X Born in 1928	3.342***	1.206
	(1.239)	(0.875)
Constant	-95.50***	41.12***
	(24.50)	(5.993)
Observations	224,022	224,022
R-squared	0.175	0.158
Birth X Enlist. year FE	YES	YES
County FE	YES	YES
State trends	YES	NO

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; Coefficients correspond to percentage point changes in the probability of enlisting in the Air Force during the enlistment period January 1940 to November 1945. Disturbances are clustered at the county and birth year levels.

with multiple iodine content data. In addition, only 7 of those states have more than two county observations. This is why we also estimate the model without state trends.

Table 7 presents estimation results. Column (1) includes state-specific trends, whereas column (2) omits them. We generally do not have enough statistical power to get statistically significant results (except for the 1925 cohort in column (2), but the pattern of coefficients reflects our earlier results. Cohorts born in high (low) iodine counties after the introduction of iodized salt have a lower (higher) probability of enlisting in the Air Force. The relationship breaks down for the 1928 cohort, similarly to our baseline results, reflecting the end of preferential treatment for the Air Force, and negative selection of the youngest cohort into joining the Army. The log of iodine ranges from 0 to 9.8 in our data, so going from the lowest to the highest iodine increases the probability of enlisting in the Air Force by 12.8 percentage points (using the point estimate for the 1925 cohort in column (2)). This estimate is very similar to the lower bound suggested by our baseline results.

8.3 Robustness check: Falsification exercise with other diseases

One might worry that a high prevalence of goiter in a region reflected general ill health, or perhaps a general low level of some particular dimension of health. Our narrative of the tight link between goiter prevalence and low iodine levels in local food and water makes us doubt such a possibility, but we proceed nonetheless. Thus, our third robustness check is to see whether a pattern of coefficients similar to those observed for goiter obtains for other medical conditions. Given that there are 271 different conditions enumerated in our data, it would not be a surprise to find several of them having interesting, statistically significant patterns of coefficients in our exercise. This being said, however, it is important to remember that our analysis of the effect of goiter prevalence on adult outcomes, and the use of 1924 as the cutoff year for pre- and post-treatment groups, was not the result of a search through the data for significant patterns. Quite the contrary, the underlying medical literature, and our a priori expectations of the effect of the policy intervention of iodization, led us to focus

on this particular specification. By contrast, we do not have a theory regarding the pattern of coefficients for other diseases.

We proceed along two paths, looking first at a set of particular health conditions, and then at some measures of broader health. Examining all 271 conditions that appear in *Defects* seems unreasonable, especially since most of them were extremely rare. Simple goiter, the condition on which we focus, was the 26th most prevalent condition, affecting 4.35 men per thousand recruits. Using this as a natural cutoff, we compare our results for goiter with the 25 most prevalent conditions. The results are shown in Tables 8, 9, and 10. For each condition, the top row of the table shows its prevalence in the *Defects* data. We run the same regression as in Column 3 of Table 4, simply substituting the prevalence of a different condition, at the section level, for that of goiter. For comparison, we also show the specification using goiter prevalence in the leftmost column of each page of the table. Conditions include pes planus (flat feet), which is the most common, with a prevalence of 109.4 per thousand; several general health descriptions, such as underweight, defective and deficient teeth, mental deficiency, and general unfitness for military; two venereal diseases; various injuries and physical deformities; and several specific diseases such as tuberculosis.

The coefficients in Tables 8, 9, and 10 do not suggest to us that other diseases in the *Defects* data predict outcomes in the World War II data in the same manner that goiter does. There are certainly significant coefficients on the interactions of different diseases with different birth years, but they generally do not follow a pattern as in the case of goiter. Partial exceptions to this rule are the venereal diseases syphilis and gonorrhea (gonococcus infection). Both have negative and significant coefficients for the birth years 1925 and 1926, but both also have negative and significant coefficients for years before iodization in 1924 (specifically 1922 for syphilis and 1921 for gonococcus infection). The same is roughly true with general unfitness for military service and, with the sign reversed, loss of fingers and flat feet.

Table 8: Falsification exercise with the most common diseases in *Defects* (part 1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Simple goiter	Pes planus	Defective vision	Under-weight	Gonococcus infection	Hyper-trophic tonsillitis	Hernia	Pulmonary tuber-culosis	Enlarged inguinal rings	Unclassified valvular lesions
Prevalence (per 1,000)	4.35	109.35	28.3	26.5	24.59	23.09	20.83	20.2	18.99	16.27
disease X Born 1920	-0.0229	-0.00483	-0.0161	-0.0375	-0.0629	-0.0440	0.385***	-0.0415	0.0243	0.0276
	(0.184)	(0.0360)	(0.0635)	(0.0431)	(0.0573)	(0.0645)	(0.117)	(0.0662)	(0.0325)	(0.0894)
disease X Born 1921	0.0491	0.0187	-0.00529	-0.0545	-0.0824**	-0.0657	0.242	-0.0423	0.0253	0.0462
	(0.137)	(0.0262)	(0.0478)	(0.0365)	(0.0417)	(0.0564)	(0.150)	(0.0560)	(0.0312)	(0.0729)
disease X Born 1922	-0.0352	0.0379**	0.00465	-0.0563	-0.0332	-0.0958**	0.328***	-0.0168	0.0368	0.0354
	(0.0707)	(0.0188)	(0.0351)	(0.0373)	(0.0346)	(0.0395)	(0.111)	(0.0483)	(0.0281)	(0.0619)
disease X Born 1923	-0.135***	-0.00748	0.0372*	-0.00327	-0.00476	0.0248*	-0.0352	-0.0296	0.0189	0.0213
	(0.0533)	(0.00774)	(0.0201)	(0.0145)	(0.0143)	(0.0144)	(0.0347)	(0.0209)	(0.0199)	(0.0296))
disease X Born 1925	0.516***	0.0809***	0.0183	-0.0647**	-0.100***	-0.0135	0.155	-0.0765	0.127***	0.0138
	(0.158)	(0.0288)	(0.0490)	(0.0317)	(0.0341)	(0.0635)	(0.159)	(0.0546)	(0.0387)	(0.0856)
disease X Born 1926	0.726***	0.122***	0.0659	-0.0579	-0.147***	-0.201***	0.0275	-0.117	0.0805	-0.0387
	(0.259)	(0.0388)	(0.0747)	(0.0666)	(0.0503)	(0.0746)	(0.197)	(0.0868)	(0.0582)	(0.115)
disease X Born 1927	0.491**	0.0446	-0.115*	-0.105**	-0.0335	-0.299***	0.102	-0.0563	-0.111**	-0.199**
	(0.209)	(0.0318)	(0.0604)	(0.0420)	(0.0380)	(0.0695)	(0.195)	(0.0394)	(0.0560)	(0.0792)
disease X Born 1928	-0.370	-0.230**	-0.435***	0.116	0.336***	-0.0886	-0.268	0.0730	-0.346***	-0.209
	(0.399)	(0.0905)	(0.134)	(0.112)	(0.103)	(0.100)	(0.167)	(0.0942)	(0.106)	(0.134)
Constant	77.03***	34.20	23.61**	64.46***	78.80***	44.92***	3.519	60.42***	33.20***	46.64***
	(16.81)	(43.14)	(11.76)	(8.194)	(10.11)	(10.78)	(10.50)	(8.872)	(7.615)	(10.28)
Observations	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444
R-squared	0.153	0.154	0.153	0.153	0.153	0.153	0.153	0.153	0.153	0.153
Birth X Enlist. year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Section FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
State trends	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; Coefficients correspond to percentage point changes in the probability of enlisting in the Air Force during the enlistment period January 1940 to November 1945. Disturbances are clustered at the section and birth year levels.

Table 9: Falsification exercise with the most common diseases in *Defects* (part 2)

	(1) Simple goiter	(2) Mental deficiency	(3) Defective and deficient teeth	(4) Mitral insufficiency	(5) Otitis media	(6) Syphilis	(7) Pronated foot	(8) Suspected tuberculosis	(9) Curvature of spine
Prevalence (per 1,000)	4.35	14.45	13.48	10.75	7.84	6.94	6.31	6.22	5.53
disease X Born 1920	-0.0229	0.0899	0.0382	0.0950	-0.0693	-0.0547	-0.0966	0.333**	0.860*
	(0.184)	(0.0738)	(0.0710)	(0.130)	(0.138)	(0.128)	(0.101)	(0.151)	(0.500)
disease X Born 1921	0.0491	-0.00244	-0.00744	0.191	0.0576	-0.157	-0.0958	0.167	0.712*
	(0.137)	(0.0874)	(0.0567)	(0.127)	(0.118)	(0.0962)	(0.0824)	(0.114)	(0.398)
disease X Born 1922	-0.0352	-0.214***	-0.0310	0.172**	0.111**	-0.177***	-0.179***	-0.108	-0.218
	(0.0707)	(0.0745)	(0.0518)	(0.0785)	(0.0529)	(0.0667)	(0.0682)	(0.130)	(0.326)
disease X Born 1923	-0.135***	-0.0468	0.0309	-0.0411	0.114**	-0.0153	0.0334	-0.135**	-0.0743
	(0.0533)	(0.0308)	(0.0191)	(0.0394)	(0.0490)	(0.0332)	(0.0268)	(0.0665)	(0.0827)
disease X Born 1925	0.516***	-0.125	0.00848	0.368***	0.0505	-0.203**	0.0955	-0.226**	0.969
	(0.158)	(0.0850)	(0.0426)	(0.114)	(0.145)	(0.0813)	(0.0772)	(0.109)	(0.603)
disease X Born 1926	0.726***	-0.213*	0.0323	0.257	0.108	-0.389***	-0.0519	-0.130	1.619**
	(0.259)	(0.124)	(0.0650)	(0.164)	(0.269)	(0.124)	(0.0861)	(0.207)	(0.780)
disease X Born 1927	0.491**	0.0150	-0.174***	0.144	-0.269*	-0.0715	-0.402***	0.105	-0.305
	(0.209)	(0.0957)	(0.0501)	(0.179)	(0.137)	(0.0975)	(0.0763)	(0.203)	(0.522)
disease X Born 1928	-0.370	0.329**	-0.293**	-0.825**	-1.300***	0.679**	-0.162	1.208***	-3.636***
	(0.399)	(0.166)	(0.123)	(0.323)	(0.314)	(0.284)	(0.160)	(0.380)	(1.382)
Constant	77.03***	64.63***	41.13***	28.14***	29.98***	64.47***	51.69***	67.52***	-18.43
	(16.81)	(11.15)	(8.095)	(9.454)	(9.376)	(7.913)	(8.201)	(9.865)	(26.61)
Observations	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444
R-squared	0.153	0.153	0.153	0.153	0.153	0.153	0.153	0.153	0.153
Birth X Enlist. year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Section FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
State trends	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; Coefficients correspond to percentage point changes in the probability of enlisting in the Air Force during the enlistment period January 1940 to November 1945. Disturbances are clustered at the section and birth year levels.

Table 10: Falsification exercise with the most common diseases in *Defects* (part 3)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Simple goiter	Defective hearing	Blindness in one eye	Epilepsy	Loss of fingers	Deformity of lower extremity	Hallux valgus	General unfitness for military	Other diseases of bones & organs of locomotion
Prevalence (per 1,000)	4.35	5.47	5.4	5.15	5.05	4.71	4.65	4.65	4.54
disease X Born 1920	-0.0229	0.0370	-0.404	0.107	1.303**	0.246	-0.0394	0.0378	0.138
	(0.184)	(0.194)	(0.484)	(0.365)	(0.628)	(0.365)	(0.187)	(0.249)	(0.376)
disease X Born 1921	0.0491	-7.39e-05	-0.359	0.0731	1.143**	0.0514	-0.0438	-0.158	0.000667
	(0.137)	(0.224)	(0.365)	(0.392)	(0.530)	(0.292)	(0.161)	(0.213)	(0.303)
disease X Born 1922	-0.0352	0.119	-0.0268	-0.238	0.231	-0.275	-0.156	-0.398*	-0.437**
	(0.0707)	(0.220)	(0.252)	(0.264)	(0.488)	(0.225)	(0.133)	(0.207)	(0.216)
disease X Born 1923	-0.135**	0.00892	-0.218***	0.211*	-0.264	-0.259**	0.108*	-0.170**	-0.0722
	(0.0533)	(0.0540)	(0.0811)	(0.109)	(0.191)	(0.108)	(0.0609)	(0.0731)	(0.0953)
disease X Born 1925	0.516***	0.234	-0.611*	-0.683**	1.308**	-0.569*	0.150	-0.455**	0.457
	(0.158)	(0.291)	(0.348)	(0.326)	(0.632)	(0.328)	(0.164)	(0.204)	(0.348)
disease X Born 1926	0.726***	0.703*	-0.923*	-0.210	2.352**	-0.964*	-0.153	-0.945***	-0.150
	(0.259)	(0.413)	(0.496)	(0.512)	(0.938)	(0.502)	(0.204)	(0.319)	(0.450)
disease X Born 1927	0.491**	-0.117	-0.00926	-0.679*	0.442	-0.433	-0.743***	-0.135	-0.961*
	(0.209)	(0.283)	(0.409)	(0.353)	(0.641)	(0.343)	(0.285)	(0.221)	(0.498)
disease X Born 1928	-0.370	-1.514*	2.124***	0.469	-3.173*	2.860***	-0.453	2.164***	-1.074*
	(0.399)	(0.878)	(0.715)	(0.684)	(1.652)	(0.758)	(0.298)	(0.557)	(0.609)
Constant	77.03***	29.19**	90.64***	64.42***	-10.61	86.40***	37.68***	81.23***	37.59***
	(16.81)	(12.01)	(14.22)	(13.68)	(32.79)	(13.66)	(11.66)	(11.25)	(12.13)
Observations	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444
R-squared	0.153	0.153	0.153	0.153	0.154	0.153	0.153	0.153	0.153
Birth X Enlist. year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Section FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
State trends	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; Coefficients correspond to percentage point changes in the probability of enlisting in the Air Force during the enlistment period January 1940 to November 1945. Disturbances are clustered at the section and birth year levels.

Table 11: Falsification exercise with the first 10 principal components of diseases in *Defects*

	(1) goiter	(2) pc1	(3) pc2	(4) pc3	(5) pc4	(6) pc5	(7) pc6	(8) pc7	(9) pc8	(10) pc9	(11) pc10
variable X Born 1920	-0.0229	0.202	-0.0332	0.306*	0.722***	0.0330	-0.0586	0.332	-0.784*	-1.315***	-0.177
	(0.184)	(0.203)	(0.243)	(0.181)	(0.274)	(0.326)	(0.362)	(0.466)	(0.439)	(0.466)	(0.499)
variable X Born 1921	0.0491	0.266*	-0.212	0.0677	0.492**	-0.163	-0.319	-0.0761	-0.479	-1.006***	-0.163
	(0.137)	(0.154)	(0.186)	(0.162)	(0.216)	(0.218)	(0.289)	(0.332)	(0.313)	(0.380)	(0.357)
variable X Born 1922	-0.0352	0.116	-0.438***	-0.0926	0.442*	-0.375**	-0.201	0.0806	0.101	-0.400	-0.116
	(0.0707)	(0.144)	(0.133)	(0.160)	(0.264)	(0.154)	(0.239)	(0.253)	(0.235)	(0.345)	(0.309)
variable X Born 1923	-0.135**	0.105	-0.0141	-0.0925	-0.306*	-0.264	-0.238	-0.0532	0.332	0.0809	0.183
	(0.0533)	(0.0950)	(0.110)	(0.108)	(0.161)	(0.182)	(0.219)	(0.224)	(0.300)	(0.282)	(0.259)
variable X Born 1925	0.516***	0.288*	-0.430***	-0.577***	-0.0492	-0.138	-0.548**	0.00717	0.123	-0.818***	0.676*
	(0.158)	(0.172)	(0.152)	(0.175)	(0.249)	(0.194)	(0.220)	(0.202)	(0.361)	(0.295)	(0.410)
variable X Born 1926	0.726***	0.335	-0.923***	-0.337	0.0614	-0.775*	-0.563	-0.389	-0.652	-1.327***	0.635
	(0.259)	(0.235)	(0.271)	(0.261)	(0.330)	(0.424)	(0.413)	(0.449)	(0.583)	(0.515)	(0.647)
variable X Born 1927	0.491**	-0.583***	-0.787***	0.127	0.833**	-0.357	0.116	-0.781**	-1.226***	-0.656*	-0.138
	(0.209)	(0.219)	(0.232)	(0.367)	(0.366)	(0.327)	(0.344)	(0.342)	(0.398)	(0.352)	(0.359)
variable X Born 1928	-0.370	-1.508***	1.134**	1.003***	1.807***	2.349**	2.258**	0.389	-0.460	1.362	-4.016***
	(0.399)	(0.404)	(0.480)	(0.388)	(0.536)	(1.017)	(1.053)	(0.876)	(1.377)	(1.502)	(1.219)
Constant	77.03***	53.16***	61.63***	56.43***	46.29***	39.87***	49.75***	41.53***	49.72***	49.93***	50.50***
	(16.81)	(6.538)	(8.261)	(7.326)	(6.618)	(4.486)	(5.981)	(5.090)	(5.883)	(7.234)	(6.023)
Observations	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444	1,935,444
R-squared	0.153	0.153	0.153	0.153	0.153	0.153	0.153	0.153	0.153	0.153	0.153
Birth X Enl.year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Section FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
State trends	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; Coefficients correspond to percentage point changes in the probability of enlisting in the Air Force during the enlistment period January 1940 to November 1945. Disturbances are clustered at the section and birth year levels.

Our second path for checking robustness to other diseases involves looking at more general measures of health. To get measures of general health for each section, we extract principal components from our set of 271 conditions that appear in *Defects*. In the interest of practicality, we focus on the first ten principal components. Cumulatively, these components account for 33.4% of the variation in the *Defects* data. The first principal component explain 6.1% of the variation[49]. The results in Table 11 show even less of a systematic pattern than that observed in the robustness exercise that looks at individual diseases.

9 Interpretation of coefficients

In this section we translate the results of our regressions in section 7 into a measure of the implied intelligence increases in the population following iodine supplementation. To do so, we construct a simple model of the selection in to the Army Air Forces. We write the model in terms of IQ scores, which are a more conventional measure than the AGCT scores that were actually used in the selection of recruits.

Consider a pool of recruits composed of two populations. A fraction $(1-\phi)$ of recruits are from regions in which there is no iodine deficiency. Let their intelligence be represented by a random variable IQ that is distributed normally with mean of 100 and standard deviation of 15. A fraction ϕ of recruits are from iodine deficient regions. Prior to treatment, their intelligence is given by the random variable IQ_d, distributed normally with a mean of X and a standard deviation of 15. The average level of IQ in the population would be $\overline{IQ} = \phi X + (1-\phi)100$. After treatment, the distributions are assumed to be identical.

As discussed above, a specific fraction of AAF recruits scored above the mean on the AGCT. Given this, we wish to examine how the probability of entering the AAF differs between the two populations before and after treatment. Let ψ be the fraction of recruits entering the AAF that had AGCT scores above the mean, and let θ be the fraction of the

[49]Kaiser's rule suggests using all principal components with eigenvalue higher than one. In our case, this would come to 72 principal components. Together, these explain 88.4% of the variation. We viewed presenting results from this many components as excessive.

overall pool of recruits that entered the AAF.

The sources that we have available do not describe the exact process by which the target percentage of above-average recruits in the AAF was achieved. Here we consider a simple and robust algorithm that would have produced this result. Suppose that initially, the AAF draws randomly from the pool of recruits until it has filled its quota for recruits *below* the mean. Since half the distribution is below the mean, this would fill a fraction $2(1-\psi)$ of the allotted slots. To fill the remaining available slots, the AAF would draw randomly from the pool of recruits with AGCT score above the mean.

Under such an algorithm, we can calculate the probability of someone from a low-iodine region entering the AAF as the sum of the initial draw from the entire distribution and the secondary draw where only the top half of the distribution is considered.

$$P(AAF|\text{iodine deficient}) = 2(1-\psi)\theta + \frac{(2\psi-1)\theta \cdot P(IQ_d > \overline{IQ})}{(1-\phi)P(IQ > \overline{IQ}) + \phi P(IQ_d > \overline{IQ})}$$

Similarly, the probability of someone from a non-deficient region entering the AAF is:

$$P(AAF|\text{non-deficient}) = 2(1-\psi)\theta + \frac{(2\psi-1)\theta \cdot P(IQ > \overline{IQ})}{(1-\phi)P(IQ > \overline{IQ}) + \phi P(IQ_d > \overline{IQ})}$$

The difference between these two probabilities corresponds to the regression coefficient in our specification. The difference is the increase in the probability of entering the AAF resulting from iodization[50]:

$$(2\psi-1)\theta \frac{P(IQ > \overline{IQ}) - P(IQ_d > \overline{IQ})}{(1-\phi)P(IQ > \overline{IQ}) + \phi P(IQ_d > \overline{IQ})}$$

We apply this model to our setting as follows. First Figure 9 suggests that, on average, across all the different recruitment periods (the 75% rule period, the period with no rule, and the period with two rules), about 70% of Air Force recruits had an AGCT score above

[50]There is one subtle issue, which is how the cutoff for entry into the AAF would have changed when treated cohorts from iodine-deficient regions began to enter the recruit pool. At this time, the average level of AGCT scores would have gone up, and if the test were continuously re-normed, the cutoff for entry into the AAF would have gone up as well. We do not have any information about whether this took place. In the calculations presented here, we make the assumption that the cutoff did not change.

Table 12: Interpretation

X **(Mean IQ)**	**Regression Coefficient**
100	0
95	.015
90	.029
85	.041
80	.052
75	.060
70	.065

the median. Thus we set $\psi = 0.70$. Second, according to our data the fraction of all recruits going to the Air Force was 14%. We use this value for θ. Finally, we need a measure of the fraction of recruits coming from iodine-deficient regions. In practice, there was a range of iodine deficiency as recorded in the *Defects* data. In column (2) table 5 we defined our "high goiter" dummy as corresponding to the one-quarter of sections with the highest prevalence of goiter. Thus we use a value of $\phi = 0.25$, and evaluate the coefficient from the regression that used this dummy variable.

Using these parameters, Table 12 shows the implied effect of iodization on the probability of entry into the AAF for a range of possible values of X, the mean level of IQ in the iodine-deficient regions.

The estimated coefficient on the high-goiter dummy in table 5 is in the range 3.8-10 percentage points. Comparing these estimates to Table 12, we see that the lower range of these estimates are consistent with iodization raising IQ by 15 points (that is, $X = 15$), which is a reasonable expectation given the work of (Bleichrodt and Born 1994). However, the higher range of estimates is larger than any reasonable estimate of the increase in IQ that would have resulted from iodization. Given the uncertainty surrounding the selection process and prior literature on iodization we do not consider the larger results plausible. In any case, our results are consistent with a substantial effect in line with the existing

literature[51].

Our estimates and their interpretation are based on a sample of male recruits. However, there is a lot of evidence suggesting that females are more prone to the development of iodine deficiency disorders than males, while Field et al. (2009) also discuss experimental evidence that, in animals, females are more likely to suffer cognitive damage in utero as a result of iodine deficiency. Both Field et al. (2009) and Politi (2010) find larger effects of iodization programs on graduation rates for females than for males. Therefore, the effect of iodization on the general population might well be larger than our estimates suggest.

10 Conclusion

Iodization of salt in the United States was one of the first instances of broad-based food fortification, a practice that continues in wealthy countries to this day and is spreading rapidly to the developing world. The experience of the US offers a useful natural experiment for identifying the long-term effects of an important micronutrient deficiency. Prior to iodization, there was significant geographic variation in iodine deficiency, resulting from identifiable soil and water conditions. This variation can be measured using data on the prevalence of goiter, allowing us to clearly distinguish between treatment and control groups. Iodization of salt was national in scope and implemented rapidly. The key period in which iodine deficiency affects cognitive development is a narrow window during gestation. Thus we are able to easily differentiate treated from non-treated cohorts.

We find that cognitive abilities of men born in iodine-deficient regions rose relative to those born in non-deficient regions for cohorts in utero after the advent of iodization. Our measure of cognitive ability is the probability of a man being assigned to the Army Air

[51] As mentioned above, we do not know the algorithm that was applied in order for the AAF to hit the target number of recruits with AGCT scores above the average. An alternative to the method presented here would be for the AAF to set a minimum AGCT score such that, if it took all recruits with scores above this minimum, it would hit the target percentage above the median. For a given value of X (the average IQ in iodine deficient areas), such a method would produce a larger effect of iodization on the probability of entering the AAF than the method that we consider.

Forces, entry into which was partially based on performance on a standardized intelligence test. Interpreting our measure in terms of IQ, our finding is that in iodine-deficient regions, iodization raised IQ scores by roughly one standard deviation, or 15 points. Given that one-quarter of the population lived in such regions, this implies a nationwide increase in average IQ of 3.5 points. In developed countries, average IQ rose at a rate of roughly three points per decade for much of the twentieth century. This is the Flynn Effect, which is thought to be largely attributable to reduced health and nutrition insults in utero and among children. Thus our results are consistent with the elimination of iodine deficiency in the United States accounting for roughly one decade's worth of the Flynn Effect.

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.*, August 2006, *114*, 672–712.

___ **and Janet Currie**, "Killing Me Softly: The Fetal Origins Hypothesis," *Journal of Economic Perspectives*, 2011, *25* (3), 153–72.

Andersson, Maria, Vallikkannu Karumbunathan, and Michael B Zimmermann, "Global iodine status in 2011 and trends over the past decade," *The Journal of nutrition*, 2012, *142* (4), 744–750.

Behrman, Jere R. and Mark R. Rosenzweig, "Returns to birthweight," *Review of Economics and Statistics*, 2004, *86* (2), 586–601.

Bishai, David and Ritu Nalubola, "The History of Food Fortification in the United States: Its Relevance for Current Fortification Efforts in Developing Countries," *Economic Development and Cultural Change*, October 2002, *51* (1), 37–53.

Black, Robert E, Lindsay H Allen, Zulfiqar A Bhutta, Laura E Caulfield, Mercedes De Onis, Majid Ezzati, Colin Mathers, and Juan Rivera, "Maternal and child undernutrition: global and regional exposures and health consequences," *The Lancet*, 2008, *371* (9608), 243–260.

Bleakley, Hoyt, "Disease and Development: Evidence from Hookworm Eradication in the American South," *Quarterly Journal of Economics*, February 2007, *122* (1), 73–117.

Bleichrodt, N and M P Born, "A meta-analysis of researh on iodine and its relationship to cognitive development," in John B. Stanbury, ed., *The damaged brain of iodine deficiency*, New York: Cognizant Communication, 1994, pp. 195–200.

Carpenter, Kenneth J., "David Marine and the Problem of Goiter," *Journal of Nutrition*, April 2005, *135*, 675–680.

Connolly, R. J., "An increase in thyrotoxicosis in southern Tasmania after an increase in dietary iodine," *The Medical Journal of Australia*, 1971, *1* (24), 1268–1271.

Copenhagen Consensus Center, "Copenhagen consensus 2008," *Fredriksberg, Denmark: Copenhagen Consensus Center*, 2008.

Curtis, George M. and M. Been Fertman, "Iodine in Nutrition," in "Handbook of Nutrition," second ed., American Medical Association, 1951, chapter 6, pp. 111–135.

de Benoist, Bruno, Erin Mclean, Ines Egli, and Mary Cogswell, eds, *Worldwide prevalence of anemia 1993-2005: WHO Global Database on Anemia*, Geneva, Switzerland: World Health Organization, 2008.

___ , Maria Andersson, Ines Egli, Bahi Takkouche, and Henrietta Allen, eds, *Iodine status worldwide: WHO Global Database on Iodine Deficiency*, Geneva, Switzerland: World Health Organization, 2004.

Ferrie, Joseph P, Karen Rolf, and Werner Troesken, "Cognitive disparities, lead plumbing, and water chemistry: Prior exposure to water-borne lead and intelligence test scores among World War Two US Army enlistees," *Economics & Human Biology*, 2012, *10* (1), 98–111.

Field, Erica, Omar Robles, and Maximo Torero, "Iodine deficiency and schooling attainment in Tanzania," *American Economic Journal: Applied Microeconomics*, October 2009, *1* (4), 140–169.

Fleischer, Michael, Richard M. Forbes, Robert C. Harriss, Lennart Krook, and Joe Kubota, "Iodine," in "The relation of selected trace elements to health and

disease," Vol. 1 of *Geochemistry and the Environment*, Washington, D.C.: National Academy of Sciences, 1974, chapter 3, pp. 26–28.

Fost, Carolyn Ann, "The Salt Industry: A Case Study in the Evaluation of Public Policy," *unpublished Ph.D. dissertation, Southern Illinois University*, 1970, pp. 49–50.

Grove, Robert D. and Alice M. Hetzel, *Vital statistics rates in the United States, 1940-1960*, U.S. Department of health, education, and welfare, Public Health Service, National Center for Health Statistics, 1968.

Katz, Frank J., *Mineral Resources of the United States 1924, Department of Commerce*, Washington: US Government Printing Office, 1927.

Kelly, F.C. and W.W. Snedden, "Prevalence and Geographical Distribution of Endemic Goitre," in "Endemic Goitre" number 44. In 'WHO Monograph Series.', Geneva, Switzerland: World Health Organization, 1960, p. 27.

Kohn, Lawrence A., "Goiter, Iodine and George W. Goler: the Rochester Experiment," *Bulletin of the History of Medicine*, 1975, *49*, 389–399.

Koutras, Demetrios A., Josip Matovinovic, and Robert Vought, "The Ecology of Iodine," in John B. Stanbury M.D. and Basil S. Hetzel, eds., *Endemic Goiter and Endemic Cretinism*, John Wiley and Sons, 1980, chapter 9, pp. 185–195.

Langer, P., "History of Goitre," in "Endemic Goitre" number 44. In 'WHO Monograph Series.', Geneva, Switzerland: World Health Organization, 1960, pp. 9–25.

Love, Albert G. and Charles B. 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.

Mannar, MG Venkatesh and Lucie Bohac, *Achieving Universal Salt Iodization: Lessons Learned and Emerging Issues*, Ottawa, Canada: Micronutrient Initiative, 2009.

Marine, David and O.P. Kimball, "The prevention of simple goiter in man," *Journal of the American Medical Association*, October 1, 1921, *77* (14), 1068–1070.

Markel, Howard, ""When it rains, it pours":Endemic Goiter, Iodized Salt and David Murray Cowie, MD," *American Journal of Public Health*, 1987, *77* (2), 219–229.

Matovinovic, J. and V. Ramalingaswami, "Therapy and Prophylaxis of Endemic Goitre," in "Endemic Goitre" number 44. In 'WHO Monograph Series.', Geneva, Switzerland: World Health Organization, 1960, pp. 385–410.

McClendon, J. F., "Inverse relation between Iodin in Food and Drink and Goiter, Simple and Exophthalmic," *Journal of the American Medical Association*, May 1924, *82* (21), 1668–1672.

McClure, Roy D., "Thyroid surgery as affected by the generalized use of iodized salt Salt in an Endemic Goitre Region - Preventive Surgery," *Annals of Surgery*, November 1934, *100* (5), 924–932.

Micronutrient Initiative, *Investing in the Future: A United Call to Action on Vitamin and Mineral Deficiencies: Global Report, 2009*, Micronutrient Initiative, 2009.

National Archives and Records Administration, "World War II Army Enlistment Records, 1938-1946," 2002.

Palmer, Robert R., Bell I. Wiley, and William R. Keast, *The Procurement and Training of Ground Combat Troops* United States Army in World War II: The Army Ground Forces, Washington, D.C.: Center of Military History, United States Army, 1948.

Politi, Dimitra, "The impact of iodine deficiency eradication on schooling: evidence from the introduction of iodized salt in Switzerland," 2010.

Scrimshaw, Nevin S., "Malnutrition, Brain Development, Learning and Behavior," *Nutrition Research*, 1998, *18* (2), 351–379.

U.S. Department of Commerce, *Mortality Statistics*, Bureau of the Census, various years.

Watson, Tara, "Public Health Investments and the Infant Mortality Gap: Evidence from Federal Sanitation Interventions on U.S. Indian Reservations," *Journal of Public Economics*, September 2006, *90*, 1537–1560.

World Health Organisation, "Is it true that lack of iodine really causes brain damage?," Online Q&A, http://www.who.int/features/qa/17/en/index.html, May 2013.