A Bit of Salt, A Trace of Life:

Gender Norms and The Impact of a Salt Iodization Program on Human Capital Formation of School Aged Children*

Zichen Deng† Maarten Lindeboom‡

Current version: Saturday 17th August, 2019

Abstract

This paper examines the effects of a massive salt iodization program on human capital formation of school-aged children in China. Exploiting province and time variation we find strong positive effects on cognition for girls and no effects for boys. For non-cognitive skills we find the opposite. We show in a simple model of parental investment that gender preferences can explain our findings. Analyses exploiting within province, village level variation in gender attitudes confirm the importance of parental gender preferences. Consequently, large scale programs can have positive (and possibly) unintended effects on gender equality in societies with son preference.

KeyVerbals: Iodine, parental investments, gender attitudes, cognitive skills, non-cognitive skills.
JEL: I15, J16, J24, 015

*The main title is borrowed from Endeavors. Thanks to Douglas Almond, Prashant Bharadwaj, Arnaud Chevalier, Hans Grönqvist, Seema Jayachandran, David Johnston, Johannes Kunz, Travis Lybbert, Peter Nilsson, Ilyana Kuziemko, Giacomo Pignataro, Hitoshi Shigeoka and Gerard van den Berg, for helpful conversations and seminar/workshop audiences at TI PhD, RES Symposium, Essen, Toulouse, NEUDC, Kyoto, Milan, Tilburg, Oslo and Uppsala for useful comments. All errors are our own.

†Vrije Universiteit Amsterdam and Tinbergen Institute, zichen.deng@vu.nl

‡Vrije Universiteit Amsterdam, Centre for Health Economics, Monash University, Melbourne, Tinbergen Institute and IZA Bonn, m.lindeboom@vu.nl

1 Introduction

Iodine deficiency early in pregnancy can have significant, irreversible effects on brain development of the fetus (Cao et al., 1994) and can therefore have important consequences for human capital formation of children and subsequent socio-economic outcomes. Indeed, the large medical literature (see Zimmermann, 2011, for a systematic literature review) has found adverse effect of iodine deficiency on later life growth and development, in particular when exposed early in life. Using historical data, recent economics papers such as Adhvaryu et al. (2016); Feyrer, Politi, and Weil (2017) found that cohorts exposed to a higher levels of iodine intake have higher labor force participation rates, higher incomes and also higher probabilities of entering top-tier occupations with higher cognitive demands. Although the benefits of micro-nutrient improvement have been documented for decades, still only around 66% of households have access to iodized salt globally.[1] Further, little is known about pathways between early in life exposure and adult productivity and earnings. This impedes our understanding of how developmental trajectories unfold over the life course and who benefits most from large scale interventions.

Our contribution to the extent literature is to consider and account explicitly for possible interactions with gender preferences in receiving parental investments early in life. These differences are deeply rooted in gender norms, which are important in the Asian context as well as in the Middle East and North Africa. The idea that gender preferences play a role in the western world cannot be excluded either. Biased gender norms may already have an impact on children very early in life and might explain heterogeneous gender effects found in the literature on long run effects of early life shocks (Field, Robles, and Torero, 2009; Maccini and Yang, 2009; Adhvaryu et al., 2016). This paper examines the effect of a massive, nationally implemented salt iodization program in China. Evidence of differential program effects by gender would not only provide new interpretation to gender differences in the impact of early in life shocks found in the literature, but also shed light on predictions of theoretical models of human capital formation that allow parents to make compensatory and reinforcing investments in different dimensions of human capital. Parental child preferences play an important role in such investments. If compensatory investments of parents are relevant, then large scale interventions may crowd out (or reinforce) private investments and may also affect skills dimensions initially not aimed at. Public policy can therefore have positive (and possibly) unintended effects on gender equality in societies where gender norms are important. Understanding mechanisms underlying program effects are crucial as

[1] In 2007, an estimated 31.5% of school-age children (266 million) had insufficient iodine intake. In the general population this amounted to 2 billion people. See also: UNICEF, WHO

any intervention would be blind without knowledge of the mechanisms underlying behavioral responses.

To fight against iodine deficiency related diseases around the country, the Chinese government implemented a national program of regulating salt to contain iodine in October, 1994. In the second year, biennial province-based monitoring was introduced to record the use and iodine content of household salt, along with urinary iodine concentrations among schoolchildren from the same households. After the introduction of the program, the urinary iodine concentration reached satisfactory levels from 1995 onward and the percentage of children who had goiter[2] dropped rapidly. Given the importance of iodine during the gestational period for brain development (Cao et al., 1994; Zimmermann, 2011) we, in first instance, focus on the potential impact of this policy on cognitive ability and school attainment of children who are affected in utero.

For this, we link iodine deficiency information across locations collected at the start of the intervention to nationally representative rural samples drawn from the *China Family Panel Studies* (CFPS). A unique aspect of the CFPS survey is that math and vocabulary ability from standardized tests were collected besides information on educational attainment and schooling. Similar to Shah and Steinberg (2017), our human capital measurements in the CFPS have the advantage that the same questions were given to each individual in the survey, no matter whether he/she is currently enrolled in school or not. With this data set, our empirical design does not suffer from selection bias caused by censoring individuals who had already left school. To identify the long-term benefits of the salt iodizing policy, we use the national salt iodizing program as a quasi-experiment and exploit geographic variation in goiter prevalence prior to the intervention. So, we essentially compare improvements in math and vocabulary ability as well as educational attainment and years of schooling of cohorts conceived before and after the salt iodization in areas with varying pre-intervention goiter prevalence. Our difference-in-differences estimates show that the salt iodization policy has strong and significant effects on cognition for girls. A one standard deviation (12%) decrease in the pre-intervention regional goiter rate is associated with an increase of 15% in female math and vocabulary test scores. We also see large increases in educational attainment and schooling of women. Yet, for men we find much smaller and insignificant effects.

Gender preferences are important in the Asian context and this may be important for our findings on gender. We consider a simple theoretical model along the lines of (Heckman, 2007; Cunha, Heckman, and Schennach, 2010; Yi et al., 2015; Almond, Currie, and Duque,

[2]Iodine is an essential component of hormones produced by the thyroid gland. Iodine deficiency can lead to an enlarged thyroid gland located at the base of the neck. This is the most visible consequence of iodine deficiency. Goiter in adulthood does not have severe consequences, iodine deficiency in utero can lead to impaired neuro-development of the child and post-birth cognitive disabilities.

2018) on the role of parental investment in human capital formation. In this model gender preferences play a role and parental investments interact with different endowments at birth. The predictions of the model are in line with our finding of strong and sizable effects for girls and small and insignificant effects for boys. According to the model, the iodization program reduces the need for private investments in cognition for boys and induce parents to divert their investments to other skill dimensions. Indeed, we find that the program has positive effects for boys on non-cognitive skill measures. We find no effects for girls in these non-cognitive skill dimensions. Similar to Dahl, Kotsadam, and Rooth (2017); Dhar, Jain, and Jayachandran (2018); Dossi et al. (2019) we proxy gender preference by gender attitudes, specifically, about the appropriate roles and rights of women and girls. Across all outcomes, we find that the gains in cognition are largest for girls born in regions with strongest son preferences. For boys, we don't find heterogeneous treatment effects of iodized salt across districts with different gender attitudes.

Our study contributes to at least four strands in the literature. Firstly, we add to the literature on the long-term effects of early life conditions. Much of this "fetal origins" work (see, among others, Almond, 2006; Van den Berg, Lindeboom, and Portrait, 2006) has focused on demonstrating the impact of extreme, traumatic experiences (disease outbreak, recessions, famines, extreme environmental shocks, etc.) in early life. Few studies have estimated gains to exposure to a purposeful large-scale distribution of resources. (Exceptions are Hoynes, Schanzenbach, and Almond (2016), food stamps and Brown, Kowalski, and Lurie (2018), medicaid.) The nationally implemented intervention in China started just after the launching of the 1993 WHO campaign and is to the best of our knowledge the largest of its kind. It is also a commonplace, moderate intervention and has therefore relevant external validity. This aids policy makers in optimizing similar policies in the future. Furthermore, while there exist a large body of studies that look at the impact of in utero exposure to the quantity of food (see Lumey, Stein, and Susser, 2011, for an excellent review of the famine literature), only a few studies (e.g. Field, Robles, and Torero, 2009; Adhvaryu et al., 2016; Feyrer, Politi, and Weil, 2017) have looked at the long run effects of food *quality* or nutrient intake.

Secondly, and related to the above, we contribute to the discussion on intermediate proxy indicators of long-term outcomes. Adhvaryu et al. (2016); Feyrer, Politi, and Weil (2017) examine the long term effect of a salt iodization program, promoted by a private firm, on lifetime income at later ages. Adhvaryu et al. (2016) find an about 10% income increase for those exposed to the iodine program. We look at the effect of a public program on childhood cognition and education and find substantial effects. Our study thus adds to a full picture of how early-life disadvantage unfold over the life course. We use measurements of human capital that include standardized numeracy tests for all children, as opposed to most of the previous literature which only focuses on school enrollment. A few recent studies such as (Figlio et al., 2014; Almond, Mazumder, and van Ewijk, 2015; Bharadwaj, Lundborg, and Rooth, 2017; Shah and Steinberg, 2017) take a similar approach as we do by examining the effects of events in early childhood on cognitive test outcomes during the school years. Most studies use administrative data of developed countries, where standardized tests cover most of the school going children at a certain age. In developing countries, however, a substantial share of the children is already out of school at young ages and therefore a similar strategy will only partly measure the effectiveness of the intervention. Moreover, if collected, most human capital measurements are self-reported performance measures, which makes the comparison across individuals difficult. Our work complements the literature with evidence from a large developing country by using a data set where the results of standardized math and verbal tests are collected for all children, in and out of school.

Thirdly, we also shed light on the literature about child gender preferences. Gender biases favoring males particularly in education are larger in developing countries like China and India. Females in those countries often receive less investments from parents (see, for example, Oster, 2009; Jayachandran and Kuziemko, 2011; Bharadwaj and Lakdawala, 2013; Barcellos, Carvalho, and Lleras-Muney, 2014) and are likely not to reach their full potential in education, health and personal autonomy. In our current study, the iodine policy has positive (possibly unforeseen) spillovers to females. Our study therefore also speaks to the relevance of early life conditions in explaining gender differences in socio-economic outcomes later in life.[3]

Finally, our heterogeneous analysis that include gender preferences is motivated by theoretical models of human capital formation. In such models often dynamic complementarities between investments at different stages of childhood are considered (Cunha, Heckman, and Schennach, 2010). A handful of studies (Adhvaryu et al., 2015; Gunnsteinsson et al., 2018; Duque, Rosales, and Sanchez, 2018; Aguilar and Vicarelli, 2018; Rossin-Slater and Wüst, 2018) attempt to identify these dynamic complementarities empirically. For this they use exogenous variation at different stages of the life cycle and generally cannot find evidence for dynamic complementarities. However, as pointed out by Malamud, Pop-Eleches, and Urquiola (2016), parents might increase investments in the child to counter adverse effects of the initial shock. This may confound the effect of subsequent shocks. They furthermore argue that human capital outcomes for children are the result of parental preferences, the family budget constraint and the shape of the child health production technology. This makes

[3]See Almond and Currie (2011) who call for work that integrates work on son preference with work on fetal origins.

it difficult to interpret reduced form effect estimates. We use parental gender attitudes as a proxy for parental preferences to identify one of the channels in the human capital formation of children and find that gender preferences are important in the formation of human capital of school aged children.

Our results point towards four observations that are relevant for the strands of literature referred to above: the relevance of parental investment responses in mitigating the effects of adverse shocks early in life; that child gender preferences are important for these investments decisions; that large scale interventions may crowd out private investments and may also affect skills dimensions initially not aimed at; large scale programs can have positive and possibly unintended effects on gender equality in societies where boy preferences are important.

The rest of the paper is organized as follows. Section 2 provides a brief overview of Iodine Deficiency Disorders (IDD), the Universal Salt Iodization (USI) and related literature. Section 3 provides a description of the data used in the analysis. Section 4 outlines empirical model and Section 5 discusses the results of the models. We zoom in on our finding of differential effects by gender in section 6. This section presents a simple model where gender preferences may differently affect parental investment in girls and boys. We introduce gender attitudes as a proxy for gender preferences and examine whether the program effects on cognitive and non-cognitive skills vary with gender attitudes. Section 7 summarizes our findings, places these findings into context and concludes.

2 Background

Iodine is an essential component of the hormones produced by the thyroid gland, and is therefore essential for human life (Zimmermann, 2011). From the fetal stage to adulthood, insufficient iodine intake causes many disorders, the most common of which is an enlargement of the thyroid gland. Although this enlargement, which is called goiter, is the most visible symptom of iodine deficiency, besides being inconvenient it has no severe consequences. However, fetal exposure to iodine deficiency may lead to impaired neuro-development. The brain damage caused by severe iodine deficiency in this stage of life is often irreversible.

The knowledge that iodine can help prevent goiter has existed since the mid-1800's (Zimmermann, 2008). It was not until 1895, however, that iodine was first discovered in the thyroid gland (Baumann, 1896). Switzerland was the first country in the world to introduce iodized salt in 1922. The United States introduced iodized salt in 1924 after the executive Council of the Michigan State Medical Society officially endorsed iodized salt. In 1993, the World Health Organization (WHO) proposed a worldwide campaign to eradicate IDD. The primary intervention strategy for IDD control is Universal Salt Iodization (USI), which is a notably simple, universally effective and particularly cheap instrument. The World Bank reports that it only cost approximately \$0.05 per child per year.

Historically, endemic goiter was found particularly in the mountain regions in China. For instance in 1940s, more than 20% of the residents of Kunming, the capital of the province of Yunnan had goiter (Simoons, 1990). The Chinese Academy of Preventive Medicine had estimated that about 450 million people lived in iodine-deficient areas, with more than 30% of the population considered at risk of IDD (see Chen and Wu, 1998). The iodine deficiency disorders problem was acknowledged as a public health threat and in response to this in 1993, the State Council of China announced the Universal Salt Iodization (USI) policy with the aim to virtually eliminate IDD by 2000.

The USI was a national strategy. The challenge of the policy was to increase salt iodine levels sufficiently to bring the median urinary iodine concentration of children into the 100–199 µg/L range and at the same time maintaining the optimal urinary iodine concentration (MUIC) levels of pregnant women (150–249 µg/L). To reach the desired intake of iodine, the State Council enacted in October 1994 the national regulation of salt iodization.[4] In addition, between 1993 and 1995, a national monitoring system was built to track trends in goiter prevalence among school children aged 8-10. The monitoring was held between March and June 1995. In our empirical analyses (see section 4), we will use the outcome of this monitoring exercise as the pre-policy distribution of iodine deficiency levels across the different provinces. Note that this is a few months after the implementation of the Salt Iodization Program. Therefore, a concern may be that the cross-province variation in goiter rates does not reflect the pre-policy distribution of iodine deficiency rates. However, the literature has documented lags of at least one year before goiter rates normalize after iodine repletion Zimmermann et al. (2003).[5]

Figure 1 shows the pre-policy spatial distribution of iodine deficiency levels of schoolchildren aged 8-10. Goiter rates among children under 10 years old does not differ significantly by gender (Sun, 2018). The dark areas (mostly western and northern provinces) indicate high prevalence rates (up to 43.3%), while the light areas (south east) indicate low prevalence rates. For our empirical analyses it is of importance to know whether the Universal Salt Iodization policy was effective in increasing the urinary iodine concentration levels in the population. We turn to this below in the data section.

[4]The level of salt iodization during the manufacturing process was set at 50 mg/kg in 1994 to ensure a level of not less than 40 mg/kg in the manufacturer's supply outlets.

[5]Zimmermann et al. (2003) documented non-significant reduction in thyroid size among children age 8-9 in Cote d'Ivoire one year after the introduction of iodized salt and even two years after the salt intervention, they only observed an 8% reduction in the goiter rate.

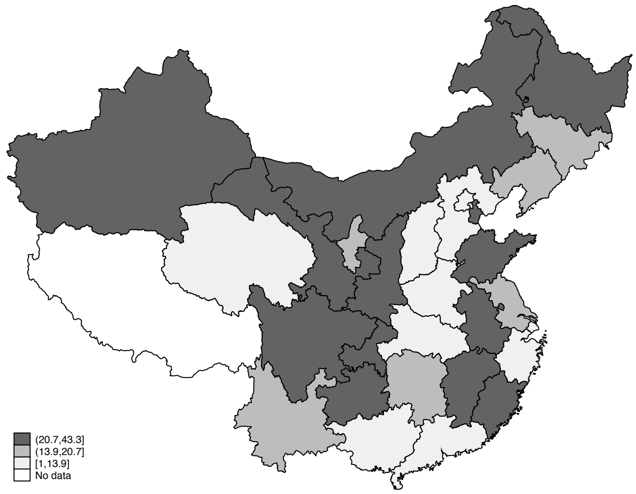

Figure 1. Goiter Distribution in 1995
Notes: Figure 1 reports goiter rates (%) (among schoolchildren aged 8-10) in 1995. Darker areas represent higher goiter rates.
Sources: National Iodine Survey

3 Data

3.1 Goiter Data

The base, pre-policy, geographic distribution of goiter prevalence before the salt iodization policy (see Figure 1) came from the 1995 national survey on goiter rates among schoolchildren. In each provincial survey, a multistage, probability proportionate to the population size cluster sample was obtained. The county served as the primary sampling unit and in each province 30 counties (clusters) were selected from a county population list. In each selected county, a school was then sampled at random. Children aged 8 to 10 years at the time of the survey served as the index population. For each cluster, 40 children were selected at random from the enrollment list. All children were examined for thyroid size by palpation and/or ultrasound. Therefore, the goiter rate is defined as the percentage of schoolchildren who have either Class I or Class II goiter.[6] The sample sizes by province ranged from 1,200 to 2,400 (mean, 1,259). Our goiter data has an important advantage over goiter measures used in some recent studies like Adhvaryu et al. (2016); Feyrer, Politi, and Weil (2017) who use goiter prevalence among military recruits, an index population that consists of young and healthy and as such may not be a representative measure of local iodine deficiency problems. Ideally we would like to have used the county level goiter rates, but as we do not observe the county of birth we use the province-level goiter rates. Along with measuring the thyroid

[6]Class I goitre in normal posture of the head cannot be seen and it is only found by palpation. Class II goitre is palpable and can be easily seen.

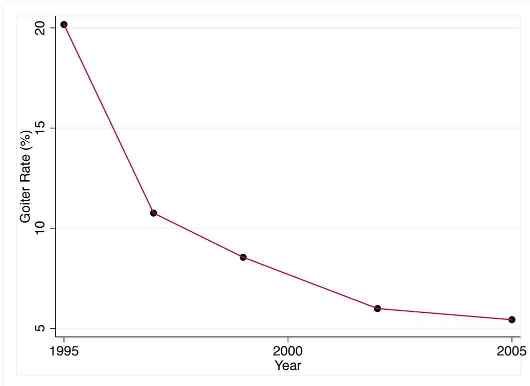

Figure 2. Goiter Rates Decline

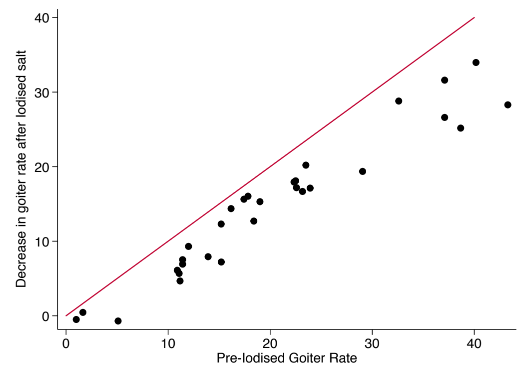

Figure 3. Before and After the Intervention

Notes: Figure 2 reports time-series data on mean goiter rate (among schoolchildren aged 8-10) across country between 1995 to 2005. Figure 3 shows the post-intervention decline in goiter rate versus pre-intervention levels across China.
Sources: National Iodine Survey

size, the survey also includes for a sub sample of the children urinary iodine concentration levels. The median urinary iodine concentration in schoolchildren progressed from 160 µg/L in 1995 to 300 µg/L in 1997 and 282 µg/L in 1999.

The survey was held every two or three years, which enables us to track the effect of the Universal Salt Iodization program over time. Indeed, the program proved to be very effective. By 2002, provinces converged to very low child goiter rates, so that provinces with high pre-eradication levels of goiter experienced the largest reductions. This is illustrated in Figure 2 that shows the (unweighted) average goiter rate (%) across the whole country over 1995-2002. The average goiter rate decreases from 20% in 1995 to around 5% in 2005. Figure 3 shows the post-campaign decline in goiter rate versus pre-campaign levels. Of importance for our empirical analyses is that this figure shows that the policy was effective in bringing down goiter rates for all provinces.

3.2 The Sample, Outcome Variables and Control Variables

3.2.1 The sample

The micro-level data used in this study come from the *China Family Panel Studies* (CFPS). The CFPS is a large-scale nationally representative panel survey conducted by the Social Science Survey Institute at Peking University. Three waves of survey have been published until 2017. The CFPS baseline wave (hereafter CFPS-2010) selected a total of 14,798 households, containing 33,600 adults and 8,990 children. A second and third wave of the CFPS followed the same individual from the households in 2012 and 2014 (hereafter CFPS-2012 and CFPS-2014). A standard math and verbal tests was carried out in CFPS-2010 and CFPS-2014. For all three waves educational attainment was recorded. Besides cognitive measures also some non-cognitive measures were collected. We provide more information on these measures in section 6. The data include accurate birth information, such as year, month of birth, place (province) of birth and whether individuals were born in a rural area. We restrict ourselves to those born in the rural areas (81% of the total population). The intervention is likely to be cleaner for those born in a rural area, as in the urban areas people had better access to micro nutrient food supplements.[7] The survey also collects respondents' migration history. Migration at young ages is very low (less than 3%). The Salt Iodization policy was implemented in October 1994. We include cohorts born between July 1991 and June 2000. At the base wave (CFPS-2010) these individuals were between 10 and 19 years old. This leaves us with 3,901 children at the baseline survey. For 3,640 children individuals we have all key information on test scores and education. The third wave (CFPS-2014) included the same math and verbal tests as the baseline wave. We therefore added this third wave to our baseline data. This resulted in 324 additional children. However, due to sample attrition between 2010 and 2014, we lost 951 individuals. We return to this issue in the sensitivity analyses of section 5.3. In the end, we have 6,653 observations on schooling (3433 males and 3220 females) and 5,889 observations on test scores (3026 male and 2873 female test scores). In the analyses we link these individual observations to the pre-policy goiter rates.

3.2.2 Variables

Educational attainment is an ordered categorical variable ranging from one to four (illiterate, graduated from primary school, middle school and high school). Schooling is measured by the number of years in school. As alternative measures of human capital we also use a math test and a verbal test designed by the CFPS team. Both tests were presented to the students, irrespective of their age. The math test consists of twenty-four mathematical problems, designed based on knowledge in textbooks used in primary and secondary schools. Questions were sorted in order of increasing difficulty and each of those questions counted for one point. Similarly, the verbal test consists of thirty-four Chinese characters designed based on the language textbooks, again sorted in order of increasing difficulty and counting for one point each. Therefore, full scores of the math and language test are 24 and 34, respectively. Note that as we measure cognition and schooling at fixed points in time, older children will perform much better. It will therefore be important to control for age in the

[7] The use of such supplements are also likely to vary by parental Socio-Economic Status (SES), which in itself correlates strongly with our cognitive and educational outcomes measures.

later regressions. We also return to this in the robustness analyses of section 5.3.

The sample also includes basic socio-demographic variables, such as age, gender, parental educations and family size.[8] In our empirical strategy it will be important to control for possible time varying confounders such as developments in public health investments. We therefore supplement the set of individual controls with some characteristics of the province at birth, obtained from the National Bureau of Statistics of China (The number of hospitals per capita and the number of hospital beds per capita).

Table 1 reports summary statistics of socio-demographic variables by gender for provinces with high initial (pre-policy) goiter rates (goiter rates above median of 17%) and low initial goiter rates (goiter rate prevalence less than 17%) for cohorts born before the implementation of the program (1991-1994).[9] The table shows that parental education is higher in regions with low initial goiter rates. The table also shows that female outcomes on test scores and education are slightly better than the outcomes for males. However, the table also shows that the differences in schooling and test scores between high and low goiter regions are larger for girls than for boys.

4 Empirical Strategy

4.1 Baseline Econometric Model

As we discussed in the previous section, salt iodization was rolled out nationwide in October 1994. Therefore, there is no province that could serve as a pure control group. As a proxy for pre-policy iodine deficiency rates we use province goiter rates among 8-10 years old children at the start of 1995 (see Figure 1). Like, among others, Bleakley (2010b); Adhvaryu et al. (2016), we use a difference-in-differences design. In this way we compare trends in various outcome measures in provinces with different levels of iodine deficiency levels prior to the implementation of the salt iodization program.

We define someone as treated if the entire gestation period is after the date of the implementation of the salt iodization program. All others are considered to be controls.[10] In section 5, as a robustness check, we also consider alternative definitions to assign treated and controls in groups. In contrast to earlier studies that used the 1924 salt iodization policy

[8] Families in the rural area were regularly exempted from the one child policy. For example, rural married couples were allowed have a second child if the first child was female (Zhang, 2017). The sex ratio in our sample of rural born children is about 1.07, which is lower than 1.12 (i.e., the sex ratio at birth in 1990s documented in Jayachandran (2015)).

[9] In the Appendix, we provide results for the cohorts born after the implementation of the salt iodization policy (Table A2).

[10] As is generally done in the literature, we use a nine month gestation period.

Table 1: Summary Statistics

	High Goiter Provinces		Low Goiter Provinces	
	Females	Males	Females	Males
Outcomes				
Educational Attainment	2.94	2.93	3.11	2.94
	[0.86]	[0.86]	[0.76]	[0.78]
Illiterate	0.059	0.063	0.016	0.018
Primary School	0.22	0.22	0.19	0.29
Middle School	0.44	0.45	0.46	0.44
High School or above	0.28	0.27	0.33	0.26
Schooling	9.86	9.71	10.4	9.91
	[3.15]	[3.21]	[2.61]	[2.68]
Math Test Scores	15.2	15.1	16.2	15.3
	[5.64]	[5.81]	[4.88]	[5.42]
Verbal Test Scores	26.3	25.3	26.9	25.1
	[7.24]	[7.60]	[5.90]	[7.30]
Demographics				
Age	18.6	18.8	18.8	18.8
	[2.39]	[2.41]	[2.39]	[2.40]
Father's Education	2.29	2.26	2.51	2.49
	[0.95]	[0.99]	[0.93]	[0.89]
Mother's Education	1.69	1.74	2.13	2.02
	[0.85]	[0.87]	[0.92]	[0.89]
Birth Order	1.60	1.57	1.60	1.80
	[0.84]	[0.77]	[0.79]	[0.95]
Family Size	4.87	4.60	4.89	4.68
	[1.46]	[1.43]	[1.47]	[1.55]
Additional Controls				
No. of hospitals per capita, 0-3 avg.	15.8	15.9	20.6	20.2
	[5.24]	[5.82]	[12.6]	[12.6]
No. of hospital beds per capita, 0-3 avg.	21.5	21.2	28.5	28.2
	[3.50]	[3.13]	[11.8]	[11.8]
Number of observations	679	748	644	728

Notes: Author's tabulations of CFPS-2010 and CFPS-2014. Sample consists individuals born in rural area between July 1991 and June 1995. We label a province as high/low goiter if its goiter rate is above/below 17%.

in the US, the Chinese iodized salt campaign was implemented rapidly across the entire country. At the start of 1995 more than 80% of families had already access to iodized salt, two years later this has increased to 95%.

During the gestational period iodine primarily affects the neuro development of fetus, with consequences for post natal cognitive functioning (Cao et al., 1994; Zimmermann, 2011). Therefore we focus in first instance on the impact of the salt iodization policy on cognitive ability and school attainment of children who are affected in utero. For this we use the following baseline regression:

$$Y_{ipt} = \beta_0 + \beta_1 Post_t \times Goiter_p + X_{ipt}\rho + \delta_p + \gamma_t + \epsilon_{ipt} \tag{1}$$

where outcome Y_{ipt} is either the logarithm of the cognitive test scores (math and verbal), educational attainment as well as the number of schooling years for individual i, who was born in province p, in year t. $Post_t$ indicates whether the individual was born after the introduction of iodized salt. $Goiter_p$ is a measure of pre-eradication endemicity in individual i's province of birth. The vector X_{ipt} includes individual characteristics such as age, parents education and family size, time varying province variables (the number of hospitals and the number of hospital beds per capita) and a mean-reversion control.[11] δ_p and γ_t are province and birth cohort fixed effects. Note that the birth cohort fixed effects are particularly important here as our outcome variables are measured in 2010 and 2014, which implies that those exposed to the salt iodization are much younger than the controls. We also control for region-specific linear trends in all models.[12] We run specification (1) separately for males and females. We also consider alternative specifications in section 5.3.

Of prime interest is the continuous treatment variable $Post_t \times Goiter_p$ that proxies potential iodine exposure. Recall from Figure 3 that the salt iodization policy was very effective in reducing goiter rates in all provinces to very low levels. So while the parameter β_1 in a strict sense is the intention-to-treat effect, the high compliance rates make it very close to the treatment effect (i.e. IV estimates). Also note that the controls (i.e. those who are not exposed in utero) are exposed to the program at later ages.

4.2 Dynamic Specification

Since our estimates use the cross-province convergence in goiter rates created by the introduction of iodized salt (Figure 2 and 3), convergent pre-trends across high and low-base goiter rate provinces prior to 1995 are a concern. Therefore, we use in addition an event study design to formally test for the common pre-trends assumption. More specifically, we run the following regression:

$$Y_{ipt} = \beta_0 + \sum_{t=1991}^{2000} \beta_t \times Goiter_p + X_{ipt}\rho + \delta_p + \gamma_t + \epsilon_{ipt}, \tag{2}$$

[11]If the oldest cohorts had high Iodine Deficiency Disorders and low human capital because of some mean-reverting shock, we might expect human capital gains for the subsequent cohorts even in the absence of a direct effect of the salt iodization policy on productivity. Following the similar logic as Bleakley (2010a), we construct the mean-reversion control by interacting provincial average educational attainment in the 1980 Census with the dummy variable $Post_t$.

[12]The regions consist of several provinces. See Table A1 of the appendix for the precise definition.

where β_t gives the cohort-specific relationship between pre-eradication endemicity and later-childhood outcomes.[13] If salt iodization affected the human capital formation of exposed cohorts, these effects should be visible in a break from pre-existing trends in β_t. This method would also shed light on the partial effects of iodine exposure in late childhood (rather than in utero), if such effects exist.[14] Note that all individuals born in 1995 or later are exposed to iodized salt from conception onward. Individuals born in 1994 experience higher iodine intake in their year of birth; thus, this cohort is partially exposed to higher iodine levels in utero and fully exposed from birth onward. Individuals born in 1993 experience higher iodine from age 1 onward and those born in 1992 experience higher iodine from age 2 onward. Since we normalize the coefficient for the 1994 cohort to be zero, our analysis essentially tests for differential effects of exposure relative to exposure at age 1 and older. If there are additional benefits to having access to iodine between conception and age 1, we would expect the coefficients β_{1995+} to be positive. Similarly, if iodine at age 1 has an additional benefit relative to iodine exposure at age 2 or older, we would expect coefficients β_{1993-} to be negative.

5 Results

5.1 Baseline Results

Table 2 reports the main results of two separate regressions of our basic model (Equation 1): one for men in Panel A and the other for women in Panel B. In all the regressions discussed in this section, the coefficients of interest are the post-by-goiter rate interaction, which represent the effect of salt iodization on our outcomes of interest. To interpret the size of the effect, the coefficients in all tables have been multiplied by 12, which is the inter-quartile range of the goiter distribution from a relatively low goiter province (at the 25th percentile) to a high goiter province (at the 75th percentile). Although the following tables only report the coefficient of interest, in all specifications we include controls for province and year of birth fixed effects, age, birth order, family size, parents' characteristics and region-specific linear trends. The introduction of iodized salt took place during a period of a rapid growth in the economy. To explore these possible confounders, we also control for several characteristics of

[13] In practice, β_{1994} represents estimates of individuals born between July 1994 and June 1995, and β_{1995} for the cohort born between July 1995 and June 1996, etc. We did such adjustments because individuals born after July 1995 were conceived after the implementation of the salt iodization policy. Similarly, the 1993 cohort are those born between July 1993 and June 1994.

[14] See Zimmermann (2011) for a comprehensive summary of the role of iodine in human growth at different stages in life. For example, neonatal iodine deficiency may cause endemic cretinism. Deficiency during childhood and adolescence may impair mental functioning and delay physical development.

the province of birth (hospitals and hospital beds per capita), measured as averages over the first three years of life. Standard errors are clustered at the province-of-birth level to allow for arbitrary correlation of the errors for individuals born in the same province. We also report two-way clustered standard errors by province and family appear in angle brackets to take into account the situation that one family has multiple children. The standard errors of two-way clustering are almost identical to the usual cluster robust standard errors. Given that our data only contain 28 provinces, we produce statistical inference based on the wild-bootstrap approach (Cameron, Gelbach, and Miller, 2008). The associated wild-bootstrap standard errors turn out to be slightly larger than the usual cluster robust standard errors. For all four human capital measures, we identify significant effects of the intervention for females. A one standard deviation decrease in the pre-intervention goiter rate is associated with a roughly 6% increase in educational attainment as well as a roughly 15% increase in math and verbal test scores. We also see a 12% increase in years of schooling for females. The estimated coefficients using the male sample are consistently smaller in magnitude and not significant for any of the outcome variables.

Table 2: Iodine Exposure and Human Capital Attainment

	(1) Math Test ln(scores)	(2) Verbal Test ln(scores)	(3) Educational Attainment	(4) Schooling ln(years)
Panel A: Males				
Post × Goiter	0.0250	0.0511	0.0453	0.0353
	[0.0736]	[0.0748]	[0.0555]	[0.0613]
	⟨0.0735⟩	⟨0.0752⟩	⟨0.0522⟩	⟨0.0686⟩
P-value	(0.882)	(0.665)	(0.676)	(0.678)
Mean of Dep. Var.	**2.46**	**3.05**	**2.41**	**2.02**
Observations	3026	3026	3433	3318
Panel B: Females				
Post × Goiter	0.158	0.138	0.0953	0.125
	[0.0392]***	[0.0563]**	[0.0235]***	[0.0214]***
	⟨0.0393⟩***	⟨0.0562⟩**	⟨0.0239⟩***	⟨0.0226⟩***
P-value	(0.0460)**	(0.322)	(0.0190)**	(0.0100)***
Mean of Dep. Var.	**2.49**	**3.14**	**2.47**	**2.04**
Observations	2873	2873	3220	3112

Notes: Each coefficient is from a separate regression. All regressions control for fixed effects specific to birth province and birth year, mean-reversion control, age, birth order, family size, parents' characteristics, hospitals per capita, hospital beds per capita and region-specific linear trends. Standard errors clustered by province appear in square brackets. Two-way clustered standard errors by province and family appear in angle brackets. P-values based on wild-bootstrap approach (Cameron, Gelbach, and Miller, 2008) with 999 replications appear in parenthesis. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.

5.2 Results from the Event Study

Our interpretation of the above coefficients relies on building a causal relationship between the introduction of iodized salt and changes in human capital. If provinces with high and low goiter rates had different pre-intervention trends, however, the effects our models capture may not be due to the salt iodization program. In order to test for differential pre-trends, we run regressions of the form specified in Equation (2). Figures 4, 5, 6 and 7 plot the estimated coefficients (β_t) of the event study for the different outcome measures for females. Graphs for males (Figure B2, B3, B4 and B5) are presented in Appendix B. Note from the discussion in section 4.2 that we normalize the coefficient for the 1994 cohort to be zero. Therefore the other coefficients essentially reflect differential effects of exposure relative to iodine exposure at age 1. The figures show that the trends leading up to the year of the intervention are identical and insignificant from the 1994 effect. This gives confidence in the validity of the common trend assumption in our difference-in-differences design. In Figure 4 and 5, math and verbal scores of females born in the year of the policy intervention increase substantially by about 15%, although the 95 percent confidence interval of a few cohorts includes zero. For educational outcomes (Figure 6 and 7), we see a similar picture of substantial increases shortly after the implementation of the salt iodization program (6% and 12% for educational attainment and years of schooling, respectively). While there is some variation in the effect size of the program across the cohorts, the 95 percent confidence intervals overlap for all cohorts. This suggests that females born in the later cohorts obtain comparable benefits from the adoption of iodized salt. The results for males in Appendix B point at insignificant effects for all outcome variables and all cohorts.

5.3 Robustness Analysis

There are a number of reasons why trends in educational outcomes across birth cohorts might differ across provinces. In this section, we explore some additional specification checks to make sure our that our estimates from the previous section can be interpreted as the causal effects of the iodization program. Table 3 presents additional specifications/checks. Since the previous section showed that the salt iodization policy only impacted females, we restrict these sensitivity analyses to this group.

The 1990's and 2000's have been years of economic growth in China and that may also have led to increasing educational expenditures that may have been especially beneficial for girls. The common time fixed effects of our base specification might not be sufficiently flexible to control for differential time trends across birth cohorts and regions. We therefore also ran regressions with region of birth by birth year interactions and in addition added

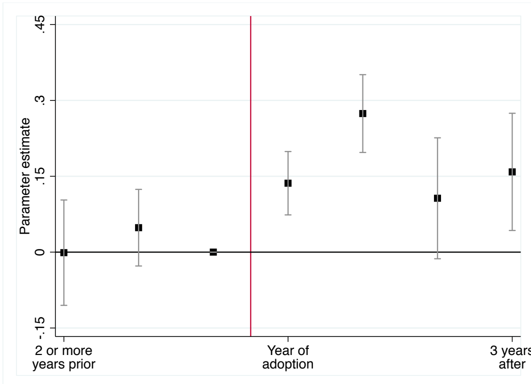

Figure 4. Math Test Scores

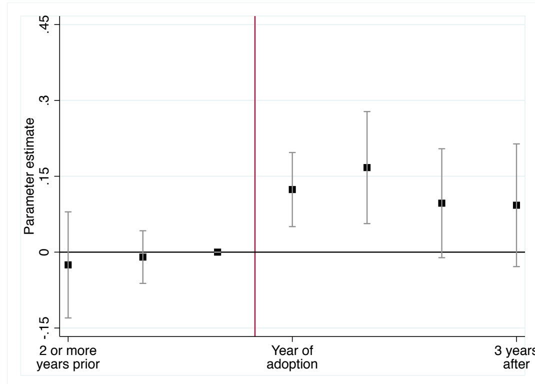

Figure 5. Verbal Test Scores

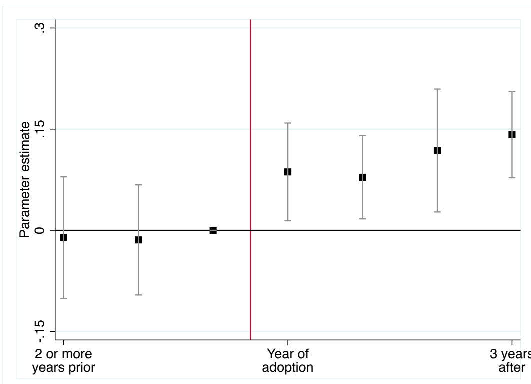

Figure 6. Educational Attainments

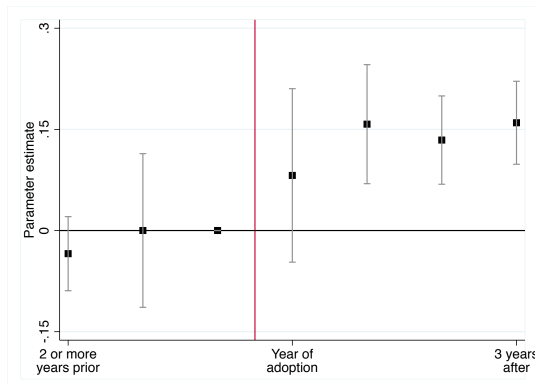

Figure 7. Schooling

Notes: The sample includes all female respondents from two waves of survey (CFPS-2010 and CFPS-2014). Each point reflects the coefficient estimated on an interaction term between the birth year (compared to 1995) and the pre-intervention (base) level of the goiter rate in the birth-province. Capped spikes represent 95 percent confidence intervals. All models condition upon birth province and birth year fixed effects and the full set of controls used in our main analysis.

the number of schools per capita at age 10. We also control for age polynomials (up to a degree of 3). The results of this exercise, reported in Panel A, show that the effects remain significant and are quantitatively similar (if not, larger) than the baseline estimates. A one standard deviation decrease in the pre-intervention goiter rate is associated with a roughly 18% increases in cognitive ability and 7-15% in educational outcomes for females.

In section 3, we assigned individuals into the exposed group if the full gestation period was after October 1994 (the date of the implementation of the program). All others, those whose were partially exposed to iodine in utero as well as those who were not exposed, were assigned to the non-exposed group. This assumes that the gestational period was exactly nine months and therefore premature births are wrongly classified as non-exposed. This also assumes that those whose gestational period partly lies after October 1994 are assumed to be not exposed in utero. This classification is perhaps a bit too conservative. Panel C reports

Table 3: Robustness Checks (Female)

	(1) Math Test ln(scores)	(2) Verbal Test ln(scores)	(3) Educational Attainment	(4) Schooling ln(years)
Panel A: baseline				
Post × Goiter	0.158	0.138	0.0953	0.125
	[0.0392]***	[0.0563]**	[0.0235]***	[0.0214]***
Mean of Dep. Var.	**2.49**	**3.14**	**2.47**	**2.04**
Observations	2873	2873	3230	3112
Panel B: additional controls				
Post × Goiter	0.166	0.176	0.0904	0.137
	[0.0411]***	[0.0573]***	[0.0264]***	[0.0295]***
Mean of Dep. Var.	**2.49**	**3.14**	**2.47**	**2.04**
Observations	2873	2873	3230	3112
Panel C: drop partial exposed group				
Post × Goiter	0.138	0.133	0.0868	0.118
	[0.0406]***	[0.0530]**	[0.0245]***	[0.0264]***
Mean of Dep. Var.	**2.48**	**3.14**	**2.46**	**2.04**
Observations	2789	2789	3132	3021
Panel D: only using baseline wave 2010				
Post × Goiter	0.146	0.131	0.0696	0.122
	[0.0349]***	[0.0592]**	[0.0328]**	[0.0317]***
Mean of Dep. Var.	**2.43**	**3.10**	**2.05**	**1.88**
Observations	1726	1726	1771	1771
Panel E: small sample window				
Post × Goiter	0.138	0.146	0.0557	0.0863
	[0.0394]***	[0.0459]***	[0.0264]**	[0.0332]**
Mean of Dep. Var.	**2.52**	**3.17**	**2.49**	**2.06**
Observations	2277	2277	2579	2469
Panel F: placebo test using poverty rates				
Post × Poverty	-0.00959	0.00580	0.0749	0.0496
	[0.0610]	[0.0415]	[0.0455]	[0.0499]
Mean of Dep. Var.	**2.49**	**3.14**	**2.47**	**2.04**
Observations	2873	2873	3230	3112

Notes: Each coefficient is from a separate regression. All regressions use the same controls as the baseline model in Table 2. In Panel B, we control for birth-region and birth-year specific interaction instead and we control for the number of schools per capita at 39 age 10 and age polynomials (up to degree 3). Standard errors clustered by province appear in square brackets. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.
Data: CFPS-2010 and CFPS-2014

the results using the sub-sample where the partial exposed children (who were born in April, May and June 1995) are dropped from the analyses. The results in Panel C show that this hardly affects the estimates.

The verbal and math test remained the same over the 2010 and 2014 wave. In the basic specification we pooled the observations from the 2010 and the 2014 wave. Individuals in our analysis are in their late teen ages/early-adulthood and therefore as we do in the regressions controlling for age is important. Still, however, using information from all waves has as a threat that individuals might learn from the test results in the first wave. This makes makes the test scores in the second wave inappropriate measurements of human capital development. As a check, we run regressions using only the data from the first wave of the CFPS. This conditioning on the first wave in 2010 is also informative on the impact of sample attrition between the first wave in 2010 and the last wave in 2014. Indeed, we lost around 30% percent of observations due to attrition (see section 3). The results in Panel D show again that the results remain close to our baseline specification and reinforces our main findings on the causal relation between iodine in utero and human capital development. We also performed the same event study design as before but now only using data from the first wave (CFPS-2010). The results are displayed in Figure B6, B7, B8 and B9 of Appendix B. Again, we can see very clear trend break at the intervention year and there do not appear to be differential pre-existing trends in the outcomes.

We also used a slightly smaller sample window by restricting individuals born between July 1992 and June 1999. The advantage is a lower attrition rate in wave 2014 as older cohorts (those born in 1991) are more likely to have left home. Moreover, the trimming of the left and right tail of the age distribution makes the sample more homogeneous. A downside of this age/cohort restriction is a smaller sample size and hence less power. Still, however, all four estimates in Panel E remain to be significant and close to the baseline estimates in Table 2.

As a last robustness check, we considered two sets of placebo analyses. The salt-iodization policy took place during a period of rapid economic growth which resulted in a substantial decline in poverty rates. This may confound the effect of the sal-iodization policy. For this reason we estimated the same baseline model using provincial poverty rates before the salt policy instead of the provincial goiter rates. The provincial poverty rates in 1993 are obtained from Woo et al. (2004) (see figure B1 of the appendix for the geographic distribution of the poverty rates). The results displayed in Panel F of Table 3 show that for all outcomes we find small and insignificant effects. For the second set of placebo tests, we use parental education (which was used as control variables in main specifications) as outcome variables. The idea of the tests are that parental education should not be affected by future exposure to iodized salt. Indeed, both mother and father's education (measured by education attainment or years of schooling) are not affected by future exposure (results available upon request).

Finally, we further verify our identification assumptions and demonstrate the statistical power of our inferences by conducting falsification tests where we assign a pseudo-treatment. More precisely, we randomly assign province of birth and thus pre-intervention goiter rates to each respondents in our sample. If our identification strategy is indeed valid, we would expect estimates using those pseudo-samples to be centered around zero. We can then confront our baseline estimates with the results from the pseudo-sample. In Figures 8, 9, 10 and 11, we plot the distribution of the t-statistics from 5,000 estimated pseudo-treatment effects on educational attainments, schooling, math and verbal ability, respectively. As expected, all four distributions are centered around zero. Taking together, these results imply that assumptions in our empirical model are unlikely be be severely violated. To access the statistical power of our model, we mark within the pseudo-treatment effect distribution the location of the t-statistic of the corresponding treatment effect using the actual pre-intervention goiter rate (baseline results). We also report the share of the pseudo-treatment t-statistics that is exceeds the actual t-statistic of the baseline model (in absolute values). These numbers (p-values) are reported at the top of each figure.[15] The p-values give confidence in our design and statistical power of our exercises.

5.4 Putting our Findings into Context: Comparison with Other Cohort-Based Iodine Studies

Some recent studies (Field, Robles, and Torero, 2009; Adhvaryu et al., 2016; Feyrer, Politi, and Weil, 2017) also analyze longer run impacts of iodine deficiency in early-life. Field, Robles, and Torero (2009) analyzes Tanzanian data to provide one of the first micro-level evidence of the influence of iodine availability in utero on cognitive development in children. Adhvaryu et al. (2016); Feyrer, Politi, and Weil (2017) exploit a nation wide salt-iodization program initiated by the public health authorities in Michigan in 1924.[16] As in our study, they use pre-program geographical information on goiter prevalence along with the time variation in the introduction of iodized salt to assess the causal effect of iodine on later life outcome. An advantage of these studies over ours is that they are able to examine the effects on a wide range of outcome variables: skill levels of drafted men (Feyrer, Politi, and Weil, 2017), years of schooling, being married, age at first marriage, income and labor force participation (Adhvaryu et al., 2016). On the other hand these studies measure goiter

[15] The p-values can be seen as alternatives to the p-values obtained from our clustered standard errors reported in Table 2.

[16] Morton's salt, the largest producer in the US at that time began selling iodized salt in the fall of 1924.

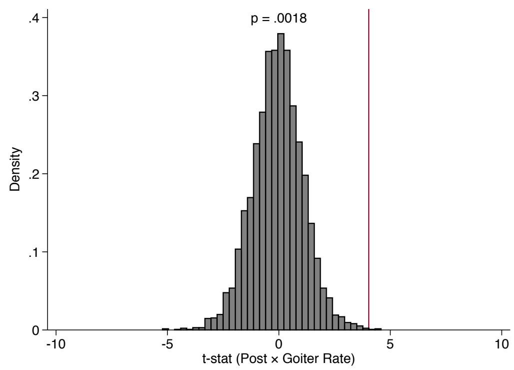

Figure 8. Math Test Scores

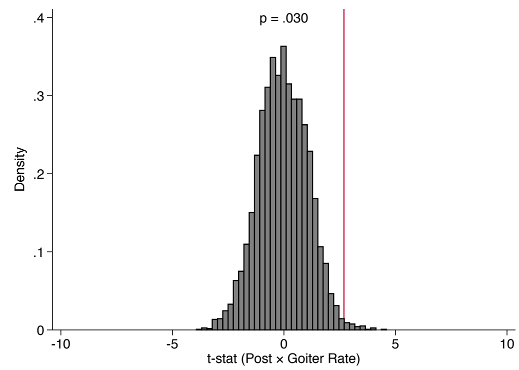

Figure 9. Verbal Test Scores

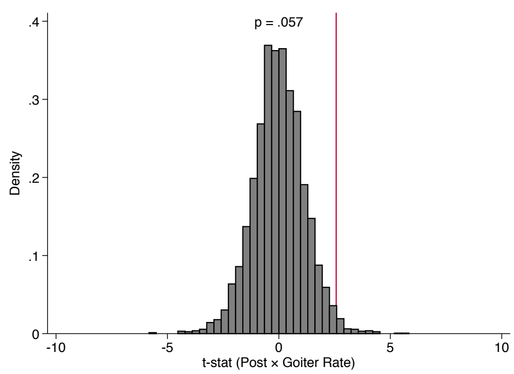

Figure 10. Educational Attainments

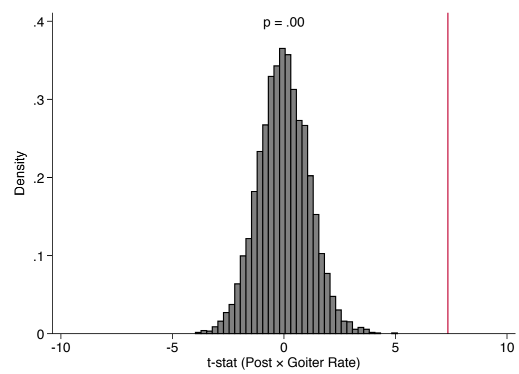

Figure 11. Schooling

Notes: Pseudo-treatment vs. actual policy intervention: the distribution of t-statistics resulting from 5,000 random assignments of treatment to individuals, as well as the t-statistics from the actual treatment through the policy intervention (red line). "p-values" report the share of the pseudo-treatment t-statistics that is larger than the actual t-statistics.

prevalence using information from military recruits, where the index population to measure goiter rates are usually young healthy men.This might introduce some measurement error. Furthermore, although the introduction of iodized salt across the US was rapid, the actual consumption rate of iodized salt (the compliance rate) was likely to be well below 100% and the compliance rate may vary across states. This may bias their difference-in-differences estimates when state compliance rates vary with (average) socio-economic outcomes. The government controlled national implementation of iodized salt in China does not suffer from that problem.

Feyrer, Politi, and Weil (2017) find strong effects for males: iodized salt in utero leads to a 15 point increase in IQ. Interestingly, Adhvaryu et al. (2016), using census data, find no effects for males, but strong increases in income (11%), labor force participation (0.68%) and full-time work (0.9%) for females. Field, Robles, and Torero (2009) also find stronger effects for girls. Like Adhvaryu et al. (2016), we only find effects for females. Relating our findings to Field, Robles, and Torero (2009) and Adhvaryu et al. (2016), we find that the introduction of iodized salt leads to 0.43 additional years of schooling (results available upon request). This estimate is in line with Field, Robles, and Torero (2009) who find that the iodine supplement program increased schooling years with 0.35-0.56 years. Wang (2013) finds for China that one year of additional schooling raises income with 15%, this translates about 6% (0.43×0.15) increases in income for females in our study.

The studies listed above concern relatively mild shocks, as opposed to more extreme nutrition shocks such as famines (see for instance in the context of China the review of Li and Lumey (2017)). In these famine studies selective mortality and fertility are more important. Selection effects are less likely to be present in mild early life shocks. Almond, Currie, and Duque (2018) provides an overview of the effects of such shocks in Table 1 of their paper. Besides iodine deficiency these mild shocks include, for instance, exposure to Ramadan fasting, nutritional supplementation, maternal health behaviors, domestic violence, elevated cortisol levels in pregnant women, etc. The findings with respect to gender are mixed. There are quite a few studies that find stronger effects for girls, but at the same time some studies find only effects for boys or no differential effects with respect to gender. The type of shock as well as contextual factors may be very important for the diverse findings of these studies. Our strong effects of iodine in utero for females is consistent with the medical literature that posits that female fetuses are more sensitive to maternal thyroid deficiency than male fetuses Zimmermann (2011). However, this does not imply absence of effects for males (see for instance, the findings of Feyrer, Politi, and Weil (2017) and Field, Robles, and Torero (2009)).

In developing countries, gender differences in socio-economic outcomes are stronger than in developed countries and cultural gender norms may contribute to gender differences in outcomes. These gender norms may already have an impact on children very early in life and might explain heterogeneous gender effects found in the literature on long run effects of early life shocks.[17] Indeed, parents' preferences for boys make these parents to desire more sons than daughters, but it may also result in parents to choose to invest more in sons than in daughters (Jayachandran, 2015). The former mechanisms may result in a male-skewed sex ratio. Almond and Edlund (2008); Abrevaya (2009) find that son preferences in Asian immigrants to the US seem to persist with changing economic environment, suggesting a strong role for culture (i.e., preferences). Important for our study is that son preferences may not only imply more material and non-material resources in boys post-birth, but also

[17] This was recently also argued by Dinkelman (2017), who refers to the effects of local shocks to the environment in the Asian context that affect resource availability. Findings for Indonesia (Maccini and Yang, 2009) and India (Pathania, 2007) point at more negative effects for girls.

that parental investments post-birth may mitigate some of the adverse effects of iodine deficiency in boys more than in girls. This would be consistent with our finding of strong effects of the salt iodization policy for girls and no effect for boys. Below we sketch the contours of a simple model on the role of parental investments which helps explaining the mixed finding in the literature with respect to gender.

6 More on Gender Differences

6.1 Conceptual Framework

To understand the interaction between gender preferences and early-life health shocks, we consider a model of child human capital formation. Our model is a simplified version of Yi et al. (2015); Almond, Currie, and Duque (2018) in which each child has two components of human capital: non-cognitive skills (N) and cognitive skills (C). The human capital production function of type $k \in \{N, C\}$ is specified as follows:

$$\theta^k = f^k(\omega^N, \omega^C, I^k, e), \tag{3}$$

where ω^k is the child's endowments of the human capital k. I^k is the parental investment in the human capital k. e, which captures the iodine deficiency in the current study, is defined as a (negative) shock affecting child's endowments of cognitive skills.

Parents value the child's outcomes and their own consumption. More specifically, parental preferences are represented by the utility function:

$$U = U(c, q) = (1 - \alpha_s)\log c + \alpha_s \log q(\theta^N, \theta^C) \tag{4}$$

where $s \in \{boys, girls\}$, c is parental consumption and q is the quality of child. Child quality is a combination of non-cognitive and cognitive skills such that $q = q(\theta^N, \theta^C)$. We consider a Cobb–Douglas utility function as Almond, Currie, and Duque (2018). In our context α_s captures the preferences for child quality. This may include cultural norms. Son preferences imply higher values of α_s for boys, $\alpha_{boys} > \alpha_{girls}$. In addition, some families may have stronger son preferences than other families, implying that there is within gender, across family, variation in α_s.[18]

[18] In the utility function of our model, the weight (α_s) for non-cognitive and cognitive skills are specified with one single parameter. However, depending on cultural norms, parents might have separate preferences for non-cognitive and cognitive skills. For example, parents could put larger weights on cognitive skills while smaller weights on non-cognitive skills. And of particular interest is that the difference can be gender-specific. However, since we are not able to separate the difference, we don't model such differences explicitly.

The budget constraint is specified as follows:

$$Y = pc + P^N I^N + P^C I^C, \tag{5}$$

where P^k is the price of investment on human capital $k \in \{N, C\}$ and p is the price for consumption. Parents adjust resource allocation in response to the shock (e) on their children. The parents' problem is to maximize the utility function (4) subject to the budget constraint (5) and the production technology (3). Therefore, the optimal human capital investment of type k is a function of a number of factors:

$$I^{k^*} = \psi^k(\omega^N, \omega^C, p, P^N, P^C, Y, e, \alpha_s). \tag{6}$$

Particularly, the optimal investment depends on the child's endowments, the price of the investment, the health shock in utero and family preferences. Note that the investments could be broadly interpreted. It could be well targeted efforts aimed to improve specific skills as well as the provision of an environment (e.g., food, housing, attention, etc.) that foster child well-being and child outcomes. In this multidimensional model of human capital formation, investment strategies of parents reflect not only parental preferences but also the production technology and the budget constraint. More importantly, parents can compensate and reinforce along different dimensions of human capital.[19]

The literature (see Almond, Currie, and Duque, 2018, for a recent review) is particularly interested in the parental response to an early-life shock (i.e., $\partial I^k/\partial e$) and how parental investments affect the formation of human capital (i.e., $\partial \theta^k/\partial I^k$). Insights into these effects can be obtained by the decomposition of the total effect of an early shock on cognition (Yi et al., 2015; Almond, Currie, and Duque, 2018):

$$\underbrace{\frac{\mathrm{d}\theta^C}{\mathrm{d}e}}_{A} = \underbrace{\frac{\partial \theta^C}{\partial e}}_{B} + \underbrace{\frac{\partial \theta^C}{\partial I^C}}_{C} \times \underbrace{\frac{\partial I^C}{\partial e}}_{D}. \tag{7}$$

The term A on the left-hand side is the total effect of an early-life shock. This corresponds usually to the reduced form estimates in the literature. In our study, prior to the salt-iodization policy, A would be the effect of iodine deficiency, i.e. the pre-intervention cognitive outcomes in late childhood due to geographical variation in iodine levels. The first term on the right-hand side (B) is the biological effect that directly operates through the production

[19] Almond and Mazumder (2013) give an example that show that compensation for the early-life shock is optimal in some settings and not in others. Moreover, the absence of behavioral responses could also be the result of a family being financial constraint. See Almond and Mazumder (2013) for more details.

function. The second term ($C \times D$) is a behavioral effect from the parental investment response post-birth. Where C is the productivity effect of the investment and D the resource allocation effect. The resource allocation effect D depends on α.[20] If parents respond with post-birth investments to counter the adversity of an early-life shock, the total effect of the shock (A) will be less than the biological effect (B). High values α_s imply more weight to child outcomes and hence imply higher levels of parental investment to counter the early-life shock. Conversely, given a biological effect B, a higher α_s for boys than for girls ($\alpha_{boys} > \alpha_{girls}$) implies more investments for boys than for girls and hence smaller reduced form effects of iodine deficiency for boys at later ages. This is consistent with findings from observational studies that girls are more sensitive to iodine deficiency (see, for instance the systematic literature review in Zimmermann (2011) or the paper by Murcia et al. (2011)).[21]

The salt iodization policy neutralizes geographically determined iodine deficiency. The policy has therefore two consequences. First, it reduces negative cognition effects of iodine deficiency for both genders. In the presence of gender preferences, the reduction in female disadvantage in cognition at birth will be larger than for boys. This is consistent with our finding of positive and significant cognition effects of the salt iodization policy for girls and small and insignificant effects for boys. Second, the salt ionization program will alter parental investment decisions. The need to invest in cognition is reduced or may even be absent and it may therefore divert (part of the) parental investments to other dimensions of human capital (notably non-cognitive skills). For example, parents may set a particular target for schooling for their children. When after the salt iodization program targets for schooling are met, parents may shift some of their attention to other non-cognitive skills. Note that this does not imply that cognitive and non-cognitive skills are substitutes and that this behavior does not contradict skill complementarity in the sense of Heckman (2007).[22] Indeed, a prediction from the model (Yi et al., 2015; Almond, Currie, and Duque, 2018) is that parents

[20] One could of course argue that the efficiency of the investment C also depends on α, for instance because families with low α do not monitor the effect of the investment. It then, however, becomes merely a matter of definition whether this should be attributed to D or C. In either case, the first order effect of preferences α will be via D.

[21] Murcia et al. (2011) looks at the effects of maternal intake of iodine supplements in Spain and finds only effects for girls.

[22] It needs to be pointed out that here that dynamic complementarities as discussed in the literature by Heckman (2007), only imposes restrictions on the functional form of the human capital production technology. Importantly, the dynamic process of human capital formation is jointly determined by parental preferences, budget constraints and the production technology. So, even if there are dynamic complementarities in the technology of human capital formation such that 'capabilities beget capabilities', parents can still respond endogenously by mitigating or reinforcing disadvantages early in life. Similar arguments have been made by Malamud, Pop-Eleches, and Urquiola (2016) and were confirmed by a series of empirical studies (Adhvaryu et al., 2015; Gunnsteinsson et al., 2018; Duque, Rosales, and Sanchez, 2018; Aguilar and Vicarelli, 2018; Rossin-Slater and Wüst, 2018).

can compensate and reinforce initial shocks along different dimensions of human capital. Or, related, via "self-productivity" (Heckman, 2007), changes in one dimension of human capital may also affect the accumulation of other dimensions.[23] Therefore, although the medical literature primarily documents neuro-developmental impairments as a consequence of in-utero exposure to iodine deficiency (Zimmermann, 2011), it cannot be ruled out that elimination of iodine deficiency also impacts non-cognitive dimensions of human capital (such as, physical health, personality traits, risk attitude, etc). The effect (sign and magnitude) of the salt iodization on non-cognitive outcomes is ultimately an empirical matter.

Ideally, one would like to have a broad set of measures for non-cognitive skills (Grönqvist et al., 2018; Ichino, Fort, and Zanella, 2019) which often use factors like: personality traits, psychological energy, emotional stability and social maturity. Unfortunately, we do not have such measures in our CFPS survey. For our cohorts born between 1991 and 2000 the CFPS survey provides information on self-reported health, a 20 items extended Centre for Epidemiological Studies Depression (CES-D) Scale and a Satisfaction index. Self-reported health is obtained from the widely used question "How good is your health". The response categories in the 2010 wave are 1 (good), 2 (fair), 3 (mild health problems), 4 (bad health), 5 (very bad health). In 2014 these are 1 (excellent), 2 (very good), 3 (good), 4 (fair), 5 (bad). From the responses in both waves we construct a bad health variable defined as fair at best (categories 2-5 in the 2010 wave) and categories 4 and 5 in the 2014 wave). A full description of the twenty items of the CES-D scale is given in Appendix D. The CES-D test score essentially measures well-being and consists of four underlying dimensions: somatic complaints (items 1, 2, 5, 7, 11, 13 and 20 range 0-21), depressed affect (items 3, 6, 9, 10, 14, 17 and 18, range 0-21), positive affect (items 4, 8, 12 and16, range 0-12) and interpersonal problems (items 15 and 19, range 0-6). The depressed affect and interpersonal problems relate to the emotional stability and social maturity scores used earlier in the literature (Grönqvist et al., 2018). The satisfaction score is derived from three questions, each with a score ranging from 1 (lowest) to 10 (highest): "How popular are you?(Popularity)"; "How satisfied are you with your life?(Happiness)"; "How well are you getting along with others?(Social Skills)". The satisfaction score was only collected in the CFPS-2014 wave and is therefore based on fewer observations. We use the baseline specification 1 to estimate the effect of the salt iodization policy on the above listed non-cognitive skills measures. The results are displayed in Table 4. The table shows that the salt-iodization policy significantly improved somatic retarded activity, depressed affect and

[23] Note, however, that in our simple model there is no role for "Self-productivity" and this is thus not captured in the decomposition (7). An example of decomposition which includes "self-productivity" can be found in Grönqvist et al. (2018).

satisfaction (though only at the 10% level) for boys. The corresponding coefficients for females are much smaller and insignificant. This finding is consistent with the idea that when the salt-iodization policy eliminates the cognitive disadvantage at birth, parental investments in boys that were initially geared towards cognitive skills are now diverted to other dimensions of human capital. However, it can not be ruled out that it is primarily the production technology of cognitive and non-cognitive skills and their interaction that drives our findings (Almond, Currie, and Duque, 2018).

We also used alternative definitions of the non-cognitive variables score. A regression using the total CES-D score has a sizable and significant coefficient of -0.934 (with a p-value of 0.045). Separate estimates for the sub-scores of the satisfaction index (see Table C1 of Appendix C) point at large and significant effects on Social Skills for boys and small and insignificant effects for girls. As to the other coefficients for males, the insignificant effects on the bad health indicator might result from the change in the response scales in the two waves. While the cut-off for the bad health indicator is taken the at the same label (fair at best), the different scales presented to the respondents may also induce them to interpret these labels differently). We therefore also estimated separate regressions by wave for this bad health indicator. These regressions show small and insignificant effects for the 2010 wave and larger and significant (at the 10% level) effects for males for the 2014 wave (coefficients of -0.034 (p-value 0.019) and -0.0011 (p-value 0.0176)) for males and females, respectively).

6.2 Gender Attitudes

To understand child human capital formation, we ideally would like to identify all three terms on the right hand side of decomposition (7) (see Yi et al., 2015, for an example). For the role of parental investment and how this interacts with shocks early in life, we can focus on the role of D. One avenue is to use direct information on parental investment, I^C and I^N. In this way we would be able to identify the term D directly. Unfortunately, this information was not collected for the respondents born between 1991 and 2000. Instead, we exploit plausibly exogenous variation on the resource allocation effect (D) generated by parental preferences (α in the utility function). Assuming that the biological effect (B) is independent of parental preferences (α), then the total effect (A) will not vary with parental preferences if there are no behavioral effects. Hence, if there are behavioral effects, then reduced form estimates that vary with gender preferences are informative on parental behavioral responses to different endowments early in life. This holds of course, assuming that $C > 0$. However, the assumption that α plays no role in the biological effect requires more discussion. With the introduction of ultra sound techniques, parents could respond

Table 4: Iodine Exposure and Non-Cognitive Skills

	(1) Somatic Complaints	(2) Depressed Affect	(3) Positive Affect	(4) Interpersonal Problems	(5) Satisfaction (index)	(6) Bad Health
Panel A: Males						
Post × Goiter	-0.381	-0.463	0.0111	-0.106	0.772	-0.00316
	[0.129]***	[0.164]***	[0.250]	[0.0792]	[0.447]*	[0.0152]
Mean of Dep. Var.	**3.35**	**2.57**	**4.81**	**0.57**	**22.3**	**0.16**
Observations	1387	1389	1389	1389	950	3438
Panel B: Females						
Post × Goiter	0.0323	-0.00457	0.116	-0.0388	-0.0150	-0.00889
	[0.174]	[0.215]	[0.162]	[0.0581]	[0.474]	[0.0183]
Mean of Dep. Var.	**3.50**	**3.23**	**4.90**	**0.65**	**22.5**	**0.18**
Observations	1327	1327	1326	1327	915	3228

Notes: Each coefficient is from a separate regression. All regressions control for fixed effects specific to birth province and birth year, mean-reversion control, age, birth order, family size, parents' characteristics, hospitals per capita, hospital beds per capita and region-specific linear trends. Standard errors clustered by province appear in square brackets. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.
Data: CFPS-2010, CFPS-2012 and CFPS-2014.

prenatally, either with selective abortion,[24] or with increased antenatal investments when a male fetus is identified (see Bharadwaj and Lakdawala, 2013). In the latter case we take the position that antenatal investments are included in parental investment decisions. Moreover, as we will show later on, our empirical model will rely on within gender variation in parental preferences. This allows us to only need the much weaker assumption that for a given gender the biological effect does not depend on α.

We do not directly observe gender preferences and therefore rely on a plausibly exogenous proxy for gender preferences of the parents. One candidate could be the sex ratio (see Edlund et al., 2013, for instance). Selective abortion may be higher in regions where preferences for boys are stronger than in other regions. A regions with a high boys to girls ratio would then point at stronger preferences for boys in that region than other regions. We do not pursue in that direction for several reasons. Firstly, to identify son preferences off the effect of the salt iodization policy we need within province variation in son preferences. Reliable within province sex ratio information is not available. Secondly and more importantly, as argued by Jayachandran (2015), the sex ratio is not a measure of son preference *per se*, but rather it is the realization of the family's son preference combined with the preferences over the family

[24] We observe a boy girl ratio of 1.07 in our sample of rural households. This is lower than the average ratio of 1.12 reported for China in Jayachandran (2015). Rural families were often exempted from the one child policy. This is also reflected in the mean family size in Table 1.

size. We therefore proceed in a different way and proxy son preferences by gender attitudes (See Dahl, Kotsadam, and Rooth, 2017; Dhar, Jain, and Jayachandran, 2018; Dossi et al., 2019).

The CFPS-2014 includes a module with a number of questions on gender equity attitudes. This module covers topics such as gender roles within the household and in public life and whether girls and boys should have equal educational opportunities. Surveyed respondents were asked to state whether they agree with six statements phrased against gender equality and women empowerment. The response categories ranged from 1 "Strongly disagree" to 5 "Strongly agree". See Table 5 for the statements and average responses for adult males and females. The gender attitude index at the two bottom rows are unweighted and weighted means of the individual responses for the six statements. A lower gender index means more gender equitable views. For the construction of the weights we refer to (Anderson, 2008).[25]

Table 5: Summary Statistics of Gender Attitudes

	Males	Females
Preferences		
Son should live together with parents.	3.58	3.69
	[1.35]	[1.36]
Every family should at least have a son.	3.73	3.72
	[1.40]	[1.44]
Husband takes care of the business, wife takes care of the family.	4.18	4.23
	[1.05]	[1.04]
Woman's marriage is more important than her career.	3.66	3.92
	[1.28]	[1.19]
Every woman should have a child.	4.18	4.39
	[1.09]	[0.95]
Disagree: Husband should do half of the housework.	1.85	1.83
	[1.04]	[1.03]
Indexes		
Unweighted gender index	3.53	3.63
	[0.66]	[0.65]
Weighted gender index	-0.079	0.076
	[1.01]	[0.98]
Number of observations	8482	8838

Notes: Sample includes all individuals born between 1905 and 1975 in CFPS-2014. Surveyed respondents were asked if they agree with these six statements. The respondents report how much they agreed with a certain statement on a scale of 1-5, with 1 being *Strongly agree* and 5 being *Strongly disagree*. Unweighted Gender index is the average of the 6 indicators. Weighted gender index is the weighted average value of the same 6 categorical variables, with weights constructed by normalizing the variables to have the same standard deviation and then recovering the weights given by the inverse covariance matrix (Anderson, 2008).

[25] The categorical variables are normalized to have the same standard deviation. The weights are subsequently recovered from the inverse of the covariance matrix (for details, see Anderson, 2008).

Table C2 of the Appendix C reports sample averages by gender for areas with strong son preferences (index exceeds the mean) and weak son preferences (index below the mean) for cohorts born before the implementation of the program (1991-1994). From the table, we can see clear gaps of school, math and verbal tests scores between females from areas with strong attitudes favoring boys and areas with weaker attitudes. For males, these gaps are much smaller, suggesting that our proxy might capture some variation of son preferences. It might also hint at the relevance of compositional differences between areas with strong and weak attitudes favoring boys. It will therefore be important to control for parental and province characteristics. Of importance for the empirical model discussed in the next section is that the within province variation of gender attitudes accounts for 2/3 of the total variation in gender attitudes in the sample. Additionally, when comparing gender attitudes in provinces with high (above median) or low (below median) pre-policy goiter rates, we find that the difference between high and low goiter regions for males and females are similar.[26]

Gender attitudes have drawn great interests in the recent literature, with some studies that explicitly link son preference to under-performance of girls (Dossi et al., 2019). Of particular importance for our study is whether these parental attitudes reflect parental preferences and relate to parental investment in children. As already mentioned above, our data do not provide information of parental investment behavior for the cohorts born between 1991 and 2000 used in the analyses of the previous sections. Fortunately, the survey does include some information about parental investments behavior for a different sample of children younger than 6 years of age. For this sample we construct a "parental investments index" by calculating an unweighted average of the z-score of 5 variables: breastfeeding practices; whether the child went to kindergarten; how often the parent read to their child; how often the parent buys books for their child; and how often they travel with their child.[27] Breastfeeding is a dichotomous variable whether the duration of breastfeeding was longer than three months. Kindergarten is an dichotomous variable indicating whether the child has attended kindergarten. The last three variables are categorical variables for the frequency of the event ranging from 0 ("never") to 5 ("every day"). We take higher values of the score to be associated with more investments in the child. It is not obvious that this also holds for the kindergarten indicator. Recent papers on the effect of day care (children aged 0-2) find mixed results. Positive effects are found for Norway (Drange and Havnes (2019) shown

[26]The difference in the weighted gender attitude index between high and low goiter regions is 0.05 and 0.07 for boys and girls, respectively.

[27]Iodine could also be obtained from food such as vegetables, eggs, fish and meat. The composition of the diet as well as the quantity of food may therefore influence the iodine intake of pregnant women and infants. The eating of healthy and nutritionally rich food could be viewed as an antenatal investment and may differ with knowledge of the gender of the unborn child. Unfortunately we do not have information on food consumption of pregnant women and infants.

significant gains in language and mathematics at age 6–7 of childcare enrollment using childcare assignment lotteries.) and Germany (Felfe and Lalive (2018) shown early child care has a strong positive effects on children's motor and socio-emotional skills.) while the recent paper by Ichino, Fort, and Zanella (2019) find negative effects, in particular for girls.[28] Note that kindergarten usually concerns older infants (age 4 and older). Also, for our sample of rural families interaction with other infants may be more beneficial and can be viewed as an investment.

For this sample of parents who have children younger than 6, we run a regression that relates the parental investment index to our proxy for gender preferences:

$$I_{ij} = \beta_0 + \beta_1 Male_i \times GA_j + \beta_2 Male_i + \beta_3 GA_j + X_i \rho + \epsilon_i, \tag{8}$$

where I_{ij} is parental investments index for child i from village/community j. $Male_i$ is dummy variable indicating the gender of the child, and GA_j is the gender attitude at the village/community j.[29] Similar to Dossi et al. (2019), we also control for a series of variables X_i such as parental education, birth order, family size and province fixed effects. Prime interest is in the parameter β_1, the association between gender preferences and parental investment in young children.

The results of the regression are reported in Table 6. In first three columns, we gradually add family background controls to the regression. In the fourth column, we also control for province fixed effects. In this specification the estimates solely rely on within province variation in gender attitudes. The estimates of β_1 are reported in the first row of Table 6 and show that gender attitudes are are strongly associated with parental investments in children and that investments in girls are lower in areas where preferences favoring boys are stronger. Adding controls does not affect the estimate of β_1. The small change in β_1 when we add province fixed effects (column 4) also underlines that there is substantial within province variation in the gender attitude variable. This is important for the next sub-section where we exploit this variation to identify the role of preferences in the effect of the iodization policy on cognitive and non-cognitive outcomes.

[28] They examine the effects of extended day care of children aged 0-2 on cognitive and non-cognitive skills. The idea is that daycare implies fewer one-to-one interactions with adults as inputs in the technology of skill formation and therefore be more harmful for infants. Girls at age 0–2 are relatively more capable of making good use of stimuli that improve their skills. Extended exposure to daycare may therefore be particularly harmful for girls.

[29] With this regression we do not aim to make causal statements. Nevertheless, in the calculation of the village/community level average we left out the individual response to minimize confounding.

Table 6: Gender Attitudes and Parental Investment

	(1)	(2)	(3)	(4)
		Parental Investment Index		
Gender Attitudes × Male	0.246	0.241	0.228	0.200
	[0.102]**	[0.0989]**	[0.0981]**	[0.0994]**
Male	-0.0779	-0.0438	-0.0411	-0.0291
	[0.0394]**	[0.0370]	[0.0368]	[0.0377]
Gender Attitudes	-0.361	-0.261	-0.214	-0.168
	[0.102]***	[0.0889]***	[0.0906]**	[0.101]*
Parental Education	No	Yes	Yes	Yes
Birth Order & Family Size	No	No	Yes	Yes
Province Fixed Effects	No	No	No	Yes
Observations	1,642	1,642	1,642	1,642

Notes: Each column is from a separate regression. Standard errors clustered by village/community appear in square brackets. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.
Data: CFPS-2010, CFPS-2012 and CFPS-2014

6.3 Gender Attitudes and the Effect of the salt Iodization Program

Assuming that the biological effect does not depend on gender preferences α, the model in section 6.1 hypothesizes that systematic variation of the reduced form estimate A with gender preferences hints at an important role for parental gender preferences in human capital formation of children. Gender preferences can be incorporated by gender specific preference weights α and to test for this we specify the following triple-difference equation:

$$Y_{ijpt} = \beta_0 + \beta_1 Post_t \times Goiter_p \times GA_j + \beta_2 Post_t \times Goiter_p + \beta_3 Post_t \times GA_j + \beta_4 GA_j + X_{ipt}\rho + \delta_p + \gamma_t + \epsilon_{ipt} \quad (9)$$

Y_{ijpt} is the human capital outcome of child i, living in village/community j of province p at time t. $Post_t$, $Goiter_p$ and GA_j are defined as before. The proxy for gender preferences GA_j is taken at the village/community level. The mean gender attitude for an individual i is calculated among adults aged 30 to 50 (approximately the same age as the parents of our target sample) for the village/community that individual i resides in.[30] We leave out the response of the parents of our main analyzing sample (i.e., cohorts born between July 1991 and June 2000) in the calculation of the village/community mean as the salt-iodization program might influence parents' gender attitudes directly. We also estimated the model with a gender attitude variable that includes responses from parents of our main analyzing sample. This did not alter the results (see Appendix C, Table C5). As in the main

[30] The distribution of the village/community level gender attitude can be seen in Figure C1.

analyses, we condition on province of residence fixed effects (δ_p). Therefore, our empirical specification absorbs differences in child human capital across provinces and solely rely on within province community/village level variation in gender preferences. By conditioning on birth cohort fixed effects (γ_t), we aim to absorb all variation across age groups. We also include X_{ipt} to control for individual, family and provincial characteristics. As robustness checks, we also consider alternative proxies of parental preferences α. Following the approach in (Dahl and Moretti, 2008; Dossi et al., 2019), we approximate son preferences by fertility stopping patterns. We find similar results. More details can be found in section C.

The main coefficient of interest is β_1 and as we estimate the model separately by gender, the coefficient measures the within gender differential effect of the iodine fortification program across families with varying levels of son preferences. Mapping β_1 to the decomposition (7), the first term (B) on the right-hand side is the biological effect, which we assume to be homogeneous within gender and thus independent from parental gender preferences (α). As a consequence, β_1 will reflect the extent to which the behavioral response ($C \times D$) depends on α. We expect preferences to primarily have a role via D, the parental investment response to the salt-iodization program. The term C is the efficiency of the investment and depends on the health production technology. While C might be related to gender preferences α, it is expected that α plays a more direct role in D.[31] Note that the Gender Attitude variable is normalized with zero mean and therefore the estimate of β_2 can be compared with the reform effect of section 5.

Table 7 reports the results for by gender for cognitive outcomes. For females, high values of GA mean lower values of α in the theoretical model. The estimates of β_1 are reported in the first row of panel A (for males) and B (for females). The estimates of β_1 show large and significant effects for females (specifically for Math and Schooling) and small and insignificant effects for males. Girls residing in villages/communities with strong preferences for boys benefit more from the universal salt iodization program than otherwise similar girls in communities/villages with less strong preferences for boys. This suggests that at least part of the differential effects by gender of the effect of the salt iodization program on cognitive outcomes is driven by son preferences of parents. The much smaller (and not significant) effects for boys and strong effects for girls also suggests that the parental investments (D) in cognition in boys are partially crowded out by the universal salt-iodization program. This, however, does not imply that the salt-iodization program did not affect parental investment behavior in boys. Parents could divert their investments in cognitive skills for boys to other

[31] The efficiency of the investment may vary with gender preferences if, for instance, parents do not monitor the effect of the investment. It should be noted, however, that when this kind of 'neglect' is important, why parents would invest in the first place in the child. It is expected that an effect of α on C (if present) should be a second order effect.

Table 7: Gradients in Long Run Impacts of Iodine Exposure by Gender Attitudes

	(1) Math Test ln(scores)	(2) Verbal Test ln(scores)	(3) Educational Attainment	(4) Schooling ln(years)
Panel A: Males				
Post × Goiter × Gender Attitudes	-0.0917	-0.0367	-0.0654	0.0518
	[0.0970]	[0.0723]	[0.0744]	[0.0998]
Post × Gender Attitudes	0.0925	0.0179	0.116	-0.0302
	[0.221]	[0.197]	[0.128]	[0.143]
Post × Goiter	0.0548	0.0740	0.0525	0.0219
	[0.0676]	[0.0690]	[0.0469]	[0.0517]
Mean of Dep. Var.	**2.46**	**3.05**	**2.37**	**2.01**
Observations	2841	2841	3204	3098
Panel B: Females				
Post × Goiter × Gender Attitudes	0.272	0.272	0.0534	0.157
	[0.132]**	[0.186]	[0.0354]	[0.0614]**
Post × Gender Attitudes	-0.295	-0.344	0.00170	-0.136
	[0.163]*	[0.191]*	[0.0842]	[0.106]
Post × Goiter	0.112	0.104	0.0709	0.100
	[0.0396]***	[0.0428]**	[0.0282]**	[0.0246]***
Mean of Dep. Var.	**2.49**	**3.14**	**2.42**	**2.02**
Observations	2696	2696	2988	2891

Notes: Each coefficient is from a separate regression. All regressions control for fixed effects specific to birth province and birth year, mean-reversion control, age, age square, birth order, family size, parents' characteristics, hospitals per capita, hospital beds per capita and region-specific linear trends. Standard errors clustered by province appear in square brackets. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.
Data: CFPS-2010 and CFPS-2014

dimensions of human capital, notably non-cognitive skills such as physical health, mental health, social skills, non-cognitive skills that enhance labor market outcomes etc. Indeed, the results in Table 4 showed that boys do benefit in some of our measures for non-cognitive skills. To see whether the program effects on non-cognitive factors vary with gender preferences, we also estimated the triple-difference model for the set of non-cognitive variables available for our cohorts (born between 1991-2000). This includes the sub scores of the CES-D test, the satisfaction index and an indicator for bad health (Table C6 of the Appendix) and the components "Popularity", "Happiness" and "Social Skills" of the Satisfaction index (C7 of the Appendix).

The estimates in the third row of Panel A (Boys) and B (Girls) in Tables C6 and C7, reflecting the effects of the implementation of the salt iodization program (β_2) are in line with the findings in Table 4. However, none of the effect estimates of β_1 in the first row of Panel A and B are significant for our set of non-cognitive skill measures, suggesting that the effect of the salt-iodization program does not vary by gender attitude. An explanation for this can be found in the primary impact of the salt-iodization program. In utero exposure to iodine deficiency has primarily effects on fetal brain development and therefore the reform's first order effects is on cognition. The reform also affects the allocation of resources over the different dimensions of human capital. This second order effect is reflected in the estimates of the program effect β_2. Therefore, to what extend the program effect on non-cognitive skills for boys (or girls) varies within gender by gender preferences (β_1) can be considered to be a third order effect. Indeed, this effect might be small. It might be small for data reasons as well. It might be too demanding for the data and the measures of non-cognitive skills that we have at our disposal.

7 Conclusion

Currently, around about one third of the world population amounting to about 2 billion people suffer from iodine deficiency. The medical literature documents that iodine deficiency can lead to neuro developmental problems when fetuses are exposed in utero. This can lead to a reduction in cognitive skills in school aged children and consequently to adverse labor market outcomes later in life. This paper is the first to evaluate the effect of a nationally implemented salt iodization program on cognition of school aged children in China. We do this in a difference-in-differences framework that measures the effect of the program by pre-program, geographically determined iodine deficiency rates as measured by provincial goiter rates. We find strong positive effects of the program for girls. A one standard deviation decrease (12%) in the pre-intervention children goiter rate is associated with math and vocabulary scores increasing by roughly 15%. We also see large increases in the educational attainment of females. Yet, we do not find any effects for boys. These findings are robust against alternative specifications and falsification tests. These findings thus support the effectiveness of an important, low costs, public health intervention. Indeed, compared to other interventions to raise education, the cost of salt iodization is extremely low. The costs associated with the intervention are about 0.05$ per person per year (WHO, 2005). This contrasts sharply with other interventions, such as class size reductions, costing over $5,000 (2010 dollars) per year per student (Chetty et al., 2011).

We proceed by further looking into these gender specific findings and suggest that gender preferences may play a role in our findings. We do this in the context of a simple model along the lines of Yi et al. (2015); Almond, Currie, and Duque (2018) that explain how parental investments respond to (adverse) effects of shocks early in life. In this model, the total effect of the early life shock can be decomposed into a biological effect and a behavioral effect. The behavioral effect includes the effects of the parental investment decision and mitigates or reinforces the biological effect. This behavioral ('resource allocation effect') depends on child preferences that may differ by gender. This simple model can explain the policy effect on cognition for girls and absence of an effect on cognition for boys. Prior to the salt iodization policy parents may have countered adverse initial shocks for boys and less so for girls. Therefore, when boy preferences are important, girls may benefit more from a nationally implemented programs. This also suggest that the salt-iodization program crowds out private parental investments in cognition. Another consequence of this may be that parents divert their investments into other skill dimensions, notably for boys. Indeed, we find program effects on non-cognitive skills for boys, but not for girls.

We next examine the role of gender preferences in explaining our findings on cognitive and non-cognitive skills. We proxy gender preferences with an index for gender attitudes and estimate a triple-difference model. The model identifies the effect of gender preferences from within gender variation in community/village gender attitudes. We find that, in line with the model and expectations, for girls program effects on cognition are stronger in communities/villages where preferences favoring boys are stronger. We do not find such effects for boys. Nor do we find such triple-difference for non-cognitive skills, both for boys and girls. Cultural or other contextual factors that feed attitudes favoring boys are not restricted to China, but also hold for other Asian countries, the Middle East, and North Africa. The idea that gender preferences play a role in the western world can not be excluded either. Our findings therefore do not only speak to the external validity of the current study, but also suggest that later life gender differences in labor market outcomes observed in a lot of countries may be rooted early in life.

Preferences favoring boys may lead to unequal investment in girls and justify policies that aim to reduce the consequences of these preferences. Improved economic circumstances may reduce such gender preferences, but as earlier work has shown (Almond and Edlund, 2008; Abrevaya, 2009) gender biases persist even with improved economic circumstances. This then calls for policies such as increasing mandatory schooling for both genders and other public programs that affect all, such as the public health campaign in this paper. Still, of course for an important part these general investments can only improve outcomes as long as the investments translates into human capital; in terms of the model, the efficiency of investment (i.e, the term C in the decomposition (7)). With low values in the efficiency of investment, the impact of any investment (public or private) remains modest. The key of understanding the factors that drive the efficiency of the investment effect lies in understanding the human capital production technology. Indeed, this is one of the topics that is in need for further development (see e.g. Cunha, Heckman, and Schennach, 2010; Agostinelli and Wiswall, 2016;

Attanasio, Meghir, and Nix, 2018). Unfortunately, with the data at hand we can not tackle this problem in a satisfactory way and therefore leave this to future research.

References

Abrevaya, Jason. 2009. "Are There Missing Girls in the United States? Evidence from Birth Data." *American Economic Journal: Applied Economics* 1 (2):1–34.

Adhvaryu, Achyuta, Steven Bednar, Teresa Molina, Quynh Nguyen, and Anant Nyshadham. 2016. "When It Rains It Pours: The Long-run Economic Impacts of Salt Iodization in the United States." *Working Paper* .

Adhvaryu, Achyuta, Teresa Molina, Anant Nyshadham, and Jorge Tamayo. 2015. "Helping Children Catch Up: Early Life Shocks and the Progresa Experiment." *Working Paper* .

Agostinelli, Francesco and Matthew Wiswall. 2016. "Estimating the Technology of Children's Skill Formation." *Working Paper* .

Aguilar, Arturo and Marta Vicarelli. 2018. "El Nino and Mexican Children: Medium-Term Effects of Early-Life Weather Shocks on Cognitive and Health Outcomes." *Working Paper* .

Almond, Douglas. 2006. "Is the 1918 Influenza Pandemic Over? Long-term Effects of in Utero Influenza Exposure in the Post-1940 U.S. Population." *Journal of Political Economy* 114 (4):672–712.

Almond, Douglas and Janet Currie. 2011. "Killing Me Softly: The fetal origins hypothesis." *Journal of Economic Perspectives* 25 (3):153–72.

Almond, Douglas, Janet Currie, and Valentina Duque. 2018. "Childhood Circumstances and Adult Outcomes: Act II." *Journal of Economic Literature* 56 (4):1360–1446.

Almond, Douglas and Lena Edlund. 2008. "Son-biased Sex Ratios in the 2000 United States Census." *Proceedings of the National Academy of Sciences* 105 (15):5681–5682.

Almond, Douglas and Bhashkar Mazumder. 2013. "Fetal Origins and Parental Responses." *Annual Review of Economics* 5 (1):37–56.

Almond, Douglas, Bhashkar Mazumder, and Reyn van Ewijk. 2015. "In Utero Ramadan Exposure and Children's Academic Performance." *Economic Journal* 125 (589):1501–1533.

Anderson, Michael L. 2008. “Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects.” *Journal of the American statistical Association* 103 (484):1481–1495.

Attanasio, Orazio, Costas Meghir, and Emily Nix. 2018. “Human Capital Development and Parental Investment in India.” *Working Paper* .

Barcellos, Silvia Helena, Leandro S Carvalho, and Adriana Lleras-Muney. 2014. “Child Gender and Parental Investments in India: Are Boys and Girls Treated Differently?” *American Economic Journal: Applied Economics* 6 (1):157–189.

Baumann, Eugen. 1896. “Ueber das normale Vorkommen von Jod im Thierkörper.(I. Mittheilung).” *Hoppe-Seyler's Zeitschrift für physiologische Chemie* 21 (4):319–330.

Bharadwaj, Prashant and Leah K Lakdawala. 2013. “Discrimination Begins in the Womb: Evidence of Sex-Selective Prenatal Investments.” *Journal of Human Resources* 48 (1):71–113.

Bharadwaj, Prashant, Petter Lundborg, and Dan-Olof Rooth. 2017. “Birth Weight in the Long Run.” *Journal of Human Resources* :0715–7235.

Bleakley, Hoyt. 2010a. “Health, Human Capital, and Development.” *Annual Review of Economics* 2 (1):283–310.

———. 2010b. “Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure.” *American Economic Journal: Applied Economics* 2 (2):1–45.

Brown, D, A Kowalski, and I Lurie. 2018. “Long-Term Impacts of Childhood Medicaid Expansions on Outcomes in Adulthood.” *Working Paper* .

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller. 2008. “Bootstrap-Based Improvements for Inference with Clustered Errors.” *The Review of Economics and Statistics* 90 (3):414–427.

Cao, Xue-Yi, Xin-Min Jiang, Zhi-Hong Dou, Murdon Abdul Rakeman, Ming-Li Zhang, Karen O'donnell, Tai Ma, Kareem Amette, Nancy DeLong, and G Robert DeLong. 1994. “Timing of Vulnerability of the Brain to Iodine Deficiency in Endemic Cretinism.” *New England Journal of Medicine* 331 (26):1739–1744.

Chen, Junshi and Huiyun Wu. 1998. “Fortification of Salt with Iodine.” *Food and Nutrition Bulletin* 19.

Chetty, Raj, John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR." *The Quarterly Journal of Economics* 126 (4):1593–1660.

Cunha, Flavio, James J Heckman, and Susanne M Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3):883–931.

Dahl, Gordon B, Andreas Kotsadam, and Dan-Olof Rooth. 2017. "Does Integration Change Gender Attitudes?" *Working Paper* .

Dahl, Gordon B and Enrico Moretti. 2008. "The Demand for Sons." *The Review of Economic Studies* 75 (4):1085–1120.

Dhar, Diva, Tarun Jain, and Seema Jayachandran. 2018. "Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India." *Working Paper* .

Dinkelman, Taryn. 2017. "Long-Run Health Repercussions of Drought Shocks: Evidence from South African Homelands." *The Economic Journal* 127 (604):1906–1939.

Dossi, Gaia, David N Figlio, Paola Giuliano, and Paola Sapienza. 2019. "Born in the Family: Preferences for Boys and the Gender Gap in Math." *Working Paper* .

Drange, Nina and Tarjei Havnes. 2019. "Early Childcare and Cognitive Development: Evidence from an Assignment Lottery." *Journal of Labor Economics* 37 (2):581–620.

Duque, V, MF Rosales, and F Sanchez. 2018. "How Do Early-Life Shocks Interact with Subsequent Human-Capital Investments? Evidence from Administrative Data." *Working Paper* .

Edlund, Lena, Hongbin Li, Junjian Yi, and Junsen Zhang. 2013. "Sex Ratios and Crime: Evidence from China." *Review of Economics and Statistics* 95 (5):1520–1534.

Felfe, Christina and Rafael Lalive. 2018. "Does Early Child Care Affect Children's Development?" *Journal of Public Economics* 159:33–53.

Feyrer, James, Dimitra Politi, and David N Weil. 2017. "The Cognitive Effects of Micronutrient Deficiency: Evidence from Salt Iodization in the United States." *Journal of the European Economic Association* 15 (2):355–387.

Field, Erica, Omar Robles, and Maximo Torero. 2009. "Iodine Deficiency and Schooling Attainment in Tanzania." *American Economic Journal: Applied Economics* 1 (4):140–169.

Figlio, David, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth. 2014. "The Effects of Poor Neonatal Health on Children's Cognitive Development." *American Economic Review* 104 (12):3921–3955.

Grönqvist, Hans, J Peter Nilsson, Per-Olof Robling et al. 2018. "Early Lead Exposure and Outcomes in Adulthood." *Working Paper* .

Gunnsteinsson, Snaebjorn, Achyuta Adhvaryu, Parul Christian, Alain Labrique, Jonathan Sugimoto, Abu Ahmed Shamim, and Keith P West Jr. 2018. "Protecting Infants from Natural Disasters: The Case of Vitamin A Supplementation and a Tornado in Bangladesh." *Working Paper* .

Heckman, James J. 2007. "The Economics, Technology, and Neuroscience of Human Capability Formation." *Proceedings of the National Academy of Sciences* 104 (33):13250–13255.

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-Run Impacts of Childhood Access to the Safety Net." *American Economic Review* 106 (4):903–934.

Ichino, Andrea, Margherita Fort, and Giulio Zanella. 2019. "Cognitive and Non-Cognitive Costs of Daycare 0-2 for Children in Advantaged Families." *Working Paper* .

Jayachandran, Seema. 2015. "The Roots of Gender Inequality in Developing Countries." *Annual Review of Economics* 7 (1):63–88.

Jayachandran, Seema and Ilyana Kuziemko. 2011. "Why Do Mothers Breastfeed Girls Less Than Boys? Evidence and Implications for Child Health in India." *The Quarterly Journal of Economics* 126 (3):1485–1538.

Li, Chihua and LH Lumey. 2017. "Exposure to the Chinese Famine of 1959–61 in Early Life and Long-term Health Conditions: A Systematic Review and Meta-analysis." *International Journal of Epidemiology* 46 (4):1157–1170.

Lumey, L.H., Aryeh D. Stein, and Ezra Susser. 2011. "Prenatal Famine and Adult Health." *Annual Review of Public Health* 32 (1):237–262.

Maccini, Sharon and Dean Yang. 2009. "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall." *American Economic Review* 99 (3):1006–1026.

Malamud, Ofer, Cristian Pop-Eleches, and Miguel Urquiola. 2016. "Interactions Between Family and School Environments: Evidence on Dynamic Complementarities?" *Working Paper* .

Murcia, Mario, Marisa Rebagliato, Carmen Iniguez, Maria-Jose Lopez-Espinosa, Marisa Estarlich, Belén Plaza, Carmen Barona-Vilar, Mercedes Espada, Jesus Vioque, and Ferran Ballester. 2011. "Effect of Iodine Supplementation During Pregnancy on Infant Neurodevelopment at 1 Year of Age." *American Journal of Epidemiology* 173 (7):804–812.

Oster, Emily. 2009. "Does Increased Access Increase Equality? Gender and Child Health Investments in India." *Journal of Development Economics* 89 (1):62–76.

Pathania, Vikram. 2007. "The Long Run Impact of Drought at Birth on Height of Women in Rural India." *Working Paper* .

Rossin-Slater, Maya and Miriam Wüst. 2018. "What is the Added Value of Preschool? Long-Term Impacts and Interactions with an Infant Health Intervention." *Working Paper* .

Shah, Manisha and Bryce Millett Steinberg. 2017. "Drought of Opportunities: Contemporaneous and Long-Term Impacts of Rainfall Shocks on Human Capital." *Journal of Political Economy* 125 (2):527–561.

Simoons, Frederick J. 1990. *Food in China: A Cultural and Historical Inquiry*. CRC Press.

Sun, Dianjun. 2018. *Endemic Disease in China*. Springer.

Van den Berg, Gerard J., Maarten Lindeboom, and France Portrait. 2006. "Economic Conditions Early in Life and Individual Mortality." *American Economic Review* 96 (1):290–302.

Wang, Le. 2013. "Estimating Returns to Education When the IV Sample is Selective." *Labour Economics* 21:74 – 85.

Woo, Wing Thye, Shi Li, Ximing Yue, Harry Xiaoying Wu, and Xinpeng Xu. 2004. "The Poverty Challenge for China in the New Millennium." *The Millennium Development Goals Project of the United Nations* .

Yi, Junjian, James J Heckman, Junsen Zhang, and Gabriella Conti. 2015. "Early Health Shocks, Intra-household Resource Allocation and Child Outcomes." *The Economic Journal* 125 (588):F347–F371.

Zhang, Junsen. 2017. "The Evolution of China's One-Child Policy and Its Effects on Family Outcomes." *Journal of Economic Perspectives* 31 (1):141–60.

Zimmermann, Michael B. 2008. "Research on Iodine Deficiency and Goiter in the 19th and Early 20th Centuries." *Journal of Nutrition* 138 (11):2060–2063.

———. 2011. "The Role of Iodine in Human Growth and Development." In *Seminars in Cell & Developmental Biology*, vol. 22. Elsevier, 645–652.

Zimmermann, Michael B, Sonja Y Hess, Pierre Adou, Toni Toresanni, Rita Wegmüller, and Richard F Hurrell. 2003. "Thyroid Size and Goiter Prevalence after Introduction of Iodized Salt: A 5-y Prospective Study in Schoolchildren in Côte d'Ivoire." *The American Journal of Clinical Nutrition* 77 (3):663–667.

Appendix A Additional Information

Table A1: Regional Classification of Provinces

Region	Provinces
North China	Beijing, Tianjin, Hebei, Shanxi and Inner Mongolia
Northeast China	Liaoning, Jilin and Heilongjiang
East China	Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi and Shandong
South Central China	Henan, Hubei, Hunan, Guangdong, Guangxi and Hainan
Southwest China	Chongqing, Sichuan, Guizhou, Yunnan and Tibet
Northwest China	Shaanxi, Gansu, Qinghai, Ningxia and Xinjiang

Notes: Author's tabulations

Table A2: Summary Statistics

	High Goiter Provinces		Low Goiter Provinces	
	Females	Males	Females	Males
Outcomes				
Educational Attainment	2.08	2.00	2.09	2.02
	[0.83]	[0.82]	[0.85]	[0.82]
Illiterate	0.27	0.31	0.28	0.29
Primary School	0.42	0.40	0.38	0.43
Middle School	0.27	0.26	0.29	0.25
High School or above	0.046	0.027	0.045	0.041
Schooling	7.34	7.10	7.47	7.33
	[2.73]	[2.61]	[2.62]	[2.47]
Math Test Scores	12.2	11.9	12.6	12.2
	[5.23]	[5.31]	[5.19]	[4.80]
Verbal Test Scores	23.9	22.2	24.2	22.5
	[7.28]	[7.79]	[7.01]	[7.29]
Demographics				
Age	14.5	14.3	14.3	14.4
	[2.52]	[2.50]	[2.47]	[2.47]
Father's Education	2.19	2.20	2.46	2.48
	[0.94]	[0.95]	[0.87]	[0.88]
Mother's Education	1.68	1.69	2.16	2.13
	[0.85]	[0.83]	[0.92]	[0.89]
Birth Order	1.56	1.69	1.62	1.72
	[0.73]	[0.87]	[0.95]	[0.92]
Family Size	5.13	4.79	5.10	4.79
	[1.57]	[1.45]	[1.67]	[1.58]
Additional Controls				
No. of hospitals per capita, 0-3 avg.	15.7	15.8	19.2	18.5
	[5.13]	[5.14]	[10.7]	[10.4]
No. of hospital beds per capita, 0-3 avg.	21.3	21.3	27.9	26.9
	[3.62]	[4.14]	[11.0]	[10.2]
Number of observations	956	1028	997	1007

Notes: Author's tabulations of CFPS-2010 and CFPS-2014. Sample consists individuals born in rural area between July 1995 and June 2000. We label a province as high/low goiter if its goiter rate is above/below 17%.

Appendix B Additional Results: Iodine and Long-Run Outcomes

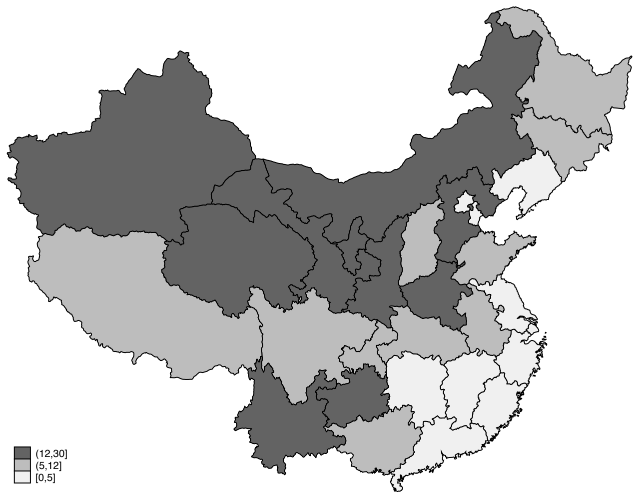

Figure B1. Rural Poverty in 1993

Notes: Figure B1 reports rural poverty rates (%) in 1993. Darker areas represent higher poverty rates.
Sources: Woo et al. (2004)

Table B1: Robustness Checks (Male)

	(1) Math Test ln(scores)	(2) Verbal Test ln(scores)	(3) Educational Attainment	(4) Schooling ln(years)
Panel A: baseline				
Post × Goiter	0.0250	0.0511	0.0455	0.0367
	[0.0736]	[0.0748]	[0.0558]	[0.0620]
Mean of Dep. Var.	**2.46**	**3.05**	**2.40**	**2.02**
Observations	3026	3026	3442	3317
Panel B: additional controls				
Post × Goiter	0.0560	0.0449	0.0673	0.0500
	[0.0476]	[0.0558]	[0.0504]	[0.0544]
Mean of Dep. Var.	**2.46**	**3.05**	**2.40**	**2.02**
Observations	3026	3026	3442	3317
Panel C: drop partial exposed group				
Post × Goiter	0.0240	0.0475	0.0450	0.0350
	[0.0770]	[0.0792]	[0.0551]	[0.0614]
Mean of Dep. Var.	**2.45**	**3.05**	**2.39**	**2.02**
Observations	2951	2951	3345	3223
Panel D: only using baseline wave 2010				
Post × Goiter	0.0291	0.0715	-0.0123	0.0244
	[0.0750]	[0.0871]	[0.0590]	[0.0580]
Mean of Dep. Var.	**2.42**	**3.01**	**2.02**	**1.89**
Observations	1830	1830	1866	1866
Panel E: small sample window				
Post × Goiter	0.0453	0.0402	0.0322	0.00806
	[0.0661]	[0.0743]	[0.0472]	[0.0555]
Mean of Dep. Var.	**2.49**	**3.08**	**2.42**	**2.05**
Observations	2336	2336	2678	2569
Panel F: placebo test using poverty rates				
Post × Poverty	-0.0203	0.0226	-0.0740	-0.162
	[0.0800]	[0.0876]	[0.0549]	[0.0494]***
Mean of Dep. Var.	**2.46**	**3.05**	**2.40**	**2.02**
Observations	3026	3026	3442	3317

Notes: Each coefficient is from a separate regression. All regressions use the same controls as the baseline model in Table 2. In Panel B, we control for birth-region and birth-year specific interaction instead and we control for the number of schools per capita at age 10 and age polynomials (up to degree 3). Standard errors clustered by province appear in square brackets. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.
Data: CFPS-2010 and CFPS-2014

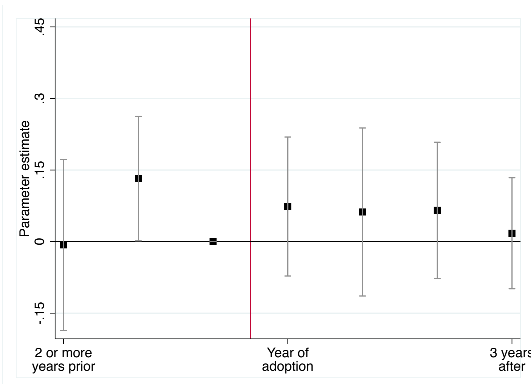

Figure B2. Math Test Scores

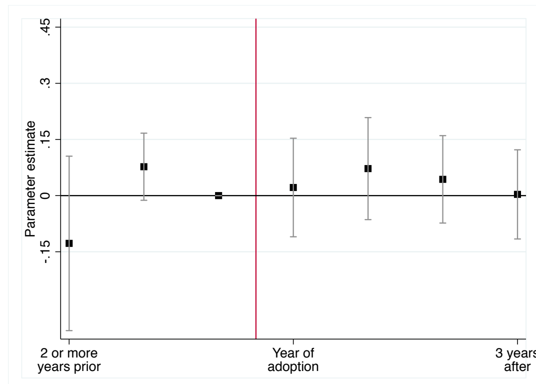

Figure B3. Verbal Test Scores

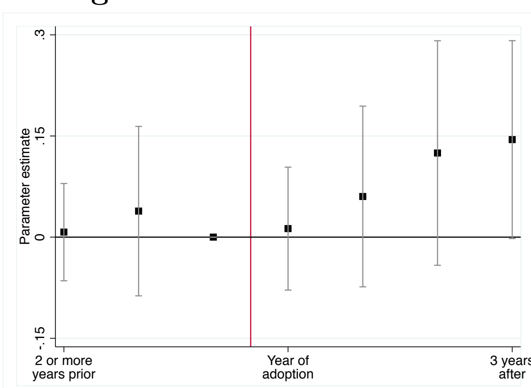

Figure B4. Educational Attainments

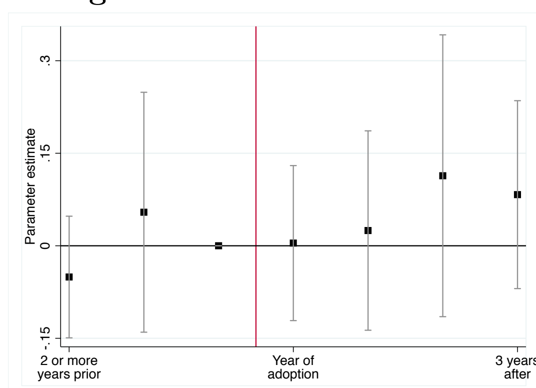

Figure B5. Schooling

Notes: The sample includes all male respondents from two waves of survey (CFPS-2010 and CFPS-2014). Each point reflects the coefficient estimated on an interaction term between the birth year (compared to 1995) and the pre-intervention (base) level of the goiter rate in the birth-province. Capped spikes represent 95 percent confidence intervals. All models condition upon birth province and birth year fixed effects and the full set of controls used in our main analysis.

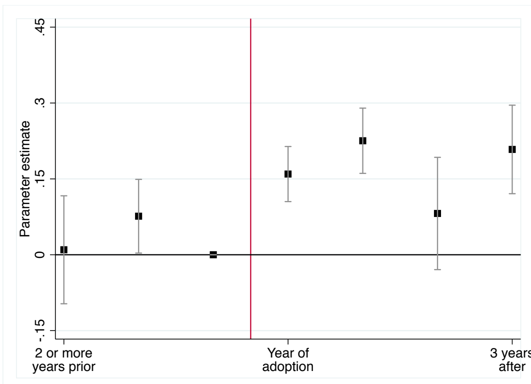

Figure B6. Math Test Scores

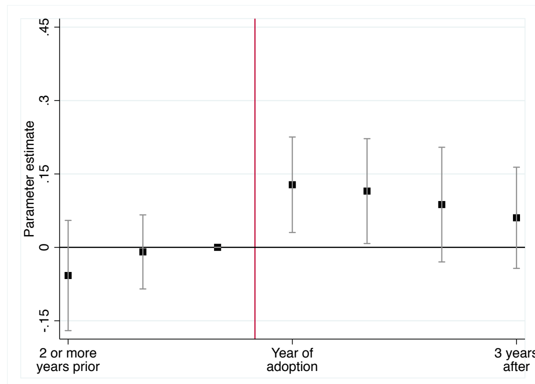

Figure B7. Verbal Test Scores

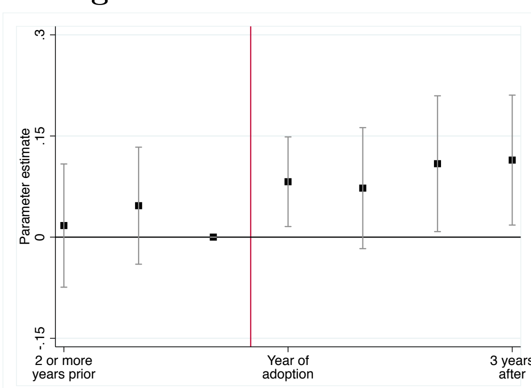

Figure B8. Educational Attainments

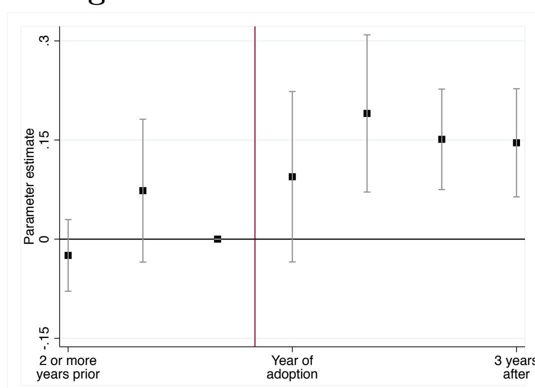

Figure B9. Schooling

Notes: The sample includes all female respondents from two waves of survey (CFPS-2010 and CFPS-2014). Each point reflects the coefficient estimated on an interaction term between the birth year (compared to 1995) and the pre-intervention (base) level of the goiter rate in the birth-province. Capped spikes represent 95 percent confidence intervals. All models condition upon birth province and birth year fixed effects and the full set of controls used in our main analysis.

Appendix C Additional Results: More on Gender

C.1 Additional Summary Statistics

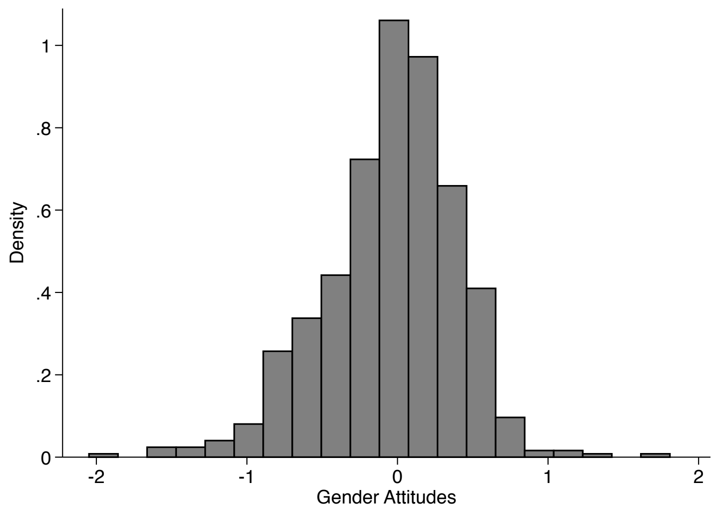

Figure C1. Distribution of Gender Attitudes across Villages/Communities

Notes: Gender attitudes across villages/communities are the mean of weighted gender index at village/community level. The response of the parents of our main analyzing sample (i.e., cohorts born between July 1991 and June 2000) are leaved out in the calculation of mean.

Table C1: Iodine Exposure and Non-Cognitive Skills

	(1) Popularity	(2) Happiness	(3) Social Skills	(4) Satisfaction (index)
Panel A: Males				
Post × Goiter	0.0252	0.302	0.462	0.772
	[0.164]	[0.218]	[0.189]**	[0.447]*
Mean of Dep. Var.	**7.32**	**7.76**	**7.25**	**22.3**
Observations	954	950	950	950
Panel B: Females				
Post × Goiter	0.0538	-0.0693	-0.00890	-0.0150
	[0.202]	[0.177]	[0.170]	[0.474]
Mean of Dep. Var.	**7.17**	**8.07**	**7.21**	**22.5**
Observations	915	918	918	915

Notes: Each coefficient is from a separate regression. All regressions control for fixed effects specific to birth province and birth year, mean-reversion control, age, birth order, family size, parents' characteristics, hospitals per capita, hospital beds per capita and region-specific linear trends. Standard errors clustered by province appear in square brackets. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.
Data: CFPS-2014.

Table C2: Summary Statistics

	Strong Son Preference		Weak Son Preference	
	Females	Males	Females	Males
Outcomes				
Educational Attainment	2.97	2.92	3.10	2.96
	[0.84]	[0.85]	[0.76]	[0.77]
Illiterate	0.044	0.047	0.019	0.023
Primary School	0.23	0.26	0.18	0.25
Middle School	0.43	0.41	0.47	0.48
High School or above	0.30	0.28	0.32	0.25
Schooling	9.83	9.60	10.5	10.0
	[3.08]	[3.22]	[2.68]	[2.62]
Math Test Scores	15.0	14.9	16.3	15.4
	[5.52]	[5.83]	[5.01]	[5.41]
Verbal Test Scores	26.0	24.7	27.2	25.6
	[7.10]	[8.05]	[6.06]	[6.83]
Number of observations	702	788	623	689

Notes: Author's tabulations of CFPS-2010 and CFPS-2014. Sample consists individuals born in rural area between July 1991 and June 1995. We label a village/community as strong/weak son preferences if its gender attitude is above/below 0.

C.2 Additional Results using Fertility Patterns

An alternative proxy of son preferences used in the literature is the fertility stopping pattern (Dahl and Moretti, 2008; Dossi et al., 2019). As a robustness check, we approximated son preferences by the fertility stopping pattern and did a similar heterogeneity analysis. Following the approach in Dossi et al. (2019), we create a dummy variable for "boy biased" families equal to 1 if the family has at least 3 children and all children are girls except for the last born in. The dummy variable equal to 0 for all the other families. Our measures of fertility patterns are noisy, as a particular fertility outcome could simply be a reflection of randomness. However, the One Child Policy (OCP) was implemented in 1979 and later on in the middle of the 1980s modified to allow rural parents a second child if the first was a daughter. Despite the fact that the OCP was difficult to enforce in rural areas, rural families wanted to have more kids did bear larger costs. Our additional restriction for "boy biased" families to have at least 3 children singles out those families who really want to have a son. Therefore, our proposed measure of fertility decisions should decrease the impact of randomness substantially. According to this rule, around 10% of our main sample were born in a "boy biased" family. We estimate the regression (9) using the dummy variable for "boy biased" as the proxy of son preferences. The results is reported in Table C3 and all the estimates are consistent with our previous findings. Following a similar logic, families which have a daughter as the only child are often thought to be "gender neutral". The fact that the only child in family often receives more investments from parents also implies a larger α in their parents' utility functions. Our previous analysis would predict that girls as the only child of families should benefit less from the policy. Heterogeneity analyses confirmed our prediction (see Table C4).

Table C3: Long Run Impacts of Iodine Exposure by Fertility Patterns

	(1) Math Test ln(scores)	(2) Verbal Test ln(scores)	(3) Educational Attainment	(4) Schooling ln(years)
Panel A: Males				
Post × Goiter Rate × Boy Biased	-0.0725	-0.253	-0.0992	-0.0141
	[0.0933]	[0.105]**	[0.0674]	[0.0504]
Post × Boy Biased	0.221	0.580	0.0643	-0.134
	[0.167]	[0.247]**	[0.180]	[0.119]
Post × Goiter	0.0149	0.0400	0.0640	0.0324
	[0.0742]	[0.0776]	[0.0597]	[0.0643]
Mean of Dep. Var.	**2.46**	**3.05**	**2.41**	**2.03**
Observations	2918	2918	3318	3197
Panel B: Females				
Post × Goiter Rate × Boy Biased	0.0956	0.0344	0.0706	0.0602
	[0.0718]	[0.0551]	[0.0428]	[0.0271]**
Post × Boy Biased	-0.325	-0.199	-0.166	-0.0595
	[0.164]*	[0.135]	[0.0882]*	[0.0531]
Post × Goiter	0.138	0.123	0.0803	0.105
	[0.0353]***	[0.0487]**	[0.0258]***	[0.0256]***
Mean of Dep. Var.	**2.50**	**3.15**	**2.48**	**2.05**
Observations	2793	2793	3137	3022

Notes: Each coefficient is from a separate regression. All regressions control for fixed effects specific to birth province and birth year, mean-reversion control, age, age square, birth order, family size, parents' characteristics, hospitals per capita, hospital beds per capita and region-specific linear trends. Standard errors clustered by province appear in square brackets. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.
Data: CFPS-2010 and CFPS-2014

Table C4: Long Run Impacts of Iodine Exposure by Fertility Patterns

	(1) Math Test ln(scores)	(2) Verbal Test ln(scores)	(3) Educational Attainment	(4) Schooling ln(years)
Panel A: Males				
Post × Goiter Rate × Only Child	0.104	0.0857	-0.00339	0.0611
	[0.0867]	[0.100]	[0.0407]	[0.0758]
Post × Gender Attitudes	-0.157	-0.111	-0.123	-0.140
	[0.137]	[0.154]	[0.123]	[0.116]
Post × Goiter Rate	0.00862	0.0351	0.0248	0.0280
	[0.0641]	[0.0612]	[0.0503]	[0.0507]
Mean of Dep. Var.	**2.46**	**3.05**	**2.41**	**2.02**
Observations	3026	3026	3433	3318
Panel B: Females				
Post × Goiter Rate × Only Child	-0.195	-0.158	0.109	0.0768
	[0.0756]**	[0.0874]*	[0.0935]	[0.150]
Post × Gender Attitudes	0.164	0.119	-0.210	-0.256
	[0.168]	[0.105]	[0.202]	[0.250]
Post × Goiter Rate	0.182	0.162	0.0472	0.118
	[0.0514]***	[0.0763]**	[0.0317]	[0.0230]***
Mean of Dep. Var.	**2.49**	**3.14**	**2.47**	**2.04**
Observations	2873	2873	3220	3112

Notes: Each coefficient is from a separate regression. All regressions control for fixed effects specific to birth province and birth year, mean-reversion control, age, age square, birth order, family size, parents' characteristics, hospitals per capita, hospital beds per capita and region-specific linear trends. Standard errors clustered by province appear in square brackets. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.
Data: CFPS-2010 and CFPS-2014

C.3 Additional Results on Gender Attitudes

Table C5: Robustness on Gender Attitudes

	(1) Math Test ln(scores)	(2) Verbal Test ln(scores)	(3) Educational Attainment	(4) Schooling ln(years)
Panel A: Males				
Post × Goiter × Gender Attitudes	-0.0808	-0.0603	-0.132	0.132
	[0.133]	[0.143]	[0.0896]	[0.176]
Post × Gender Attitudes	0.105	0.0928	0.258	-0.0906
	[0.277]	[0.308]	[0.151]	[0.220]
Post × Goiter	0.0473	0.0746	0.0596	0.0148
	[0.0548]	[0.0589]	[0.0443]	[0.0473]
Mean of Dep. Var.	**2.46**	**3.05**	**2.37**	**2.01**
Observations	2841	2841	3204	3098
Panel B: Females				
Post × Goiter × Gender Attitudes	0.318	0.248	0.0619	0.211
	[0.117]**	[0.133]*	[0.0418]	[0.0793]**
Post × Gender Attitudes	-0.398	-0.379	-0.0111	-0.233
	[0.200]*	[0.181]**	[0.111]	[0.0860]**
Post × Goiter	0.117	0.117	0.0721	0.0981
	[0.0432]**	[0.0567]**	[0.0285]**	[0.0256]***
Mean of Dep. Var.	**2.49**	**3.14**	**2.42**	**2.02**
Observations	2696	2696	2988	2891

Notes: Each coefficient is from a separate regression. All regressions control for fixed effects specific to birth province and birth year, mean-reversion control, age, age square, birth order, family size, parents' characteristics, hospitals per capita, hospital beds per capita and region-specific linear trends. Standard errors clustered by province appear in square brackets. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.
Data: CFPS-2010 and CFPS-2014

Table C6: Gradients in Long Run Impacts of Iodine Exposure by Gender Attitudes

	(1) Somatic Complaints	(2) Depressed Affect	(3) Positive Affect	(4) Interpersonal Problems	(5) Satisfaction (index)	(6) Bad Health
Panel A: Males						
Post × Goiter × GA	0.193	-0.227	-0.373	-0.00379	0.565	-0.0765
	[0.305]	[0.460]	[0.528]	[0.142]	[1.001]	[0.0387]*
Post × GA	-0.348	0.860	0.817	0.135	-0.284	0.0965
	[0.597]	[0.753]	[1.365]	[0.302]	[2.241]	[0.0825]
Post × Goiter	-0.487	-0.505	0.0620	-0.112	0.520	0.0201
	[0.143]***	[0.180]***	[0.256]	[0.0794]	[0.431]	[0.0189]
Mean of Dep. Var.	**3.33**	**2.57**	**4.80**	**0.57**	**22.3**	**0.17**
Observations	1332	1334	1334	1334	807	3200
Panel B: Females						
Post × Goiter × GA	-0.195	-0.0700	-0.0576	-0.218	0.832	-0.0437
	[0.379]	[0.303]	[0.231]	[0.185]	[0.862]	[0.0295]
Post × GA	1.030	0.775	0.571	0.608	-0.103	0.00297
	[0.676]	[0.626]	[0.558]	[0.335]*	[1.642]	[0.0551]
Post × Goiter	-0.0750	-0.0843	0.132	-0.00469	-0.0594	-0.0124
	[0.249]	[0.279]	[0.180]	[0.0775]	[0.439]	[0.0242]
Mean of Dep. Var.	**3.49**	**3.21**	**4.89**	**0.66**	**22.3**	**0.18**
Observations	1271	1271	1270	1271	769	2985

Notes: Each coefficient is from a separate regression. All regressions control for fixed effects specific to birth province and birth year, mean-reversion control, age, birth order, family size, parents' characteristics, hospitals per capita, hospital beds per capita and region-specific linear trends. Standard errors clustered by province appear in square brackets. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.
Data: CFPS-2010, CFPS-2012 and CFPS-2014.

Table C7: Gradients in Long Run Impacts of Iodine Exposure by Gender Attitudes

	(1) Popularity	(2) Happiness	(3) Social Skills	(4) Satisfaction (index)
Panel A: Males				
Post × Goiter × GA	0.255	-0.265	0.606	0.565
	[0.433]	[0.457]	[0.481]	[1.001]
Post × GA	-0.150	0.467	-0.700	-0.284
	[0.773]	[0.721]	[1.094]	[2.241]
Post × Goiter	-0.164	0.358	0.346	0.520
	[0.171]	[0.245]	[0.188]*	[0.431]
Mean of Dep. Var.	**7.31**	**7.74**	**7.22**	**22.3**
Observations	811	807	807	807
Panel B: Females				
Post × Goiter × GA	0.0493	0.232	0.524	0.832
	[0.375]	[0.272]	[0.359]	[0.862]
Post × GA	0.392	-0.0119	-0.491	-0.103
	[0.723]	[0.670]	[0.633]	[1.642]
Post × Goiter	0.0806	-0.155	0.00374	-0.0594
	[0.179]	[0.162]	[0.166]	[0.439]
Mean of Dep. Var.	**7.10**	**8.06**	**7.14**	**22.3**
Observations	769	772	772	769

Notes: Each coefficient is from a separate regression. All regressions control for fixed effects specific to birth province and birth year, mean-reversion control, age, birth order, family size, parents' characteristics, hospitals per capita, hospital beds per capita and region-specific linear trends. Standard errors clustered by province appear in square brackets. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.
Data: CFPS-2014.

Appendix D CES-D Questions

CES-D includes 20 questions, which were aggregate to four categories: somatic complaints (Q1, Q2, Q5, Q7, Q11, Q13, Q20); depressed affect (Q3, Q6, Q9, Q10, Q14, Q17, Q18); positive affect (Q4, Q8, Q12, Q16)[32]; interpersonal problems (Q15, Q19).

Questions: Below is a list of the ways you might have felt or behaved. Please tell me how often you have felt this way during the past week with a score: 1 for rarely or none of the time (less than 1 day); 2 for some or a little of the time (1-2 days); 3 for occasionally or a moderate amount of time (3-4 days); 4 for most or all of the time (5-7 days).

1. I was bothered by things that usually don't bother me.
2. I did not feel like eating; my appetite was poor.
3. I felt that I could not shake off the blues even with help from my family or friends.
4. I felt I was just as good as other people.
5. I had trouble keeping my mind on what I was doing.
6. I felt depressed.
7. I felt that everything I did was an effort
8. I felt hopeful about the future.
9. I thought my life had been a failure.
10. I felt fearful.
11. My sleep was restless.
12. I was happy.
13. I talked less than usual.
14. I felt lonely.
15. People were unfriendly.
16. I enjoyed life.
17. I had crying spells.
18. I felt sad.
19. I felt that people dislike me.
20. I could not get "going".

[32] We reverse the score of each question in the positive affect according to the literature (i.e., 1 means most or all of the time (5-7 days) and 4 means rarely or none of the time (less than 1 day); 2 for occasionally or a moderate amount of time (3-4 days); 3 for some or a little of the time (1-2 days)).