FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

PRISCILA NATACHA DE OLIVEIRA

CHRONIC DISEASE BURDEN AND HUMAN CAPITAL INVESTMENT: EVIDENCE FROM THE CHAGAS DISEASE CAMPAIGN IN BRAZIL

São Paulo
2016

PRISCILA NATACHA DE OLIVEIRA

CHRONIC DISEASE BURDEN AND HUMAN CAPITAL INVESTMENT: EVIDENCE FROM THE CHAGAS DISEASE CAMPAIGN IN BRAZIL

Dissertação apresentada à Escola de Economia de São Paulo da Fundação Getulio Vargas como requisito para obtenção do título de Mestre em Economia de Empresas

Campo de Conhecimento:
Economia da Saúde

Orientador: Prof. Dr. Rodrigo Soares

São Paulo
2016

Oliveira, Priscila Natacha de.
Chronic disease burden and human capital investment: evidence from the Chagas disease campaign in Brazil / Priscila Natacha de Oliveira. - 2016.
52 f.

Orientador: Rodrigo Reis Soares
Dissertação (mestrado) - Escola de Economia de São Paulo.

1. Expectativa de vida. 2. Capital humano. 3. Chagas, Doença de. 4. Fecundidade. 5. Educação – Brasil. 6. Mercado de trabalho. I. Soares, Rodrigo Reis. II. Dissertação (mestrado) - Escola de Economia de São Paulo. III. Título.

CDU 37(81)

PRISCILA NATACHA DE OLIVEIRA

CHRONIC DISEASE BURDEN AND HUMAN CAPITAL INVESTMENT: EVIDENCE FROM THE CHAGAS DISEASE CAMPAIGN IN BRAZIL

Dissertação apresentada à Escola de Economia de São Paulo da Fundação Getulio Vargas como requisito para obtenção do título de Mestre em Economia de Empresas.

Campo de Conhecimento:
Economia da Saúde

Data de Aprovação:
___/___/_____

Banca examinadora:

_____
Prof. Dr. Rodrigo Soares (Orientador)
FGV-EESP

_____
Prof. Dr. Vladimir Ponczek
FGV-EESP

_____
Profa. Dra. Fernanda Estevan
FEA-USP

ABSTRACT

We investigate the effects of augmented life expectancy and health improvements on human capital investment, labor supply and fertility decisions. Our main motivation is the prediction of human capital theory that a longer and healthier life encourages educational investment and female labor force participation, while discouraging fertility. To assess the magnitude of these effects, we explore a national campaign against Chagas disease in Brazil as an exogenous source of adult mortality decline and improvement in health conditions. We show that, relative to non-endemic areas, previously endemic regions saw higher increases in educational investment, measured by literacy, school attendance and years of schooling, following the campaign. Additionally, we find that labor force participation increased in high prevalence areas relative to low prevalence ones. Furthermore, we estimate a substantially higher effect on female labor force participation relative to male, suggesting that longevity gains and health improvements affected women's incentives to work, encouraging women to join the labor force. We do not find significant effects on fertility decisions.

Keywords: Chagas disease; chronic diseases; mortality; human capital; fertility; difference-in-differences estimator.

RESUMO

Investigamos os efeitos de aumentos na expectativa de vida e de melhorias na saúde sobre investimento em capital humano, oferta de trabalho e decisões de fecundidade. Nossa principal motivação provém da predição da teoria de capital humano de que uma vida mais longa e saudável encoraja investimentos em educação e participação feminina no mercado de trabalho, ao mesmo tempo em que desencoraja fecundidade. Para avaliar a magnitude desses efeitos, exploramos a campanha nacional contra doença de Chagas no Brasil como fonte exógena de diminuição na mortalidade adulta e de melhorias nas condições de saúde. Mostramos que, em relação a áreas não endêmicas, regiões previamente endêmicas tiveram aumentos maiores no investimento educacional, medido por taxa de alfabetização, matrícula escolar e anos de escolaridade, após a campanha. Adicionalmente, encontramos aumentos na participação na força de trabalho em áreas de alta prevalência em relação às de baixa prevalência. Ademais, estimamos um efeito substancialmente maior na participação feminina no mercado de trabalho em relação à masculina, sugerindo que ganhos de longevidade e melhorias na saúde alteram os incentivos de mulheres para trabalhar, encorajando-as a entrar na força de trabalho. Não identificamos efeitos significantes em decisões de fecundidade.

Palavras-chave: Doença de Chagas; doenças crônicas; mortalidade; capital humano; fecundidade; estimador de diferença-em-diferenças.

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

1 Introduction

Global life expectancy has increased substantially in the last century, and health conditions have improved considerably in the past decades throughout the world (Lee, 2003; WHO, 1998). These improvements in general health conditions of the population may have impacts that go well beyond direct gains in well-being, as they are likely to affect major life decisions such as human capital investment, labor supply and fertility.

Over the past years, there has been an increasing interest in the potential effects of health conditions on educational outcomes and other economic decisions. One possible explanation for the relationship between health and education comes from the assumption that an investment that pays out a fixed amount each period is more valuable if the stream of payments lasts longer. This implies that increased longevity should foster human capital investment. This theoretical prediction has been extensively explored in the past years, as well as the role of health in affecting incentives to work and to have children (see Becker, 1964; Ben-Porath, 1967; Cervellati and Sunde, 2013; Kalemli-Ozcan, 2003; Soares, 2005 for further discussion).

Recently, an empirical branch of this literature has emerged to discuss whether this theoretical prediction finds empirical support. For instance, Jayachandran and Lleras-Muney (2009) and Oster et al. (2013) estimate a positive causal effect of longevity on human capital investment, thus providing empirical evidence for the human capital theory prediction. Another related literature analyzes whether better health conditions (in a broader sense) affect educational attainment and income or household consumption, as in Bleakley (2007, 2010), Cutler et al. (2010) and Lucas (2010). Regarding the role of health on fertility decisions, there is still little causal empirical evidence (e.g., Bleakley and Lange, 2009). In general, the literature suggests that health indeed plays a significant role in human capital and fertility decisions.

We focus on mortality and morbidity reductions as a result of a successful campaign against Chagas disease implemented in Brazil in the mid-1970s, a source of exogenous variation in health conditions for endemic areas. Chagas disease is a parasitic disease that affects mainly adults during the productive lifetime, with symptoms often related to cardiac problems and disability. Life expectancy of an individual with Chagas disease is currently estimated to be 50 years in Brazil, relative to 75 years for the general population. By virtually eliminating the main mode of transmission of Chagas disease through insecticide-spraying of houses, the program represented a sharp reduction in infection rates and adult mortality.

Our main contribution is to provide evidence on the effects of augmented healthy life expectancy on economic decisions (e.g., human capital investment, labor supply and fertility). We underline our focus on healthy life expectancy, in contrast to longevity. Indeed, most major economic decisions depend on the quality of health during the later years of life, which makes

this an important dimension to be considered. As we analyze a Chagas disease eradication campaign in Brazil, we are able to identify the economic consequences of healthy longevity, and our results could reinforce theoretical predictions of human capital theory concerning the role of health in shaping individual incentives to invest in education, to join the labor force and to have children. Furthermore, although Chagas disease was recognized a major health problem in Brazil and in several Latin American countries in the 1970s, the disease is still endemic in some countries[1]. Therefore, estimating Chagas disease impact on economic decisions is particularly important in Latin America.

Using Chagas disease eradication campaign in a quasi-experimental framework has several advantages, as Chagas disease characteristics can be used to validate our identification strategy. First, the most severe Chagas disease symptoms usually occur during the prime productive years of the life cycle. Thus, they directly affect the duration of the stream of wages without strongly interfering in the individual's capacity to attend school and to invest in education. Therefore, increased healthy longevity is the major channel through which education is expected to be affected (more precisely, increased expected length of the stream of payments). Second, Chagas disease has a mortality distribution highly concentrated in young adults, so that its mitigation translates into considerable life expectancy gains. Finally, since Chagas disease was endemic in some areas in Brazil but not in others, the campaign was supposed to affect mortality and morbidity rates differently according to the baseline prevalence rate, so that individuals in zero prevalence regions can be used as a comparison group to individuals in endemic regions.

We explore the variation in pre-treatment infection rates within a difference-in-differences strategy, adapted to an intensity of treatment technique. The first difference is over time, whereas the second is across geographical areas (minimal comparable areas, MCAs). Our identification strategy is thus based on heterogeneity in baseline prevalence rates across MCAs and on the exogeneity of the beginning of the campaign.

We find that, relative to non-endemic areas, previously endemic regions saw a higher increase in educational investment. Our estimates suggest that a 7.4 p.p. reduction in Chagas disease prevalence (which corresponds to eradication in an average MCA) increased years of schooling by 0.07 year (or 3.6%). Similarly, literacy rate increased by 1.6 p.p. (or 2.7%), and school attendance increased by up to 5.5 p.p. (or 8.2%), thus suggesting a significant role of health in human capital decisions. Furthermore, our results indicate an implied elasticity of years of schooling and literacy with respect to longevity of 1.3 and 1.0, respectively, in line with the quasi-experimental evidence on this field. We also find suggestive evidence that reduced adult mortality and improvements in health conditions fostered a substitution from work towards study for younger cohorts, while increasing overall labor force participation and work hours.

[1] It is estimated that approximately 6 to 8 million people are currently infected and an average of 12,000 deaths per year are caused by Chagas disease in the Americas (PAHO, 2014).

The effects on labor-related decisions were considerably higher for women relative to men, indicating that better health conditions and reduced adult mortality affected women's incentives to work, thus shifting women from household chores and raising children towards the labor market. We do not find statistically significant effects on the decision to have children or on the total number of children.

Furthermore, as an alternative to the main mechanism from human capital theory, we discuss other mechanisms that might explain our findings. For instance, Chagas disease may affect investments in children through the health conditions of their parents. Investments in children's education may be reduced through diminished family income if sickness and disability of parents reduce work hours, which in turn might also induce intrahousehold reallocation of work time. This channel may be leading our estimates of the effects on labor force participation for children and young adults, and it may partially explain our findings for educational outcomes. Another possible mechanism leading to our educational effects is related to a scarring effect, since the initial phase of the disease could be quite severe during early childhood. Child development may be affected by severe infection during early childhood, thus impacting educational outcomes in the future. Nevertheless, we understand that our results are not primarily driven by this channel, as the occurrence of severe symptoms in the initial phase was very rare.

The paper is organized as follows. The next section describes the main characteristics of Chagas disease. Section 3 discusses the campaign against Chagas disease in Brazil. Section 4 describes the data. The empirical strategy is discussed in Section 5. Section 6 presents the results. Section 7 concludes with a discussion of our findings and their relation to the existing literature.

2 Chagas disease

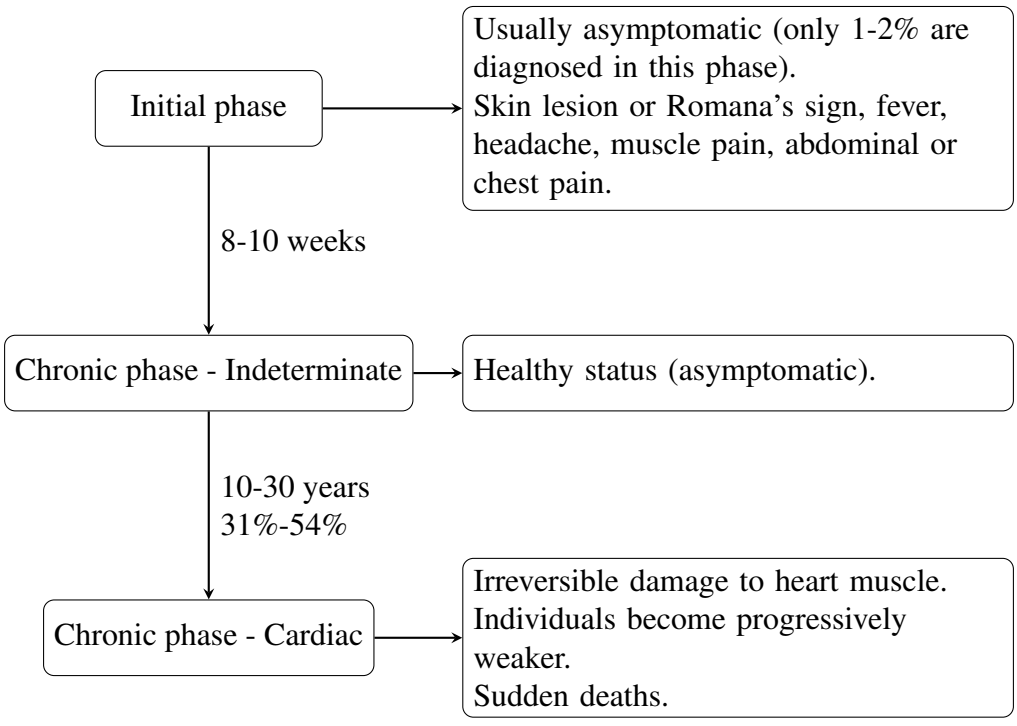
Chagas disease (*American Trypanosomiasis*) is a tropical parasitic disease caused by the protozoan *Trypanosoma cruzi*, and is often related to cardiac and digestive complications up to 30 years after infection. The disease is transmitted mainly by kissing bugs, specially *Triatoma infestans* in Brazil. Although there are other modes of transmission such as blood transfusion, organ transplants, contaminated food and congenital transmission, the vectorial mechanism accounted for the vast majority of Brazilian cases.

Chagas disease evolves into two distinct phases: acute (initial) and chronic, as shown in Figure 1. The acute phase starts immediately after infection and is characterized by high concentration of the protozoan in the blood. Usually 8 to 10 weeks after infection, the concentration of the protozoan in the blood is reduced, marking the start of the chronic phase.

The acute phase is usually asymptomatic, which hinders early diagnosis of the disease. According to Prata and Macedo (1984), less than 2% of all infected individuals are diagnosed in this phase. When presented in a symptomatic form, the acute phase is characterized by mild symptoms, such as skin lesion, fever, headache, muscle pain, abdominal or chest pain, and only in rare cases meningoencephalitis and myocarditis. When manifested, the acute phase is rarely lethal for adolescents and adults, presenting itself in a more severe form in early childhood.

After the acute phase subsides, the chronic phase starts in its indeterminate form, in

Figure 1 – Epidemiology of Chagas Disease



which the infected individual recovers a healthy status, with no organ damage. Although these patients are apparently healthy and asymptomatic, they may be considered as belonging to the category of potential cardiac disease. This asymptomatic phase usually lasts from 10 to 30 years, but it is important to emphasize that many infected individuals may never leave this phase, dying from causes other than Chagas disease[1].

Nevertheless, for up to 54% of all infected individuals[2], the chronic phase evolves to the cardiac form, usually 10 to 30 years after infection. This phase is characterized by irreversible damage to heart muscle, which makes individuals progressively weaker and may lead to disability and death from heart failure or other cardiac complications. Although less common in Brazil, some infected individuals may develop severe digestive complications due to chagasic infection[3].

Treatment for Chagas disease is very limited. There are two drugs available to treat infection with *T. cruzi* (nifurtimox and benznidazole), which were proved to be effective in treating the acute phase of the disease, with many successful cases of cure in this stage. Nevertheless, when administered in the chronic phase, the drugs appear to have little effect on the course of the cardiopathy and to be incapable of preventing cardiac complications. Since less than 2% of all infected individuals are diagnosed in the acute phase, the majority of individuals is not cured using the available antiparasitic treatment and is unable to reverse the course of the disease, thus remaining infected until their death.

Life expectancy of an individual with Chagas disease is currently estimated to be 50 years in Brazil (Taylor and Bestetti, 2009), relative to 75 years for the general population. More precisely, the authors find that having positive serology per se does not affect longevity, in the sense that individuals who never develop the cardiac or digestive forms of Chagas disease usually have the same life expectancy as uninfected ones. Nevertheless, life expectancy is sharply reduced for people with the cardiac form of the disease. Patients with electrocardiographic abnormalities have a slight decrease in life expectancy relative to non-chagasic patients, and patients with heart failure have a marked decrease in survival (only about 30% live five years after symptom onset).

Similarly, a 10-year follow-up study conducted in Venezuela in 1973-1983 pointed to an overall mortality of 36% for chagasic patients aged 22 to 78 (Espinosa et al., 1985). The authors find a significant life expectancy decrease for chagasic patients relative to non-chagasic ones, with a sharp reduction in survival rates for infected individuals across time. The survival

[1] See Laranja et al. (1956) and Prata and Macedo (1984) for further information on epidemiology of Chagas disease in Brazil.

[2] The fraction of infected individuals that develop cardiac problems related to Chagas disease varies according to the geographic region of infection, and some studies carried out in Brazil showed that it ranged from 31.2% to 54.2% in Central Brazil.

[3] In Brazil, the most common form of the chronic phase is the cardiac form, which was found to be up to 54.2% of all infected individuals in the municipality of Luz in Minas Gerais. Estimates vary according to geographic location, but the incidence of digestive form in Brazil is estimated to be less than 10% of all cases.

rate for chagasic patients was 88% for the first year of the study, 63% after 5 years and 40% after 10 years. It is important to emphasize that reduced life expectancy was driven mainly by patients with electrocardiographic abnormalities and heart failure, with no death registered for chagasic patients without cardiac abnormalities. Although these findings are partially explained by the high fraction of patients with the cardiac form of the disease in their sample, they are helpful to illustrate the poor prognosis of the disease once the cardiac form starts.

Hence, Chagas disease is associated with high adult morbidity and mortality, specially during the productive years of the life cycle, as can be seen in Table 1 and Figure 2. Table 1 shows deaths due to Chagas disease as a percentage of total deaths in 1980 for several age groups and according to different baseline prevalence rates of Chagas disease. In areas with baseline prevalence rate higher than 30%, deaths by Chagas disease accounted for more than 5% of all adult deaths in 1980, reaching 11.1% of all deaths in the age group of 45 to 60 years old. Interestingly, Chagas disease accounted for a low share of child deaths, with low variance across baseline prevalence rates, confirming the conjecture that Chagas disease was indeed strongly related to adult mortality and had only minor effects on child mortality.

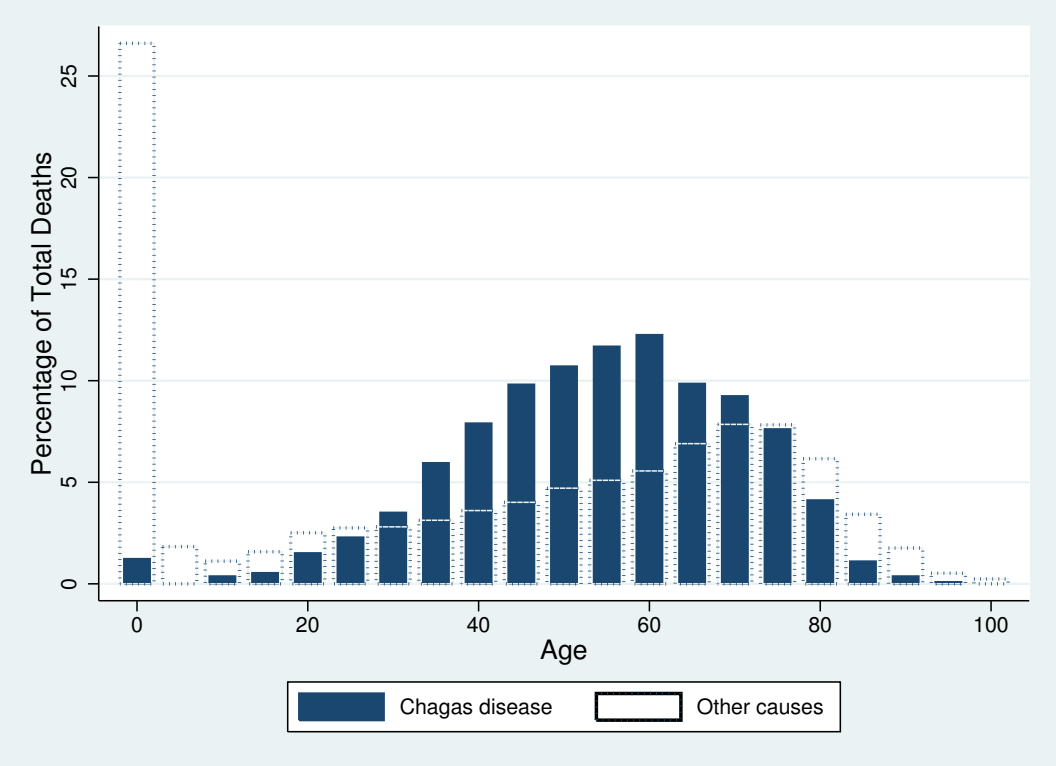
Table 1 – Share of total deaths due to Chagas disease in 1980 (%)

		Baseline Prevalence Rate			
		$< 1\%$	1%-10%	10%-30%	$> 30\%$
Age Group	0-15	0.03	0.16	0.16	0.11
	15-30	0.06	1.71	2.33	5.58
	30-45	0.54	4.96	6.21	7.08
	45-60	0.47	5.96	6.98	11.11
	60+	0.16	2.94	3.29	6.46
Observations (MCAs)		298	358	187	44

Notes: Mortality data from 1980, aggregated at the minimal comparable area (MCA) level. The sample includes MCAs from the former Brazilian Central West states and from the state of Minas Gerais.

Figure 2 shows two different histograms of age of death distribution in 1980: the age distribution of deaths due to Chagas disease (blue bars) and the age distribution of deaths due to all other causes (white bars). The age pattern of deaths caused by Chagas disease greatly differs from the pattern of general mortality. As can be seen in Figure 2, deaths by Chagas disease were more frequent during the productive years of the life cycle than were deaths due to other causes. For instance, almost 30% of all deaths not caused by Chagas disease happened to individuals younger than 20 years old, whereas the percentage of deaths due to Chagas disease in this age group was less than 5%.

Figure 2 – Age of death distribution in 1980 (%)



Notes: Histogram of the age distribution of deaths due to Chagas disease in solid blue bars. Histogram of the age distribution of deaths due to all of the other causes of death in white bars. Mortality data from 1980.

3 The program

In 1975, the Ministry of Health started a national program against vector-borne transmission of Chagas disease in Brazil, motivated by the discovery of effective pesticides against kissing bugs, namely BHC and pyrethroids. Although BHC was discovered in the late 1940s, its efficacy in eliminating kissing bugs was only confirmed after its use in large scale campaigns as the ones implemented in the state of Sao Paulo (Brazil) and in Venezuela in the 1960s. The effectiveness of the pyrethroids against kissing bugs, in turn, dates back to the 1970s. Hence, after the successful campaigns against kissing bugs in Sao Paulo and Venezuela, the Brazilian Ministry of Health implemented a similar nationwide[1] program, aiming to eliminate vectorial transmission of the disease.

The campaign comprised three phases: preparatory, attack and surveillance. The preparatory phase consisted on the implementation of serological and entomological surveys, aiming to map and provide an overview of the endemicity of Chagas disease in Brazil. This early stage was followed by the attack phase, which was based on insecticide-spraying on houses and surrounding areas. The length of the attack phase greatly differed across municipalities, as new rounds of insecticide-spraying were carried out until no kissing bug was found in the location. Once the attack phase ended, the area entered the surveillance phase, characterized by the notification of the presence of kissing bugs by the community to local authorities, followed by selective fumigation in case any kissing bug was found. The surveillance phase continues active until today in Brazil.

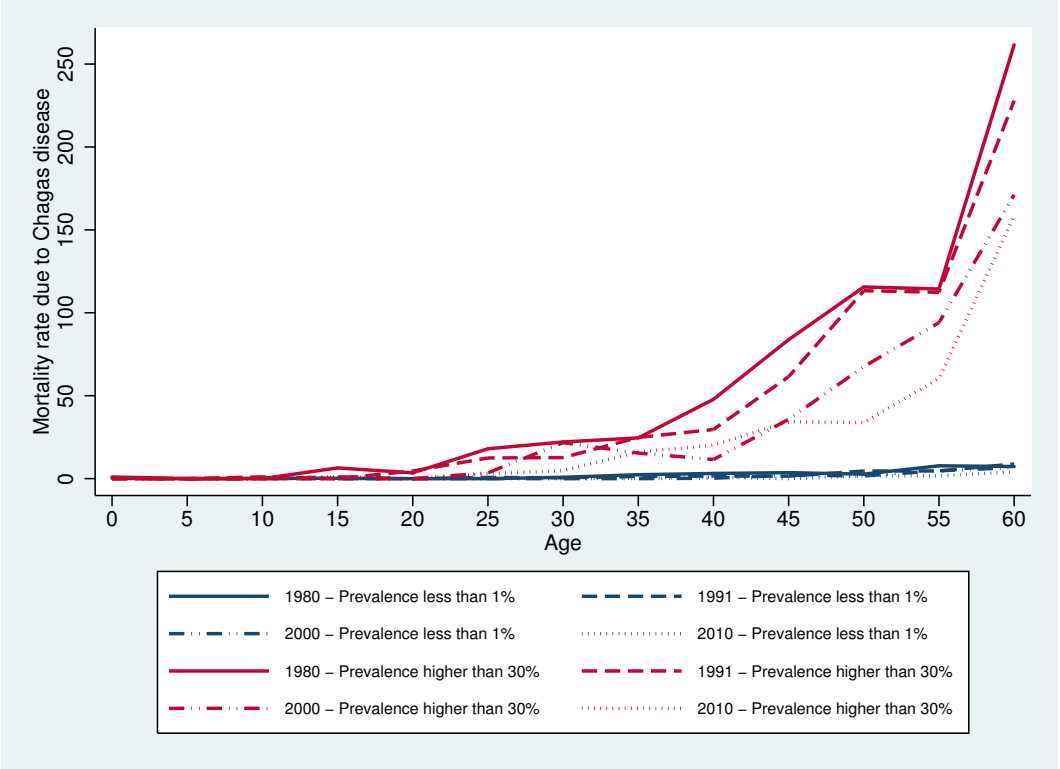
By 1980, the preparatory phase was completed in all areas. Notwithstanding, the attack phase was implemented in a phase-in system, so that insecticide-spraying did not start immediately after the first phase in every locality. In fact, there was a substantial delay in fumigation efforts, reaching all endemic regions only in 1983.

The campaign proved effective in interrupting vectorial transmission of Chagas disease, as confirmed by the certification of the country as free of vectorial transmission by *T. infestans* in 2006 and by follow-up surveys. A follow-up survey conducted in some Brazilian states between 1987 and 1999 estimated a prevalence rate of 0.04% in children aged 7 to 14, relative to an infection rate of 10.65% in children aged 5 to 14 in 1975. Between 2000 and 2005, a survey conducted in a sample of children aged up to 5 years old estimated a prevalence rate of 0.03% in this age group (compared to an infection rate of 2.21% in children aged 0 to 4 years old in 1975), with the majority of cases due to congenital transmission.

Therefore, the almost complete interruption of Chagas disease transmission represented

[1] Only Brasilia (DF) and the state of Sao Paulo were not included in the program, since those areas were already carrying on their local campaigns against Chagas disease transmission.

Figure 3 – Mortality rate by Chagas disease (per 100,000 residents)



Notes: Age-specific mortality rate due to Chagas disease (per 100,000 residents). Data from Sistema de Informacoes de Mortalidade for years 1980, 1991, 2000 and 2010.

a sharp reduction in infection rates and, over time, in adult mortality in previously endemic areas. For instance, age-adjusted mortality rate due to Chagas disease (per 100,000 residents) dropped from 53.7 in 1980 to 14.6 in 2010 in the state of Goias, and from 26.3 in 1980 to 5.9 in 2010 in the state of Minas Gerais.[2] This represented a sharp decline in adult mortality. Figure 3 shows age- and cause-specific mortality rates per 100,000 residents in initially high and low prevalence areas (areas with baseline prevalence rate higher than 30% and lower than 1%, respectively) in 1980, 1991, 2000 and 2010. Indeed, there was a sharp decline in adult mortality rates in endemic areas relative to low prevalence regions. The fact that mortality rate by Chagas disease has not yet dropped to zero reflects almost entirely the absence of cure for previously infected individuals.

[2] Goias and Minas Gerais figured among the most endemic states in Brazil.

4 Data and Descriptive Statistics

The micro-level data used in this study are drawn from the census of 1970, 1980, 1991, 2000 and 2010, whereas mortality data come from the *Sistema de Informacoes sobre Mortalidade*, from the Brazilian Ministry of Health. Vital statistics are available at the micro- and municipality-level from 1979 onward. Data on the baseline prevalence of Chagas disease are drawn from the seroprevalence survey of human Chagas' infection 1975-1980, conducted by the Brazilian Ministry of Health.

Originally, prevalence data were aggregated at the municipality level. Notwithstanding, Brazilian geographic division was reshaped[1] during the time period considered, with more than 1500 municipalities being created. Therefore, in order to make data comparable across time, we further aggregate all municipality-level data into minimal comparable areas (MCAs), which mirrors the original geographical division from 1970.

We match individuals to the baseline prevalence rate of Chagas disease in their place of residence. Although some migration concerns may arise with this approach, we conduct some robustness checks by restricting our sample to nonmovers only, in an attempt to address these potential issues.

Our sample is composed by a group of states with high pre-intervention endemicity: Goias (including Tocantins), Mato Grosso (including Mato Grosso do Sul) and Minas Gerais. Pre-treatment infection rates greatly varied within these states, ranging from 0% to 53%. The average baseline prevalence rate in our sample was 7.4%, compared to a national prevalence of 4.2%.

We restrict our sample in this way to focus on regions where Chagas disease was a relevant public health problem and that were more homogeneous in other socioeconomic characteristics, and thus least likely to differ on relevant unobservables. In our view, these states were more likely to have followed a similar path over the years in terms of education and health related policies, diminishing the probability that confoundedness of different public policies biases our results. Nonetheless, considerable variation in infection rates is observed within these states, providing enough heterogeneity for our estimation strategy.

Table 2 shows descriptive statistics of all variables used in our estimation, with observations grouped according to the baseline prevalence rate of Chagas disease. The average prevalence rate among the top quartile of the baseline prevalence distribution (23.15%) was considerably higher than the one in the bottom quartile (0.12%). Literacy rate, male labor force

[1] Between 1970 and 2010, two states were created (Tocantins, which originally was part of Goias, and Mato Grosso do Sul, constituted from a partition of Mato Grosso), and the number of municipalities increased from 3952 to 5561.

participation and total number of children were lower in the top quartile in comparison to the bottom one, whereas female labor force participation was substantially higher in the top quartile relative to the bottom one. Furthermore, we find a higher fraction of households with a disable adult in the top quartile of prevalence distribution relative to the bottom quartile.

Table 2 – Descriptive Statistics

		Baseline Prevalence Rate			
	Whole Sample	Bottom quartile	Second quartile	Third quartile	Top quartile
Prevalence Rate	8.791	0.118	1.720	5.932	23.15
	(10.45)	(0.146)	(0.830)	(2.025)	(8.551)
Fraction of Rural Population	0.503	0.568	0.534	0.502	0.441
	(0.500)	(0.495)	(0.499)	(0.500)	(0.496)
Fraction of Male Population	0.503	0.504	0.507	0.505	0.496
	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
Age	21.90	22.66	21.90	21.42	21.98
	(17.95)	(18.63)	(18.04)	(17.60)	(17.83)
Employment Share in Agriculture	0.534	0.616	0.564	0.525	0.470
	(0.287)	(0.229)	(0.235)	(0.271)	(0.353)
Employment Share in Manufacturing	0.134	0.108	0.134	0.137	0.146
	(0.105)	(0.0791)	(0.111)	(0.109)	(0.105)
Share of Households with Electricity	0.338	0.351	0.303	0.311	0.394
	(0.248)	(0.196)	(0.214)	(0.224)	(0.309)
Household Income per capita (log)	4.185	4.054	4.193	4.143	4.284
	(0.537)	(0.434)	(0.462)	(0.436)	(0.700)
Fraction of Households with Disable Adult	0.0767	0.0780	0.0709	0.0781	0.0805
	(0.266)	(0.268)	(0.257)	(0.268)	(0.272)
Years of Education	1.953	1.929	1.840	1.809	2.216
	(2.638)	(2.507)	(2.503)	(2.498)	(2.931)
Literacy	0.589	0.606	0.586	0.573	0.599
	(0.492)	(0.489)	(0.493)	(0.495)	(0.490)
School Attendance (7-14)	0.670	0.696	0.651	0.666	0.681
	(0.470)	(0.460)	(0.477)	(0.472)	(0.466)
School Attendance (15-18)	0.366	0.333	0.357	0.388	0.371
	(0.482)	(0.471)	(0.479)	(0.487)	(0.483)
School Attendance (18-25)	0.147	0.13	0.135	0.153	0.161
	(0.354)	(0.336)	(0.342)	(0.360)	(0.368)
Male Labor Force Participation	0.713	0.710	0.721	0.713	0.706
	(0.452)	(0.454)	(0.448)	(0.452)	(0.456)
Male Work Hours	2.892	2.888	2.902	2.886	2.891
	(0.344)	(0.355)	(0.337)	(0.354)	(0.339)
Female Labor Force Participation	0.149	0.123	0.134	0.144	0.181
	(0.356)	(0.329)	(0.341)	(0.351)	(0.385)
Female Work Hours	2.688	2.659	2.701	2.667	2.703
	(0.515)	(0.530)	(0.511)	(0.533)	(0.498)
Fertility - At least one child	0.602	0.603	0.615	0.612	0.581
	(0.489)	(0.489)	(0.486)	(0.487)	(0.493)
Fertility - Total Children	3.283	3.434	3.349	3.290	3.140
	(3.867)	(4.015)	(3.877)	(3.829)	(3.814)

Notes: Variable means displayed to the right of the variable name. Standard deviations are reported in parentheses, below the means. Data from 1970 census. Work hours is a categorical variable that ranges from 1 to 3, where 1 corresponds to working less than 15 weekly hours, 2 corresponds to working between 15 and 40 weekly hours and 3 corresponds to working at least 40 weekly hours. Work hours are relative to the main job.

A potential problem in our empirical exercise may arise if pre-intervention Chagas disease prevalence is correlated with poverty or economic development. In that case, our results might simply reflect a catch-up in income and development of the poorest areas (and initially endemic ones) instead of the exogenous shock in health conditions we aim to investigate.

Table 3 shows the results of a OLS regression of baseline prevalence rates on MCAs' characteristics. In the first column, we only include as control variables geographical characteristics, such as temperature, soil types and distance to the sea. The results suggest that higher prevalence rates are correlated with higher temperatures and larger distances to the sea. In the

second column, we include the covariates used in our estimations. After this addition, the correlation between distance to the sea and average temperature with Chagas disease prevalence rate vanished. We also find that baseline prevalence rate is strongly correlated with some soil types, although this should not be a problem to our estimation as we include MCA fixed effects in our specifications.

Table 3 – Correlations - Baseline Prevalence Rate

Dependent variable: Baseline Prevalence Rate		
	(1)	(2)
Distance to the sea	0.32**	0.084
	(0.14)	(0.34)
Average temperature	0.71**	0.92
	(0.29)	(0.71)
Acrisol	0.95	-2.60*
	(0.97)	(1.37)
Cambisol	1.70	-0.54
	(1.20)	(2.71)
Latosol	-0.51	-1.12
	(0.98)	(1.28)
Neosol	2.90*	4.70*
	(1.56)	(2.58)
Plinthosol	-4.47***	-10.9***
	(1.63)	(2.79)
Spondosol	-8.64***	-13.0***
	(1.19)	(4.10)
Nitosol	-10.1***	-14.7***
	(2.01)	(4.37)
Planosol	-9.22***	-10.6***
	(2.02)	(2.25)
Fraction of Rural Population		4.49
		(9.45)
Employment Share in Agriculture		-20.8
		(14.1)
Employment Share in Manufacturing		-25.0
		(18.9)
Share of Households with Electricity		-3.41
		(6.75)
Average Household Income per capita (log)		-0.094
		(4.12)

Notes: This table reports estimates of the OLS regression of baseline prevalence rate (on a 0-100 scale) on the variables indicated in the rows, using census year of 1970. The unit of observation is the MCA level, weighted by its respective population. Standard errors (reported in parentheses) are clustered at the MCA level.
*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Furthermore, we do not find a correlation between Chagas disease and poverty, as the coefficient of household income per capita is not statistically different from zero. We also do not find any specific pattern between prevalence rates and sectoral composition of the labor force, as indicated by the lack of statistical significance of the baseline share of the labor force employed in the agriculture and in the manufacturing sectors. Although Chagas disease is commonly associated with rural areas, there is no clear relation between Chagas disease dispersion and rural areas in our sample.

Hence, the results in Table 3 suggest no consistent correlation between variables related to economic development and pre-intervention prevalence of Chagas disease, once we include

our main controls. Accordingly, a number of alternative explanations for our findings are likely to be excluded by these estimates.

5 Empirical strategy

Our empirical strategy consists of a difference-in-differences approach adapted to intensity of treatment, which closely follows the identification strategy of Bleakley (2007) and Jayachandran and Lleras-Muney (2009). The first difference is over time, whereas the second is across geographic areas (MCAs). We use baseline Chagas disease infection rates as a source of variation in the expected response to the eradication campaign.[1] We compare outcomes (e.g., human capital investment, labor force participation and fertility) of cohorts across time and MCAs with different baseline prevalence rates, in order to estimate the contribution of health improvements to the observed changes in outcomes.

Our identification strategy is thus based in a two-fold argument: exogeneity of the beginning of the campaign and heterogeneity in pre-treatment infection rates across MCAs. Indeed, the implementation of the program was related to advances in health technologies, specially the discovery of effective pesticides against kissing bugs. The first large scale campaigns based on insecticide-spraying of BHC to eliminate kissing bugs were carried in the state of Sao Paulo (Brazil) and in Venezuela in the 1960s. When its effectiveness against Chagas disease main vector was confirmed in the 1970s, the Brazilian government engaged on a national campaign aiming to cease Chagas disease transmission. Therefore, the implementation of the campaign was due to factors external to the MCAs.

We also argue that the intensity of treatment was proportional to baseline prevalence rate. Individuals living in highly infected areas were more likely to benefit from Chagas disease eradication than residents of areas previously free of the disease. As suggested by Figure 4, more infected areas indeed saw greater prevalence declines.

We explicitly show that the campaign against Chagas disease indeed reduced adult mortality rates in our sample. In order to do so, we estimate the effects on mortality rate through the following equation:

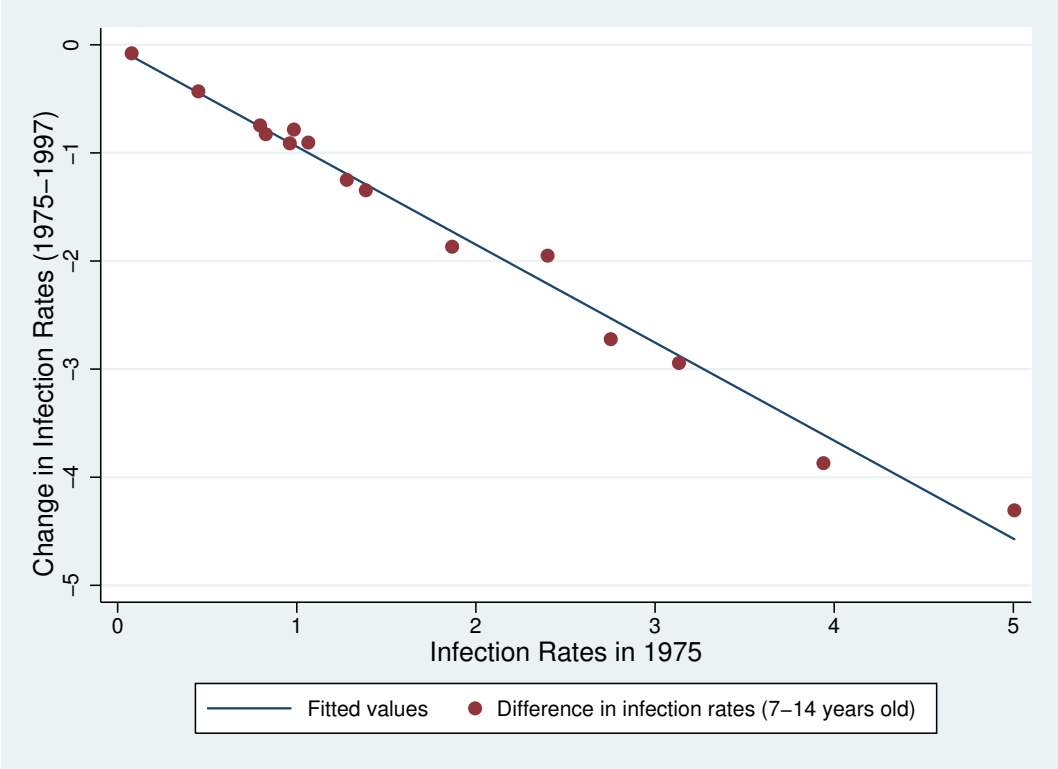
$$Y_{jst} = \beta Prev_{j,pre} \cdot Post_t + \Theta Z_{j,pre} \cdot \delta_t + \delta_{st} + \delta_j + \varepsilon_{jst} \tag{5.1}$$

where Y_{jst} is the mortality rate in MCA j in state s in year t, $Prev_{j,pre}$ denotes baseline Chagas disease prevalence in MCA j, $Post_t$ is an indicator variable for whether the year is after 1980[2], $Z_{j,pre}$ is a vector of baseline MCA-level controls, δ_t is a vector of year dummies, δ_{st} and δ_j are state-year and MCA fixed effects, respectively. Standard errors are clustered at the MCA level. The parameter of interest is thus β.

[1] As suggested by Figure 4 (below), baseline prevalence rates can be interpreted as a proxy for declines in chagasic infection.

[2] Since our micro- and municipality-level mortality data starts at 1979, we use 1980 as a pre-treatment period and then consider years from 1980 onwards as post-treatment.

Figure 4 – Decline in Infection Rates from 1975 to 1997 across States (p.p.)



Notes: Each dot represents a state. The line is the fitted values from a regression of changes in infection rate between 1975 and 1997 on the initial 1975 level. The follow-up survey was conducted in 15 states between 1989 and 1997, among children aged 7-14 years old enrolled in school.

The interaction between $Z_{j,pre}$ and δ_t is included in order to control for pre-treatment differences across regions that might affect our results. For instance, $Z_{j,pre}$ may include the share of households in MCA j with electricity in 1970, the average household income per capita in MCA j in 1970, and the share of the labor force employed in agriculture and in manufacturing sectors in MCA j in 1970.

The inclusion of a nonlinear trend on the average baseline household income per capita and on the pre-intervention share of households with electricity to our specification aims to address the potential concern that our estimates may be driven by pre-treatment differences in income or development levels across regions, with a possible catch up effect. If the poorest and least developed MCAs were also the ones with higher pre-treatment infection rates, our estimates could be reflecting a convergence in income or economic development over time, and not exclusively better health conditions. The addition of these variables rules out these mechanisms as potential explanations for our results. The addition of a nonlinear trend on pre-treatment sectoral composition of the labor force, in turn, is motivated by the possibility that, throughout the period (1970-2010), Brazilian municipalities might have been differently hitten by economic shocks.

Concerning the effects on human capital investment, labor force participation and fertility, we develop two empirical strategies, one that closely follows the strategy outlined above

("standard approach") and another based on the fraction of life exposed to the campaign efforts. The standard empirical strategy can be summarized by the following equation:

$$Y_{ijst} = \beta Prev_{j,pre} \cdot Post_t + \Gamma X_{ijt} + \Phi X_{ijt} \cdot Post_t + \Theta Z_{j,pre} \cdot \delta_t + \delta_{st} + \delta_j + \varepsilon_{ijst} \quad (5.2)$$

where Y_{ijst} is the outcome of interest for individual i in MCA j in state s and census year t, $Post_t$ equals 1 for census year of 1980, 1991, 2000 and 2010, and equals 0 for 1970, and X_{ijt} is a vector of individual level controls that might affect economic decisions (gender, rural dummy and age fixed effects). Standard errors are clustered at the MCA level. We are interested in the parameter β.

Our second identification strategy consists on a characterization of pre/post-treatment periods that considers the likelihood of an individual being infected in a given year. Because there is no cure for Chagas disease after the acute phase subsides, a person who had already been infected by the time the campaign started would not benefit (or would only extract minor benefits) from eradication efforts. This heterogeneous dimension of the treatment is disregarded by the standard difference-in-differences approach.

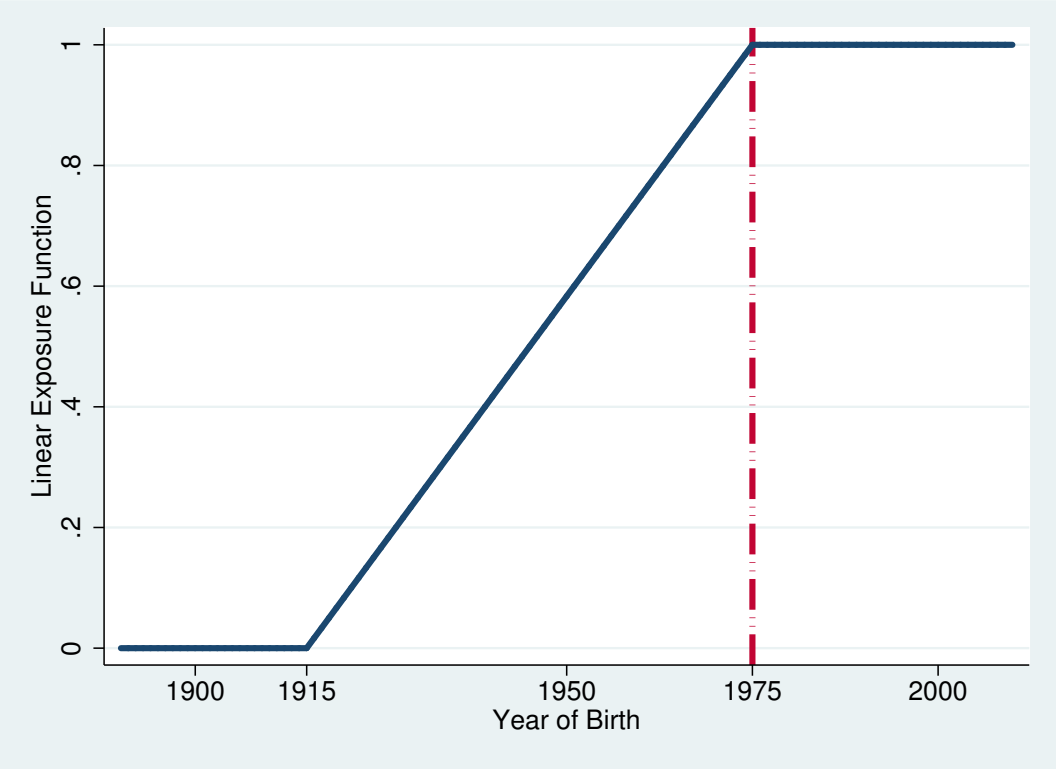
For instance, consider two people interviewed by the 1991 census, one who was 10 and another who was 30 years old at the time. The 30-year old adult in 1991 likely had considerable exposure to kissing bugs, whereas a 10-year old child in 1991 was probably born after eradication and so she had zero or very low potential exposure to the disease. Therefore, the younger individual, by having a lower probability of being infected, was more likely to benefit from the campaign than the older person. Similarly, two 20-year old individuals, but one in 1991 and the other in 2010, were likely to have been differently exposed to kissing bugs and accordingly to have been differently affected by the campaign. By attaching equal weights to all individuals sampled in the post-treatment census years, the standard pre/post empirical strategy disregards any possible differential gains from the program across individuals after the intervention.

Thus, in our second empirical strategy, we attach different weights across the population, according to their potential benefit from the program. Particularly, instead of giving unitary weight to all individuals in the census years of 1980 until 2010, we construct an exposure function that corresponds to the fraction of the individual's life exposed to the campaign efforts, which reflects the likelihood that the individual would directly benefit from the campaign. To build this function, we consider a life expectancy of 60 years, which is close to the life expectancy at birth in Brazil in 1970 (57.6 years). The exposure function thus reflects the individual degree of exposure to the program and is defined as

$$min\left\{ max\left\{ \frac{BirthYear - 1915}{60}, 0 \right\}, 1 \right\} \quad (5.3)$$

which is graphically represented in Figure 5.

Figure 5 – Life Exposure to Eradication Efforts



Notes: This graph displays on the fraction of life that is exposed to the campaign as a function of year of birth.

Hence, all individuals born after the beginning of the campaign (i.e. after 1975) were fully exposed to eradication efforts, so that their exposure function reaches the maximum value of 1 as they mostly likely had a life free of Chagas disease. Similarly, individuals older than 60 in 1975 were not young enough to have benefited from the campaign in order to change their educational, labor supply and fertility decisions by the time the program was adopted, and so they were the least exposed cohort (zero exposure). For those individuals who were already born but were younger than 60 years old when the program started, the exposure function is linear and increasing in the year of birth.

Accordingly, our second empirical strategy consists of estimating the following equation:

$$Y_{ijst} = \beta Prev_{j,pre} \cdot Exp_i + \Gamma X_{ijt} + \Phi X_{ijt} \cdot Post_t + \Theta Z_{j,pre} \cdot \delta_t + \delta_{st} + \delta_j + \varepsilon_{ijst} \quad (5.4)$$

where the post-treatment period dummy ($Post_t$) interacted with the baseline prevalence rate in equation (2) was replaced by the exposure function (Exp_i), which depends only on the year of birth of individual i. Standard errors are clustered at the MCA level.

6 Results

6.1 Effects on Mortality

Table 4 shows the results from the estimation of equation (1). Each column presents the estimated effect on mortality rate for a different age group. Specifications are indicated in each row, so that each pair row-column shows the result of a different regression.

Our estimates in Panel A in Table 4 indicate that highly infected areas indeed saw larger decreases in adult mortality rate due to Chagas disease relative to non-endemic regions. We find negative and statistically significant effects on mortality rate due to Chagas disease for people older than 15 years old, and no significant effect for the youngest cohort.

The results are robust to the inclusion of several controls, as suggested by Panel A. The first row shows the estimated effect using a specification that controls only for state-year and MCA fixed effects. Additionally to the fixed effects, we include the share of households in the MCA with electricity in 1970 (interacted with year-specific dummies) in the second row, which composes our basic specification. The third and fourth rows add the share of the labor force employed in agriculture and manufacturing and the log of household income per capita in the MCA in 1970 (interacted with year-specific dummies), respectively, to our basic specification. The last row of Panel A includes all of the above controls simultaneously, and the results remain close to the ones from the others specifications.

The estimated effects closely correspond to the theoretical expectations of the program's impact on mortality, given Chagas disease morbidity and mortality characteristics. Particularly, since no significant effect for the 0-15 age group is expected according to the symptoms and time pattern of the disease, the estimates for this age group can be interpreted as a placebo exercise. The absence of impact on child mortality thus reassures us of the validity of our empirical strategy.

As another placebo test, we assess the campaign's effects on deaths due to causes that were not supposed to be affected by the program, as reported in Panels B-F in Table 4. Using our most complete specification, we indeed find no significant effect on child and adult mortality rates due to causes of death other than Chagas disease.

Since Chagas disease may evolve into fatal cardiac complications, it is possible that some deaths primarily caused by Chagas disease may be reported as deaths due to heart diseases. Hence, in order to assess the relevance of misreporting in our sample, we estimate the intervention's impact on deaths due to heart diseases. Although we find a negative effect on deaths for the 15-30 age group, no statistically significant effect is estimated for the other age

Table 4 – Chagas Disease and Mortality Rate

Age Group:	0-15	15-30	30-45	45-60	60+
			Panel A: Chagas disease		
FE only	-0.0089	-0.12*	-0.45**	-1.00*	-1.41
	(0.011)	(0.067)	(0.21)	(0.57)	(1.18)
Basic specification	-0.011	-0.13*	-0.50**	-1.32**	-2.31*
	(0.012)	(0.068)	(0.20)	(0.54)	(1.19)
Add sectoral composition	-0.010	-0.13*	-0.48**	-1.27**	-2.31*
	(0.013)	(0.070)	(0.20)	(0.54)	(1.19)
Add income per capita	-0.011	-0.13*	-0.50**	-1.30**	-2.31*
	(0.012)	(0.068)	(0.20)	(0.54)	(1.19)
Full controls	-0.010	-0.13*	-0.48**	-1.28**	-2.31*
	(0.013)	(0.071)	(0.20)	(0.54)	(1.19)
			Panel B: Malaria		
Full controls	0.0057	0.0023	0.0037	0.034	-0.041
	(0.0068)	(0.0032)	(0.0037)	(0.037)	(0.035)
			Panel C: Helminthiases (include schistosomiasis, filarial infection and ancylostomiasis)		
Full controls	0.015	-0.0027	0.015	0.042	0.036
	(0.0097)	(0.018)	(0.030)	(0.028)	(0.049)
			Panel D: Intestinal infectious diseases (include cholera, typhoid fever, salmonella, food poisoning, amebiasis)		
Full controls	0.36	0.0043	-0.0043	0.020	0.14
	(0.26)	(0.0071)	(0.050)	(0.098)	(0.31)
			Panel E: Other infectious or parasitic diseases		
Full controls	0.43	0.0095	0.015	0.074	0.15
	(0.26)	(0.021)	(0.058)	(0.11)	(0.32)
			Panel F: Respiratory diseases		
Full controls	0.29	0.013	-0.18	-0.20	-1.61
	(0.18)	(0.067)	(0.18)	(0.28)	(1.10)
			Panel G: Heart diseases		
Full controls	-0.032	-0.18*	0.34	0.39	0.30
	(0.072)	(0.10)	(0.31)	(0.67)	(2.22)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and post dummy in equation (1) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the MCA level. The dependent variables are cause- and age-specific mortality rates (the causes of death considered are listed in the panels), per 100,000 residents. The age groups considered are listed in the column headings. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include state x year and MCA fixed effects. Additionally, basic controls specification includes the baseline share of households in the MCA with electricity. Income per capita adds the log of the baseline average household per capita income to the basic controls specification. Sectoral composition adds the baseline share of the labor force employed in agriculture and manufacturing to the basic controls specification. Full controls includes all of the above control variables.

*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

groups, as can be seen in Panel G in Table 4. Therefore, misreporting does not seem to be a primary issue in our case.

In Table 5, we investigate the effect on overall mortality. Panel A reproduces the effect on mortality due to Chagas disease from Table 4, whereas Panel B in Table 5 reports the impact of the intervention on overall mortality. We find a positive effect on child mortality and negative (although not statistically significant) point estimates for deaths during the productive years of the life cycle.

Nevertheless, as limitations in the treatment for Chagas disease may induce a substantial delay between the intervention and its impact on adult mortality, we also investigate the effect on overall mortality using only the year of 2010 as post-treatment period. The results are presented in Panel C in Table 5, and suggest that, giving enough time, the campaign indeed had a negative and statistically significant effect on adult mortality for the age group of 45 to 60.

Table 5 – Chagas Disease and Overall Mortality Rate

Age Group:	0-15	15-30	30-45	45-60	60+
Panel A: Chagas disease					
Full controls	-0.010	-0.13*	-0.48**	-1.28**	-2.31*
	(0.013)	(0.071)	(0.20)	(0.54)	(1.19)
Panel B: All causes of death					
Full controls	3.36***	-0.40	-0.26	-2.60	5.68
	(0.93)	(0.38)	(0.97)	(1.97)	(6.04)
Panel C: All causes of death (1980 and 2010 only)					
Full controls	3.91***	-0.61	-0.91	-5.17**	4.44
	(1.29)	(0.56)	(1.42)	(2.48)	(7.82)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and post dummy in equation (1) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the MCA level. The dependent variables are cause- and age-specific mortality rates (the causes of death considered are listed in the panels), per 100,000 residents. The age groups considered are listed in the column headings. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include state x year and MCA fixed effects, baseline share of households in the MCA with electricity, log of baseline average household per capita income, and baseline share of the labor force employed in agriculture and manufacturing.

*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Concerning child mortality, the estimated positive effect on overall mortality for children aged 0 to 15 suggests they were worse off in previously highly infected areas relative to low prevalence regions. This could insert a downward bias in our estimated effects on human capital decisions.

6.2 Effects on Education

In our first empirical strategy, we compare cohorts across time and MCAs in order to estimate the effects on education. For instance, we restrict our sample to individuals aged 10-20 and 20-30 when analyzing the effects on years of schooling and literacy[1], and to people aged 7-14, 15-18 and 18-25 to assess the effects on school attendance for school-aged individuals (elementary and middle school, high school, and college, respectively). Because the majority of the Brazilian population completes only high school education, we believe that these are the relevant age groups to be considered.

Table 6 shows the results from our first empirical strategy, based on a standard pre/post comparison. The outcomes and the age groups considered are indicated in column headings. Our estimates suggest that highly infected areas saw larger increases in educational attainment relative to non-endemic regions. In the first row, we only control for individual characteristics that might affect educational decisions (gender, an indicator for rural areas and age fixed effects), additionally to state-time and MCA fixed effects. The effects on education in this specification are positive and statistically significant at the 1% level, except for literacy in the age group of 20 to 30 years old.

The results are robust to the inclusion of several variables that control for possible pre-treatment differences across regions that might partially explain the changes in outcomes over time, as shown in the additional rows in Table 6. In the second row in Table 6, we include the share of households in the MCA with electricity in 1970 (interacted with time-specific dummies) to the individual controls specification, thus constructing our basic controls specification. The results are robust to the inclusion of this covariate, suggesting that pre-treatment differences in development levels across MCAs are not leading our estimates. In the third row in Table 6 we include a nonlinear trend on the average MCA baseline household income per capita to our basic controls specification. The robustness of our results indicate that our estimates are not simply reflecting a convergence in income over time across MCAs. In the fourth row in Table 6 we add a nonlinear trend on pre-treatment sectoral composition of the labor force, and our estimates remain positive and statistically significant.

Adding all of the above covariates simultaneously, we estimate a positive and statistically significant effect on school attendance, literacy and years of schooling, as shown in the last row in Table 6. Our estimates suggest that a 7.4 p.p. reduction in Chagas disease prevalence (which corresponds to eradication in an average MCA) would have increased years of schooling by 0.03 year, which represents a 1.5% increase. Similarly, the literacy rate would have increased by 1.0 p.p. or 1.7%, and school attendance would have increased by up to 1.2 p.p. or 1.8%, thus

[1] The reason why we split the age groups of 10-20 and 20-30 instead of merging them into one group is based on the different degrees of exposure to the campaign efforts of each group. A 30-year old individual in 1991 already had some considerable potential exposure to kissing bugs throughout his life, whereas a 10-year old child in 1991 was probably born after eradication and so she had zero or very low potential exposure to the disease. The knowledge of this fact has motivated our second empirical strategy, discussed in Section 5.

suggesting a significant role of health in human capital decisions.

Table 6 – Chagas Disease and Human Capital: Main Effects

Dependent variables:	Years of education		Literacy		School attendance		
Age groups:	10-20	20-30	10-20	20-30	7-14	15-18	18-25
Individual controls	0.0062***	0.0054***	0.0019***	0.00067	0.0024***	0.0019***	0.00056***
	(0.0021)	(0.0015)	(0.00070)	(0.00084)	(0.00038)	(0.00055)	(0.00019)
Basic controls	0.0043***	0.0039***	0.0017***	0.00062**	0.0019***	0.0016***	0.00060**
	(0.0011)	(0.0015)	(0.00041)	(0.00027)	(0.00047)	(0.00040)	(0.00026)
Income per capita	0.0046***	0.0043***	0.0019***	0.00087***	0.0019***	0.0015***	0.00052**
	(0.0011)	(0.0016)	(0.00045)	(0.00031)	(0.00046)	(0.00036)	(0.00023)
Sectoral composition	0.0032***	0.0027*	0.0011***	0.00015	0.0015***	0.0015***	0.00064**
	(0.0011)	(0.0016)	(0.00039)	(0.00030)	(0.00042)	(0.00044)	(0.00031)
Full controls	0.0037***	0.0033**	0.0014***	0.00043	0.0016***	0.0015***	0.00060**
	(0.0011)	(0.0017)	(0.00041)	(0.00029)	(0.00043)	(0.00042)	(0.00028)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and post dummy in equation (2) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings, such as the age groups considered. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for male, rural, age fixed effects, state x year and MCA fixed effects; and the interactions of the demographic controls with $Post_t$. Additionally, basic controls specification includes the share of households in the MCA with electricity. Income per capita adds the average household per capita income to the basic controls specification. Sectoral composition adds the share of the labor force employed in agriculture and manufacturing to the basic controls specification.

*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table 7 presents additional sensitivity tests. First, we exclude data from the 1980 census, as it can be difficult to characterize this year into pre/post-treatment categories. As discussed in Section 3, insecticide-spraying was implemented in a phase-in system, so that some areas were not reached by 1980. For those areas, the year of 1980 might be seen as a pre-treatment period instead of post. The results are presented in Panel B and they are not statistically different from the ones obtained in panel A.

In order to mitigate some migration concerns due to matching individuals to the baseline prevalence rate in their place of residence and to the lack of information on municipality of birth, we restrict the sample to nonmovers, therefore reassuring that the assigned prevalence rate was indeed the relevant one. The results are shown in Panel C in Table 7, and they do not differ statistically from the estimates in Panel A.

As will be discussed in our analysis of possible mechanisms in Section 7, in addition to the channel suggested by human capital theory, there may be another mechanism behind the results, as the health of the parents may affect investments in children through income and intrahousehold allocation of time. In order to infer how important this channel is to our results,

Table 7 – Chagas Disease and Human Capital: Sensitivity Tests

Dependent variables:	Years of education		Literacy		School attendance		
Age groups:	10-20	20-30	10-20	20-30	7-14	15-18	18-25
			Panel A: Main Effects				
Full controls	0.0037***	0.0033**	0.0014***	0.00043	0.0016***	0.0015***	0.00060**
	(0.0011)	(0.0017)	(0.00041)	(0.00029)	(0.00043)	(0.00042)	(0.00028)
			Panel B: Robustness - Exclude census year of 1980				
Full controls	0.0032**	0.0043**	0.0015***	0.00053*	0.0017***	0.0016***	0.00076**
	(0.0013)	(0.0020)	(0.00044)	(0.00031)	(0.00046)	(0.00045)	(0.00030)
			Panel C: Robustness - Nonmovers				
Full controls	0.0033***	0.0025	0.0014***	0.00040	0.0015***	0.0014***	0.00045**
	(0.0012)	(0.0023)	(0.00041)	(0.00033)	(0.00042)	(0.00030)	(0.00020)
			Panel D: Robustness - Add disability dummy (1970-1991)				
Full controls	0.0019*	-0.0019	0.00093***	-0.000047	0.0015***	0.0015***	0.00024
	(0.0010)	(0.0016)	(0.00026)	(0.00022)	(0.00039)	(0.00048)	(0.00026)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and post dummy in equation (2) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings, such as the age groups considered. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for male, rural, age fixed effects (and the interactions of the demographic controls with $Post_t$), state x year and MCA fixed effects, share of households in the MCA with electricity, average household per capita income, share of the labor force employed in agriculture and manufacturing. For panel B, the original sample is restricted to 1970, 1991, 2000 and 2010 census years. For panel C, the original sample is restricted to nonmovers. For panel D, the original sample is restricted to 1970, 1980 and 1991 census years, and an indicator variable for whether there is a disable adult in the household is added to the full controls specification.

*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

we conduct an indirect test by including a dummy for whether the individual lives in a household with a disabled adult to the full controls specification.

Information on disability was available only in the censuses of 1970, 1980 and 1991. Therefore, only those census years were used to estimate the effects in Panel D. This might explain the reduced magnitude and the lack of significance of the education effects for older cohorts, as they were considerably less exposed to the campaign efforts. Indeed, there is no significant difference from the estimates in Panel D in Table 7 and the ones from Panel C in appendix Table A1, where we restrict the sample to the census years of 1970 to 1991, in order to produce a more reliable comparison. Our estimates thus suggest this is not the leading channel of the estimated effects on education.

As a placebo exercise for years of education and literacy, we estimate exactly the same specifications for individuals that, by being older, were less likely to alter their educational choices in response to the eradication campaign. We conduct the same exercise as above but restrict our sample to people aged from 60 to 80 years old. The placebo results, reported in Panel B in Table 8, are not statistically significant, and suggest there was no prior difference in trends across MCAs.

Nevertheless, it is not possible to conduct such analysis for school attendance, as there are very few older people attending school and they might not be representative of the overall population. We then conduct an indirect test for school attendance, which consists on estimating the effects using only 1970 and 1980 censuses. Given that those individuals had only a minor exposure to the campaign efforts (when compared to 1991, 2000 and 2010 censuses), we would expect the magnitude of the estimated effects to be lower than in our main findings. Indeed, we find significantly lower effects on school attendance when restricting our sample to 1970 and 1980 only, as described in Panel C in Table 8.

We also analyze the effects on educational outcomes over time, as reported in Table 9. The estimated effects on years of education and literacy for older cohorts are increasing over time, whereas the effects on school attendance reach their maximum in 2000 in general. These findings suggest that it may take longer in order to fully impact educational decisions, specially for older cohorts. This is consistent with Chagas disease epidemiology and the absence of cure for most infected individuals, and it might indicate that our second empirical strategy is more appropriate for assessing educational effects.

Using our second empirical strategy, which explores the fraction of the lifetime exposed to the campaign efforts, we find significantly larger effects on educational outcomes (although the effects on years of education are not statistically significant), as reported in Table 10. This may indicate that we were underestimating our results by assigning equal weights to every person interviewed in the 1980-2010 censuses, as if the individual degree of exposure to the campaign was homogeneous within and between census years. Considering that individual gains

Table 8 – Chagas Disease and Human Capital: Placebo Tests

Dependent variables:	Years of education		Literacy		School attendance		
			Panel A: Main Effects				
Age groups:	10-20	20-30	10-20	20-30	7-14	15-18	18-25
Full controls	0.0037***	0.0033**	0.0014***	0.00043	0.0016***	0.0015***	0.00060**
	(0.0011)	(0.0017)	(0.00041)	(0.00029)	(0.00043)	(0.00042)	(0.00028)
			Panel B: Placebo Test - Years of Education and Literacy: Age Groups 60-70 and 70-80				
Age groups:	60-70	70-80	60-70	70-80			
Full controls	0.0019	0.0019	-0.000082	0.000038			
	(0.0014)	(0.0015)	(0.00041)	(0.00044)			
			Panel C: Test - School Attendance: Using only 1980 as Post-treatment				
Age groups:					7-14	15-18	18-25
Full controls					0.00069*	0.0012**	0.00010
					(0.00039)	(0.00051)	(0.00033)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and post dummy in equation (2) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings. The age groups considered are stated in each panel. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for male, rural, age fixed effects (and the interactions of the demographic controls with $Post_t$), state x year and MCA fixed effects, share of households in the MCA with electricity, average household per capita income, share of the labor force employed in agriculture and manufacturing. For panel B, the effect is estimated for individuals aged 60 to 70 and 70 to 80 years old (separately). For panel C, the original sample is restricted to 1970 and 1980 census years, and the age groups considered are the same as the original analysis.
*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

from Chagas disease eradication might differ across individuals in the post-treatment period enable us to provide a more accurate measure of the effects of health on educational outcomes.

Our estimates based on the exposure function indicate that a 7.4 p.p. reduction in Chagas disease prevalence (which corresponds to eradication in an average MCA) increased years of schooling by 0.07 year (3.6%), literacy rate by 1.6 p.p. (2.7%) and school attendance by up to 5.5 p.p. (8.2%). Therefore, our results are robust to several specifications and even to a modified empirical strategy, and so do not raise severe concerns about the internal validity of our estimates.

6.3 Effects on Labor Force Participation

Tables 11 and 12 show the estimated effects on male and female labor force participation using equation (2), respectively, for each age group in 5-year intervals indicated in the rows. In

Table 9 – Chagas Disease and Human Capital: Year-Specific Effects

Dependent variables:	Years of education		Literacy		School attendance		
Age groups:	10-20	20-30	10-20	20-30	7-14	15-18	18-25
1980	0.0051***	-0.00031	0.00073***	-0.00011	0.0013***	0.0013***	-0.000031
	(0.0014)	(0.0030)	(0.00028)	(0.00030)	(0.00037)	(0.00035)	(0.00021)
1991	0.00088	-0.00050	0.0011***	0.000082	0.0018***	0.0018***	0.00040
	(0.0012)	(0.0023)	(0.00032)	(0.00024)	(0.00043)	(0.00053)	(0.00028)
2000	0.0027*	0.0024	0.0018***	0.00061*	0.0017***	0.0020***	0.0011***
	(0.0014)	(0.0022)	(0.00050)	(0.00032)	(0.00050)	(0.00046)	(0.00033)
2010	0.0067***	0.0091***	0.0017***	0.00084**	0.0015***	0.00080*	0.00066*
	(0.0018)	(0.0023)	(0.00055)	(0.00041)	(0.00053)	(0.00045)	(0.00035)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and year-specific dummies in equation (2) using OLS. Each column reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings, such as the age groups considered. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for male, rural, age fixed effects (and the interactions of the demographic controls with $Post_t$), state x year and MCA fixed effects, share of households in the MCA with electricity, average household per capita income, share of the labor force employed in agriculture and manufacturing.
*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

this empirical exercise we consider individuals that were at least 10 years old (those for whom the question on labor force participation was asked). Due to major differences in male and female labor supply incentives and decisions, we examine those groups separately.

Our estimates suggest that high prevalence areas saw larger decreases in male labor force participation during young adulthood, as can be seen in Table 11. The results are robust to the inclusion of several control variables. Nevertheless, when we change the dependent variable to be the activity indicator (defined as an indicator for whether the individual works or studies), we estimate a robust positive and statistically significant effect for the youngest cohort (10 to 15 years old) whereas the effect for young adults vanishes, as described in the last column of Table 11. Therefore, we observe a substitution from work towards study by male children, which might be interpreted as higher investments in children's education.

When considering the effect on overall male adults, we estimate a positive impact on labor force participation, specially for people older than 40 years. Considering the epidemiology of Chagas disease, we would indeed expect more pronounced effects of the campaign for older cohorts. Given that the disease usually presents itself in a more severe form at later stages, a higher impact on labor force participation would be expected for older people.

Although we find a similar pattern of impact across age for female labor force participation, the only statistically significant effects are for girls aged 10 to 15 years old and for women

Table 10 – Chagas Disease and Human Capital: Linear Life Exposure Function

Dependent variables:	Years of education	Literacy	School attendance		
Age groups:	10-30	10-30	7-14	15-18	18-25
Individual controls	0.013	0.0026	0.0093***	0.0041***	0.0017***
	(0.021)	(0.0023)	(0.0015)	(0.0011)	(0.00065)
Basic controls	0.012	0.0027***	0.0084***	0.0037***	0.0018**
	(0.016)	(0.00077)	(0.0022)	(0.0012)	(0.00082)
Income per capita	0.012	0.0032***	0.0084***	0.0035***	0.0016**
	(0.017)	(0.00080)	(0.0022)	(0.0010)	(0.00068)
Sectoral composition	0.0085	0.0017**	0.0072***	0.0035***	0.0019**
	(0.017)	(0.00078)	(0.0020)	(0.0013)	(0.00096)
Full controls	0.0092	0.0022***	0.0075***	0.0035***	0.0018**
	(0.017)	(0.00074)	(0.0021)	(0.0012)	(0.00083)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and linear exposure function in equation (4) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings, such as the age groups considered. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for male, rural, age fixed effects, state x year and MCA fixed effects; and the interactions of the demographic controls with $Post_t$. Additionally, basic controls specification includes the share of households in the MCA with electricity. Income per capita adds the average household per capita income to the basic controls specification. Sectoral composition adds the share of the labor force employed in agriculture and manufacturing to the basic controls specification.
*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

aged 60 to 65 years old, as reported in Table 12. We also estimate a negative effect on work and a positive and significant effect on the activity indicator for girls aged 10 to 15 years old, indicating that a faster substitution from work-related activities towards study in more endemic areas, relative to non-endemic ones, was not gender-specific.

As a placebo test we estimate our most complete specification restricting our sample to individuals that, by being old enough, were less likely to have altered their labor decisions in response to the campaign. This corresponds to people older than 55 years old (usual retirement age) by the time the program was implemented in 1975. Table 13 reports the results from the placebo test. In Panels A through D we restrict the sample to census years of 1970 and 1980, and each panel reports the coefficient for a different age group, for which we would not expect major effects on labor decisions. Although the estimates regarding female labor force participation are not statistically significant, we find some positive and significant effects on male labor force participation. This may indicate that we might be overestimating the effect on male labor force participation for older cohorts and underestimating it for children and young adults.

Nevertheless, it may be even more important to use our second empirical strategy when

Table 11 – Chagas Disease and Male Labor Force Participation

Dependent variables:	Labor Force Participation					Activity
	Individual controls	Basic controls	Income per capita	Sectoral Composition	Full controls	Full controls
age 10-15	-0.0011**	-0.0010**	-0.0010**	-0.00077*	-0.00082**	0.0011***
	(0.00043)	(0.00046)	(0.00043)	(0.00041)	(0.00041)	(0.00031)
age 15-20	-0.0015***	-0.0014***	-0.0013***	-0.0012***	-0.0012***	0.000071
	(0.00044)	(0.00047)	(0.00043)	(0.00044)	(0.00043)	(0.00021)
age 20-25	-0.00036	-0.00038*	-0.00037*	-0.00029	-0.00031	-0.00014
	(0.00028)	(0.00022)	(0.00023)	(0.00023)	(0.00024)	(0.00018)
age 25-30	-0.00028	-0.00028*	-0.00028*	-0.00017	-0.00020	-0.00012
	(0.00021)	(0.00016)	(0.00016)	(0.00015)	(0.00016)	(0.00015)
age 30-35	-0.000096	-0.000083	-0.00014	0.0000042	-0.000055	0.000016
	(0.00022)	(0.00015)	(0.00016)	(0.00015)	(0.00015)	(0.00014)
age 35-40	0.000049	0.000066	0.0000052	0.00012	0.000077	0.000088
	(0.00027)	(0.00017)	(0.00018)	(0.00016)	(0.00017)	(0.00016)
age 40-45	0.00026	0.00032**	0.00020	0.00036**	0.00027*	0.00027*
	(0.00041)	(0.00014)	(0.00014)	(0.00015)	(0.00015)	(0.00014)
age 45-50	0.00027	0.00039**	0.00032**	0.00033*	0.00029*	0.00031*
	(0.00048)	(0.00016)	(0.00015)	(0.00017)	(0.00017)	(0.00017)
age 50-55	0.00091	0.0011***	0.0010***	0.0010***	0.00099***	0.0010***
	(0.00070)	(0.00027)	(0.00024)	(0.00027)	(0.00026)	(0.00026)
age 55-60	0.0011	0.0014***	0.0012***	0.0012***	0.0011***	0.0011***
	(0.00082)	(0.00043)	(0.00034)	(0.00037)	(0.00032)	(0.00033)
age 60-65	0.0014	0.0017***	0.0015***	0.0015***	0.0014***	0.0015***
	(0.00097)	(0.00043)	(0.00037)	(0.00039)	(0.00036)	(0.00037)
age 65-70	0.0020	0.0020***	0.0016***	0.0015***	0.0014***	0.0014***
	(0.0018)	(0.00053)	(0.00041)	(0.00047)	(0.00042)	(0.00042)
age 70-75	0.0013	0.0012**	0.00099**	0.0012**	0.0011**	0.0011**
	(0.0015)	(0.00053)	(0.00049)	(0.00057)	(0.00053)	(0.00054)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and post dummy in equation (2) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings, such as the specifications considered. The age groups considered are listed in the rows. The activity indicator is defined as an indicator variable for whether the individual works or studies. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for rural, age fixed effects, state x year and MCA fixed effects; and the interactions of the demographic controls with $Post_t$. Additionally, basic controls specification includes the share of households in the MCA with electricity. Income per capita adds the average household per capita income to the basic controls specification. Sectoral composition adds the share of the labor force employed in agriculture and manufacturing to the basic controls specification.

*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table 12 – Chagas Disease and Female Labor Force Participation

Dependent variables:	Labor Force Participation					Activity
	Individual controls	Basic controls	Income per capita	Sectoral Composition	Full controls	Full controls
age 10-15	-0.00018	-0.00028*	-0.00026*	-0.000099	-0.00013	0.0011***
	(0.00034)	(0.00016)	(0.00015)	(0.00014)	(0.00014)	(0.00039)
age 15-20	-0.00031	-0.00049	-0.00033	0.000094	0.000069	0.00047
	(0.00074)	(0.00047)	(0.00040)	(0.00036)	(0.00036)	(0.00030)
age 20-25	0.000042	-0.00026	-0.00018	0.000018	0.000017	0.00020
	(0.00021)	(0.00037)	(0.00035)	(0.00031)	(0.00031)	(0.00024)
age 25-30	0.00019	-0.000035	0.000020	0.00015	0.00015	0.00028
	(0.00030)	(0.00029)	(0.00028)	(0.00025)	(0.00025)	(0.00023)
age 30-35	-0.000050	-0.00030	-0.00030	-0.000087	-0.00013	-0.000037
	(0.00024)	(0.00028)	(0.00028)	(0.00024)	(0.00025)	(0.00023)
age 35-40	-0.000073	-0.00032	-0.00023	-0.00012	-0.00010	0.000045
	(0.00019)	(0.00035)	(0.00032)	(0.00031)	(0.00030)	(0.00027)
age 40-45	-0.00012	-0.00037	-0.00036	-0.00018	-0.00020	-0.00010
	(0.00022)	(0.00036)	(0.00034)	(0.00032)	(0.00032)	(0.00031)
age 45-50	0.000089	-0.00011	-0.00013	0.00010	0.000044	0.00016
	(0.00041)	(0.00024)	(0.00026)	(0.00025)	(0.00026)	(0.00026)
age 50-55	0.00019	-0.000013	0.000011	0.00012	0.00011	0.00015
	(0.00047)	(0.00023)	(0.00025)	(0.00026)	(0.00026)	(0.00025)
age 55-60	0.00016	-0.000052	-0.000090	-0.00014	-0.00014	-0.000041
	(0.00062)	(0.00022)	(0.00023)	(0.00023)	(0.00024)	(0.00025)
age 60-65	0.00067	0.00050**	0.00058***	0.00044**	0.00050**	0.00060***
	(0.00046)	(0.00020)	(0.00022)	(0.00022)	(0.00023)	(0.00023)
age 65-70	0.00025	0.000019	0.000084	0.000011	0.000056	0.00014
	(0.00056)	(0.00023)	(0.00025)	(0.00025)	(0.00026)	(0.00027)
age 70-75	0.00039	0.00024	0.00018	0.00022	0.00019	0.00020
	(0.00041)	(0.00021)	(0.00022)	(0.00023)	(0.00024)	(0.00024)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and post dummy in equation (2) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings, such as the specifications considered. The age groups considered are listed in the rows. The activity indicator is defined as an indicator variable for whether the individual works or studies. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for rural, age fixed effects, state x year and MCA fixed effects; and the interactions of the demographic controls with $Post_t$. Additionally, basic controls specification includes the share of households in the MCA with electricity. Income per capita adds the average household per capita income to the basic controls specification. Sectoral composition adds the share of the labor force employed in agriculture and manufacturing to the basic controls specification.

*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table 13 – Chagas Disease and Labor Force Participation: Placebo Tests

Dependent variables:	Male Labor Force Participation	Female Labor Force Participation
Panel A: 60 to 65 years old, census years 1970-1980		
Full controls	0.0013**	0.00029
	(0.00056)	(0.00031)
Panel B: 65 to 70 years old, census years 1970-1980		
Full controls	0.00090	0.000049
	(0.00072)	(0.00036)
Panel C: 70 to 75 years old, census years 1970-1980		
Full controls	0.0016**	-0.000060
	(0.00076)	(0.00030)
Panel D: 75 to 80 years old, census years 1970-1980		
Full controls	0.00063	-0.00052
	(0.00097)	(0.00036)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and post dummy in equation (2) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings. The age groups considered are stated in each panel. The sample is restricted to 1970 and 1980 census years. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for rural, age fixed effects (and the interactions of the demographic controls with $Post_t$), state x year and MCA fixed effects, share of households in the MCA with electricity, average household per capita income, share of the labor force employed in agriculture and manufacturing.

*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

investigating the effects on labor decisions than educational outcomes, given that decisions on labor force participation and labor supply are a function of exposure during the entire adult life of an individual, whereas educational decisions are usually concentrated in the first decades of an individual's life. Therefore, the standard pre/post comparison might underestimate considerably the effects on labor decisions.

Table 14 shows the results of our parameter of interest in equation (4) for labor force participation, activity indicator and work hours, when we consider the fraction of individual life exposed to the intervention. We estimate a positive and statistically significant effect on both male and female labor force participation, as indicated in the first and fourth columns in Table 14. This suggests that the standard pre/post comparison indeed induce a considerable downward bias in our estimates.

Our estimates based on the exposure function suggest that a 7.4 p.p. reduction in Chagas disease prevalence (which corresponds to eradication in an average MCA) increased male and female labor force participation by 0.4 p.p (or 0.6%) and 1.7 p.p. (or 11.4%), respectively.

Furthermore, we estimate a positive and statistically significant effect on the intensive margin of labor supply. These results are shown in columns 3 and 6 in Table 14, and indicate that besides encouraging more people to join the labor force, the campaign also increased individual's incentives to work more hours.

Hence, our results suggest that high prevalence areas saw larger increases in male and female labor supply (in both intensive and extensive margins), relative to non-endemic regions. Furthermore, the substantially higher effects on female labor force participation relative to male suggest that longevity gains and health improvements affected women's incentives to work, encouraging women to engage in the labor force.

6.4 Effects on Fertility

Table 15 shows the estimated effects on fertility, measured by an indicator variable for having at least one child and by total number of children. We restrict our sample to women that were at least 15 years old, since questions concerning fertility were not addressed to girls younger than 15 years old in the census[2].

Our results indicate mild effects on the fraction of women aged 15 to 25 years old with at least one child, and no effect on women aged 25 to 35 years old, as suggested by columns 1 and 2 in Table 15. Therefore, the change in health and mortality induced by the program does not seem to have strongly affected the extensive margin of fertility decision. Nevertheless, when considering the intensive margin of fertility decision, we find a positive and statistically

[2] In the census years of 1970 and 1980, fertility questions were not addressed to girls younger than 15 years old. In 1991, those questions started to be directed at girls that were 10 years old or older.

Table 14 – Chagas Disease and Labor Force: Linear Life Exposure Function

Dependent variables:	Male Labor Force Participation	Male Activity Indicator	Male Work Hours	Female Labor Force Participation	Female Activity Indicator	Female Work Hours
Individual controls	0.00058	0.00074	0.00037	0.0021***	0.0022***	0.0031***
	(0.00042)	(0.0017)	(0.00066)	(0.00037)	(0.00082)	(0.00073)
Basic controls	0.00057**	0.00077	0.00066*	0.0021***	0.0022**	0.0036***
	(0.00026)	(0.0015)	(0.00037)	(0.00036)	(0.00092)	(0.00050)
Income per capita	0.00054**	0.00072	0.00097***	0.0022***	0.0022**	0.0041***
	(0.00026)	(0.0015)	(0.00031)	(0.00034)	(0.00096)	(0.00054)
Sectoral composition	0.00059**	0.00072	0.0010***	0.0023***	0.0024**	0.0045***
	(0.00026)	(0.0015)	(0.00037)	(0.00031)	(0.00099)	(0.00052)
Full controls	0.00056**	0.00069	0.0011***	0.0023***	0.0023**	0.0045***
	(0.00026)	(0.0015)	(0.00035)	(0.00031)	(0.00100)	(0.00054)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and linear exposure function in equation (4) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings, and the age group considered in all regressions are individuals aged 10 to 80 years old. Work hours is a categorical variable that ranges from 1 to 3, where 1 corresponds to working less than 15 weekly hours, 2 corresponds to working between 15 and 40 weekly hours and 3 corresponds to working at least 40 weekly hours. Work hours are relative to the main job. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for rural, age fixed effects, state x year fixed effects, MCA fixed effects; and the interactions of the demographic controls with $Post_t$. Additionally, basic controls specification includes the share of households in the MCA with electricity. Income per capita adds the average household per capita income to the basic controls specification. Sectoral composition adds the share of the labor force employed in agriculture and manufacturing to the basic controls specification.

*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

significant effect, suggesting that high prevalence areas saw larger increases in fertility (total number of children) relative to less infected regions.

As a placebo test we estimate our preferred specification (full controls) restricting our sample to women older than 49 years old by the time the campaign was implemented. This corresponds to women older than 54 years old in the 1980 census and older than 65 years old in the 1991 census. Those women were not young enough to have altered their fertility decisions in response to the eradication efforts. As reported in Table 16, the estimates are not statistically significant in general, and so do not raise major internal validity concerns.

Analyzing the effects on fertility over time, we observe a fading out throughout the decades for older cohorts, as reported in Table 17. The estimated effects on the total number of children and on the indicator for having at least one child for women aged 25-35 suggest that these effects were significantly reduced over time.

As fertility decisions are carried out during a significant fraction of woman's adult life, it may take more time to the campaign fully impact every woman. Table 18 presents the results using our second empirical strategy, which considers the fraction of life potentially free of Chagas disease transmission. By considering this individual heterogeneity dimension, we find negative point estimates, although the estimated effects on fertility are not statistically significant. Comparing the results on Table 18 with the ones presented on Table 15 may indicate

Table 15 – Chagas Disease and Fertility

Dependent variables:	At least one child		Total Children	
Age groups:	15-25	25-35	15-25	25-35
Individual controls	0.00018	-0.00011	0.00077*	0.0047***
	(0.00031)	(0.00054)	(0.00041)	(0.0017)
Basic controls	0.00018	-0.00012	0.00084**	0.0052***
	(0.00020)	(0.00027)	(0.00041)	(0.0016)
Income per capita	0.00020	-0.00014	0.00094**	0.0054***
	(0.00020)	(0.00027)	(0.00042)	(0.0016)
Sectoral composition	0.00022	-0.000035	0.00081*	0.0049***
	(0.00021)	(0.00028)	(0.00046)	(0.0015)
Full controls	0.00023	-0.000061	0.00089**	0.0050***
	(0.00020)	(0.00028)	(0.00045)	(0.0015)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and post dummy in equation (2) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings, such as the age groups considered. At least one child is an indicator variable for whether the women had any children born alive, and total children denotes the total number of children born alive. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for rural, age fixed effects, state x year and MCA fixed effects; and the interactions of the demographic controls with $Post_t$. Additionally, basic controls specification includes the share of households in the MCA with electricity. Income per capita adds the average household per capita income to the basic controls specification. Sectoral composition adds the share of the labor force employed in agriculture and manufacturing to the basic controls specification.
*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

that younger (and more exposed) cohorts are tilting the effect on fertility towards having fewer children relative to older and less exposed ones.

Table 16 – Chagas Disease and Fertility: Placebo Tests

Dependent variables:	At least one child	Total Children
Panel A: 55 to 65 years old, census years 1970-1980		
Full controls	-0.00047**	0.00091
	(0.00023)	(0.0033)
Panel B: 65 to 75 years old, census years 1970-1980		
Full controls	0.00034	0.0021
	(0.00037)	(0.0041)
Panel C: 65 to 75 years old, census years 1970-1991		
Full controls	0.000011	0.0027
	(0.00027)	(0.0038)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and post dummy in equation (2) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings. The age groups considered are stated in each panel. In panels A and B, the sample is restricted to 1970 and 1980 census years, and in panel C the sample is restricted to 1970, 1980 and 1991 census years. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for rural, age fixed effects (and the interactions of the demographic controls with $Post_t$), state x year and MCA fixed effects, share of households in the MCA with electricity, average household per capita income, share of the labor force employed in agriculture and manufacturing.

*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table 17 – Chagas Disease and Fertility: Year-Specific Effects

Dependent variables:	At least one child		Total Children	
Age groups:	15-25	25-35	15-25	25-35
1980	0.00040*	0.00053**	0.00092*	0.0055***
	(0.00021)	(0.00026)	(0.00047)	(0.0018)
1991	0.00010	0.00020	0.0011**	0.0067***
	(0.00027)	(0.00030)	(0.00050)	(0.0017)
2000	0.00012	-0.00014	0.00080	0.0061***
	(0.00025)	(0.00033)	(0.00054)	(0.0017)
2010	0.00034	-0.00047	0.00084*	0.0027*
	(0.00021)	(0.00034)	(0.00046)	(0.0016)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and year-specific dummies in equation (2) using OLS. Each column reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings, such as the age groups considered. At least one child is an indicator variable for whether the women had any children born alive, and total children denotes the total number of children born alive. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for rural, age fixed effects (and the interactions of the demographic controls with $Post_t$), state x year and MCA fixed effects, share of households in the MCA with electricity, average household per capita income, share of the labor force employed in agriculture and manufacturing.

*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table 18 – Chagas Disease and Fertility: Linear Life Exposure Function

Dependent variables:	At least one child	Total Children
Individual controls	-0.00080	-0.0029
	(0.0011)	(0.010)
Basic controls	-0.00079	-0.0017
	(0.00098)	(0.011)
Income per capita	-0.00071	-0.0013
	(0.00095)	(0.012)
Sectoral composition	-0.00068	-0.0022
	(0.00100)	(0.011)
Full controls	-0.00066	-0.0019
	(0.00098)	(0.011)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and linear exposure function in equation (4) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings, and the sample considered in all regressions are composed by women aged 15 to 80 years old. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for rural, age fixed effects, state x year fixed effects, MCA fixed effects; and the interactions of the demographic controls with $Post_t$. Additionally, basic controls specification includes the share of households in the MCA with electricity. Income per capita adds the average household per capita income to the basic controls specification. Sectoral composition adds the share of the labor force employed in agriculture and manufacturing to the basic controls specification.

*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

7 Discussion

The theoretical foundations of our interpretation of the results borrow from the human capital theory, according to which longer life expectancy fosters educational attainment by raising the rate of return to human capital investment (e.g. Becker, 1964; Soares, 2005).

Since Chagas disease burden is manifested mainly during the productive years of the life cycle, the end of its transmission sharply reduces adult mortality and morbidity. As adults live longer and healthier lives[1], their productive horizon is augmented, thus increasing the rate of return to investments in human capital. Therefore, we would expect educational investment to rise.

Despite the main mechanism from human capital theory stated above, the results might also be driven by another channel, as Chagas disease may affect investments in children through the health conditions of their parents. Investments in children's education may be reduced through diminished family income if sickness and disability of parents reduce work hours, which in turn might also induce intrahousehold reallocation of work time. It is thus possible that children may have to work in order to provide for their family or receive lower investments.

In order to investigate how important this mechanism is to explain our results, we estimate equation (2) using a disability indicator for whether there is a disabled adult in the household and household income per capita as dependent variables. The results can be found in Table 19. We do not find robust effects on income, but we do find a negative effect on disability, indicating that high prevalence areas saw larger decreases in disability relative to low prevalence regions. Although disability of the parents may be an important mechanism to explain our estimates of the effects on labor force participation for children and young adults, it does not seem to be leading our results for educational outcomes (as indicated by a comparison between Panel D in Table 7 and Panel C in appendix Table A1).

Another possible mechanism leading to our educational effects is related to a scarring effect, since the initial phase of the disease could be quite severe during early childhood. Child development may be affected by severe infection during early childhood, thus impacting educational outcomes in the future. Notwithstanding, we understand that our results are not primarily driven by this channel, as symptom onset of the initial phase was relatively rare (with an occurrence rate lower than 2%) and the occurrence of severe symptoms was even rarer.

[1] As adult morbidity of Chagas disease is closely related to disability, lower morbidity translates into increases in the productive horizon of the life cycle. We believe that mortality adjusted by disability is the relevant dimension of health to be considered in the particular case of Chagas disease. A mortality rate that does not consider potential losses from disability misses a relevant dimension regarding life cycle decisions such as labor supply and retirement, which play a determinant role in education investment decisions. Therefore, although we cannot separate morbidity and mortality effects in our analysis, we believe that this is in fact the relevant dimension of health to be considered in this particular case.

Table 19 – Chagas Disease and Education: Mechanisms

Dependent variables:	Disability Indicator	Household Income per capita (log)
Individual controls	-0.000045	-0.0011
	(0.00016)	(0.00080)
Basic controls	-0.000096	-0.0014*
	(0.000091)	(0.00073)
Income per capita	-0.00021***	0.00064
	(0.000076)	(0.00092)
Sectoral composition	-0.00017*	-0.0012
	(0.000092)	(0.00081)
Full controls	-0.00024***	0.00015
	(0.000085)	(0.00077)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and post dummy in equation (2) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the household level in the disability regression and the head of the household in the income regression. The dependent variables are listed in the column headings, and the sample considered in income regressions are composed by people aged 10 to 80 years old. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for rural, state x year fixed effects, MCA fixed effects; and the interactions of the demographic controls with $Post_t$. Additionally, income regressions include gender dummy and age fixed effects (and its interactions with $Post_t$). Additionally, basic controls specification includes the share of households in the MCA with electricity. Income per capita adds the average household income per capita to the basic controls specification. Sectoral composition adds the share of the labor force employed in agriculture and manufacturing to the basic controls specification.

*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Regarding fertility, theory suggests that gains in life expectancy due to adult mortality reductions discourage fertility (Soares, 2005). As adult mortality decreases, the benefits from large families diminish, moving the quantity-quality trade-off towards fewer and better educated children. Although we do not find a statistically significant impact on fertility, the point estimates are negative, as suggested by demographic theory.

Concerning the impact on labor supply, there is a direct channel explaining our estimates: as individuals live healthier lives and are less likely to suffer from disability, more people are able to work, so that labor force participation and work hours rise. If that was the only mechanism leading to our results, we would expect the effect on labor supply to be similar across genders, since Chagas disease affects men and women equally. Nevertheless, we estimate significantly higher impact on female labor supply in both intensive and extensive margins. The difference between effects on female and male labor force participation might be at least partially explained by changes in the incentives for women to work due to extended life expectancy and improved overall health conditions. As reduced adult mortality raises the return to human capital investment and reduces fertility, women reallocate time from raising children and household chores towards the labor market (Falcao and Soares, 2008).

Our findings are in line with the quasi-experimental evidence on this field. Jayachandran and Lleras-Muney (2009) analyze a sudden drop in maternal mortality in Sri Lanka between 1946 and 1953, and assess that for every extra year of life expectancy, literacy and years of schooling increase by 0.7 p.p. and 0.11 years, with an implied elasticity of education with respect to life expectancy of 0.6 and 1.0, respectively. Oster et al. (2013) explore variation in life expectancy within a population at risk for Huntington disease (an inherited neurological disorder) in order to estimate the effects of reduced life expectancy on educational attainment. Their estimates suggest that individuals with reduced life expectancy complete less education and an elasticity of demand for college completion with respect to life expectancy ranging from 0.78 to 1.31. Meltzer and Stoler (2013) also explore Huntington's disease to investigate the hypothesis that high risk of mortality during young adulthood decreases investment in human capital. The authors estimate a strong negative and statistically significant relation between educational attainment and expected longevity.

Our results indicate a positive causal relationship between healthy life expectancy and human capital investment also in the case of Brazil. Considering an average MCA in the top quartile of the baseline Chagas disease prevalence distribution, complete eradication would represent a reduction in Chagas disease prevalence rate from 22% to zero. Considering a life expectancy of infected individuals of 50 years (Taylor and Bestetti, 2009) and an overall life expectancy of uninfected individuals of 75 years, the average longevity gain from eradicating Chagas disease in the average endemic municipality would be of 5.5 years (or 7.9%). Furthermore, according to our estimates, a 22 p.p. reduction in Chagas disease prevalence rate would increase years of schooling by 0.2 years (or 10.2%) and literacy by 4.8 p.p. (or 8.2%), whereas

male and female labor force participation would increase by 1.2 p.p. (or 1.7%) and 5.1 p.p. (or 34.2%), respectively. These results suggest an elasticity of years of schooling and literacy with respect to longevity of 1.3 and 1.0, respectively. Similarly, we estimate male and female labor force participation elasticities of 0.2 and 4.3, respectively.

Given that our back of the envelope elasticity calculations involve estimates that reflect health improvements in addition to longevity increases, we might expect an implied elasticity of education with respect to life expectancy larger than the one found in Jayachandran and Lleras-Muney (2009) and more similar to the one from Oster et al. (2013), where morbidity was also considered. Despite the simplicity of our elasticity computation, these estimates are useful to provide comparable quantitative interpretations of our findings, as they enable us to put our results in perspective to other findings in the literature. Particularly, the similarity of our education elasticities to those found by Jayachandran and Lleras-Muney (2009) and Oster et al. (2013) is encouraging and reinforces the role of health in human capital investment decisions.

Therefore, we provide evidence that health conditions indeed impact major economic decisions. As life expectancy and overall health conditions affect individuals' incentives to invest in human capital, to supply labor and to have children, health-related policies may have long run consequences that can be as important as their direct short-run effect on well-being.

Appendix

Table A1 – Chagas Disease and Human Capital: Additional Sensitivity Tests

Dependent variables:	Years of education		Literacy		School attendance		
Age groups:	10-20	20-30	10-20	20-30	7-14	15-18	18-25
Panel A: Effects up to 2010 (1970-2010)							
Full controls	0.0037***	0.0033**	0.0014***	0.00043	0.0016***	0.0015***	0.00060**
	(0.0011)	(0.0017)	(0.00041)	(0.00029)	(0.00043)	(0.00042)	(0.00028)
Panel B: Effects up to 2000 (1970-2000)							
Full controls	0.0023**	0.000062	0.0013***	0.00028	0.0017***	0.0018***	0.00058**
	(0.0011)	(0.0017)	(0.00036)	(0.00025)	(0.00042)	(0.00044)	(0.00027)
Panel C: Effects up to 1991 (1970-1991)							
Full controls	0.0020*	-0.0016	0.00095***	-0.000010	0.0015***	0.0015***	0.00025
	(0.0010)	(0.0016)	(0.00026)	(0.00022)	(0.00039)	(0.00048)	(0.00026)

Notes: This table reports estimates of the coefficient on the interaction of baseline prevalence rate and post dummy in equation (2) using OLS. Each cell reports the coefficient from a separate regression. The unit of observation is the individual level. The dependent variables are listed in the column headings, such as the age groups considered. In panels A, we have the original sample (that includes 1970, 1980, 1991, 2000 and 2010 census years). In panel B the sample is restricted to 1970, 1980, 1991 and 2000 census years. In panel C the sample is restricted to 1970, 1980 and 1991 census years. Standard errors (reported in parentheses) are clustered at the MCA level. All regressions include controls for male, rural, age fixed effects (and the interactions of the demographic controls with $Post_t$), state x year and MCA fixed effects, share of households in the MCA with electricity, average household per capita income, share of the labor force employed in agriculture and manufacturing.

*** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

References

Becker, G. S. (1964). Human capital theory. *Columbia, New York*.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *The Journal of Political Economy*, 352-365.

Bleakley, H. (2007). Disease and Development: Evidence from Hookworm Eradication in the American South. *The Quarterly Journal of Economics*, 73-117.

Bleakley, H., & Lange, F. (2009). Chronic Disease Burden and the Interaction of Education, Fertility, and Growth. *The Review of Economics and Statistics*, 91(1), 52-65.

Bleakley, H. (2010). Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure. *American Economic Journal: Applied Economics*, 1-45.

Cervellati, M., & Sunde, U. (2013). Life expectancy, schooling, and lifetime labor supply: theory and evidence revisited. *Econometrica*, 81(5), 2055-2086.

Cutler, D., Fung, W., Kremer, M., Singhal, M., & Vogl, T. (2010). Early-life malaria exposure and adult outcomes: Evidence from malaria eradication in India. *American Economic Journal: Applied Economics*, 72-94.

Espinosa, R., Carrasco, H. A., Belandria, F., Fuenmayor, A. M., Molina, C., Gonzalez, R., & Martinez, O. (1985). Life expectancy analysis in patients with Chagas' disease: prognosis after one decade (1973–1983). *International Journal of Cardiology*, 8(1), 45-56.

Jayachandran, S., & Lleras-Muney, A. (2009). Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines. *The Quarterly Journal of Economics*, 124(1), 349-397.

Kalemli-Ozcan, S. (2003). A stochastic model of mortality, fertility, and human capital investment. *Journal of Development Economics*, 70(1), 103-118.

Laranja, F. S., Dias, E., Nobrega, G., & Miranda, A. (1956). Chagas' disease A clinical, epidemiologic, and pathologic study. *Circulation*, 14(6), 1035-1060.

Lee, R. (2003). The Demographic Transition: Three Centuries of Fundamental Change. *The Journal of Economic Perspectives*, 17(4), 167-190.

Lucas, A. M. (2010). Malaria Eradication and Educational Attainment: Evidence from Paraguay and Sri Lanka. *American Economic Journal: Applied Economics*, 46-71.

Oster, E., Shoulson, I., & Dorsey, E. (2013). Limited life expectancy, human capital and health investments. *The American Economic Review*, 103(5), 1977-2002.

PAHO (2014). Fact Sheet: Chagas in the Americas for the General Public. Washington,

United States of America: Pan American Health Organization; 2014. Available at http://www.paho.org/hq/index.php?option=com_topics&view=readall&cid=5831&Itemid=40743&lang=en.

Prata, A., & Macedo, V. (1984). Morbidity of Chagas' heart disease. *Memorias do Instituto Oswaldo Cruz*, 79, 93-96.

Soares, R. R. (2005). Mortality reductions, educational attainment, and fertility choice. *American Economic Review*, 580-601.

Soares, R. R., & Falcao, B. L. (2008). The Demographic Transition and the Sexual Division of Labor. *Journal of Political Economy*, 116(6), 1058-1104.

Stoler, A., & Meltzer, D. (2013). Mortality and morbidity risks and economic behavior. *Health economics*, 22(2), 132-143.

Taylor, J., & Bestetti, R. B. (2009). Chagas' disease and its toll on the heart. *European Heart Journal*, 30(17), 2063-2065.

World Health Organization. The World Health Report 1998 - Life in the 21st Century: A Vision for All. Geneva, Switzerland: World Health Organization; 1998. Available at http://www.who.int/whr/1998/en/whr98_en.pdf.