The Impact of Health on Labor Market Outcomes: Experimental Evidence from MRFIT*

Melvin Stephens Jr.
University of Michigan and NBER

Desmond Toohey
University of Delaware

This Version: November 10, 2016

Draft Prepared for NBER Labor Studies Meeting, Fall 2016

*Stephens: Department of Economics, University of Michigan, 611 Tappan St., Ann Arbor, MI, 48109, e-mail: mstep@umich.edu and National Bureau of Economic Research, Cambridge, MA. Toohey: Department of Economics, University of Delaware, 420 Purnell Hall, Newark, DE, 19716, e-mail: dtoohey@udel.edu. We thank seminar participants at the Michigan Labor Lunch for helpful comments and suggestions. Research reported in this publication was supported by the National Institute On Aging of the National Institutes of Health under Award Number R03AG046503. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The Impact of Health on Labor Market Outcomes:

Experimental Evidence from MRFIT

Abstract

While economists have long posited that investments in health lead to higher earnings, isolating the causal effect of health on earnings has been challenging both due to reverse causality and unobserved heterogeneity. In this paper we examine the labor market effects of the Multiple Risk Factor Intervention Trial (MRFIT), a randomized controlled trial in which a bundle of treatments was provided to determine their joint impact on coronary heart disease mortality. Nearly 13,000 US men were enrolled in the trial and monitored for more than six years. We find that the MRFIT intervention led to higher earnings and family income. We also find, consistent with the Grossman model prediction that health investments increase earnings through reductions in days missed due to illness, a decrease in major physical injuries and work-limiting disabilities.

1 Introduction

Economists have long recognized a strong connection between health and economic outcomes. Measures of health, both self-reported and objectively measured, are positively associated with human capital, earnings, income, and wealth. The direction of causality in these relationships is unclear. Better health can lead to higher productivity, less working time lost to illness, and lower mortality, which further incentivizes human capital investment. Higher productivity and financial resources can facilitate access to care, avoidance of harmful environmental factors, and access to higher-quality food and drugs.

In this paper, we exploit health improvements from the Multiple Risk Factor Intervention Trial (MRFIT), a randomized trial studying the combined effects of multiple health interventions aimed at reducing coronary heart disease (CHD) mortality risk, to estimate the effect of health on earnings and income. MRFIT began screening men aged 35 to 57 in 1974 with high risk for CHD mortality. After three rounds these screenings, nearly 13,000 study participants were randomized into two treatment groups. A "special intervention" (SI) group received interventions aimed at lowering cholesterol, lowering blood pressure, and quitting smoking, while a "usual care" (UC) group was instructed to continue seeking standard medical care in the community. The intervention generated meaningful relative decreases in CHD risk for the SI group relative to the UC group when measured one year after enrollment and were sustained during the six (or more) years that each participant was monitored. These changes are driven by decreases in serum cholesterol, blood pressure, and smoking rates, but the impacts of the experiment can be seen in body weight and other biomarkers. We find that the improved health of the SI group yielded higher individual earnings and total family income measured six years after enrollment.

A key contribution of this paper is that it estimates the impact of health on earnings and family income while avoiding concerns about reverse causality and unobserved heterogeneity. Recent randomized controlled trials have found similar relationships in developing economies, but we believe our work to be unique in its examination of a developed economy. The distinction between these contexts matters beyond the economic environment: while many of the interventions in developing economy settings are aimed at increasing nutrient or consumption levels, MRFIT is largely aimed

at improving health through reducing overconsumption. In this way, our paper provides evidence of the effects of developed countries' health improvements on earnings.

We also examine the mechanisms through which the health improvements lead higher earnings outcomes. A number of possible avenues could drive the effect of improved health on labor market outcomes: changes to productivity per unit of time, decreased discrimination, re-optimized (non-health) human capital, less productive time lost to illness and morbidity. Using MRFIT survey responses on major life events and changes to physical appearance and activity, we find that the SI group experiences less time lost to illness and disability. This evidence is consistent with predictions of the seminal framework of Grossman (1972), which highlights the role of good health in producing a flow of healthy time available for market and non-market work. While we cannot rule out the role of other mechanisms, particularly changes in productivity per unit of time and re-optimized human capital, we argue these effects are likely to be small in our context.

The concept of health capital and its impact on economic outcomes has been of theoretical interest since at least the pioneering work of Becker (1962). The difficulties of identifying causal effects of health on economic outcomes are well-established in the literature and remain a challenge. Thomas (2009) provides a recent review of this literature, highlighting common themes and methods. A large portion of research in this area is given over to examining the effects of resources and nutrients before birth or at young ages. Using variation such as the 1918 flu pandemic (Almond, 2006), the World War II Dutch famine (Roseboom et al., 2001), the availability of prenatal iodine supplementation (Field, Robles, and Torero, 2009), or prenatal healthcare in the form of a midwife (Frankenberg, Suriastini, and Thomas, 2005; Frankenberg et al., 2009), the existing research demonstrates long-term effects of in utero health on human capital and economic outcomes. A related literature examines childhood nutrition or treatment against infection, with some measured effects persisting into adulthood.[1]

The existing research on changes to adult health is somewhat thinner. Some studies exploit the effects of short-run environmental factors (often, pollution) on productivity and labor supply, in

[1]For the effects of childhood nutrition, see Alderman, Hoddinott, and Kinsey (2006), Glewwe and Miguel (2008), Pollitt et al. (1995), Martorell et al. (2005), Maluccio et al. (2009), and Hoddinott et al. (2008). For discussion of treatment of infectious diseases, particularly intestinal helminth infections (worms), see Dickson et al. (2000) and Miguel and Kremer (2004).

some cases focusing on intrahousehold effects (Graff Zivin and Niedell 2012; Adhvaryu et al. 2016; Aragón, Miranda, and Oliva, 2016). Others leverage the impacts of medical interventions aimed at preventing or treating diseases (Thirumurthy, Graff Zivin, and Goldstein, 2008; Fink and Masiye, 2015; Dillon et al., 2015). Compelling evidence of one aspect of health on productivity comes from experimental evidence of the effects of iron supplmentation, particularly on anemic workers (Thomas et al., 2006).[2] While this work provides important evidence in developing contexts, it falls short of demonstrating causal effects of health on earnings in developed countries that face different health challenges. Health is naturally a multidimensional characteristic, and the effects of variation in one dimension informs relatively less about effects along another dimension.

Adult health interventions in developed countries are relatively rarer than those in developing economies. Some examples of plausibly causal studies examine the impacts of variation in the costs of either health insurance or healthcare itself (Newhouse et al., 1993; Dow et al., 1997; Gruber and Hanratty, 1995). However, this variation tends to induce adjustments not just in health itself but in the consumption of health services and the self-perception of health. Evidence on the effects of body composition, as measured by body-mass index (BMI), is summarized by Cawley and Ruhm (2011). In terms of the relevant health conditions, this work is most closely related to ours. However, the evidence is plagued by concerns about unobserved heterogeneity, and the effects appear to vary by demographic group.

This paper proceeds as follows. Section 2 describes the specifics of MRFIT, its sample, and the data. Section 3 details the empirical strategies. After briefly documenting the impact of the intervention on health in section 4, we estimate the effect of health on earnings and family income in section 5. Section 6 explores potential underlying mechanisms driving the results and section 7 concludes.

2 Experimental Design

In this paper, we leverage the experimental impact of MRFIT to identify causal effects of health on labor market outcomes. MRFIT was designed to understand the impact of CHD on mortality

[2]Other studies of iron supplementation on productivity include Haas and Brownlie (2001), Li et al. (1994), Edgerton et al. (1979), and Basta et al. (1979).

where CHD mortality risk was experimentally manipulated by a bundle of treatments. The trial was not specifically designed to affect or measure our outcomes of interest. This section describes the background, intervention, and data created by MRFIT. The information presented below draws heavily on the eight papers making up volume 10, issue 4 of *Preventative Medicine* (1981), which all detailed aspects of the implementation of and presented early results from MRFIT.

2.1 Development of MRFIT

Prior to the 1960s, epidemiological research had already shown connections between coronary heart disease and several risk factors. Existing research had shown both that serum cholesterol levels were associated with incidence of myocardial infarction and that cholesterol levels could affected by modifying dietary intake (Zukel et al. 1981). Blood pressure and smoking had also been established as risk factors for CHD. Interest grew throughout the 1960s in a large-scale demonstration of the effects of risk factor modification on CHD and mortality. While single-factor trials were considered by epidemiologists and public health officials, a 1970 task force, organized by the director of the National Heart and Lung Institute to address arteriosclerosis, recommended against trials modifying diet alone and suggested the a multiple risk-factor trial as the way forward.

In response to these recommendations, the National Heart and Lung Institute's Clinical Applications Program undertook planning for a trial addressing the multiple risk factors of serum cholesterol, blood pressure, and smoking. The Framingham Heart Study, an ongoing observational study of factors associated with heart disease, was used to determine the necessary sample sizes to measure the expected effects. Three rounds of enrollment awarded grants to 22 clinics across 16 metropolitan areas that would identify study participants and implement the trial. Study organizers provided detailed information to each of the clinics and ensured that information and techniques were consistent across them.

2.2 Screening and Randomization

The MRFIT clinics screened initially 361,662 men in total, of which 12,866 were randomized into the SI and UC groups (Sherwin et al.,1981). The study targeted men aged 35 to 57 in the upper

end of the CHD risk distribution as determined by the Framingham Study's results. This upper end of the age range was chosen so as to avoid participants moving at retirement. There was not a unified method for sampling possible participants, so clinics were free to enroll participants in different ways, provided that they met the age and CHD risk requirements. Initially, the intention was to exclude men with diastolic blood pressure readings over 110 mm Hg or who were taking antihypertensive medication, but both of these requirements were ultimately relaxed.

Potential participants were enrolled through three separate screening visits separated by three to four weeks. At each screening, measures of CHD risk were taken and the potential respondents were analyzed for the likelihood of responding to intervention. Only those with elevated CHD risk and a willingness to change risk factors were contacted again after each screening. Respondents who expected to move from the area, had previously been hospitalized for more than two weeks due to a heart attack, or who had been described diabetes medication were excluded, as were men with very high levels of serum cholesterol (greater than 350 mg/dl) and diastolic blood pressure (greater than 115 mm Hg). Following the third screening, participants were randomized into the SI and UC groups.

2.3 Intervention

MRFIT was a non-blind randomized trial. The SI group was subject to the intervention, which is described in detail throughout this section, while the UC group was advised to seek their usual avenues of care in the community. Information on their medical conditions and risk factors were disclosed to the UC group and their medical providers. As the MRFIT organizers were aware, the UC group was not a control group *per se*. UC participants were informed of their elevated risk, a number of medical measures, and were followed-up with throughout the study. However, the interventions described below apparently induced differences in the two groups, which were effectively identical at baseline. The SI and UC means for key variables are displayed in Table 1. The means for all of these variables are statistically equivalent between the two groups.

The intervention for the SI group had multiple arms and aspects of implementation that were used throughout the six years of the study. After initial meetings and screenings at baseline, SI group members participated in a series of 10 meetings with MRFIT staff and other SI participants

over the first two or three of months of the study (Benfari, 1981). These sessions were aimed at communicating to the participants the specific risks associated with various risk factors, giving information on changing behaviors, and providing support for doing so. Participants were encouraged to bring their wives or "homemakers", which was intended to be especially helpful in affecting changes in diet and smoking behavior. Following this intensive period of meetings, depending on their progress in modifying risk factors, the participants entered a period of extended intervention or maintenance for each. Extended intervention involved continued efforts at changing behavior, while maintenance provided support for changes already observed.

2.3.1 Cholesterol Intervention

MRFIT organizers hoped to effect decreased serum cholesterol levels for all men in the SI group, but a particular focus was on achieving a 10 percent decrease for participants with baseline levels greater than 220 mg/dl (Caggiula et al., 1981). Based on evidence from prior studies, it was thought that such a reduction could be achieved by recommending a diet with ten percent of calories each from saturated and polyunsaturated fats and a limit on dietary cholesterol of 300 mg per day.[3] The diet also set a goal for total fat intake as 35 percent of calories. Weight loss was also a target for some men in the SI group, which was a key factor in developing appropriate diet interventions. In particular, men who were over 1.2 times "ideal body weight," which is determined as a function of height, were targeted for weight loss.[4]

When MRFIT participants turned out to have healthier diets than those seen in the prior studies, even before intervention, these targets were made more ambitious. Compared to existing evidence, MRFIT participants in both treatment groups consumed fewer calories, a lower percentage of calories from total fat and saturated fat, and a higher percentage of calories from polyunsaturated fat. Diets were particularly better-than-expected among participants with high baseline cholesterol

[3] It should be noted that more recent evidence shows no impact of dietary cholesterol on serum cholesterol, which has resulted in the USDA dropping dietary cholesterol recommendations from its guidelines. See p. 17 of the USDA's "Scientific Report of the 2015 Dietary Guidelines Advisory Committee." Regardless, as shown in the results of this paper, MRFIT did achieve serum cholesterol reductions, presumably due to the fat intake recommendations and changes to participant weight.

[4] "Ideal weight" in this context is 0.9 times the average height-specific weight for men aged 18–34 in the National Health Survey, 1960-1962. This makes ideal weight for a six-foot man approximately 162 pounds, and 1.2 times ideal weight is just over 194 pounds. For most heights, 1.2 times ideal body weight amounts to a BMI in the range of 26 to 28.

levels. Caggiula et al. (1981) hypothesize that in response to screenings showing high cholesterol levels, participating men had already begun to adjust their diets. In 1976, some two to three years into the study, the saturated fat and dietary cholesterol limits were lowered to eight percent of calories and 250 mg per day, respectively. The weight loss targets were extended to include men down to 1.15 times ideal body weight, and a goal of reducing bodyweight by at least 10 pounds was added for most of the men in the SI group.

The specifics of the MRFIT intervention for modifying diet amounted to targeting particular levels of intake for various food groups. A written manual was distributed to SI participants, categorizing food types according to whether were "OK" or should be avoided. SI sessions were aimed at providing this kind of relevant information to participants and their wives or homemakers. In addition to basic information on what foods to target and avoid, MRFIT sessions highlighted shopping skills, label-reading, food demonstrations, and tastings. Over the long term, the intensity of follow-up with each participant was a function of their cholesterol response. Throughout the study, the SI group was tracked for cholesterol levels at least every four months.

2.3.2 Smoking Intervention

Smoking cessation for the 59 percent of smokers was a key design goal of MRFIT. Given the relatively high CHD risk among MRFIT participants due to other factors, smoking was viewed as a particularly high risk to this population (Hughes et al., 1981).[5] MRFIT clinics were staffed with smoking specialists, psychological consultants, and health counselors to aid participants in quitting. As in the dietary aspects of the intervention, participants' wives were also engaged to support smoking cessation. An initial intervention at baseline involved participants receiving a "strong antismoking message from a physician," who provided tailored information based on the participant's health measures. This was followed by a meeting with a smoking specialist, from which the specialist could identify preferred cessation techniques and whether the participant appeared prepared to attempt cessation.

The smoking intervention from this point forward was structured around documentation cre-

[5]However, because the sample participants were selected to meet a CHD risk threshold, there is a negative within-sample correlation between smoking intensity and the other risk factors.

ated specifically for MRFIT participants. The intervention group meetings highlighted the risks of smoking and the benefits of cessation and encouraged participants to examine details of their smoking behaviors. Meetings involved behavioral modification techniques and group discussions. The smoking specialists followed up regularly to offer support to participants who reported they had stopped smoking. Participants received feedback on a number of objective medical measures including serum thiocyanate and expressed carbon dioxide. Over the extended intervention period, those who had not quit or who had relapsed were considered for additional intervention. These interventions were tailored more specifically to each individual and potentially involved more information from physicians, additional types of group meetings, or further cessation therapies.

2.3.3 Blood Pressure Intervention

The blood pressure intervention in MRFIT targeted reduced diastolic blood pressure for hypertensive participants (Cohen et al., 1981). Blood pressure readings were taken throughout the screening process, as part of sample selection, but categorization as hypertensive was initially based on the third screening, when systolic and diastolic blood pressure readings were both taken. Participants were categorized as hypertensive if they had a diastolic blood pressure reading of 90 mm Hg at third screening and again at a confirmation follow-up. If a participant's readings later exceeded these thresholds at a regular visit and a confirmation follow-up, they were categorized as hypertensive at that point. Any participant on antihypertensive medication was considered hypertensive throughout. For participants not initially on medication, the specific target individual was 89 mm Hg or a 10 mm Hg reduction, whichever was lower. Those who were taking such medication were given a target of 80 mm Hg.

Hypertensive participants were treated with a "stepped care" approach. This involved steady increases in level of hypertensive medication, in a way that was standardized across clinics. During a period of close monitoring, participants were put on increasingly potent blood pressure medications if the desired blood pressure reductions were not observed. If blood pressure readings fell consistently below 80 mm Hg or weight loss was achieved, participants' medications were eligible to be stepped down. In addition to the medication, dietary advice on weight reduction and reduced sodium intake were counseled for hypertensive treatment.

2.4 Data and Measurement

Men in the UC group were invited back for annual visits and examinations after randomization. Men in the SI group were invited for these visits as well as interim follow-ups approximately every four months (Sherwin et al., 1981). These visits allowed MRFIT to record new information on medical history, 24-hour dietary recall, leisure activity, smoking history, and other behaviors. The annual examinations of all participants included, in addition to a physical examination, the recording of a number of biomarkers. For tracking changes to health, this paper uses data from all of these annual visits. Specific labor force information was only recorded at baseline and the six-year follow-up. This includes reported earnings from participants' main jobs as well as family income, both of which are reported in brackets. Other labor force information, including layoff, firing, and disability over the prior year, are reported at annual visits.

Participation throughout the experimental period was relatively high. Sherwin et al., (1981) note that a "large majority" of men participated in the group sessions and that those who did not were "usually" willing to participate in individual sessions. Nearly 75% of wives of the SI group participated in at least some of the group sessions. Through the fourth year of the experiment, over 91% were either attending their annual visits or known to be deceased with participation nearly equal for the SI and UC groups.

2.5 Anticipated Intervention Effects

In order to perform power calculations and determine the necessary sample size, MRFIT organizers anticipated the effects of the intervention on the SI and UC groups. They made these predictions based on evidence from prior interventions. The predicted effects are shown in Table 2. The predicted serum cholesterol effects were informed by experimental results from three prior studies.[6] The expected 10 percent reduction in diastolic blood pressure came from the Hypertension Detection and Follow-up Program, which had shown that hypertensive medications could generate reductions of that magnitude. Notably, for both cholesterol and blood pressure levels, reductions for the SI group were only anticipated among those whose baseline levels were high. This is largely

[6]These were the National Diet-Heart Study, the New York Anti-Coronary Club, and the Chicago Coronary Prevention Evaluation Program.

owing to the data on the effects of the anti-smoking intervention were less firm. The predictions were informed, however, by prior studies suggesting that greater percentage reductions among lighter smokers. The MRFIT investigators anticipated that behaviors would change more or less immediately at the beginning of the study, but that adherence to the changes would decay over time. They further assumed that actual effects on risk itself would develop more gradually.

These risk scores are generated using estimates from the Framingham Study as reported in Neaton et al. (1981).[7] Risk is estimated with a logit model for death due to CHD within 6 years as a function of serum cholesterol, diastolic blood pressure, and cigarettes per day.[8] We calculate risk scores at screening and at each followup year.[9]

Kernel-smoothed densities of baseline CHD risk by treatment status are displayed in Figure 1a. In addition to being statistically equivalent at the mean between treatment and control, as shown in Table 1, this figure demonstrate similar densities across the entire distribution. Further, while the mean risk scores are both just above two percent, the distributions exhibit a noticeable positive skew.

We also use the predicted impacts to generate predicted risk scores for all study participants. The means and densities for these predicted CHD risk distributions are shown in Figure 1b. The experiment is predicted to reduce mean CHD risk for both the UC and SI groups, but with a much larger mean impact for the SI group.[10] With a predicted mean CHD risk score of 1.16, this represents a one percentage point or 47 percent reduction in CHD risk. The predicted SI risk score distribution still exhibits a right skew. The expected risk reductions for the UC group are much more modest, at 0.11 percentage points or five percent.

[7]We deviate slightly from Neaton et al (1981) in that we do not adjust self-reported smoking levels for the respondent's measured serum thiocyanate, which is a biomarker for smoking. To the extent that individuals in the SI might be more likely to underreport their smoking behavior (perhaps because they are expected to reduce their smoking levels), we will tend to overstate the CHD risk differences between the SI and UC groups. However, we show below the thiocyanate levels are lower in the SI group than the UC group.

[8]The logit coefficients in this model are 0.0088 on serum cholesterol, 0.0464 on diastolic blood pressure, and 0.0286 on cigarettes per day, with an intercept of 11.0336.

[9]All three variables were not recorded at the same screening visits, the estimates reported here use first-screening serum cholesterol and third-screening blood pressure and cigarette intake.

[10]As shown in Table 2, these average UC group changes were expected to be entirely due to changes in smoking behavior.

3 Empirical Strategy

We measure the impact of the experiment using linear regressions of the form

$$y_{it} = \alpha_t + \beta_t SI_i + \boldsymbol{\gamma_t} \mathbf{X_i} + \varepsilon_{\mathbf{it}}, \tag{1}$$

where y_{it} is an outcome for participant i in year t, SI_i is an indicator equal to 1 if participant i is in the SI group, and $\mathbf{X_i}$ is a vector of controls for participant i measured at baseline. These baseline controls include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator. Because treatment status is randomly assigned, it is uncorrelated with these controls and their inclusion should have no impact on estimated treatment effect, in expectation. The balance demonstrated in Table 2 bears out that this is true in the sample and, in practice, dropping the controls has negligible impact on the treatment effect point estimates. However, since these covariates can reduce the variance of the error term, ε_{it}, we include them in the analysis to improve the precision of our estimates. We cluster standard errors at the level of the 22 clinics involved in the experiment.

As described in Section 2.4, earnings and income in MRFIT are reported in categories. We use specifications of the form of equation (1) to measure the effects of treatment on individual points of the CDFs associated with these categorical measures. We also estimate the impact of a measure of health, specifically CHD mortality risk, on earnings and income using ordered probits due to the categorical nature of the measured outcome. These models assume a latent measure of earnings or income, inc^*_{it}, which is a linear function of the CHD risk, the controls, and an error term,

$$inc^*_{it} = \lambda_t + \delta_t CHDrisk_{it} + \boldsymbol{\theta_t} \mathbf{X_i} + \nu_{\mathbf{it}}, \tag{2}$$

where the error term ν_{it} is assumed to be normally distributed. The latent measure maps to the J

observed categories as in

$$
\begin{aligned}
inc_{it} = 1 &\iff inc^*_{it} \leq \mu_1 \\
inc_{it} = j &\iff \mu_{j-1} < inc^*_{it} \leq \mu_j \ \forall \ j \in \{2, \ldots, J-1\} \\
inc_{it} = J &\iff \mu_{J-1} < inc^*_{it},
\end{aligned}
$$

where the μ parameters are cut points delineating the observed categories. The coefficients and cut points are estimated by maximum likelihood. The estimate of δ_t describes the effect of CHD risk on the latent measure, which itself has no interpretation, but summarizes the effect of treatment and can be used to estimate marginal effects on the probability of appearing in each category. Because CHD risk is likely endogeneously determined with earnings, we instrument CHD risk with MRFIT treatment status. Because of the nonlinearity of this model, approaches analogous to two-stage least squares (2SLS) are inconsistent. Instead, we account for the endogeneity of CHD risk by including a control function in equation (2) where the control function is the residual from an auxiliary regression of CHD risk on treatment assignment and $\mathbf{X_i}$.[11] We estimate standard errors for these models with 1000 repetitions of a block bootstrap at the clinic level.

4 The Impact of MRFIT on Health Outcomes

Before turning to job related outcomes, we first document the impact of the MRFIT intervention on health outcomes that were directly impacted by the three treatments: serum cholesterol reduction, smoking cessation, and lowering blood pressure. The longitudinal impact of MRFIT on serum cholesterol (Caggiula et al 1981), smoking behavior (Hughes et al 1981), and blood pressure (Cohen et al 1981) has been previously shown for the first four years of the experiment. We extend these findings to cover the first six years of the experiment and also present findings for the longitudinal impact on CHD Risk scores.

[11] See, e.g., Rivers and Vuong, 1988; Terza et al., 2008; and Wooldridge, 2015.

4.1 The Impact on CHD Risk Scores

Figure 2 displays the predicted CHD risk scores as well as actual CHD risk scores during the first year of the experimental period. For both the UC and SI groups, average CHD risk reductions exceeded the predicted levels. As described in prior research, this reflects smaller-than-expected initial reductions in serum cholesterol and diastolic blood pressure, but larger-than-expected reductions in smoking.

The actual reductions in CHD risk were large and sustained. After the first year of the intervention, as shown in in Figure 3a, average CHD risk in the SI group is over 40 percent lower than that found in the UC group. Consistent with the prediction that the cholesterol and blood pressure interventions will affect those with the highest initial levels, the long right tail of the SI distribution is particularly diminished. By year six of the experimental period, as displayed in Figure 3b, the average risk scores for both groups had fallen further. Decreases in the UC group, however, outpaced those in the SI group, leading to a smaller effect of SI treatment on CHD mortality risk by year six.[12]

The evolution of the average experimental CHD mortality risk differences is shown in Figure 4. Each point in this figure represents the coefficient on the treatment indicator from a regression based on equation (1), which controls for baseline characteristics, where CHD risk for different years is the outcome. The effects of treatment on risk scores are largest after one year and decay over time. The coefficients are reasonably-tightly estimated and indicate a relative decrease in CHD risk on the order of one-third for the SI group. The treatment differences fade slightly over the subsequent years, but remain precisely estimated and meaningfully different across the groups.

4.2 The Impact on Cholesterol and Dietary Intake

In order to achieve the ten percent reduction in serum cholesterol for men with baseline serum cholesterol above 220 mg/dl (Table 2), the MRFIT intervention developed a food pattern which it delivered to the SI group (Caggiula et al 1981). As discussed above, this food pattern contained a

[12]It should be noted that the sample had MRFIT participants had aged six years beyond the sample used to estimate the coefficients for CHD risk score. The CHD risk scores may no longer reflect actual six-year CHD mortality group for the sample as it ages, however, the measure remains a time-consistent summary of the effects of treatment on key variables.

number of guidelines: maintain saturated fat at ten percent of total calories, maintain polyunsaturated fat at ten percent of total calories, and keep dietary cholesterol at 300 mg/day. In addition, the SI group was encouraged to keep total fat intake at 35% of total calories.

The effect of the intervention on serum cholesterol throughout the sample period is displayed in Figure 5. As shown in Figure 5a, we find that the treatment reduced serum cholesterol by approximately 8 mg/dl with the impact falling slightly in the later years of the experimental period. Figure 5b displays the estimated treatment effects on serum cholesterol split by baseline levels based on the predicted impacts shown in Table 2. Among participants with baseline serum cholesterol levels of 220 mg/dl, the intervention reduced serum cholesterol by 10 mg/dl, which is lower than the 10% reduction predicted for this group. This effect fades slightly during the last two experimental years. The treatment significantly also lowered serum cholesterol levels for those under 220 mg/dl at baseline, although the magnitude of the response is smaller for this group relative to those with higher serum cholesterol levels at baseline.

The food intake changes related to these cholesterol effects are displayed in the four panels of Figure 6. The upper left panel shows that the percent of calories from saturated fat falls significantly in the SI group. From a starting level of 13.7 percent, the year-one treatment effect is estimated to be a reduction in this share of roughly 3.5 percentage points, such that the average level of saturated fat intake is very close to the target of ten percent of calories. As shown in the upper right panel, the share of calories from saturated fat increases by nearly two percentage points. However, since the average of this outcome at baseline of 6.3 percent, it is consistent with the intervention moving the SI group closer to their target of ten percent of calories coming from polyunsaturated fats. The SI group also reports dietary cholesterol (bottom left panel of 6) and total fat intake as a share of overall calories (bottom right panel) past their dietary targets.

Overall, the SI group self-reports a large initial reduction in total daily caloric intake of approximately 300 calories which grows to approximately 350 calories by year three of the experimental period (Figure 7a). Given the baseline average caloric intake of 2,369 calories, these differentials represent intake reductions of 12 to 15 percent. The calorie reductions remain in this range over the entire sample period, suggesting long-term food-intake changes for the SI group relative to the UC group. However, in spite of the large self-reported reduction in calories, the SI group only reports

a small, but significant, decline in weight of two pounds (Figure 7b). Although this last result suggests that the SI group may be overstating the extent to which they adhere to their food pattern, the serum cholesterol results clearly indicate that the SI group is improving their cholesterol levels.

4.3 The Impact on Smoking

The estimated impact of the MRFIT intervention on smoking outcomes is displayed in Figure 8. The effects of the treatment can be seen most pointedly at year one where the probability of smoking falls by nearly 20 percentage points. The difference in smoking quit rates across the UC and SI groups was equal to about one third of the original smoking population. The treatment effect shrinks over subsequent years. In results not shown here, it can be seen that this is due to more-rapidly decreasing smoking rates in the UC group. This general pattern is also seen in the difference between average daily cigarettes consumed by the two groups. On a base of 19 cigarettes per day, relative consumption in the SI group fell by more than 40 percent. This represents a greater percentage decrease than the change in percent of participants who are current smokers, indicating either that those who did not quit reduced the intensity of their consumption or that the quitters were drawn from the high end of the consumption distribution.

The third panel of Figure 8 displays the treatment effects on serum thiocyanate which is a biomarker for smoking levels. Although thiocyanate is present even among those who never smoke, its concentrations are dramatically higher among current smokers, increasing with the intensity of consumption (Hughes et al 1981). The effect of treatment on thiocyanate is negative and immediately apparent in the first year of the experiment. Although the difference shrinks slightly by year two, there is less decay in the effects over time. Thiocyanate levels remain relatively lower for the SI group than the UC group at all measured horizons.

Figures 9a and 9b explore how these effects differ across the distribution of baseline smoking intensity. The splits in baseline smoking intensity match those shown in Table 2 that are used in determining the predicted treatment responses. The relative decrease among the lightest smokers (fewer than 20 cigarettes per day) is over 40 percentage points. Among the heaviest smokers (at least 40 cigarettes per day), the effect is less than one quarter. This general pattern is in line with the predicted smoking effects described in Table 2: smoking decreases are largest among the lightest

smokers. While the treatment effect fades for all three groups, this happens most dramatically for the light-smoking group, as it becomes indistinguishable from the moderate smoking effect in later years. In terms of relative magnitudes across baseline smoking intensity, the thiocyanate results in Figure 9b tell the opposite story. These estimates are relatively less precisely estimated, but the point estimates indicate the smallest thiocyanate decreases for the lightest smokers, despite their relatively high quit rates. The largest initial decreases are seen among the very heavy smokers, with thiocyanate decreases larger in magnitude than 20 μmol/liter. In subsequent years, the moderate and heavy smoking groups are estimated to have the same effects.

4.4 The Impact on Blood Pressure

Basic blood pressure effects are displayed in Figure 10a. Baseline diastolic blood pressure is at 91 mm Hg, indicating that the average participant is in at the low end of stage 1 hypertension. With the intervention, the SI group's diastolic blood pressure falls by more than 3 mm Hg at all years. The effect size peaks at close to -4.5 in year two before decreasing to near -3 by year six. This outcome, in particular, shows heterogeneous effects across the baseline distribution based on the splits given in Table 2. As shown in Figure 10b, the intervention appears to be successful in lowering blood pressure levels for those with the highest starting blood pressure although the decrease does not quite reach the ten percent reduction predicted for this group. Both groups have effects that peak in the middle years of the study and fade slightly in the later years.

A similar overall pattern appears in the systolic blood pressure effects displayed in Figure 10c. The treatment effect peaks in size at year two, with a decrease of approximately 7 mm Hg on a base of 135 mm Hg. Much of these effects is likely driven by variation in the likelihood of taking hypertension medication over time. At baseline, just under one-fifth of the sample is taking hypertension medication, despite the fact that the sample is specifically selected to be relatively hypertensive. The intervention elevates medication rates at all years, peaking in year two as shown in Figure 10d. At that point, the differential effectively doubles the medication rate in the SI group relatively to baseline levels.

5 The Impact of MRFIT on Earnings and Income

5.1 Reduced Form Effects on Earnings

The results reported in the preceding section show improvements in health along a number of dimensions for the SI group relative to the UC group. We next turn to the question of whether there are corresponding differentials in labor market outcomes. Table 3 reports estimated treatment effects for basic labor force outcomes measured at year six. The first outcome is employment at year six. Most of the study participants are still of traditional working ages by year six, which is reflected in the 88.2 percent of the UC group still working at that time. As indicated by the estimated treatment effect, which is statistically and substantively zero, the treatment does not have any impact on employment at this horizon. A small minority (5.8 percent) of UC participants report that they are retired at this time. With this small share of retirees, the treatment effect of 0.003 amounts to a five percent increase in retirement, but this impact is still small and is not statistically different from zero. Similarly, while 60 percent of working UC participants report a change in job title or kind of work over the study period, there is no difference between this group and the SI workers.

As noted above, earnings and income data in MRFIT are recorded as categorical variables. The baseline PDF for the discrete earnings measure is displayed in the top panel of Figure 11 for individuals who are employed at baseline. For all but the highest and lowest categories, the earnings categories are labeled by their midpoints. The modal category at baseline is labeled as \$13,500, indicating annual earnings between \$12,000 and \$15,000. The bottom panel of 11 displays baseline differences between the SI and UC groups in the CDF of the earnings variable. These estimates are generated from linear probability models of the form of equation (1), where the outcome is an indicator for having earnings that at least as high as the given category. The negative point estimates in this figure reflect slightly lower baseline earnings for the SI group.[13] Across most of the distribution, the point estimates are negative. As measured by these points, the baseline earnings distribution of the SI group is shifted leftward slightly relative to that of the UC group.

[13] If $F(x)$ is the discrete CDF for the earnings variable, the reported coefficients give the difference between the UC and SI groups on $1 - F(x)$. Fewer participants exceeding a particular earnings threshold means that the CDF at that point is higher.

This finding stands in contrast to the balance in previously-reported baseline variables. In some cases, the differences are of meaningful magnitudes: with fewer than five percent of respondents in the highest income category, a difference of nearly one percentage point between SI and UC is potentially important. For this reason, we will account for these baseline differences in estimating the impact of the intervention on earnings.

The study recorded earnings at year six using the same income categories as at baseline. Due to nominal wage growth, higher real earnings for the participants, or both, the year-six distribution for all participants is shifted markedly to the right. This can be seen in the left panel of Figure 12, which reports the PDF for participants who are employed at year six. One quarter of participants have earnings that exceed $35,000 and approximately two-thirds have earnings exceeding $22,500. Measurably-larger shares of SI respondents have earnings exceeding $18,000 and $22,500 at year six.

The impact of the intervention on the CDF of year six earnings is shown in the remaining two panels of Figure 12. We find a significant impact of the MRFIT intervention on the CDF at the second and third largest earnings categories in year six (Figure 12a). Given that we find that the SI group has slightly lower earnings at baseline, we also present additional results in which we account for these initial differences. Restricting to participants employed at both baseline and year six, each outcome is an indicator for exceeding an earnings threshold at year six minus an indicator for exceeding the same earnings threshold at baseline. The estimates from this exercise are displayed in Figure 12c. Using this "difference-in-differences" type approach, we find that the MRFIT intervention increased the relative fraction of individuals in the top three earnings categories.

In order to assess the magnitude of these effects, in preliminary analysis we have run OLS regressions in which we assign each individual a numeric earnings value which is the midpoint of their earnings category.[14] Using this specification, we find that average year six years are $434 higher for the SI group. When we regress changes in earnings from baseline to year six on the treatment indicator, we find that average year six earnings are $701 higher for the treatment

[14]For the highest earnings category, earnings exceeding $35,000, which does not have an upper bound, we assign a value of $45,000.

group. Given that earnings for the UC group at year six are roughly $28,600, we estimate that the intervention increased earnings by 1.5-2.5%.[15]

5.2 Reduced Form Effects on Family Income

The baseline and year-six measures of family income are collected using the same numerical ranges as the earnings variables. The baseline PDF and treatment effects on the CDF are displayed in the two panels of Figure 13. Because family income is presumably weakly greater than main-job earnings, it is not surprising to see a greater concentration of respondents in the highest in the highest categories for this measure given the baseline earnings results. The CDF differences at baseline show no meaningful differences between the SI and UC groups. However, much like with main-job earnings, if there is any difference between the groups, it appears that baseline income is potentially lower in the SI group.

At year six, the family income distribution is particularly concentrated in the top two categories, as shown in the left panel of Figure 14, with a plurality of respondents' families appearing in the highest category. The right panel shows measurable differences between the two groups at four points of the CDF. The effects of treatment in moving families across these income groups is even more noticeable than in moving the participants across the distribution of main-job earnings. It may be that these larger measured effects in family income are just the result of the particular points at which the income and earnings distributions are grouped. However, it is also possible that SI treatment induces effects on the participants' earnings beyond their main jobs or that the treatment induces intrahousehold changes. These changes could be the result of intrahousehold reallocation of time and resources, as in Thirumurthy, Graff Zivin, and Goldstein (2008), or they could reflect improved health and earnings of other household members. Given that wives were to be heavily involved in the intervention itself, this latter explanation is particularly reasonable. Unfortunately, we do not have any additional measures that would allow us to explore this possibility.

Figure 14c displays the difference-in-differences effects on the CDF of family income following the approach discussed above for earnings. We find that the intervention raised the family income

[15]These estimates of the percentage increase rest heavily on the assumption that the value assigned to the top category in order to approximate average earnings and the use of OLS with censored categorical data. We acknowledge concerns with these assumptions and will refine these estimates in future iterations of the paper.

CDF with the impact on the second, third, and fifth largest income categories being statistically significant. In preliminary analyses, we also examine the impact on family income after replacing the categorical value with its numeric midpoint in order to quantify of the impact. We find that the intervention raised family income by $697 using the year six total income alone while we estimate the impact to be $850 when using the difference-in-differences approach. Given that the UC group average total income in year six is $31,700, we estimate that the intervention increased family income by 2.2%-2.7%.[16]

5.3 Causal Effect of CHD Risk on Earnings and Income

To further explore the relationship between health and labor market outcomes, we examine the impact of CHD risk as described in Section 2.5 on main-job earnings and family income. The multidimensional nature of health creates a challenge in precisely measuring the effects of health on labor market outcomes. To the extent that the labor market impacts of SI participation operate through CHD risk, we are able to identify the causal impact of CHD risk on earnings and income. If other impacts of treatment in the SI group (e.g., changes in BMI) have positive effects on labor market productivity, we will overestimate the relationship between CHD and income. The only way to avoid this would be to plausibly account for all possibly health results of the intervention that could have effects on earnings and income. This does not appear to be feasible, so we focus on CHD as a useful summary measure of health improvements in MRFIT.[17]

The relationship between CHD risk and the income measures is displayed in the estimates of Table 4. Part A of Table 4 shows estimates from ordered probit models as in equation (2), without accounting for the endogeneity of health. The estimates in this row indicate a *positive* relationship between CHD risk and both earnings and income. To the extent that CHD risk is a measure of negative health, this is an instance where, observationally, health and economic outcomes are negatively correlated. This result is potentially similar to estimates of the relationship between BMI and earnings that show a positive correlation for some population subgroups. Part B of Table

[16]The concerns with these estimates are the same as those that we raises above for the analogous estimates for the impact on earnings.

[17]For the sake of comparison to existing research, it would be ideal to have measures like self-reported health that would plausibly unify the many dimensions of health. Unfortunately, these were not recorded as part of MRFIT.

4 shows the relationship between status in the SI group and baseline earnings and income. The estimates in both columns are negative. This reflects the same (weak) negative relationship between baseline income measures and treatment status, as is seen in Figures 11 and 13.

The focus of this analysis is displayed in the lower panel of Table 4, which shows the estimated effects on outcomes at year six. Parts C and D repeat the same specifications from parts A and B, but use year-six outcomes as the dependent variables. Without explicitly controlling for endogeneity, CHD risk remains positively associated with earnings and income, though the point estimates are much closer to zero.[18] This is especially true for the family income measure. Part D displays the effects of SI-group treatment at year six on the earnings and income measures, as measured by ordered probits. The estimates are positive and significant, reversing the baseline relationships shown in part B and instead reflecting the positive relationship between treatment and year-six income as seen in Figures 12 and 14.

The last two rows of Table 4 display ordered probit estimates which include a control function to control for selection and estimate the causal impact of CHD risk on earnings and income. The table reports the estimated coefficients on CHD risk and on the control function—the first-stage residuals from estimating CHD risk on treatment and other controls. For both earnings and income, the CHD risk coefficients are negative and significant. The experimentally-induced variation in CHD risk is associated with measurable decreases in earnings and income. This dramatically reverses the endogeneous effects measured in part A of the table. At the same time, the control function, which measures the variation in CHD risk not accounted for by the experiment or the controls, is associated with higher earnings and income. This quantifies the unobserved heterogeneity that is correlated with both CHD risk and the income variables, inducing a positive correlation between them.

[18] At year six, some of the variation in CHD risk is due to the intervention and some of it is due to factors that endogenously determine it with earnings and income. The estimates in part C of Table 4 most likely reflect the effects on earnings and income from each of these sources. Conceptually, this is partially a positive relationship between CHD risk and income, as seen in part A, and partially a negative relationship, as seen in the CHD risk estimates of part E. These same relationships are manifested more directly in the negative and positive estimates of the two rows of part E.

6 Potential Mechanism

The results described in the previous sections show that the MRFIT experiment had impacts on health, earnings, and income of the treatment groups. To this point, the exact mechanisms of MRFIT are a black box. Lacking exact measures of worker productivity, non-health human capital, and wages, we examine participant responses to questions about changes in their health and labor market outcomes over time. Results relating to two of these questions are displayed in Figure 15. In each case, at baseline and over the first five years of follow-up, participants answered retrospective questions about health outcomes over the prior year. As in the previous section, the figures display the estimated coefficients on an indicator for being in the SI group in regressions of the form of equation (1). The top panel displays estimated effects on major physical illness or injury over the prior year, as measured by a week or more of hospitalization or in bed. At baseline, there is no measurable difference between the SI and UC groups in this outcome. Over the subsequent five years, the SI group is roughly ten to sixteen percent less likely to report these long periods of hospitalization. In three of the years, the differences are statistically significant at the five percent level. This evidence indicates decreased rates of illness and injury for the SI group.

The bottom panel of Figure 15 examines whether the respondent reported experience a work-limiting disability. We find that respondents in the SI group experience 15 to 30 percent lower rates of disabilities that prevent work. At years two through four, these differences are significant at the five percent level. However, it should be noted that at there were small differences in the rates of work-preventing disability at baseline.

Both panels of Figure 15 indicate decreased work-limiting health conditions for the SI group. This is most marked for the outcome of hospitalization or bed-restriction over the prior year. While we cannot rule out other avenues through which the intervention would affect earnings and income, these results are consistent with improved health providing a greater flow of productive time for market work. This evidence is consistent with models like that of Grossman (1972), in which health capital manifests itself in labor market outcomes by providing a flow of healthy time.

Much of the development literature on health and economic outcomes explores the role that health plays in changing other measures of human capital. If, for example, the SI group is not more

productive but simply accumulates more human capital in anticipation of lower future mortality, the results presented in this paper would not indicate the effect of health on labor market outcomes. There are two reasons to think that this is unlikely in our context. First, men in the age range of MRFIT are relatively less likely to make human capital adjustments than, say, school-aged individuals. Second, as argued by Bleakley (2010), re-optimization of non-health human capital should produce only second-order effects on labor market outcomes. If a MRFIT participant has already optimized his level of human capital as a function of market forces and health, the envelope theorem would imply that innovations to health should not produce first-order effects on economic outcomes through human capital re-optimization.

7 Conclusion

This paper examines the effects of a randomized health intervention, MRFIT, on labor market outcomes for a population of working-age men in the United States. This intervention succeeded in improving the health of the treatment group along several dimensions. We find that the intervention also yielded significant increases in earnings and family income. In addition, although we find a positive correlation between CHD risk and both earnings and family income, we exploit the experimental variation to show that CHD risk has a negative causal effect on earnings and family income. We believe our findings to be unique in that they demonstrate a causal effect of health on economic outcomes for adults in a developed economy. We further show that a possible mechanism for the effect of health on earnings is an increase in the availability of healthy time for market work. This finding is consistent with the basic prediction of seminal model of Grossman (1972).

References

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham (2016) "Management and Shocks to Worker Productivity" *Unpublished Manuscript.*

Alderman, Harold, John Hoddinott, and Bill Kinsey (2006) "Long Term Consequences of Early Childhood Malnutrition," *Oxford Economic Papers,* 58(3):450-474.

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

Aragón, Fernando M., Juan Jose Miranda, and Paulina Oliva (2016) "Particulate Matter and Labor Supply: The Role of Caregiving and Non-Linearities" *Unpublished Manuscript.*

Basta, Samir S., Soekirman, Darwin Karyadi, and Nevin S. Scrimshaw (1979) "Iron deficiency anemia and the productivity of adult males in Indonesia," *The American Journal of Clinical Nutrition,* 32(4): 916-925.

Becker, Gary S. (1962) "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy,* 70(5): 9-49.

Benfari, Robert C. (1981) "The Multiple Risk Factor Intervention Trial (MRFIT): III. The Model for Intervention," *Preventive Medicine,* 10(4):426-442.

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

Caggiula, Arlene W., George Christakis, Marilyn Farrand, Stephen B. Hulley, Ruth Johnson, Norman L. Lasser, Jeremiah Stamler, Graham Widdowson (1981) "The Multiple Risk Factor Intervention Trial (MRFIT): IV. Intervention on Blood Lipids," *Preventive Medicine,* 10(4):443-475.

Cawley, John, and Christopher J. Ruhm (2011) "The Economics of Risky Health Behaviors," *Handbook of Health Economics,* 2:95-199.

Cohen, Jerome D., Richard H. Grimm, Jr., and W. McFate Smith (1981) "The Multiple Risk Factor Intervention Trial (MRFIT): VI. Intervention on Blood Pressure," *Preventive Medicine,* 10(4):501-518.

Dickson, Rumona, Shally Awasthi, Paula Williamson, Colin Demellweek, and Paul Garner (2000)

"Effects of Treatment for Intestinal Helminth Infection on Growth and Cognitive Performance in Children: Systematic Review of Randomised Trials," *BMJ: British Medical Journal,* 320(7251):1697-1701.

Dillon, Andrew, Jed Friedman, and Pieter Serneels (2014) "Health Information, Treatment, and Worker Productivity: Malaria Testing and Treatment among Nigerian Sugarcane Cutters," Discussion Paper no. 8074, IZA, Bonn, Germany.

Dow, Will, Paul Gertler, Robert F. Schoeni, John Strauss, and Duncan Thomas (1997) "Health Care Prices, Health and Labor Outcomes: Experimental Evidence," Working Paper no. 97-01, RAND, Santa Monica, CA.

Edgerton, V. R., G. W. Gardner, Y. Ohira, K. A. Gunawardena, and B. Senewiratne (1979) "Iron-deficiency anaemia and its effect on worker productivity and activity patterns," *British Medical Journal,* 2(6204):1546-1549.

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

Fink, Günther, and Felix Masiye (2015) "Health and agricultural productivity: Evidence from Zambia," *Journal of Health Economics*, 42:151-164.

Frankenberg, Elizabeth, Alison Buttenheim, Bondan Sikoki, and Wayan Suriastini (2009) "Do Women Increase Their Use of Reproductive Health Care When It Becomes More Available," *Studies in Family Planning*, 40(1):27-38.

Frankenberg, Elizabeth, Wayan Suriastini, and Duncan Thomas (2005) "Can Expanidng Access to Basic Healthcare Improve Children's Health Status? Lessons from Indonesia's 'Midwife in the Village' Programme," *Population Studies*, 59(1):5-19.

Glewwe, Paul, and Edward A. Miguel (2008) "The Impact of Child Health and Nutrition on Education in Less Developed Countries," *Handbook of Development Economics,* 4:3561-13606.

Graff Zivin, Joshua, and Matthew Niedell (2012) "The Impact of Pollution on Worker Productivity" *American Economic Review,* 102(7):3652-3673.

Grossman, Michael (1972) "On the Concept of Health Capital and the Demand for Health" *Journal of Political Economy,* 80(2):223-255.

Gruber, Jonathan, and Maria Hanratty (1995) "The Labor-Market Effects of Introducing Na-

tional Health Insurance: Evidence from Canada" *Journal of Business & Economic Statistics,* 13(2):163-173.

Haas, Jere D., and Thomas Brownlie IV (2001) "Iron Deficiency and Reduced Work Capacity: A Critical Review of the Research to Determine a Causal Relationship " *The Journal of Nutrition,* 131(2):676S-690S.

Hoddinott, John, John A. Maluccio, Jere R. Behrman, Rafael Flores, and Reynaldo Martorell (2008) "Effect of a nutrition intervention during early childhood on economic productivity of Guatemalan adults" *Lancet,* 371:411-416.

Hughes, Glenn H., Norman Hymowitz, Judith K. Ockene, Nathan Simon, and Thomas M. Vogt (1981) "The Multiple Risk Factor Intervention Trial (MRFIT): V. Intervention on Smoking," *Preventive Medicine,* 10(4):476-500.

Li, Ruowei, Xuecun Chen, Huaicheng Yan, Paul Deurenberg, Lars Garby, and Joseph G.A.J. Hautvast (1994) "Functional consequences of iron supplementation in iron-deficient female cotton mill workers in Beijing, China," *The American Journal of Clinical Nutrition,* 59(4): 908-913.

Maluccio, John A., John Hoddinott, Jere R. Behrman, Reynaldo Martorell, Agnes R. Quisumbing, and Aryeh D. Stein (2009) "The Impact of Improving Nutrition during Early Childhood on Education among Guatemalan Adults" *The Economic Journal,* 119(537):734-763.

Martorell, Reynaldo, Jere R. Behrman, Rafael Flores, and Aryeh D. Stein (2005) "Rationale for a follow-up study focusing on economic productivity" *Food and Nutrition Bulletin,* 26(2):S5-S14.

Miguel, Edward, and Michael Kremer (2004) "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities" *Econometrica,* 72(1):159-217.

Neaton, James D., Steven Broste, Louis Cohen, Eugene L. Fishman, Marcus O. Kjelsberg, and James Schoenberger (1981) "The Multiple Risk Factor Intervention Trial (MRFIT): VII. A Comparison of Risk Factor Changes between the Two Study Groups," *Preventive Medicine,* 10(4):519-543.

Newhouse, Joseph P., and the Insurance Experiment Group (1993) *Free for All? Lessons from the RAND Health Insurance Experiment,* Harvard University Press, Cambridge.

Pollitt, Ernesto, Kathleen S. Gorman, Patrice L. Engle, Juan A. Rivera, and Reynaldo Martorell (1995) "Nutrition in Early Life and the Fulfillment of Intellectual Potential," *The Journal of*

Nutrition, 125(4):1111S-1118S.

Rivers, Douglas, and Quang H. Vuong (1988) "Limited Information Estimators and Exogeneity Tests for Simultaeneous Probit Models," *Journal of Econometrics,* 39:347-366.

Roseboom, Tessa J., Jan H.P. van der Meulen, Anita C. J. Ravelli, Clive Osmond, David J.P. Barker, and Otto P. Bleker (2001) "Effects of Prenatal Exposure to the Dutch Famine on Adult Disease in Later Life: An Overview," *Twin Research,* 4(5):293-298.

Sherwin, Roger, Charles T. Kaelber, Paul Kezdi, Marcus O. Kjelsberg, and H. Emerson Thomas, Jr. (1981) "The Multiple Risk Factor Intervention Trial (MRFIT): II. The Development of the Protocol," *Preventive Medicine,* 10(4):402-425.

Terza, Joseph V., Anirban Basu, and Paul J. Rathouz (2008) "Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling," *Journal of Health Economics*, 27:531-543.

Thirumurthy, Harsha, Joshua Graff Zivin, and Markus Goldstein (2008) "The Economic Impact of AIDS Treatment: Labor Suply in Western Kenya," *Journal of Human Resources*, 43(3):511-552.

Thomas, Duncan (2009) "The causal effect of health on social and economic prosperity: Methods and findings" *Unpublished Manuscript.*

Thomas, Duncan, Elizabeth Frankenberg, Jed Friedman, Jean-Pierre Habicht, Mohammed Hakimi, Nicholas Ingwersen, Jaswadi, Nathan Jones, Christopher McKelvey, Gretel Pelto, Bondan Sikoki, Teresa Seeman, James P. Smith, Cecep Sumantri, Wayan Suriastini, and Siswanto Wilopo (2006) "Causal Effect of Health on Labor Market Outcomes: Experimental Evidence," Working Paper no. 070-06, California Center for Population Research, Los Angeles, CA.

Wooldridge, Jeffrey M. (2015) "Control Function Methods in Applied Econometrics," *Journal of Human Resources*, 50(2):420-445.

Zukel, William J., Oglesby Paul, and Harold W. Schnaper (1981) "The Multiple Risk Factor Intervention Trial (MRFIT): I. Historical Perspectives," *Preventive Medicine,* 10(4):387-401.

Table 1: Balance of Baseline Characteristics

	SI	UC	Difference
Age	46.44	46.36	0.09
			(0.13)
White	0.896	0.901	-0.005
			(0.007)
HS Grad	0.212	0.209	0.004
			(0.008)
College Grad	0.268	0.277	-0.009
			(0.007)
Married	0.886	0.888	-0.002
			(0.007)
Serum Cholesterol	254	253	0.30
			(0.65)
Smoker	0.593	0.590	0.003
			(0.008)
Cigs/Day (w/zeroes)	19.2	19.3	-0.014
			(0.33)
Diastolic Blood Pressure	90.7	90.7	0.01
			(0.13)
CHD Mortality Risk (%)	2.15	2.16	-0.017
			(0.026)

Notes: Standard errors clustered at clinic level in parentheses. See the text and Neaton et al. (1981) for the calculation of CHD mortality risk.

Table 2: Predicted Percentage Change in Baseline Outcomes

	SI	UC
Serum Cholesterol:		
≥220 mg/dl	10	0
<220 mg/dl	0	0
Diastolic Blood Pressure:		
≥95 mm HG	10	0
<95 mm HG	0	0
Cigarettes Smoked:		
1-19 Cigarettes/Day	55	15
20-39 Cigarettes/Day	40	10
40+ Cigarettes/Day	25	5

Notes: Sourced from Sherwin et al. (1981, Table 1). Table presents the percentage changes in key CHD risk factors anticipated by MRFIT organizers as a function of baseline levels of the risk factors. See text and Sherwin et al. (1981) for details.

Table 3: Job-Related and Mortality Outcomes

	UC Mean	Treatment Effect
Employed at Year Six	0.882	0.001
		(0.008)
Retired at Year Six	0.058	0.003
		(0.005)
(If Working) Changed Job Title or	0.602	0.001
Kind of Work Past Six Years (Retrospect)		(0.011)
Died by 1982	0.041	0.0006
		(0.0039)
Died by 1985	0.083	-0.006
		(0.004)
Coronary Heart Disease	0.035	-0.0037
Mortality by 1985		(0.0030)

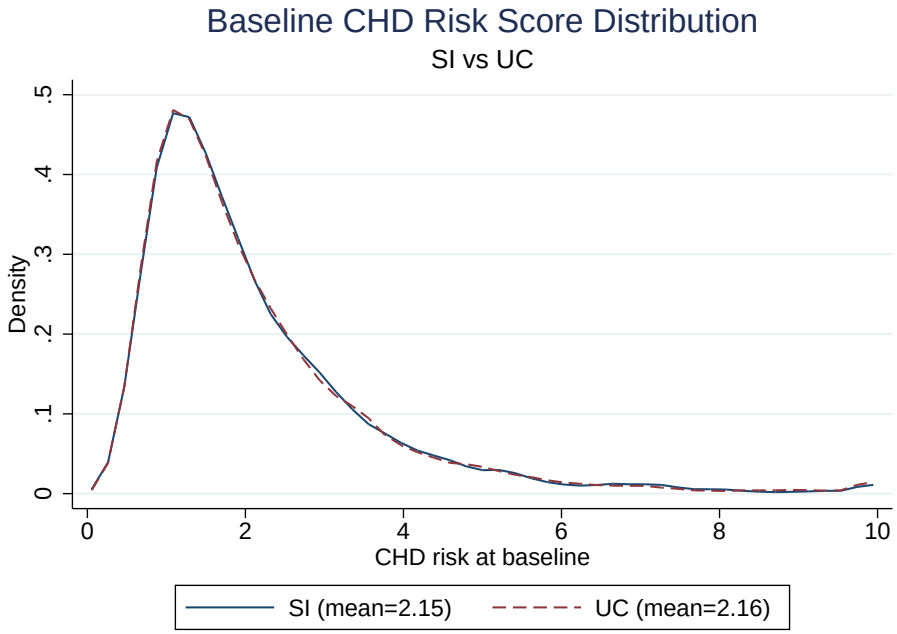
Notes: Standard errors clustered at clinic level in parentheses. Treatment effect estimates are the estimated coefficients on an indicator for being in the SI group from linear regressions of the form of equation(1). The regression controls include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator.

Table 4: Structural Earnings and Family Income Regressions

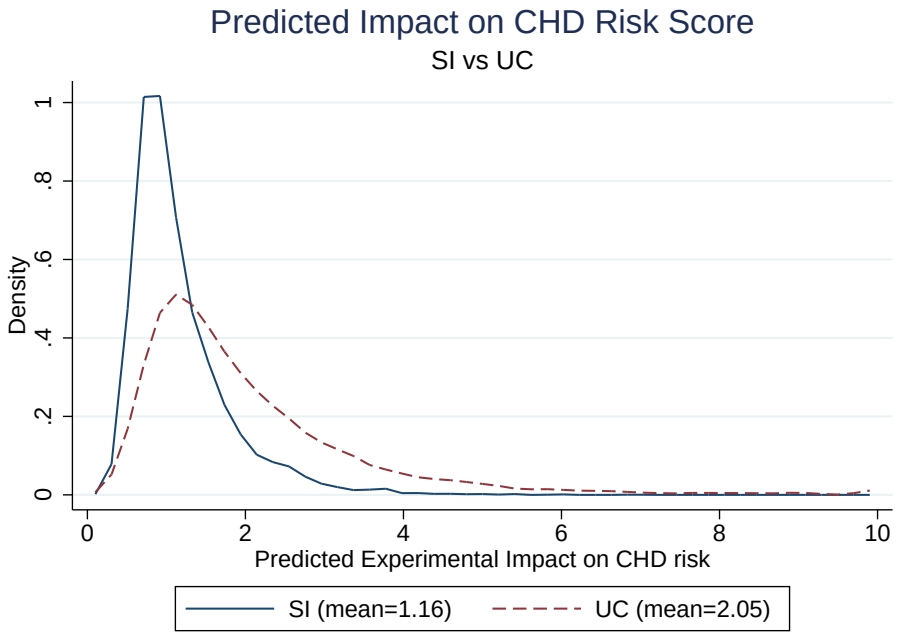
Outcome:	Baseline Earnings	Baseline Family Income
A. Ordered Probit with Endogenous Regressor		
CHD Risk	0.022	0.032
	(0.008)	(0.008)
B. "Reduced Form" Ordered Probit		
Treated	-0.027	-0.025
	(0.015)	(0.017)
Outcome:	Year Six Earnings	Year Six Family Income
C. Ordered Probit with Endogenous Regressor		
CHD Risk	0.010	0.002
	(0.007)	(0.005)
D. "Reduced Form" Ordered Probit		
Treated	0.053	0.081
	(0.019)	(0.025)
E. Ordered Probit Using Control Function		
CHD Risk	-0.128	-0.183
	(0.042)	(0.055)
Control Function	0.140	0.188
	(0.042)	(0.055)

Notes: Standard errors generated by 1000 repetitions of a block bootstrap at the clinic level are presented in parentheses. The table presents estimated coefficients from ordered probit models of the form of equation (2). The regression controls are baseline measures and include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator. The outcomes are nine-group categorical earnings and income measures with cut points at $4200, $7200, $10,000, $12,000, $15,000, $18,000, $22,500, and $35,000. See the text and Neaton et al. (1981) for the calculation of CHD risk. C.F. refers to the control function: residuals from a first-stage linear regression of CHD risk on treatment status and the controls.

Figure 1: Distributions of CHD Risk



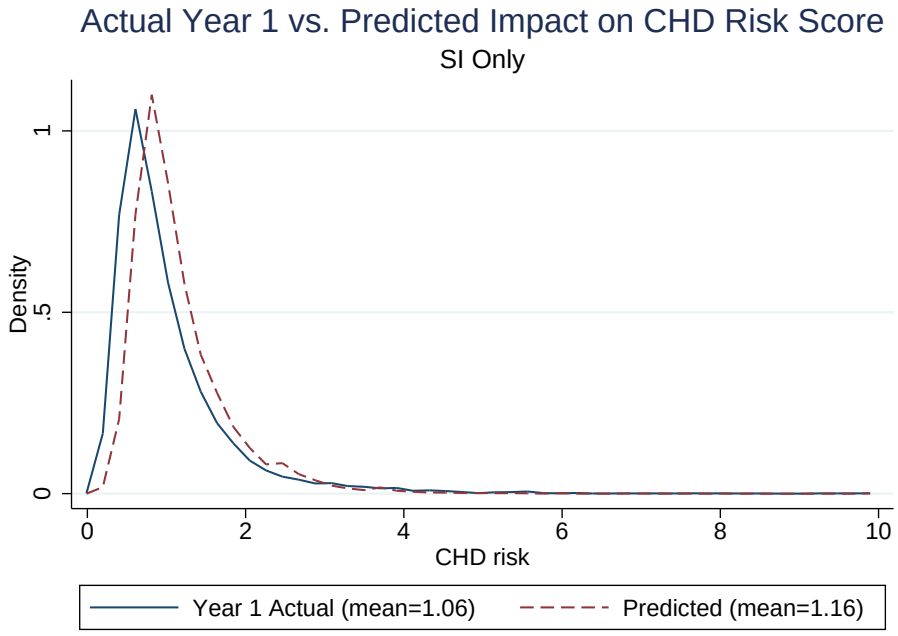
(a) Baseline Framingham CHD Risk



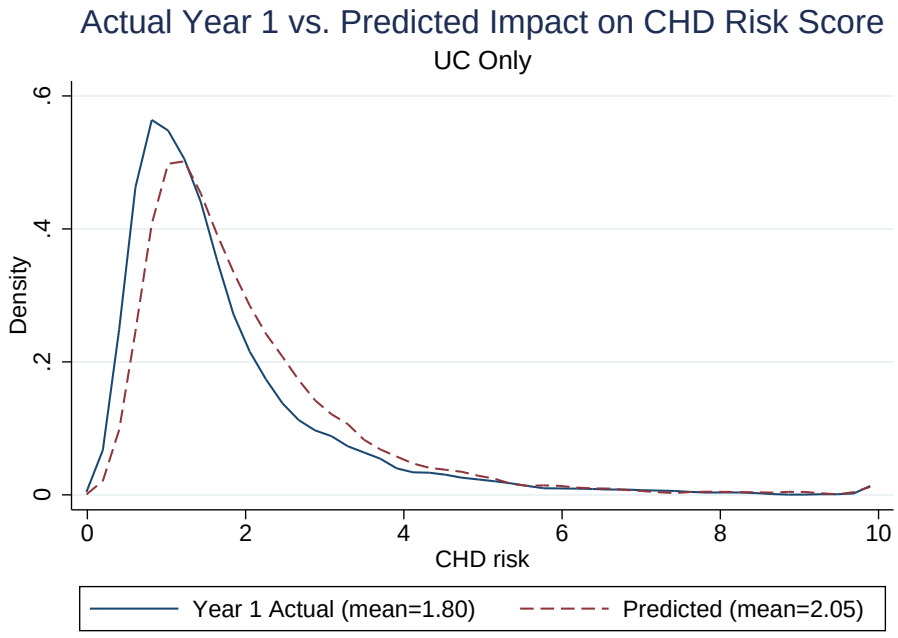
(b) Predicted Impact on CHD Risk

Notes: See the text and Neaton et al. (1981) for the calculation of CHD mortality risk. The displayed risk score levels are based on baseline risk factor levels adjusted by predicted treatment effects as described in Table 2 and Sherwin et al. (1981). The densities are smoothed with an Epanechnikov kernel.

Figure 2: Impact on Distributions of CHD Risk



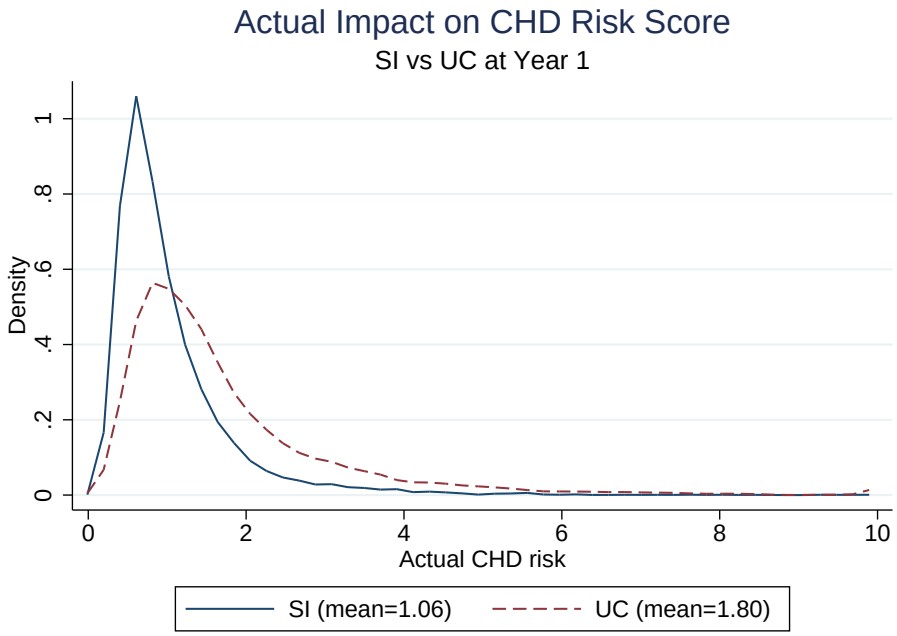
(a) SI group



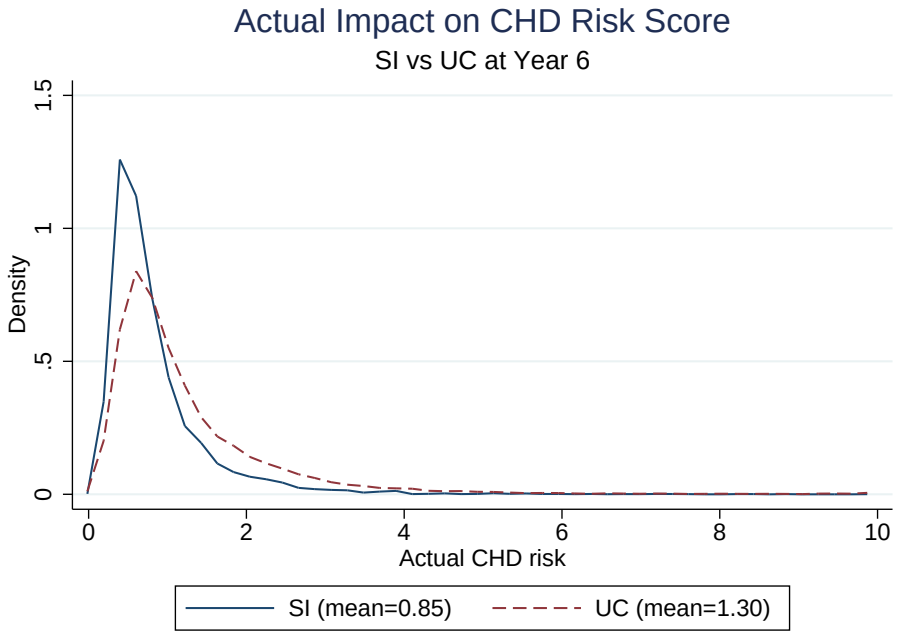
(b) UC group

Notes: See the text and Neaton et al. (1981) for the calculation of CHD mortality risk. Predicted risk score levels are based on baseline risk factor levels adjusted by predicted treatment effects as described in Table 2 and Sherwin et al. (1981). Actual year 1 risk score levels are calculated using observed risk factors at year 1. The densities are smoothed with an Epanechnikov kernel.

Figure 3: Actual CHD Risk - SI vs UC



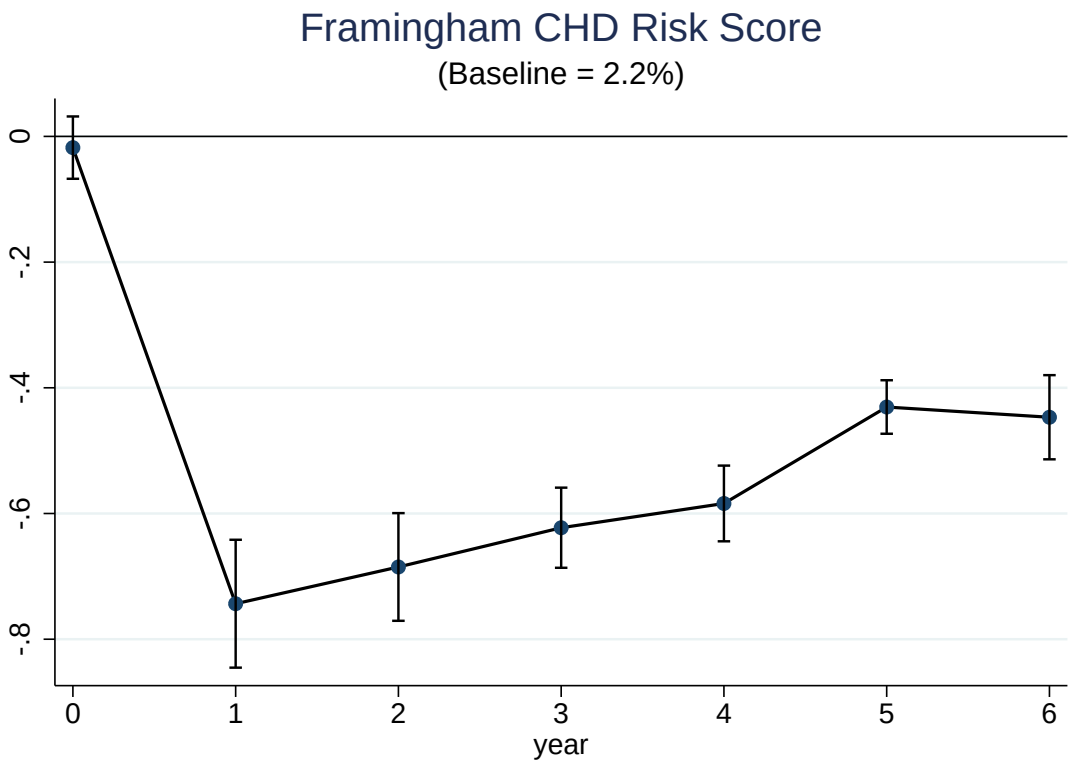
(a) Year One



(b) Year Six

Notes: See the text and Neaton et al. (1981) for the calculation of CHD mortality risk. Actual year 6 risk score levels are calculated using observed risk factors at year 6. The densities are smoothed with an Epanechnikov kernel.

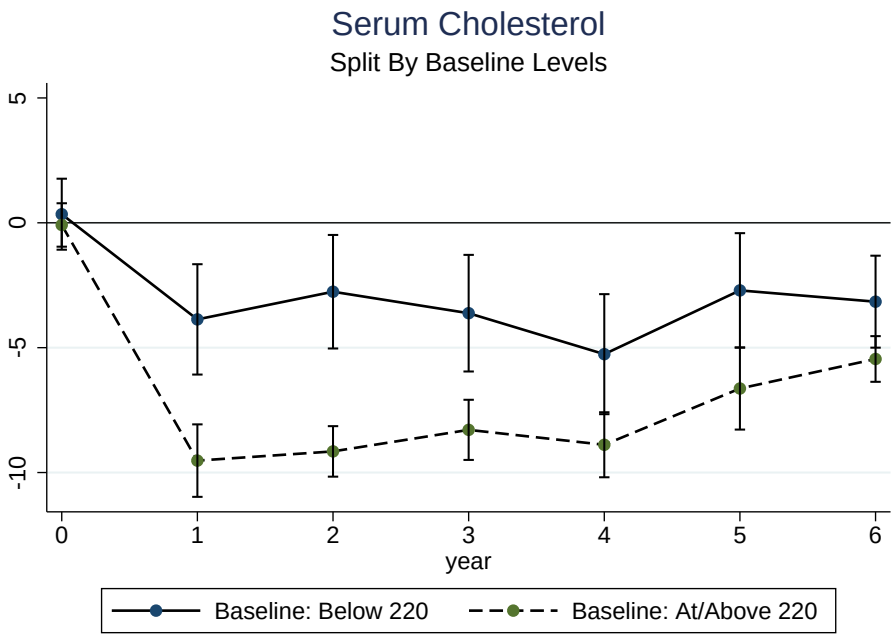
Figure 4: Experimental Impact on CHD Risk



Notes: Standard errors clustered at clinic level in parentheses. Treatment effect estimates are the estimated coefficients on an indicator for being in the SI group from linear regressions of the form of equation(1). The regression controls include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator.

Figure 5: Experimental Impact on Serum Cholesterol

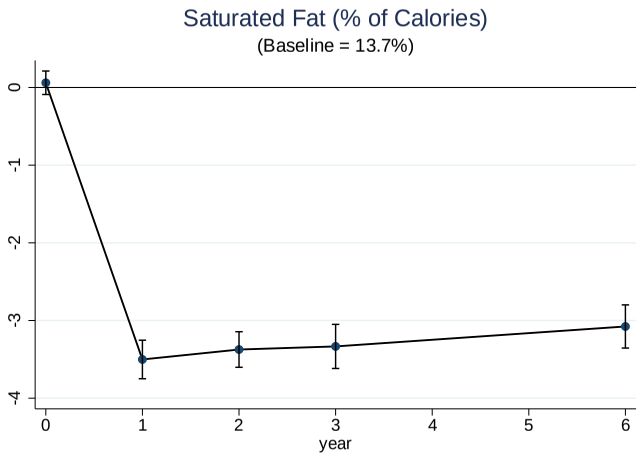
(a) Overall



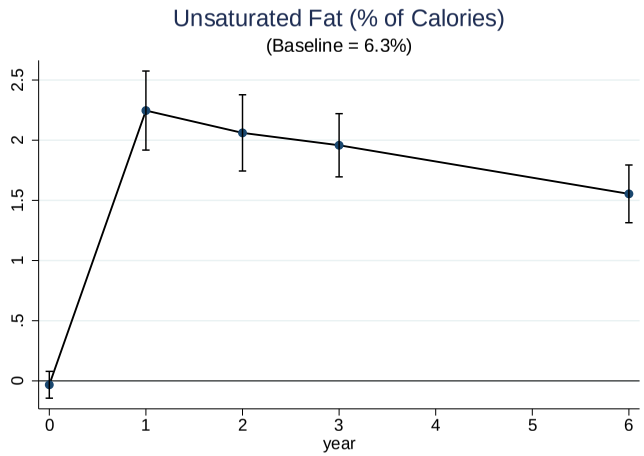
(b) Split by Baseline Levels

Notes: Each point is coefficient from a different regression of the form of equation (1). The displayed 95 percent confidence interval bars cluster the observations at the level of the 22 clinics. The regression controls are baseline measures and include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator.

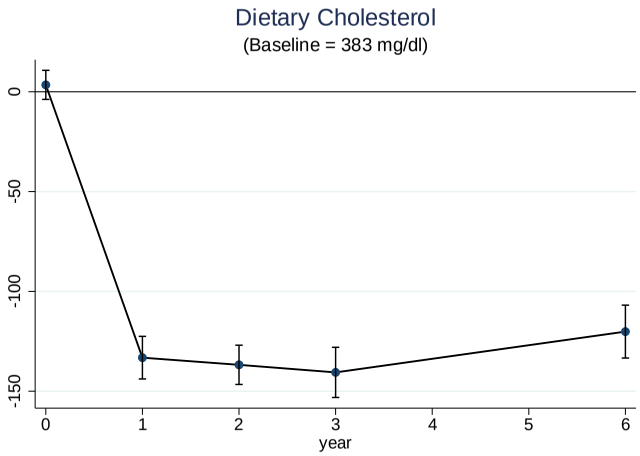
Figure 6: Experimental Impact on Cholesterol-Related Food Intake



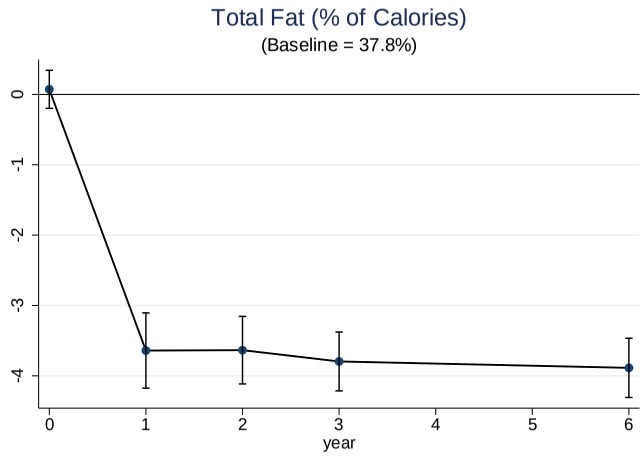
(a) Saturated Fat



(b) Unaturated Fat



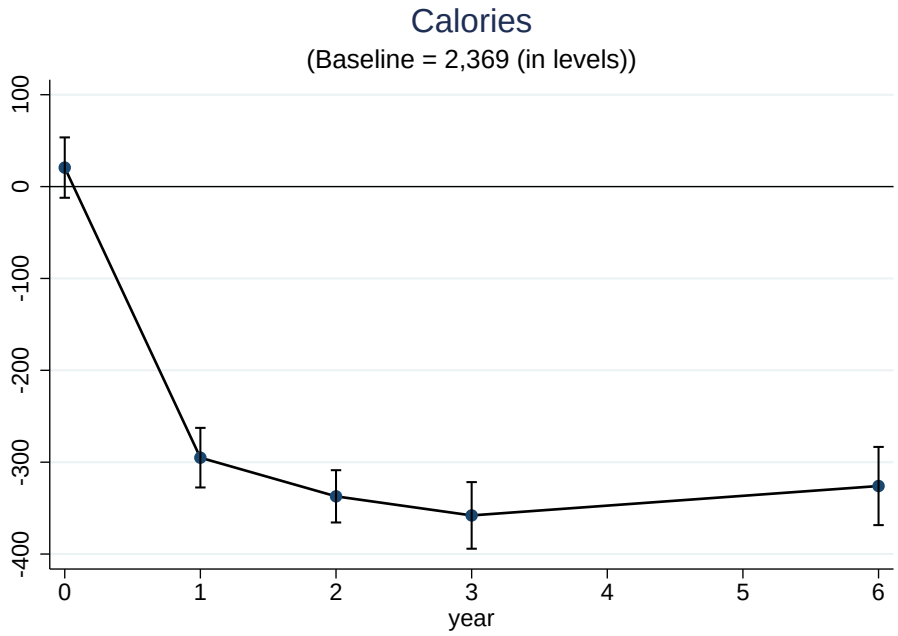
(c) Dietary Cholesterol



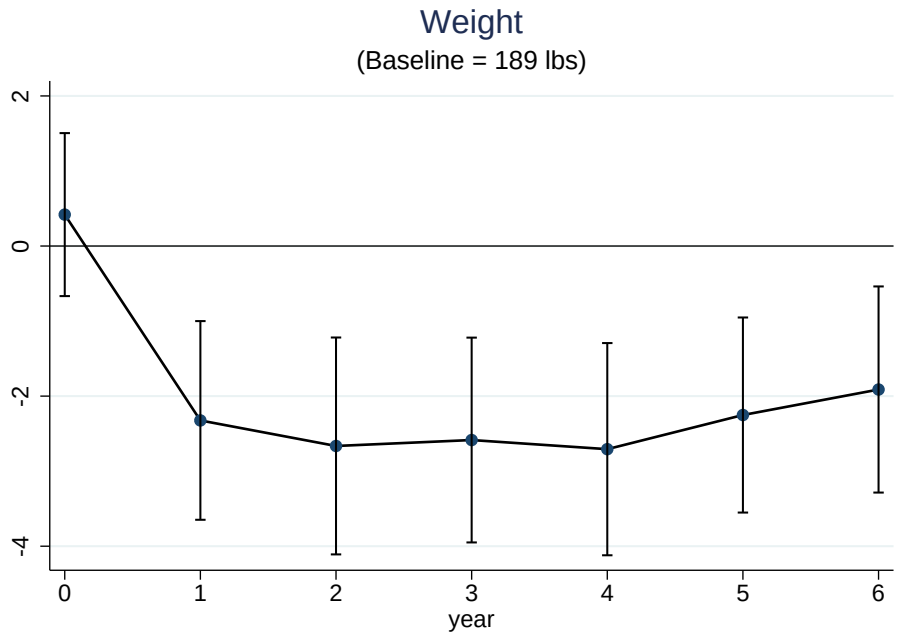
(d) Total Fat

Notes: Each point is coefficient from a different regression of the form of equation (1). The displayed 95 percent confidence interval bars cluster the observations at the level of the 22 clinics. The regression controls are baseline measures and include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator.

Figure 7: Experimental Impact on Calories and Weight



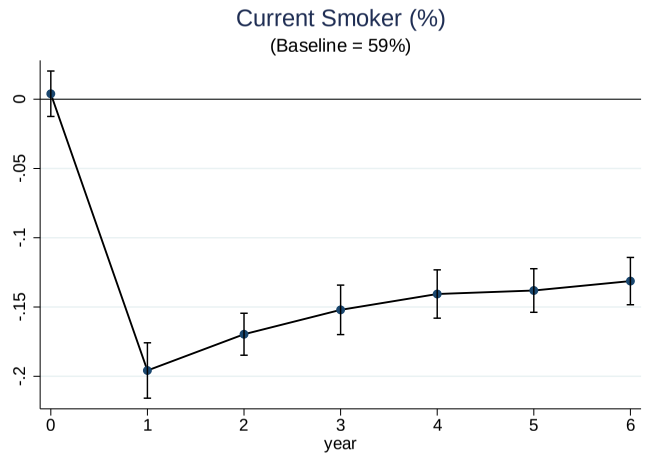
(a) Calories



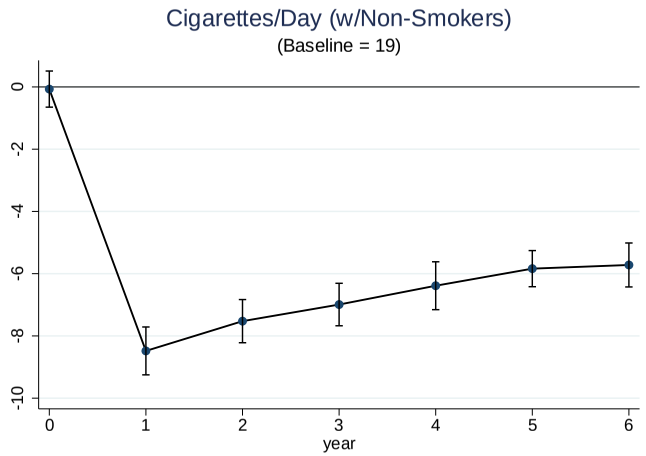
(b) Body Weight

Notes: Each point is coefficient from a different regression of the form of equation (1). The displayed 95 percent confidence interval bars cluster the observations at the level of the 22 clinics. The regression controls are baseline measures and include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator.

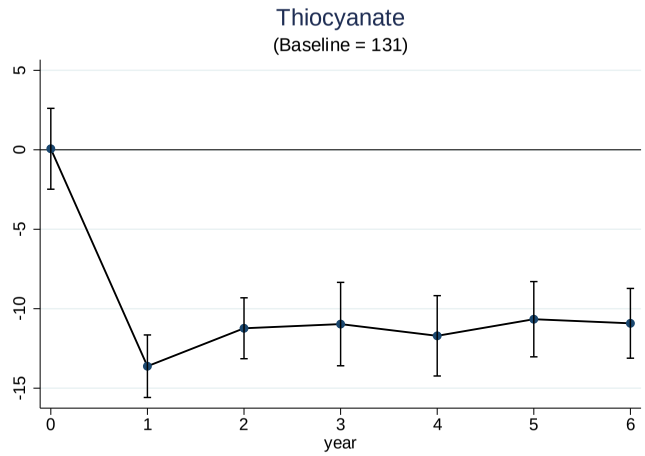
Figure 8: Experimental Impact on Smoking I



(a) Smoker



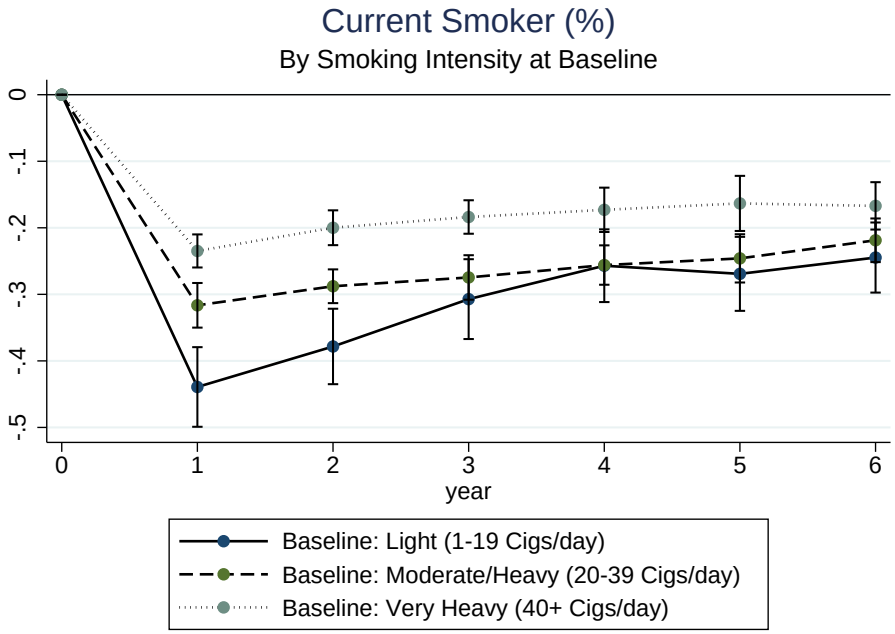
(b) Cigarettes/Day



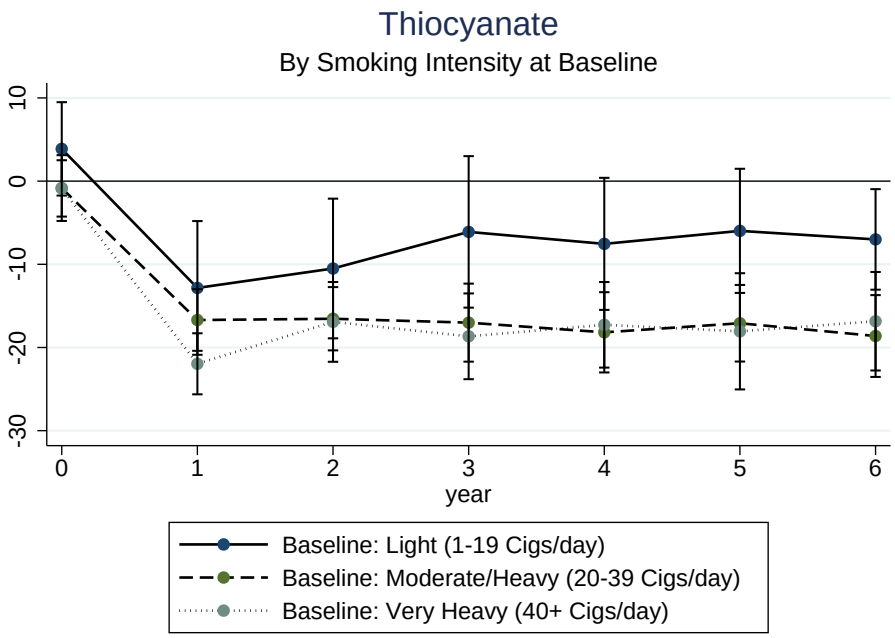
(c) Thiocyanate

Notes: Each point is coefficient from a different regression of the form of equation (1). The displayed 95 percent confidence interval bars cluster the observations at the level of the 22 clinics. The regression controls are baseline measures and include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator.

Figure 9: Experimental Impact on Smoking II



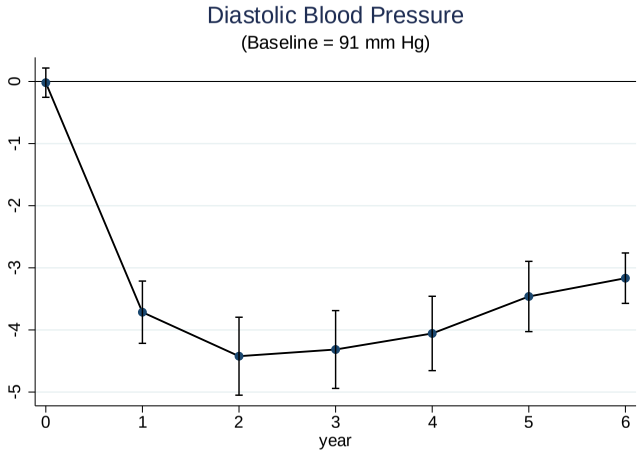
(a) Smoker



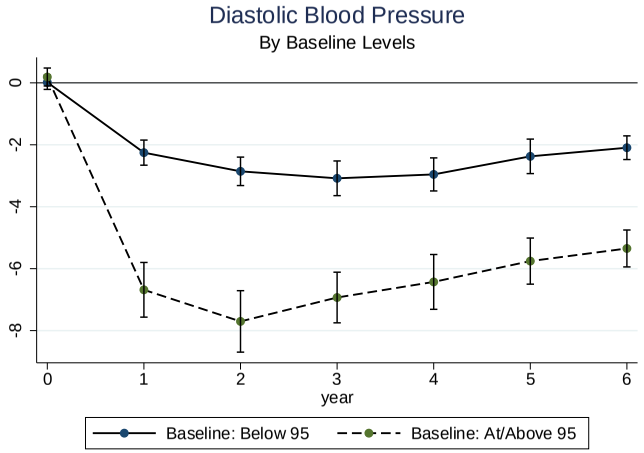
(b) Thiocyanate

Notes: Each point is coefficient from a different regression of the form of equation (1). The displayed 95 percent confidence interval bars cluster the observations at the level of the 22 clinics. The regression controls are baseline measures and include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator.

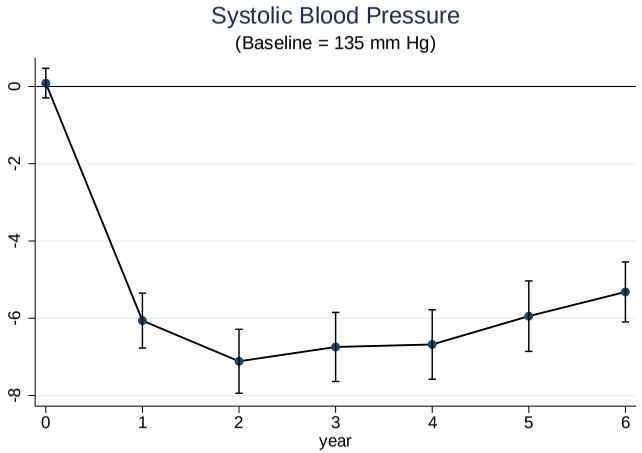
Figure 10: Experimental Impact on Blood Pressure



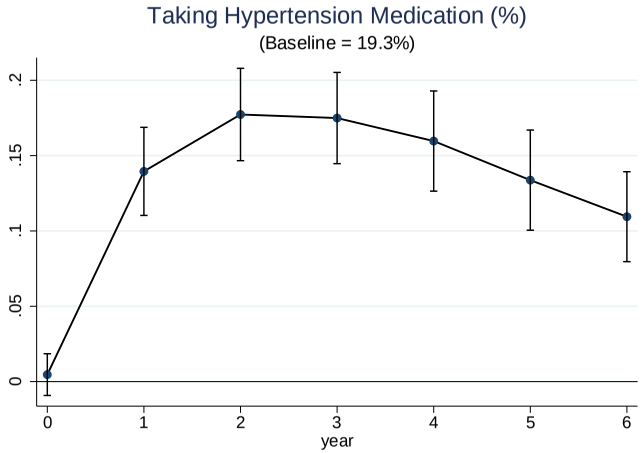
(a) Diastolic Blood Pressure



(b) Diastolic Blood Pressure by Baseline



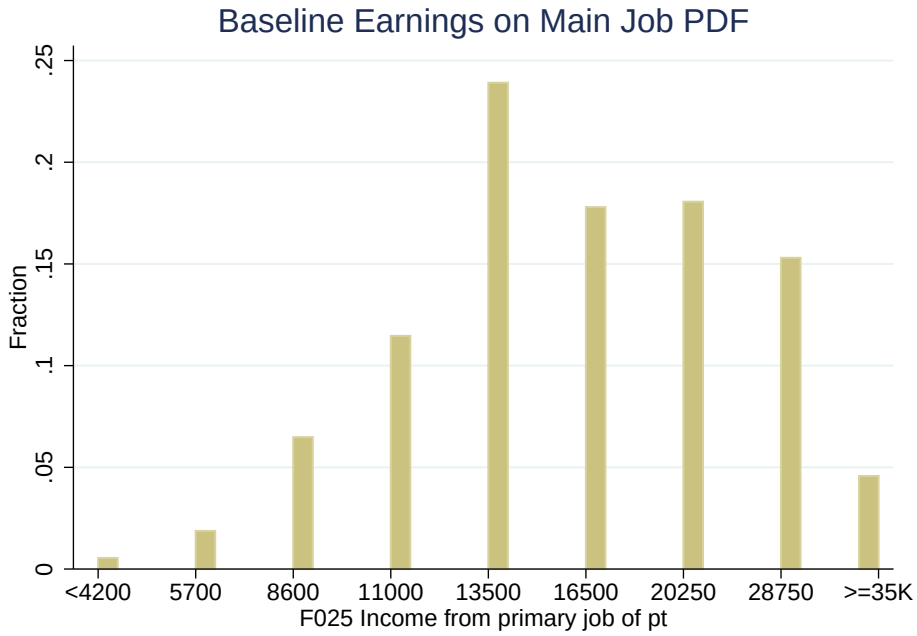
(c) Systolic Blood Pressure



(d) Taking Blood Pressure Medication

Notes: Each point is coefficient from a different regression of the form of equation (1). The displayed 95 percent confidence interval bars cluster the observations at the level of the 22 clinics. The regression controls are baseline measures and include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator.

Figure 11: Baseline Earnings



(a) Earnings PDF

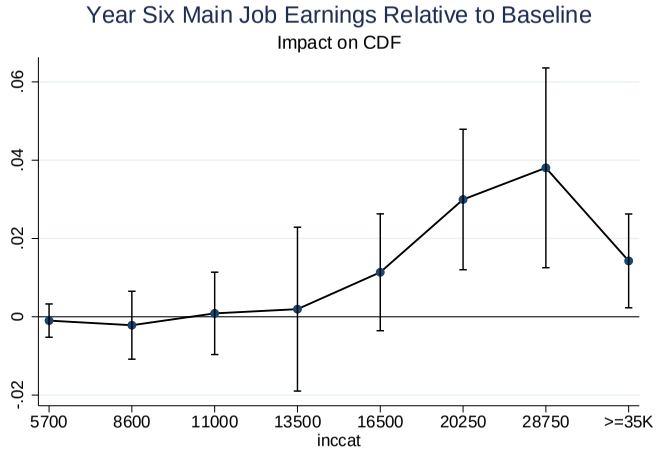
(b) CDF Differences by Experimental Treatment

Note: The top panel displays the PDF of a categorical measure of baseline main-job earnings. The measure has nine categories, with cut points at $4200, $7200, $10,000, $12,000, $15,000, $18,000, $22,500, and $35,000. The interior categories are labeled on the horizontal axis by their midpoints. The bottom panel employs the same labeling to display the estimated coefficients on an indicator for being in the SI group. Each point is coefficient from a different regression of the form of equation (1), where the outcome is a binary variable for having earnings at or above the given grooup. The displayed 95 percent confidence interval bars cluster the observations at the level of the 22 clinics. The regression controls are baseline measures and include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator.

Figure 12: Year Six Earnings

(a) Earnings PDF

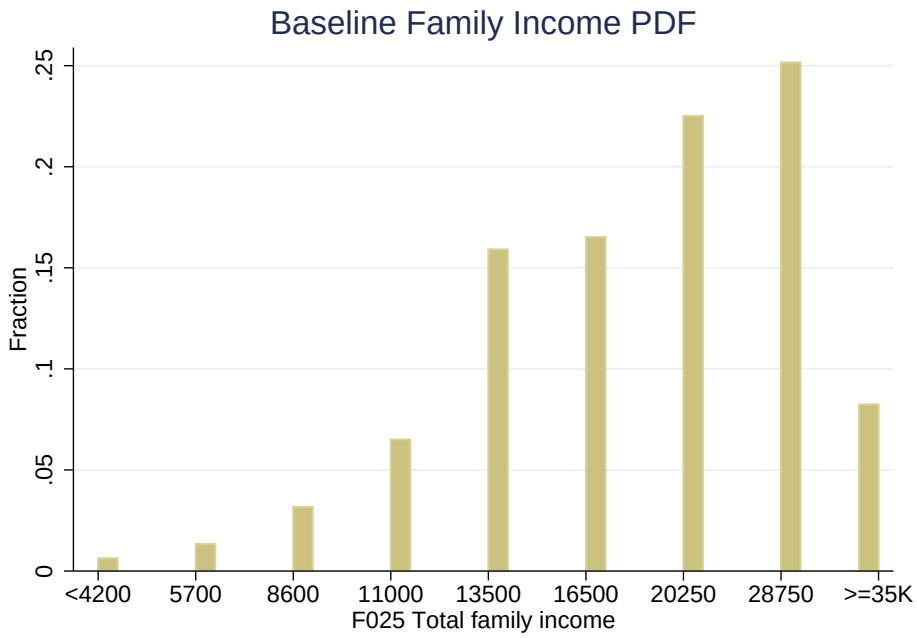
(b) CDF Differences by Experimental Treatment



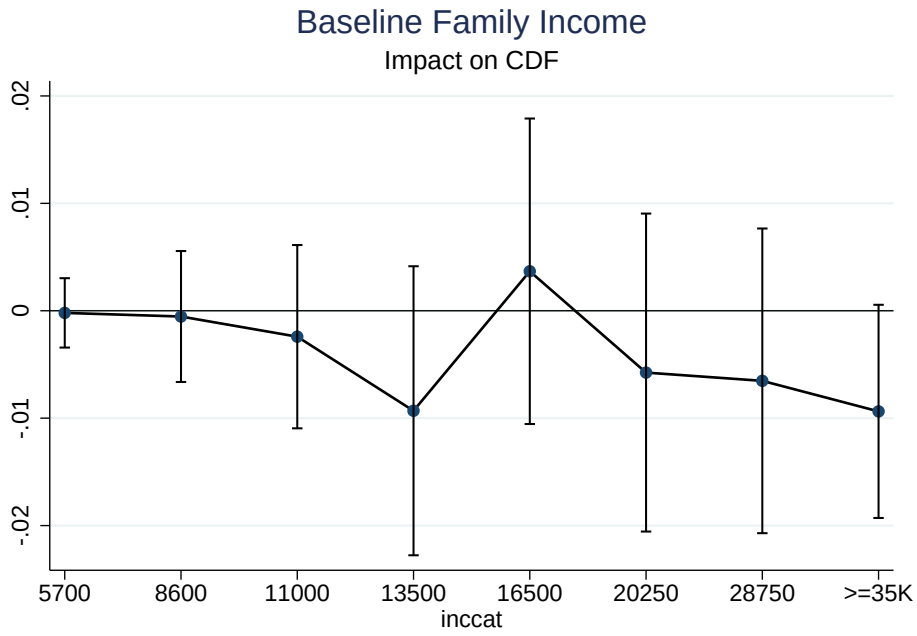
(c) Differences in Differences Impact on CDF

Note: The top left panel displays the PDF of a categorical measure of year-six main-job earnings. The measure has nine categories, with cut points at $4200, $7200, $10,000, $12,000, $15,000, $18,000, $22,500, and $35,000. The interior categories are labeled on the horizontal axis by their midpoints. The top right panel employs the same labeling to display the estimated coefficients on an indicator for being in the SI group. Each point is coefficient from a different regression of the form of equation (1), where the outcome is a binary variable for having earnings at or above the given grooup. The bottom panel shows is a first-differenced measure of having main-job earnings at least as high as the given group. It is an indicator for having earnings at least as high as the given group at year six minus the same measure at baseline. The displayed 95 percent confidence interval bars cluster the observations at the level of the 22 clinics. The regression controls are baseline measures and include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator.

Figure 13: Baseline Family Income



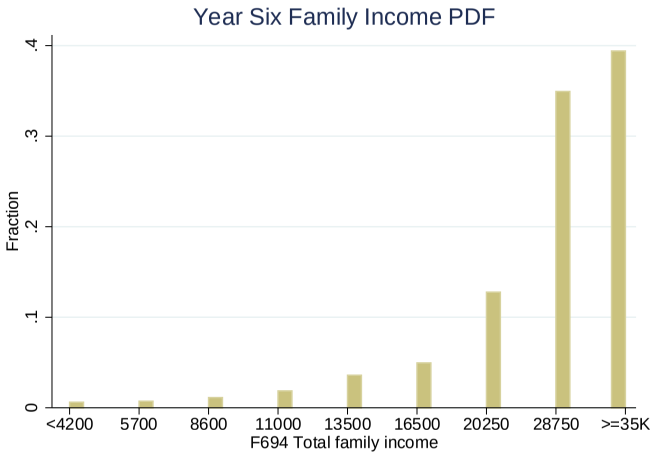
(a) Total Income PDF



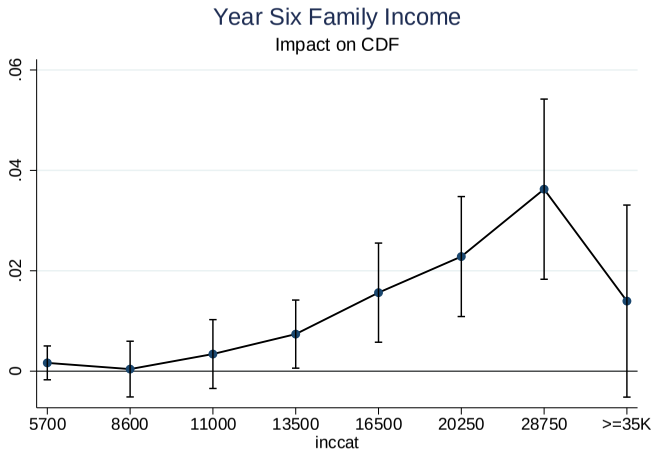
(b) CDF Differences by Experimental Treatment

Notes: The top panel displays the PDF of a categorical measure of baseline family income. The measure has nine categories, with cut points at $4200, $7200, $10,000, $12,000, $15,000, $18,000, $22,500, and $35,000. The interior categories are labeled on the horizontal axis by their midpoints. The bottom panel employs the same labeling to display the estimated coefficients on an indicator for being in the SI group. Each point is coefficient from a different regression of the form of equation (1), where the outcome is a binary variable for having income at or above the given grooup. The displayed 95 percent confidence interval bars cluster the observations at the level of the 22 clinics. The regression controls are baseline measures and include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator.

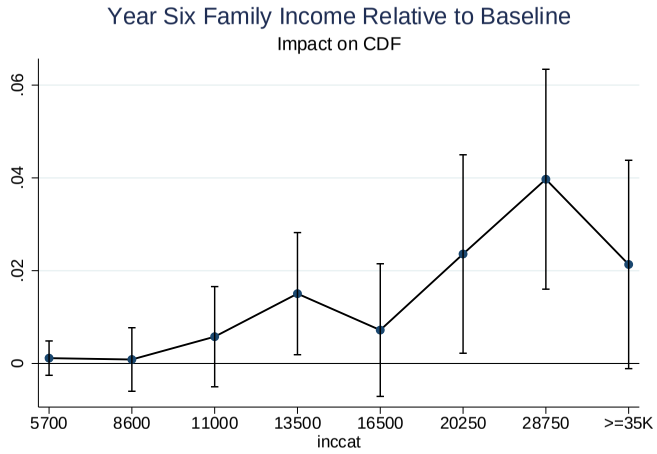
Figure 14: Year Six Family Income



(a) Family Income PDF



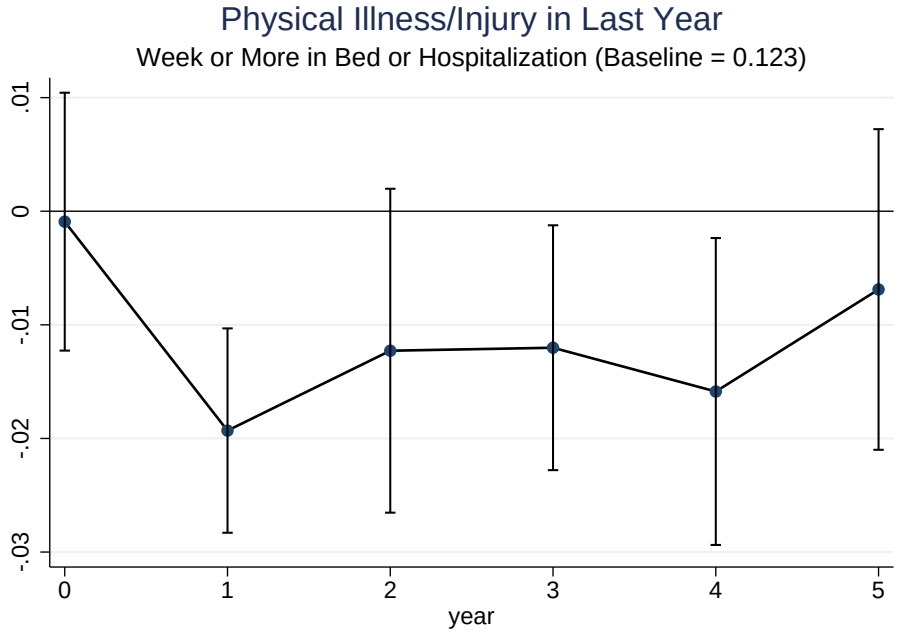
(b) CDF Differences by Experimental Treatment



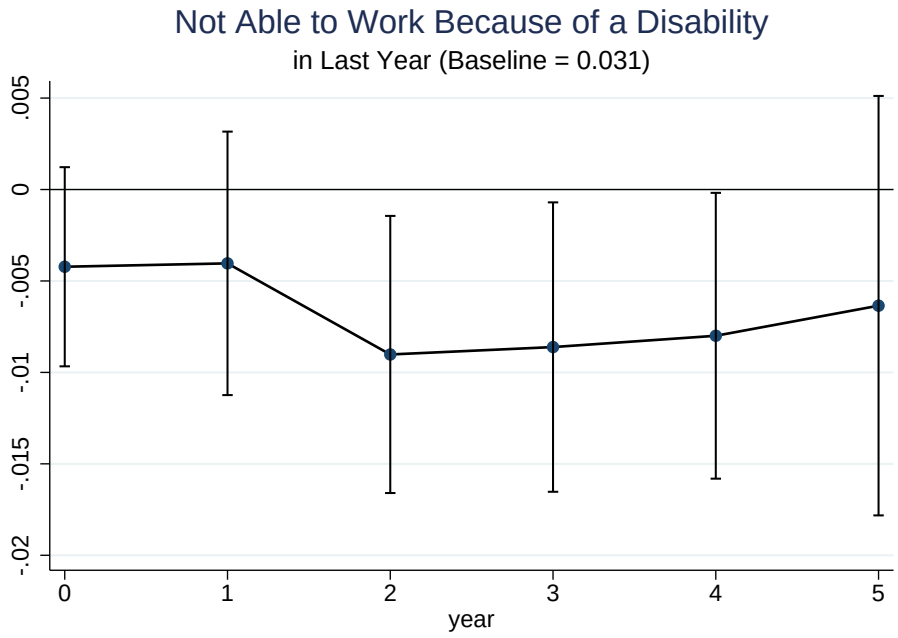
(c) Differences in Differences Impact on CDF

Notes: The left panel displays the PDF of a categorical measure of year-six family income. The measure has nine categories, with cut points at $4200, $7200, $10,000, $12,000, $15,000, $18,000, $22,500, and $35,000. The interior categories are labeled on the horizontal axis by their midpoints. The right panel employs the same labeling to display the estimated coefficients on an indicator for being in the SI group. Each point is coefficient from a different regression of the form of equation (1), where the outcome is a binary variable for having income at or above the given group. The bottom panel shows is a first-differenced measure of having family income at least as high as the given group. It is an indicator for having earnings at least as high as the given group at year six minus the same measure at baseline. The displayed 95 percent confidence interval bars cluster the observations at the level of the 22 clinics. The regression controls are baseline measures and include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator.

Figure 15: Experimental Impact on Health Outcomes in Last Year



(a) Physical Illness



(b) Disabled

Notes: Each point is coefficient from a different regression of the form of equation (1). The displayed 95 percent confidence interval bars cluster the observations at the level of the 22 clinics. The regression controls are baseline measures and include a full set of indicators for age, an indicator for being white, indicators for four education groups, and a marital status indicator.