2015

Essays on natural resources and labor economics

Mohammad Mainul Hoque
Iowa State University

Essays on natural resources and labor economics

by

Mohammad Mainul Hoque

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
Catherine L. Kling, Co-Major Professor
Peter F. Orazem, Co-Major Professor
Joseph Herriges
John Schroeter
Artz Georgeanne

Iowa State University

Ames, Iowa

2015

DEDICATION

To my parents, and all taxpayers in the US and Bangladesh who contributed to finance my education from kindergarten to the graduate school.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to thank my co-major professors, Catherine Kling and Peter Orazem, for their valuable thoughts and insights and the numerous hours they spent guiding me in the right direction by discussing research, demonstrating the difference between good and bad research, and emphasizing research integrity. I would also like to offer a special thank you to professor Orazem for allowing me to pursue research in development economics, an area that is not of relative strength in the Iowa State Economics Department, but is one of great personal interest. Without professor Orazem's dedicated involvement in arranging me access to World Bank's restricted data, I would never have been able to explore such interesting areas in health and economic development. I would like to express my unbounded gratitude to professor Kling for the generous research support she provided during my last year at CARD. Professor Kling introduced me to new potential interdisciplinary research areas concerning the linkage between agriculture and the environment that enriched my exposure and skills. I also appreciate the patience and unconditional assistance with all my limitations professors Kling and Orazem demonstrated throughout my dissertation period. My gratitude extends to professor Joseph Herriges for countless sessions with insightful suggestions that were invaluable to moving this research forward. Thanks to Dr. Artz for advising me on several occasions and providing financial support. I am grateful to my committee members, professors Brent Kreider and John Schroeter, for their valuable comments that helped to improve this research. Fellow graduate students in the department, Donggyu Yi, Yongjie Ji, and Hocheol Jeon were always helpful whenever I wanted to discuss something related to my research. Thanks are due to the Resource and Environmental Policy group in CARD, especially to Dr. Philip Gassman, for providing useful information and feedback while I was working on chapter six in this dissertation. Thanks to Nathan Cook for editorial services on several occasions.

I would like to thank my lovely wife, Auditya Shamsuddin, who accompanied me to the US, sacrificing her career in Bangladesh, and always stayed by my side during all the pains and struggles I experienced during my PhD journey at Iowa State. I do not have enough words to express my thanks to her unsung contributions. Thanks to my parents, siblings, and parents-in-law for the unconditional love, inspiration, and for believing in me and in my passion for graduate studies. Finally, I want to show my appreciation to my son, Ishaar Hoque, for being a source of constant inspiration since he was born in January, 2014, even though I was depriving you of enough daddy time while I focused on my research and completed this dissertation.

ABSTRACT

This dissertation consists of five empirical chapters spanning the areas of natural resource economics and labor economics. After a general introduction in chapter one, the next four chapters deal with how households respond to exogenous changes to economic opportunities such as shocks to employment or to life expectancy at birth. The fifth chapter investigates the linkage between agricultural management and ecosystem services. The dissertation makes extensive use of household survey data, both from the US and from a large number of cross-country surveys. The first two chapters show that unemployment during recessions may lower households' recreation expenditure but increase households' participation in local outdoor recreation activities. The findings from the third and fourth chapters suggest that rising life expectancy at birth increases years in school as well as lifetime earnings, which reinforces the role of health in economic development. The final chapter provides an estimate of the environmental benefits associated with the set of agricultural conservation practices identified in Iowa nutrient reduction Strategy 2013. The economic value from local recreation and aesthetics, drinking water purification, reduced soil erosion, and reduced greenhouse gas emissions are sizable and under some assumptions are of same order of magnitude as the estimated costs.

CHAPTER 1. GENERAL INTRODUCTION

This dissertation consists of five essays. In the second chapter, we investigate how the recession of 2008-2009 affected Iowans' outdoor recreation behavior. The U.S. economy was hit hard by a recession during 2008-2009. During periods of high unemployment, many households suffered income losses which resulted in lower spending on normal goods. However, with changes in employment status, members of some households also experienced a lower opportunity cost of time, and may therefore undertake more household activities that are time intensive. The opposing effects of lower income and cheaper time associated with unemployment motivate the research question of how effects from the recession alter household recreation behavior. To study effects of this type requires detailed household- level data both before and after a recessionary event. In this study, we utilize a panel data set that is uniquely suited to studying the effects of recession on micro decision making in the context of household recreational choices. Specifically, utilizing a panel comprised of both pre-recession and post-recession data on household employment status, usage of recreational sites, and a suite of socioeconomic variables, this paper investigates how changes in employment status during the recession alters lake-based recreation demand. The findings suggest that outdoor recreation in Iowa, in general, is recession-proof.

In the third chapter, we investigate the relationship between household recreation expenditure and job loss during a national cyclical downturn.. We utilize the Panel Study of Income Dynamics (PSID), a longitudinal data set with information on household yearly trip and vacation expenditure, that constitutes better socioeconomic and occupation data to model changes in employment status during a recession. We use a household-level fixed effect model and a difference-in-differences estimator to control for possible selection into unemployment or retirement. Our results suggest that both local economic conditions and individual unemployment during a recession affect recreation expenditure.

The fourth chapter draws empirical evidence from cross-country household surveys on the relationship between life expectancy at birth, human capital accumulation, and lifetime labor market earnings. Life expectancy at birth has improved dramatically over time and across countries during the last century. In the standard Ben-Porath framework, greater life expectancy should increase human capital investment both by extending the period in which an individual devotes full time to training in school and by increasing the fraction of time devoted to training after starting to work. Both types of training should increase lifetime earnings, as would extending the number of productive work years. Motivated by Heckman's (1976) lifecycle analysis, we investigate the causal relationship between life expectancy at birth and years of schooling by exploiting cross-birth-cohort and cross-country variation from a pool of 194 household surveys from 115 countries. We treat the country-cohort life expectancy at birth as the health endowment that parents use to plan out the investments in their children's' education. A gain of 10 years in life expectancy at birth leads individuals to increase their completed schooling by 1.1 years and raises lifetime earnings by 1 percent. To put this in perspective, life expectancy at birth in the U.S. rose 28 years from 1880 to 1980, but birth cohorts and years of schooling rose by about 6.5 years. Our estimates suggest that rising life expectancy in the U.S. explains fifty percent of this increase in schooling.

The fifth chapter tests the robustness of the link between life expectancy at birth and time spent in school across 919 cross-country household surveys. In 95 percent of the surveys, the effect of life expectancy at birth on years of schooling turns out to be positive and statistically significant.

The sixth chapter focuses an economic valuation of ecosystem benefits from nutrient reduction strategies. With the aim of improving water quality, the Iowa Nutrient Reduction Strategy 2013 set a goal of reducing agricultural non-point-source generated nitrogen load by 41 percent and phosphorus load by 29 percent in Iowa's waterways. Various combinations of nutrient reduction technologies are proposed including widespread adoption of conservation practices in farming, land retirement, and wetland restoration can meet the specified target reduction. This study identifies the dollar value of ecosystem benefits resulting from the conservation practices adopted under each of the scenarios. The ecosystem services generated from the nutrient reduction practices include carbon sequestration, increased opportunity for recreation, reduced cost for drinking water purification, aesthetic value of cleaner lakes and streams, reduced soil erosion, enhanced habitat for wildlife, and increased biodiversity. A conservative monetization of these benefits suggests that the benefits of the nutrient reduction practices can exceed the implementation costs.

CHAPTER 2. IS OUTDOOR RECREATION RECESSION PROOF? AN EMPIRICAL INVESTIGATION OF THE 2009 RECESSION

1. Introduction

The US economy was hit hard by a long recession during 2008–2009, which is considered the longest and most severe economic crisis since the end of the Great Depression. The recession affected individual economic well-being through unemployment, stock market crashes, and falling real estate prices, all of which generated low consumer confidence. While much is known about the effect of recessions on macro-level variables, much less is known about how the effects of recession alter household-level consumption behavior. Specifically, during periods of high unemployment, many households will experience lower income, which results in lower spending on normal goods. However, with changes in employment status, members of some households will also experience a lower opportunity cost of time, and may therefore undertake more household activities that are time intensive. To study effects of this type requires detailed household-level data both before and after a recessionary event.

In this paper, we utilize a panel data set that is uniquely suited to studying the effects of recession on micro decision making in the context of household recreational choices. Specifically, utilizing a panel from the "Iowa Lakes Project" comprised of both pre-recession and post-recession data on household income, usage of recreational sites, and a suite of socioeconomic variables, this paper investigates how employment status changes during the recession affects lake-based recreation demand.

Quasi-experimental studies for impact evaluation have become popular in the economics literature. However, causality analysis relating a recessionary shock to recreation behavior has not been undertaken. Using the 2009 great recession as a natural and exogenous event, we fill this gap. To our knowledge this is the first study investigating the effects of a recession on micro decision making in the context of household recreational choices.

Recession can affect an individual's recreation demand through several opposing paths. Even if employment status remains unchanged during a recession, an individual may demand less recreation due to uncertainty and therefore increase precautionary savings. Both income and the opportunity cost of time can be affected for an individual experiencing an employment change. When recession hits the economy, an individual previously employed full time may get fewer paid work hours or be forced into retirement, resulting in a fall in income. However, this change offers more time for leisure and recreation. A change in employment status during recession, therefore, may influence one's outdoor recreation demand through two opposing effects: a substitution effect from cheaper time and an income effect from a fall in income. Further, employment change might lead an individual to revise plans for exotic vacations and trips, which, in turn, might increase demand for cheap local recreation activities.

The Outdoor Foundation's aggregate statistics reveal that, compared to 2008, total participation in outdoor recreation across the United States increased slightly in the recession year 2009. However, nearly 42% of respondents reported that the recession affected their outdoor recreation participation to some extent. At the state level, Iowans' lake visitation rates increased in 2009 relative to 2005 (Iowa Lake Survey Report 2011). Almost 60% of Iowans participated in some form of lake-based recreation activities in 2009, taking around fifteen single-day lake trips on average. In contrast, the consumer-expenditure survey statistics show that expenditures on pleasure and non-business traveling declined during the recession year of 2008–2009 (Bureau of Labor Statistics 2012).

The literature investigating the relationship between recession and recreation demand is limited. Utilizing two intercept surveys conducted in 2006 and 2009 on Quandary Peak, a very popular hiking place in southeast Denver, Loomis and Keske (2012) find no significant impact of recession on total number of visits, travel expenditure, and willingness to pay for visits. However, the respondent groups studied before and after the recession are different. Thus, it is not clear whether the survey respondents experienced any employment or wealth shock during the recessionary period.

The Iowa Lakes Project survey data contains individual recreation demand behavior (participation and number of trips) and employment status both before and during the recession. In this random population survey, a rich set of information on Iowans' lake visitation patterns at 132 lakes was collected, as were demographics including employment status. The survey has been administered five times in total, including once each in 2005 and 2009. The 2005 and 2009 surveys together provide a panel of 2,773 individuals who are observed both before and during the recession. We exploit this panel to investigate how the individuals who move from full-time employment status in 2005 to part-time employment, unemployment, or retirement status in 2009, change their outdoor lake recreation usage, both at the extensive and intensive margins.

In our setting, the treatment group individuals are those who experience an employment shock during the recession year 2009. Assignment to this treatment group is non-random due to both observable and unobservable factors, also known as a selection problem. We use the propensity score matching (PSM) method (Rosenbaum and Rubin 1983) to address the selection problem. Since we have an individual-level panel, we can control for time-invariant unobservable factors utilizing the methodology of Heckman, Ichimura, and Todd (1997) and Smith and Todd (2005).

Following the non-experimental treatment effect literature, we adopt both semi-parametric cross-sectional and difference-in-difference matching strategies to conduct empirical analysis.

In our empirical design, we define the treatment and control groups based on employment status. Individuals who were employed full time both before and during the recession constitute the controls. Individuals who were employed full time before the recession year but had been unemployed, employed part time, or retired during the recession year form the first treatment group. Since the retired individuals might have different recreation preferences compared to those of the unemployed and part-time employed group, we also consider a treatment group that excludes retired people. In the first stage of the PSM analysis, we estimate an individual's propensity to experience an employment shock based on pre-recession information on individual demographics and recreation usage. Using the estimated propensity scores, we conduct the treatment effect analysis using both levels, where we compare recreation behavior of treated and matched controls in 2009, as well as differences (i.e., a difference-in-difference approach) to control for time-invariant unobservable factors. We apply five different matching methods to check consistency of the results. As a robustness check, we conduct a placebo exercise, include a subset of covariates to estimate the propensity score, and conduct exact matching based on location.

The main results from this analysis reveal that employment change during a recession does not affect outdoor recreation. Households who became unemployed either did not change or increase participation in outdoor lake recreation during the recession. On the intensive margin, they visited lakes as frequently as before the recession. However, people going into retirement during the 2009 recession did not exhibit any systematic differences in recreation behavior compared to what they would have done were they employed full time. Our placebo exercise confirms that our findings are not driven by a pre-existing differentiated trend for the treatment and control group. Incorporating county-level unemployment rate as a proxy for aggregate economic condition, we extend the analysis in an individual fixed-effect framework. The results suggest that households residing in counties with high unemployment during a recession participated more in outdoor lake recreation.

The insensitivity of recreation demand to recession implies that there are stable economic benefits from nature-based economic activities. This finding is of direct policy relevance for decisions by public officials concerning nature-based public amenities. Improving water quality and public facilities appears to provide social benefits that are resilient to recessions - the stability of returns to this form of public good provision may raise its value relative to other local public goods.

2. Background and Theoretical Motivation

Two important components that determine recreation behavior are income and the opportunity cost of time [Bockstael and Hanemann 1987; Cesario 1976; Englin and Shonkwiler 1995; Feather and Shaw 1999; Larson and Shaikh 2004; McConnell 1992]. Like any other economic good, income determines an individual's purchasing power of recreation services. If recreation is a normal good, the impact of a rise in income is positive, and vice versa if it is an inferior good. Time spent for recreation services has two components: travel time and time spent on site. Phaneuf and Requate (2013) provides a useful recreation demand model to motivate an individual's optimization between consumption of non-recreation necessities, and recreation goods and services. The individual is naturally endowed with T units of time, out of which she works for H hours in the market for an hourly wage of w, and allocates the remaining time, $T - H$, between recreation(R) and leisure(l) so that her utility from consumption of R, l, and the *numeraire* good(z) is maximized. For simplicity, we are assuming that the hours of work, H, is determined outside the model independent of choices for *R*, *l*, and, *z*. Formally, the individual wants to maximize the utility function *U(z, R, l; q)*, where *q* represents taste parameters, subject to two separate constraints

i) Money income constraint: $wH = cR + z$, where c is the \$ cost of a trip, and

ii) Time resource constraint: $= H + l + t * R$, where time remaining after work hours, $T - H$, is used for leisure and recreation, and t is the time cost for consumption of each unit of *R*.

The individual solves the following 2-constraint, utility maximization problem

$$\max_{z,R,L,\mu,\lambda} U(z, R, l; q) + \lambda(wH - cR - z) + \mu(T - H - l - tR).$$

Manipulation of the first order conditions results in $\frac{U_R}{U_Z} = c + \frac{\mu}{\lambda} t = c + \phi t$. At the optimum, the marginal monetary benefit from one unit of recreation trip $\left(\frac{U_R}{U_Z} = \frac{\delta z}{\delta R}\right)$ must equate with the marginal cost $(c + \phi t)$of the trip. The recreation price consists of an explicit part, c, and an implicit part ϕt. Solving the first order conditions with specific functional form for utility would give us demand equation for each of $z^*(c, t, w, H, T, q)$, $R^*(c, t, w, H, T, q)$, $l^*(c, t, w, H, T, q)$ and, $\phi^*(c, t, w, H, T, q)$.

The model succinctly identifies the possible pathways through which recession might influence recreation demand behavior. If the recession affects the recreationist directly through a reduction in working hours, or job loss, the individual experiences a fall in money income but have more available time for leisure and recreation. Thus, the opportunity cost of time to be spent for recreation (ϕ) decreases. However, an opposing effect takes place through the decrease in working hours and resulting fall in income. These two opposing effects are comparable to substitution and income effect resulting from a price change. Whether the time effect dominates the income effect will determine the overall effect of recession on recreation demand behavior.

Unemployment or, fall in working hours during recession and resulting income loss might lead an individual to demand more cheap local recreation activities. In modeling recreation demand, the choice set often includes an element "stay-at-home" option [e.g., Egan, Herriges and Kling 2009]. This "stay-at-home" option captures everything outside the model including options for other recreation activities such as exotic vacations or an international trip. If a recreationist has plan for such a trip but experiences a fall in income due to a recession, s/he is less likely to make those expensive tours. In such cases, the "stay-at-home" option becomes less appealing, and might induce an increase in demand for local recreation activities. In the model specified above, this is equivalent to saying that corner solution, *i.e.*, $\frac{U_R}{U_Z} < c + \phi t$, is less likely to occur.

Similar to the static model above, in a dynamic setting (Hoque, Kling, and Herriges 2013) we showed that it is difficult to predict the change in recreation behavior in response to a recession. During a recession, everyone is subject to uncertainty, and one may experience unemployment and fall in income. In the latter case, individuals try to smooth consumption of leisure and recreation across periods by reallocating time-resource within periods and monetary resources across and within periods. While in the former case, precautionary motive sets in which, depending on risk attitude, may alter one's demand for recreation in any direction. The combined effect of uncertainty and unemployment may go in any direction, and depends on the relative strength of precautionary motive against the consumption smoothing effects. Individual exposure to recession will vary by type and intensity, and, in accordance, responses to such shocks will vary as well. The implication is that the net effect is ambiguous.

Literature search suggests that Loomis and Keske (2012) is the only study that investigates the impact of an exogenous shock, such as a recession, on recreation behavior, either empirically or theoretically. Their study relies on two intercept surveys conducted in 2006 and 2009 on a single location-Quandary Peak, a popular hiking spot in Southeast Denver, Colorado. They found no significant changes in average number of visits, visitation expenditure, and willingness to pay across periods. Since they did not observe any significant difference in hikers' income between the two periods, it is not clear whether the survey respondents in their study experienced any employment shock during the recession. If not, the individuals did not face the tradeoff between time resources and income in choosing recreation demand. A longitudinal recreation data incorporating pre-recession and post-recession period will help to figure out correctly whether an individual was affected by recession, and how the affected individual alters recreation behavior during a recession.

In contrast to the recreation demand literature, studies in other applied microeconomic fields explored changes in economic behavior during a recession. For example, health economics literature demonstrates a negative association between business cycle and mortality [Ruhm 2000; 2005]. Other examples include studies on the relationship between recession and child health care [Dehejia and Lleras-Muney 2004; Baird, Friedman, and Schady 2011], recession and food-at-home [Dave and Kelly 2010]. Many of these studies focus on economic goods that have both monetary prices and time costs, and thus resemble recreation demand to some extent.[1]

[1] Time intensive activities (for example, child care) exhibits interesting implication in the context of recession. A depressed wage during recession reduces the time cost in taking various caregiving activities such as more preventive health visits, breastfeeding, cooking healthy meals, or improving general cleanliness. However, during such contraction income also falls, which might affect parents' ability to purchase nutritious food or health augmenting inputs. It seems that two opposing effects work simultaneously: substitution effect from cheaper time, and income effect from fall in income.

3. Econometric Framework

Given the multiple influences of recession on recreation usage in theory, the overall effect is purely an empirical question. We study the impact of employment change during recession on lake recreation in a non-experimental setting. In our case, the treatment group includes those who experience a change in employment status facing a recession, and assignment to this treatment group is non-random. This non-random treatment assignment is also known as a selection problem, which can be due to both observables and unobservable factors. The selection problem can hide the true causal effect of a change in employment status during a recession on recreation behavior, and there might be confounding factors that affect both selection into the treatment (experience of a change in employment status) and the outcome variable (trip taking to lake). Propensity Score Matching (PSM) method, due to Rosenbaum Rubin (1983), is one approach to overcome the selection problem. PSM is widely used in the program evaluation literature [Dehejia and Wahba 2002; Ravallion 2005; Ravallion and Jalan 2003; List *et al*. 2003; Ferret and Subervie 2013]. Under certain assumptions, the method solves the problem of missing counterfactual in non-experimental setting. [2]

3.1 Empirical design and strategy

In this study, we investigate empirically how the individual's trip-taking behavior to lake changes in response to a change in employment status during a recession. The relevant periods for the analysis are 2005, the pre-recession year, and 2009, the recession year. As identified in the theoretical model, there are several ways an individual may respond to a recessionary shock.

[2] In the first step of a two-step procedure, the method estimates a propensity score (one's probability of being included in the treatment group) for each individual in the treatment and control groups based on observed covariates, and based on that in the second step it matches the treatment observations with the appropriate control to estimate the impact of treatment.

An individual who used to visit lakes before recession may choose "stay at home" option during the recession depending on how the combination of income effect, substitution effects, and consumption smoothing effect resulting from unemployment, and attitude towards risk and uncertainty works. In contrast, depending on the relative strength of these effects, an individual may switch from "stay at home" option and start visiting lakes during recession.[3] There are confounding factors that may affect both of an individual's chance of experiencing a change in employment status, as well as lake recreation behavior during a recession. For example, facing a depressed wage during the recession, an avid trip-taker might choose voluntary unemployment. To control for such confounding factors, a semi-parametric approach, such as PSM method, seems appealing.

Treatment group: The treatment group consists of the set of individuals whose employment status has changed during the recession. In a recession, a previously fulltime employed individual might become unemployed, part-time employed, or retired.[4] The control group in all cases consists of the recreationists who are full-time employed in both of 2005 and 2009. The three treatment and control groups we study are:

$$T_{1i} = \begin{cases} 1 \ \ if \text{ "i" is fulltime employed in 2005 but Unemployed/Part time/Retired in 2009} \\ 0 \quad\quad if \text{ "i" is fulltime employed in year 2005 and 2009 .} \end{cases}$$

(1)

$$T_{2i} = \begin{cases} 1 \ if \text{ "i" is fulltime employed in 2005 but Unemployed/Part time employed in 2009} \\ 0 \quad\quad if \text{ "i" is fulltime employed in year 2005 and 2009 .} \end{cases}$$

(2)

$$T_{3i} = \begin{cases} 1 \ if \text{ "i" is fulltime employed in 2005 but Retired in 2009} \\ 0 \quad\quad if \text{ "i" is fulltime employed in year 2005 and 2009 .} \end{cases}$$

(3)

---

[3] By similar reasoning, a recreationist who do not alter lake recreation at the extensive margin may respond at the intensive margin by increasing or decreasing number of trips.

[4] Recognizing the possible differences between unemployed and retirees, we form three treatment groups including and excluding the retirees.

Outcome variable: Let the treatment group indicator is $T = \{0,1\}$, and year indicator is $t = \{Y09, Y05\}$. The first outcome variable of interest is a binary variable, $Trip_{T,t}$, indicating whether an individual in group T takes any trip in year t. The second outcome variable is $NTrip_{T,t}$ which denotes the total number of trips for group T in year t.[5]

Propensity of experiencing a change in employment status during the recession: We first estimate one's probability of being unemployed or retired during the 2009 recession. There is no clear set of standards on what variables to include in the propensity score equation. However, program evaluation literature [(Heckman, Ichimura, and Todd 1997; Smith and Todd 2005; and Caliendo and Kopeinig 2008] suggests to incorporate all important and necessary variables that may influence both outcome and treatment variables to reduce bias. Accordingly, economic theory, previous research, and institutional setting can help to characterize the covariates. In our setting, previous research is limited to guide us specify the propensity score equation. Utilizing the treatment status, as defined in equation (1)-(3), and the covariate $\boldsymbol{X}$, we estimate the probability of one's being unemployed or retired using separate logistic regression models for each treatment group *l*:

$$\Pr(T_l = 1|\boldsymbol{X}) = P(\boldsymbol{X}) = \frac{\exp(\boldsymbol{X}'\boldsymbol{\beta})}{1+\exp(\boldsymbol{X}'\boldsymbol{\beta})}. \tag{4}$$

Identification assumption: once we control for propensity score, $P(\boldsymbol{X})$, the treatment and the control groups satisfy the *ignorability* condition, as stated in equation (5) and (6) below.

$$Trip_{1,Y09}, Trip_{0,Y09} \perp T|P(\boldsymbol{X}). \tag{5}$$

$$NTrip_{1,Y09}, NTrip_{0,Y09} \perp T|P(\boldsymbol{X}). \tag{6}$$

[5] $Trip_{1,Y09}$ indicates whether an individual experiencing a change in employment status in 2009 take any trip at all in the year 2009 while $Trip_{0,Y09}$ is the similar indicator for individuals who do not experience any such changes in employment. Similarly, $NTrip_{1,Y09}$ denotes total number of trips for treatment group while $NTrip_{0,Y09}$ denotes total number of trips for the control group in the recession year 2009.

The above states that conditional on the propensity score, exposure to unemployment or retirement during the recession is independent of contemporaneous recreation outcome. One implication of the *ignorability* condition is the mean equivalence condition which states that once the propensity score is controlled for, the treatment and the control group have similar distribution for the covariate vector X: $T \perp X|P(X)$. In other words,

$E[X|P(X), T = 1] = E[X\ |P(X), T = 0]$.[6]

Since we are interested in estimating the impact of a change in employment status during a recession on outdoor recreation, we need a counterfactual estimate on what the recreationists in the treatment group would have done were they not affected by an employment shock during the recession. The weak *ignorability* assumptions below, a weaker assumption compared to that stated in equation 5 and 6, imply that conditional on the propensity of being in the treatment group, there is no difference in recreation behavior between the treatment and control absent the treatment occurs. Accordingly, recreation behavior of households in the control group in 2009 will be the counterfactual recreation for households in the treatment group, both at the intensive and extensive margin.

$$E[Trip_{0,Y09}|P(\boldsymbol{X}), T = 1] = E[Trip_{0,Y09}\ |P(\boldsymbol{X}), T = 0] = E[Trip_{0,Y09}|P(\boldsymbol{X})]. \qquad (7)$$

$$E[NTrip_{0,Y09}|P(\boldsymbol{X}), T = 1] = E[NTrip_{0,Y09}\ |P(\boldsymbol{X}), T = 0] = E[NTrip_{0,Y09}|P(\boldsymbol{X})]. \qquad (8)$$

***Estimation*:** We estimate the impact of a change in employment status during the recession on recreation adopting the following average treatment effect on the treated (ATT) estimators-

$$\widehat{ATT}_{extensive} = \frac{1}{N_T}\left[\sum_{i \in I_1 \cap S_p}[Trip_{1,Y09} - \sum_{j \in I_0}\left(\widehat{W}_{ij}\right) \widehat{Trip_{0,Y09}}]\right] \qquad (9)$$

and,

$$\widehat{ATT}_{intensive} = \frac{1}{N_T}\left[\sum_{i \in I_1 \cap S_p}[NTrip_{1,Y09} - \sum_{j \in I_0}\left(\widehat{W}_{ij}\right) N\widehat{Trip_{0,Y09}}\ ]\right]. \qquad (10)$$

[6] This is also termed as balancing of covariates, $\boldsymbol{X}$, which also indicates the quality of the matching estimator. We perform this balancing test after each round of matching done to check if the *ignorability* condition is satisfied.

where I_1 is the set of treated observations, I_0 is the set of control observations, S_p is the region of common support, and N_T is the number of observations who belong to the set $I_1 \cap S_p$. $\widehat{Trip}_{0,Y09}$ is the matched outcome of control observation for treatment "i", which actually is constructed as the weighted average of all of the matched non-treatment outcomes. Similarly, $\widehat{W}_{ij}$ is the weight assigned to each matched control "j" corresponding to the treatment observation "i". Weight will depend on the distance between the propensity scores of treatment "i" and match "j", and the number of matches as well. For unmatched observations, weight is zero. We applied five different matching algorithms.[7]

Difference-in-difference matching to control for time-invariant unobservables: The estimators stated in equation 9 and 10 will give the true estimate of a recessionary change in employment status on one's recreation behavior if selection into such employment change is due to observable factors X. However, still there may exist unobservable factors, both time-variant as well as time-invariant in nature, that affect both the likelihood of an individual's exposure to employment change during a recession as well as her recreation behavior. For example, in our context, geographic factors (such as distance to lake, local amenities, local labor market conditions etc.) might confound the results.

Households residing near lakes but not used to taking any lake trip for recreation before a recession may find it relatively easier and cheaper to take some trips after experiencing an employment change during the recession due to more available time and negligible cost of a local trip. On the other hand, one who is living at a place with no lakes in the surrounding

[7] In nearest neighbor matching, for each treatment, we pick the control with the closest propensity score, both with and without replacement. Nearest five neighbors matching picks the five controls with the closest propensity score. Radius matching: for each exposed individual, we pick all the controls whose propensity score lies within a radius distance of ½ and 1/4th of standard deviation of the estimated propensity score.

amenities, but was used to taking trips before recession may find it relatively expensive to make trips after being affected by employment change during a recession. Let us call the former individual as type A, and the later as type B. Without taking into account the influences of location and distance, if we match a type B treatment with a control that lives in a lake-rich locality and is used to taking lake trips anyway, we will not capture true changes in recreation behavior from change in employment status. Similarly, we may end up matching a type A treatment with controls who are dissimilar in terms of locational attributes. Note that we cannot include all potential time-invariant controls, such as one's residence amenities or local attributes, in the propensity score estimation stage. Heckman, Ichimura, and Todd (1997), and Smith and Todd (2005) strongly recommends using difference-in-difference approach when geographic and other individual specific fixed factors might play potentially confounding role. Since we have a panel, by applying difference (DID) matching estimators we are able to difference out time-invariant unobservable factors.

Compared to a simple propensity score matching estimator, the DID matching estimator will estimate the treatment effect on the differences of outcome variable, which requires redefining the outcome variables by taking differences in recreation pattern across pre-recession and recession years.[8] Next, the DID estimators for out setting are:

$$DID_{extensive} = \widehat{ATT^{Y09}}_{extensive} - \widehat{ATT^{Y05}}_{extensive}, \quad (11)$$

$$DID_{intensive} = \widehat{ATT^{Y09}}_{intensive} - \widehat{ATT^{Y05}}_{intensive}, \quad (12)$$

where $\widehat{ATT^{Y09}}_{extensive}$ and $\widehat{ATT^{Y05}}_{intensive}$ have similar interpretation as in equation 9 and 10.

[8] The difference in participation in lake recreation (extensive margin) for the treatment group is $\Delta Trip_1 = Trip_{1,Y09} - Trip_{1,Y05}$ while that for the control group is $\Delta Trip_0 = Trip_{0,Y09} - Trip_{0,Y05}$. Total number of lake trips (intensive margin) is similarly redefined for the treatment group as $\Delta NTrip_1 = NTrip_{1,Y09} - NTrip_{1,Y05}$ and for the control group as $\Delta NTrip_0 = NTrip_{0,Y09} - NTrip_{0,Y05}$.

4. Iowa Lake Survey

In this study we utilize data from the Iowa lake survey, a random population survey, which collects a rich set of information on Iowan's lake visitation pattern as well as demographics on gender, age, education, employment status, income, and household composition. The survey has been administered five times in total, once in each of the four consecutive years 2002-2005, and the latest is in 2009.[9] The surveys in 2005 and 2009 together comprises a panel of 2773 individuals whom we can observe both before and during the recession in terms of their recreation behavior (both participation and number of trips) and relevant demographics. We first identify the group of people who have experienced a change in employment status during the recession besides those who have not to construct the treatment and control group for our study. Table one presents the employment information across the year 2005 and 2009.

In the 2005-2009 panel, 64.12% of the respondents provided the employment status information for both years. Approximately 6.5% of the people, who were full time employed in 2005, have reported either unemployment or a fall in working hours in 2009. In addition, 10% of the previously full-time employed people have retired in 2009. In the sample, 32.5% of the respondents (900 individuals) do not provide any information on employment status in 2009, which is quite high compared to similar nonresponse in 2005 (5.27%). Again, 52% of these 900 individuals were full-time employed in the pre-recession year 2009.[10] However, a simple mean comparison reveals that total number of trips in 2009 of the group with missing employment

[9] The survey in 2009 was sent to 10,000 people out of which 4500 were those who responded to a similar survey conducted in 2005. The survey response rate in 2009 was around 60%.

[10] There is possibility that individuals who have experienced employment shock during the recession are unwilling to share this information. Since in this study, we construct our sample based on individuals' employment status, significant numbers of respondents are dropping out due to this missing data on employment.

information is not statistically different from that with non-missing employment information.[11] This gives us confidence of not being trapped into sample selection bias due to missing data.

Following the definitions given in equations 1-3, we construct the treatment and control groups for our analysis. Table 2 shows the composition of the control and three treatment groups.[12] In the analysis we include those who report at most 52 trips in either of the years.[13] Among the three, treatment group 1 is the largest in size consisting of 155 observations in total, as it includes retired people besides unemployed and part-time employed. Information on participation, average number of trips, and demographics across treatment and control groups are reported in Table 3. Participation on average remains unchanged for the control group people across the years of 2005 and 2009. However, unlike the treatment group 3, treatment group 1 and 2 increase participation in lake recreation in 2009. For total number of trips, the pattern is little different. For the control group, treatment group1, and treatment group 3, mean number of trips fall in 2009 compared to 2005. In contrast, treatment group 2 exhibits an increase in mean number of trips in 2009. This gives us an indication of possible differences in recreation behavior across the retired and unemployed people.

Based on the available information in Iowa Lake Survey, the covariates we include while estimating the propensity score are age, polynomials of age, education, gender, number of children in the household, interaction terms between education, age, and gender, recreation patterns in the previous years, and boat ownership. All covariates assume values from the pre-recession period**,** the survey round in 2005. Education, age, and their interaction terms are

[11] Similar comparison between the two groups in year 2003, 2004, and 2005 reveal the same pattern as well.
[12] For comparison, an investigation into a similar panel from 2004-2005 shows that 3.5% of the full time employed people in 2004 becomes unemployed/ part time employed in the year 2005, and 2.5% of the full-employed people in 2004 have retired in 2005.
[13] Restriction of 52 trips in one year is to account for explicit day trips. Because some survey respondents might live near a lake, casually visit the lake while passing, and report inflated number of total trips.

motivated by the earning function estimation in the labor economics literature [Mincer 1974; Heckman, Lochner and Todd 2007].[14] We assume that factors that determine one's earnings are also strong predictors for his/her labor market status as well. An individual with a college degree and considerable experience is less likely to be exposed to an employment shock during the recession compared to an individual of similar experience but with only a high school degree. Iowa lake surveys also contain information on households' residence county and zip code. We match this with the rural-urban commuting area (RUCA) codes maintained by ERS, USDA to classify households by four different types of residence location.[15]The bottom panel of Table 3 presents summary statistics on various demographics, residence location, and recreation preference variables observed in pre-recession year for each of the treatment and control groups.

Some potentially useful variables to capture the intensity of exposure to recessionary shock, such as household income and work hours, are excluded due to widespread non-response. The idea of habit persistence [Adamowicz 1994; Moeltner and Englin 2004] suggests that pre-recession recreation behavior might influence one's recreation choices during the recession. Information on past recreation usage, such as total number of day trips and overnight trips taken in the pre-recession year are chosen to group households with similar preferences for recreation or common interest in activity types. Some lake recreation activities, such as fishing, boating and

[14] In the labor economics literature experience is often captured by a quadratic of age variable.
[15] Detail documentation o RUCA code are available at http://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/documentation.aspx#.Uu8l9fldWlI (*last accessed on July20th, 2015*). As the site notes: "*The rural-urban commuting area (RUCA) codes, a detailed and flexible scheme for delineating sub-county components of rural and urban areas, have been updated using data from the 2010 decennial census and the 2006-10 American Community Survey (ACS). RUCA codes are based on the same theoretical concepts used by the Office of Management and Budget (OMB) to define county-level metropolitan and micropolitan areas. We applied similar criteria to measures of population density, urbanization, and daily commuting to identify urban cores and adjacent territory that is economically integrated with those cores.*"

hunting are included since these will capture preference similarities among recreationists in more subtle manner.[16]

5. Results & Discussions

We start with a simple OLS model for each of the three treatment groups.[17] The OLS estimates reveal that for treatment group 2(unemployed and part-time employed), treatment status and participation in lake recreation are positively associated, which is statistically significant. In contrast, no such association is observed for treatment group 1 and treatment group 3. However, since we are concerned about selection problem with changes in employment status during the recession, we cannot interpret the estimates from OLS in a causal manner.

5.1 Propensity score estimation

Table 4 reports propensity score estimation results for each of the three treatment groups. For treatment group one, education, number of children, age, interaction between age and education, rural area residence, participation in fishing and boat activities turn out to be significant predictor of ones probability of experiencing a change in employment status during a recession. For treatment group 2, number of children, participation in fishing, and total number of trips taken in pre-recession year exhibit statistical significance. Similarly, for treatment group 3, education, interaction between age, education and gender, number of children, rural or small-town residence location, participation in fishing and boat activities turn out to be statistically significant predictors of one's chance of being retired during the recession.[18]

[16] Boating is a dummy variable which assumes value of 1 if a household owns a boat or participates in any of these boating activities such as jet skiing, canoeing, boating and sailing.
[17] In OLS exercises, the outcome variables are participation and total number of trips taken in 2009 whereas the explanatory variables include the same set of variables used in the matching exercises in addition to the treatment group indicator. We do not report the OLS results here to save space.
[18] Since the purpose of these regression estimates is to obtain propensity scores, based on which we will conduct matching, we are not focusing here interpreting the parameters.

Based on the estimated propensity score, in each of the cases, we match the treatment with the control applying five matching algorithms (i) nearest neighbor matching without replacement, (ii) nearest neighbor with replacement, (iii) nearest 5 neighbors, (iv) radius matching within a caliper of 1/4th of standard deviation of propensity score, and (v) radius matching within a caliper of 1/2 of standard deviation of propensity score.[19] For each matching algorithm, the balancing of covariates is assessed based on two criterions: (i) the difference between mean of treated and matched control group, and (ii) standardized mean difference of the covariates between the treatment and control group.[20] Prior to matching, as seen in Table A1, statistically significant differences across the treatment and control group are common. After matching is completed, the covariates balance well. As reported in Table A2-A4 in the *appendix* A, across treatment groups and matching algorithms, after a matching is conducted, more than 99.5% covariate balance well.[21]

To satisfy the overlapping condition, while estimating the treatment effects we exclude the treatments that are out of the common support. Table E1 in the *appendix* A shows the number of matched as well as non-matched treatment in each matching process for each of the three treatment groups. In 95% or more cases, treatments lie in common support region, or, find a comparable counterfactual from the control group.[22]

---

[19] All of the matching estimation is conducted utilizing package "*psmatch2*" in STATA 12.

[20] The standardized difference of means is calculated as: $= \frac{Mean_{Treated} - Mean_{control}}{\sqrt{\frac{1}{2} * (Variance_{treated} + Variance_{control})}}$. Following Rosenbaum and Rubin (1985) we consider a standardized difference of means of 20 as large.

[21] For the treatment group two, one covariate, gender, did not balance when nearest neighbor matching is conducted. Similarly, for the retired group, the covariate education exhibit large standardized difference in the case of nearest neighbor matching without replacement.

[22] For treatment group 2 and 3, total 3 treated observations, while for treatment group 1, 2 to 5 treatment observations do not find any counterfactual in the common support region.

5.2 Impact on participation

Table 5 presents the impact of change in employment status during a recession on participation in lake recreation. Note that a simple mean comparison across the treatment and unmatched control shows statistically no significant difference in lake recreation at the extensive margin. For treatment group 1, out of the five matching estimators used, two (nearest five neighbors matching and radius matching within 0.5 SD of propensity score) show that the treatment group participates more in outdoor lake recreation during the recession. These estimates suggest that households who become unemployed or retired during the recession are 8.3-10.9 percentage points more likely to participate in at least one lake-trip compared to the households who remain full-time employed across the pre-recession and recession period.[23]

The retired individuals may have distinct recreation preference compared to the unemployed. Panel b in Table 5 reveals that for the treatment group with employed and part-time employed people, five matching techniques indicate statistically significant positive impact of unemployment during the recession on participation in lake recreation. The estimates imply that an average household that was employed in 2005 but become unemployed in 2009 is 14 to 25 percentage points more likely to recreate in any of the Iowa lakes compared to what s/he would have done if were still full-time employed during the recession year. The bottom panel in Table 5, panel c, reports the results for the retired people. All of the five ATT estimates turn out to be statistically insignificant, which suggests that people who become retired during the recession do not start participating more in lake recreation. Note that in the analysis with treatment group two,

[23] In this paper we report the treatment effect (ATT) is statistical significant only if the *p value* is at least less than or equal to 0.1. In calculating the *p-value*, the standard errors are constructed based on 1000 replication of bootstrapping sample. Each bootstrap sample calculates the propensity score and matching in that sample is done based on that score.

all of the mean differences between the treatment and the control group are bigger in size compared to those we observe for treatment group one. A comparison of estimates across the three treatment groups suggests that the statistically significant impact we obtain for treatment group 1 is driven by the stronger and larger effect from the unemployed and part-time employed group, i.e., treatment group 2.

In the matching results discussed above, although we assume no selection on observables, there can still be unobservable time-invariant confounding factors hiding the true causal relationship between employment change during a recession and participation in lake recreation, for example distance of lakes from one's residence. In difference-in-difference matching, we will use the information on a household's participation in lake recreation both before and during the recession, which will net-out the effects of such time-invariant unobservable factors.

The difference-in-difference matching results for participation are presented in last two columns in Table 5. For treatment group 1, the results are similar to those for participation on the level. From nearest five neighbors and radius matching difference-in-difference estimates, we notice that households who experience a change in employment status during the recession take more lake visits. When we exclude the retired group and conduct difference-in-difference matching on unemployed and part-time employed people only (treatment group 2), all nearest neighbor matching processes show significant positive impact. However, in contrast to the case of matching on the level, radius matching algorithms do not show statistical significance. The bottom panel in Table 5 depicts that none of the matching processes indicate statistically any significant impact of retirement during recession on participation in lake recreation.

This positive effect on participation in lake recreation during the recession by households in treatment group 2 might be attributed to a couple of factors discussed in section 2 and 3. These households may consider lake recreation as an inferior good, or may have switched from stay at home option to outdoor lake recreation due to reduced opportunity cost of time. But we cannot exactly disentangle which factors are working and to what extent.

5.3 Impact on total number of trips

From our arguments presented in sections on literature review and theoretical motivations, we infer that total number of trips may increase, decrease, or remain unchanged. However, in the propensity score matching analysis, none of the treatment groups show any significant impact of employment change during the recession on total number of trips. Table 6 reports the findings. Although the differences across the treatment and control group are not statistically significant, the positive estimates of average treatment effect on the treated indicates that mean number of trips for the treatment group one and two are higher compared to their counterfactual number of trips. In contrast, the negative estimates for retired households indicate that their total number of lake visits during a recession is lower compared to their counterfactual frequency of visits.

Similar to the arguments presented for participation, we suspect the confounding effects from unobservable factors for total trips as well. To wipe out the mean effects from individually varying but time-invariant unobservable factors, we conduct DID matching estimator. The last two columns in Table 6 report the DID matching results. For any of the three treatment groups, DID estimators does not show any statistically significant impact of employment change during the 2009 recession on frequencies of outdoor lake trips. The DID matching does not change this pattern that we observe for matching on the level.

Contrary to the case for participation, unemployed or partially unemployed households do not increase frequencies of lake trips during the recession. Similar to the analysis for participation, the estimates do not suggest any impact of retirement during the recession on the total number of trips. However, although the estimates are statistically insignificant, retired households take outdoor lake trips at low frequencies compared to its counterfactual outcome. The finding is consistent across the matching estimators used.

6. Robustness

We conduct three robustness checks. First, we use a placebo recession year to check if our general assumption of no differential trend for treatment and control groups for DID matching estimator is valid in our setting. Second, we change the specification for propensity score estimation including a subset of covariates previously used: we exclude recreation preference variables. Third, we match each treatment observation with controls from the same geographic region to control for unobservable factors that are time-variant in a spatial manner.[24]

6.1 Placebo exercise

The objective of the placebo exercise is to check whether it is unemployment during the recession or some pre-existing unobservable factors working differently across the treatment and control group are driving our results. If the treatment and control group exhibit differentiated trend in the pre-recession years, and recession truly has no impact on recreation, the DID matching estimator picks up this difference in trend as impact of the change in employment during recession. For the placebo exercise, we assume year 2003 as placebo recession year.[25] Table B1-B3 in the *appendix* A report that balancing of covariate is satisfied in all cases except for only one covariate in one matching process for treatment group one. We report the estimates from the Placebo exercise in Table 7 and 8. In all matching processes, neither participation nor frequencies of trips

[24] Rural and urban areas may be affected differently during a recession.

[25] Although Iowa lake project survey was conducted in 2003 and 2004 as well, we have a matched non-missing sample for all of our treatment and control group observations in year 2003. In the survey year 2004 , we have missing information for 8 treatments and 52 controls from the sample of 971 observations that we are using for our base estimation

in lake recreation turned out to be statistically different across the treatment and control groups in 2003. This finding gives us confidence in saying that our analysis based on matching exercises as reported in the previous section are not contaminated due to differential group trend.

6.2 Specification without recreation preference variable

We estimate the propensity score excluding the recreation preference variables and including only demographics and type of region for residence in the pre-recession year. Table 9 reports the estimates for participation.[26] For treatment group 1 and 3, the treatment effect estimates on participation in lake recreation follows the same pattern that we observe previously in Table 5. For these two treatment groups, the difference-in-difference matching estimates are also robust to this different set of observables covariates. Panel b in Table 9 reveals that unemployed and part-time employed households (treatment group 2) participate more in lake recreation during the recession compared to what they would have done had they been employed. Note that, for the DID matching, previously in Table 5, radius matching estimators did not show any statistical significance but under the new setting, all five matching algorithms exhibit statistical significance. For matching on the level of participation, only radius matching estimators show statistical significance whereas previously in Table 5 all five matching estimators turned out to be statistically significant.

For the frequency of lake trips, the results are reported in Table 10. With the new set of covariates, none of the matching estimators across the three treatment groups exhibit statistically any significant effect of a change in employment status during recession on frequencies of trips. Our previous finding that frequencies of lake trips do not change due to unemployment or retirement during the recession is robust to the choice of covariates.

[26] Covariates balancing results, as reported in table C1-C3 in the *appendix*, reveal that quality of the match is good. Except for one variable for treatment group one in one estimation process, all other covariates balance well.

6.3 Matching within RUCA cell

Although we have accounted for the effects of mean time-invariant unobservable through matching on the differences, we recognize that we still might end up finding estimates confounded by unobservable fixed factors that vary across regions with time. For example, rural and metropolitan areas may be affected differently during a recession year and exhibit different economic environment. Employment statistics in a rural agricultural county may not change during the recession while employment in the metropolitan area usually drops sharply during the economic crisis. Although we incorporate information on one's residence location while estimating the propensity score, we still may end up matching a rural treatment with an urban control. Our DID matching estimators cannot control for such region-specific time-variant unobservable confounding factors. So matching individuals within region can control for such time-variant confounding effects.

To control for such possible regionally time-variant confounders, we match each treatment observation with controls from the same RUCA region. An individual from a small town experiencing employment shock during the recession is matched with counterfactuals from a small town area rather than from a metropolitan or rural area. Since we will match exactly within RUCA cell, in the first step, we estimate propensity score excluding variables on geographic regions. The results are reported in Table 11 and Table 12. Table D1-D3 report covariate balance for the cell matching. In contrast to the previous exercises, quality of the matches is not satisfactory here for the treatment group one and three since some covariates do not balance after matching. However, covariates balance well for the treatment group two- the unemployed group.

The estimates in Table11 reveals that when matching is done within the RUCA cell, only one out of five matching estimators for treatment group 1, and none for treatment group 3 show statistically significant effect of change in employment status on lake trips at the extensive margin. For treatment group two, two out of five matching estimators turn out to be statistically significant. However, once we apply the difference-in-difference matching, only for treatment group 1 the estimates display statistical significance. For treatment group 2 and 3, all five matching estimators are statistically insignificant. It suggests that for the unemployed and part-time employed group, positive impact of the recession on participation in lake recreation is not robust once we control for spatially time-variant confounders. In the case of frequencies of lake trips, as can be seen in Table 12, matching within RUCA cell generates similar estimates as before. However, since matching within rural-urban region causes quality of matching to be low, we are cautious in interpreting the estimates for treatment group one and three.

7. Extension

We extend the analysis in two ways. First, we include time-variant lake specific water quality measures to check if improved water quality is not actually driving the rise in lake recreation in 2009, which we have attributed to unemployment during the recession. Second, we investigate the lake recreation during recession by exploiting cross county-cross period variation in county unemployment rate.

7.1 Water quality, employment status, and lake recreation behavior during recession

Water quality varies across lakes and time periods. Water quality can be a major determinant of Iowans' choice of lake for outdoor recreation [Egan *et al.* 2009]. Detail water quality data on 131 major lakes in Iowa are available from Iowa State University's limnology lab.[27] The water quality data is well coordinated as well as temporally and spatially matched with

[27] http://www.card.iastate.edu/lakes/ (*last accessed on June 30th, 2015*).

the recreation data that we are utilizing for this analysis. In general, the measures reveal that average water quality has improved in 2009 compared to 2005 across the lakes in our sample. The objective of this extension is to examine whether water quality improvement is playing a confounding role and biasing our estimates reported in the previous sections. Following Egan *et al*. (2009), we have considered six water quality indicators: *secchi* depth, total nitrogen, total phosphorus, inorganic suspended solid, volatile suspended solid, and chlorophyll. Among all of these indicators, *secchi* depth, a measure of water clarity, is the most perceptible and direct water quality indicator to the recreationists.

We estimate the following specification including six water quality indicators

$$Trip_{ijt} = \alpha_0 \sum_{k=1}^{6} WQ_{kjt} + \alpha_1 * Treatment_{ij} + \alpha_2 * Treatment_{ij} * Recession + \alpha_3 * Recession + \alpha_4 \boldsymbol{X} + \in_{ijt},$$

where i =1,2,…,971 denotes households, $j=1,2,3,\ldots 131$ denotes lakes, and $t=2005,2009$ denotes time periods. Recession is an indicator variable that takes a value of 1 in 2009, and the vector $\boldsymbol{X}$ contains a set of demographics. For each of the three treatment groups, we estimated the above specification twice: (a) including the demographics, and (b) including individual fixed effects. In the fixed effects model α_1 will not be identified. However, the statistical significance of α_2 will reveal if unemployment effect during recession is robust to the water quality improvement.

The results are reported in Table 13. In general, water quality indicators turn out to be small in magnitude and always statistically insignificant. The key coefficient, interaction between treatment status and recession indicator, is consistently positive and statistically significant for combined treatment group's participation and count of total trips. After controlling for the water quality changes, the treatment group consisting of unemployed households seems to participate more and take more trips during the recession year compared to their counterfactual case. The retired group does not exhibit different participation behavior during the recession year. However, their frequency of trips dropped during the recession. Overall, our findings in the previous section that unemployment during recession lead a household to participate more in lake recreation is not altered when we address the water quality improvement in lakes in Iowa.[28]

7.2 County unemployment and recreation

In this section, we adopt an alternative approach to investigating the impact of the recession on participation in lake recreation. Although Iowa lake project surveys were conducted in 2002-2005 and 2009, in the matching exercises we utilized surveys from the recession year and the nearest pre-recession year including individuals who provide complete information on employment status. In all of the survey rounds, many respondents did not respond to the questions on employment status and household income.[29] One way, we may still use their information is by using some proxy for their employment/economic status. Economics literature investigating the relationship between individual health behavior and recession have utilized group variable such as state level unemployment rate to represent business cycle [Ruhm (2000, 2005), Dehejia and Lleras-Muney (2004)].

We utilize a panel spanning the years 2002-2005 and 2009 to estimate a fixed effect model on how county level unemployment rate affect individual participation in lake recreation. In our setting, county unemployment rate is a good proxy for local economic condition. Since one cannot influence economic activities at the county level, household's trip participation is less

[28] The estimates obtained in this subsection are not directly comparable with the matching estimates since the later are more conservative. However, the smaller magnitude of water quality coefficients from a relatively less-restricted model imply that improved water quality is not playing a confounding role in our case.
[29] In 2009, 853 individuals were silent about employment status although they provided relevant employment information in year 2005.

likely to affect county unemployment rate. The fixed effect model takes care of all individual specific time-invariant unobservable.[30] In the lake surveys, 3040 observations from year 2009 have a matching observation in at least one of the year 2002-2005. Out of them, 1498 individual responded across all the years 2002-2009 to form a balanced panel. We estimate the following specification

$$Trip_{ict} = (RecYr)\beta_1 + (Un_{ct})\beta_2 + (RecYr * Un_{ct})\beta_3 + \gamma_{ct} + \gamma_i + \epsilon_{ict},$$

where $Trip_{ict}$ is a binary variable indicating whether individual "i" in county "c" takes any lake trip in year "t" or not, Un_{ct} indicates unemployment in county "c" in year "t", $RecYr$ is an indicator variable assuming a value of 1 if year "t" is a recession year, γ_i are individual specific fixed effects which take care of time-invariant demographics, such as race, gender, education, preference for recreation or work, risk attitudes etc., $\boldsymbol{\gamma_{ct}}$ are county-specific time trends.[31] We are interested in the sign of the parameter β_3 on the interaction term between county level unemployment and indicator for recession. After controlling for the level effect of recession and county level unemployment along with individual fixed effect as well as various trends, if we find $\beta_3 > 0$ and statistically significant, we interpret it as a positive effect of unemployment during recession on participation in outdoor lake recreation. Since county level economic condition might exhibit correlation cross years, standard errors are clustered at county-year level [Wooldridge (2002)].

Table 14 shows the fixed effect estimates on the impact of county unemployment rate on recreation participation. Panel (a) in the Table reports results for the sample we use in matching exercises in previous section, while panel b and c report results for the unbalanced and a balanced panel respectively. We gradually increase controls. Column I does not incorporate any trend while

[30] For instance if one's preference for recreation is time-invariant; FE model would control for it.
[31] We have tried to control time trends by including both of county specific linear and quadratic time trends. Instead of including year specific fixed effects we include a dummy variable for the recession year to disentangle the recession year effect from a normal year effect.

column II and III include linear and quadratic trend. Instead of general linear or quadratic trends, columns IV and V incorporate county specific linear and quadratic trends. The coefficients of county unemployment rate and recession turn out to be negative and statistically significant. The estimates in columns II-IV in Table 14 reveal that in a recession year, participation in outdoor lake recreation decreases in the range of 14 to 20 percentage points compared to a non-recession year.

Similarly, in a particular year, individuals from a high unemployment county participate less in outdoor lake recreation compared to an individual from a low unemployment county. A one percentage point increase in local county level unemployment rate decreases participation in lake recreation by 1.72 to 4.73 percentage points based on the specifications we have used. This pattern is common across all specifications except that with county level quadratic trend (specification V). Note that in specification V, none of the variables display statistical significance-it seems like all variation in lake participation is absorbed by the county specific quadratic trend.

The coefficient of the interaction term between county level unemployment and recession year, after controlling for level effect of unemployment and recession, turns out to be positive and statistically significant in most of the cases(except for specification V consistently in all three samples). We interpret the statistically significant, and positive coefficient of the interaction term as a positive association between recessionary unemployment and outdoor recreation participation. An individual from a county with high unemployment rate during the recession year of 2009 is more likely to participate in lake recreation compared to one from a low unemployment county.

Although unemployment, on its own, reduces participation in lake recreation, the recessionary unemployment affects participation in outdoor recreation in an opposite manner. This pattern is consistent across the balanced sample as well, as can be observed in panel c. Our findings overall suggest that participation in lake recreation responds to unemployment in a different manner depending on whether the time is recessionary.

8. Conclusion

In this paper we utilize a panel from Iowa lake Survey to investigate how change in employment during a recession alters lake recreation behavior. Exploiting semi-parametric matching techniques, including *difference-in-difference* matching that utilizes the same individual's information before and during the recession, our analysis shows that retirement during the recession has no impact on recreation behavior, either at the extensive or intensive margin. In contrast, there is some evidence that people who become unemployed during the recession participate more in lake recreation. However, such increases are not consistent once we control for spatial factors- geographic and region specific unobservables- through conducting an exact matching within rural, urban, micropolitan and metropolitan cell. The extension of the analysis, where we replace individual employment status by county level unemployment to capture local economic condition, in an individual fixed effect framework, reveals that households from counties with high unemployment during a recession participate more in lake-based recreation. It reinforces the findings from matching exercises that participation in lake recreation at least did not decrease during the 2009 recession. Facing unemployment during the recession households might have substituted relatively exotic trips for local lake-based recreation that might be one possible reason behind observing an increase in lake recreation pattern during the 2009 recession.

Overall, the findings suggest that changes in employment status during the 2009 recession did not affect demand for lake-based recreation trips: Iowans are visiting lakes as frequently as they were before the recession. The finding implies that the demand for outdoor recreations in lakes is stable and unaffected by business cycle shocks such as a recession. It will inform public officials that public amenity-based, rural, non-farm economic activities and employment are resilient to a recession.

Outdoor recreation plays a significant role in Iowa's economy as it generates local spending of approximately $3.1 billion annually and 31000 jobs (Otto, Tylka, and Erickson, 2012). Besides, there are indirect benefits from improved health and new investments through increased opportunities for recreational activities. Being a predominantly farm state, Iowa often faces the challenge in encouraging other natural resource based industries compromising agriculture. The finding from this study reemphasizes the importance of outdoor recreation, another natural resource based stable non-agricultural sector that can be affected by negative spillover effects from agriculture. Natural-resource centered development programs including investment in public goods such as water quality improvement by controlling emission from various point and non-point sources, and development of public facilities such as local roads, lakes and parks may generate higher net social return compared to other local public goods.

9. References

Adamowicz, W. L., 1994. "Habit formation and variety seeking in a discrete choice model of recreation demand." *Journal of Agricultural and Resource Economics* 19(1): 19-31.

Baird, S., J. Friedman, and N. Schady. 2011. "Aggregate Income Shocks and Infant Mortality in the Developing World." *The Review of Economics and Statistics* 93(3): 847–856.

Bockstael, N., I. Strand, and W. M. Hanemann. 1987. "Time and the recreational demand model." *American Journal of Agricultural Economics* 69:293–302.

Cesario, F. J. 1976. "Value of time in recreation benefit studies." *Land Economics* 52:32-41.

Caliendo, M., and S. Kopeinig. 2011. "Some practical guidance for the implementation of propensity score matching." *Journal of Economic Surveys* 22(1): 31-72.

Chabé-Ferret, S., and J. Subervie. 2013. "How Much Green for the Buck? Estimating Additional and Windfall Effects of French Agro-environmental Schemes by DID-matching." *Journal of Environmental Economics and Management* 65(2013): 12–27.

Dehejia, R., and Adriana Lleras-Muney. 2004. ''Booms, Busts, and Babies' Health.'' *Quarterly Journal of Economics* 119:1091–1130.

Dehejia, R. H., S. Wahba. 2002. "Propensity score-matching methods for non-experimental causal studies." *Review of Economics and Statistics* 84:151–161.

Dhaval, M.D., and I.R. Kelly. 2010. "How Does the Business Cycle Affect Eating Habits?" *NBER Working Paper No. 16638*

Egan, Kevin J., J. A. Herriges, C.L. Kling, and J.A. Downing. 2009. "Valuing Water Quality as a Function of Water Quality Measures." *American Journal of Agricultural Economics* 91:106-123.

Englin, J., and J.S. Shonkwiler. 1995. "Modeling recreation demand in the presence of unobservable travel costs: toward a travel price model." *Journal of Environmental Economics and Management* 29: 368–377.

Evans, K., J.A. Herriges, and C.L. Kling. 2011. "The Iowa Lakes Valuation Project 2009 Summary and Findings." A Report to the Iowa Department of Natural Resources. Available at: http://www.card.iastate.edu/lakes/iowa-lakes-report-2009.pdf (*accessed on July 15th, 2013*)

Feather, P., and W.D. Shaw. 1999. "Estimating the cost of leisure time for recreation demand models." *Journal of Environmental Economics and Management* 38:49–65.

Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd. 1998. "Characterizing selection bias using experimental data." *Econometrica* 66(5):1017–1098.

Heckman, J. J., L.J. Lochner, and P.E. Todd. 2006. Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation And Beyond. In Eric A. Hanushek and Finis Welch (Eds.), *Handbook of the Economics of Education*, Amsterdam: North-Holland.

Hoque, M. M., J. A. Herriges, and C. L. Kling. 2013. "Is Outdoor Recreation Recession-proof? An Empirical Investigation on Iowan's Lake Recreation Behavior during 2009 Recession?" Paper presented at Agricultural and Applied Economics Association Annual Meeting, Washington DC, August 4-6, 2013. Available at http://purl.umn.edu/150640 (last accessed on January 17th, 2014).

Jalan, J., and M. Ravallion. 2003. "Does Piped Water Reduce Diarrhea for Children in Rural India?" *Journal of Econometrics* 112:153-173.

Larson, D., and S. Shaikh. 2004. "Revealed leisure time values." *Economic Inquiry* 42:264-278.

List, John A., Daniel Millimet, Per Fredriksson, and Warren McHone. 2003. "Effects of Environmental Regulations on Manufacturing Plant Births: Evidence from a Propensity Score Matching Estimator," *Review of Economics and Statistics* 85(2003): 944–952.

Loomis, J., and C. Keske. 2012. "Did the great recession reduce visitors spending and willingness to pay for nature-based recreation? Evidence from 2006 and 2009." *Contemporary Economic Policy* 30:238-246.

McConnell, K. 1992. "On-site time in the demand for recreation." *American Journal of Agricultural Economics* 74:918–924.

Mincer, J. 1974. *Schooling, Experience, and Earnings*. New York: Columbia University Press.

Moeltner, K., and J. Englin. 2004. "Choice behavior under time-variant quality." *Journal of Business and Economic Statistics* 22(2):214-224.

Otto, D., T. Kristin, and S. Erickson. 2012. "Economic Value of Outdoor Recreation Activities in Iowa", Iowa state University. Available at http://www.card.iastate.edu/environment/items/DNR-AmenityRevised_9-25-12.pdf (*last accessed on July20th, 2015*).

Outdoor Industry Foundation (OIF). 2010. *Outdoor Recreation Participation Report*. Available at: http://www.outdoorindustry.org/images/researchfiles/TOF_ResearchParticipation2010.pdf?121. (*last accessed on July20th, 2015*).

Phaneuf, D.J., and V.K. Smith. 2005. "Recreation Demand Models," in K. Mäler and J. Vincent (Eds.), *Handbook of Environmental Economics*, Volume 2 (North-Holland).

Phaneuf, D.J., and T. Requate. 2013. *A Course in Environmental Economics: Theory, Empirics, and Practice*. Cambridge University Press.

Ravallion, M. 2005. "Evaluating anti-poverty programs," *The World Bank Policy Research Working Paper Series: 3625*

Rosenbaum, P.R., and D.B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70:41-55.

Rosenbaum, P.R., and D.B. Rubin. 1985. "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score." *The American Statistician* 39(1):33–38.

Ruhm, C.J. 2000. "Are recessions good for your health"? *Quarterly Journal of Economics* 15:617–650.

Ruhm, C.J. 2005. "Healthy living in hard times." *Journal of Health Economics* 24:241–364

Smith, V.K., W. Desvousges, and M. McGivney. 1983. "The opportunity cost of travel time in recreation demand models." *Land Economics* 59:259–278.

Smith, J., and P. Todd. 2005. "Does Matching Overcome Lalonde's Critique of Non-experimental Estimators?" *Journal of Econometrics* 125(2005): 305–353.

Wooldridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data*, First edition. MIT Press: Cambridge, MA.

Table 1: Number of survey respondents by employment status in 2005 and 2009

	Employment Status in 2009					
Employment Status in 2005	Full-time	Part-time	Student	Unemployed	Retired	Total
Full-time	848	43	4	23	100	1,018
Part-time	29	69	2	8	37	145
Student	8	0	2	1	1	12
Unemployed	17	6	1	20	13	57
Retired	14	24	1	1	506	546
Total	916	142	10	53	657	1778

Table 2: Size and decomposition of Treatment and Control Groups

Treatment group	Employment Status in year 2009	Number of Treatment Observations	Number of Control Observations
1	Unemployed	42	
	Part-time Employed	21	
	Retired	92	
	Total	155	816
2	Unemployed	42	
	Part-time Employed	21	
	Total	63	816
3	Retired	92	816
	Total	92	816

Note: *In table 2 we exclude all respondents who report more than 52 trips in one single year. Compared to table 1, this reduces the control group size from 848 to 816. Similarly we adjust the treatment group sizes as well.*

Table 3: Participation, Total Trips, Demographics and Recreation Activities across Treatment Groups by Baseline and Treatment Years

	Control Group		Treatment Group 1		Treatment Group 2		Treatment Group 3	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Participation								
Participation in 2005	0.675	0.469	0.600	0.491	0.571	0.499	0.620	0.488
Participation in 2009	0.675	0.469	0.665	0.474	0.746	0.439	0.609	0.491
Total Trip								
Total Trip in 2005	7.354	10.188	7.071	11.635	4.698	7.370	8.696	13.620
Total Trip In 2009	6.933	9.927	6.806	9.771	5.619	7.458	7.620	11.046
Pre-recession Year: 2005								
Demographics								
Age	4.433	0.750	5.161	0.802	4.746	0.842	5.446	0.635
Gender	1.246	0.431	1.355	0.480	1.429	0.499	1.304	0.463
Education	3.384	1.002	3.271	1.089	3.175	1.025	3.337	1.132
Number of Children in the household	0.939	1.210	0.310	0.717	0.492	0.840	0.185	0.592
Rural Residence	0.153	0.360	0.103	0.305	0.095	0.296	0.109	0.313
Small Town Residence	0.211	0.408	0.213	0.411	0.302	0.463	0.152	0.361
Micropolitan Residence	0.132	0.339	0.148	0.357	0.175	0.383	0.130	0.339
Metropolitan Residence	0.504	0.500	0.535	0.500	0.429	0.499	0.609	0.491
Recreation Preference Variables								
Boat Activities	0.553	0.498	0.426	0.496	0.429	0.499	0.424	0.497
Hunting	0.065	0.247	0.039	0.194	0.016	0.126	0.054	0.228
Fishing	0.512	0.500	0.503	0.502	0.492	0.504	0.511	0.503
Total Number of Trips	7.354	10.188	7.071	11.635	4.698	7.370	8.696	13.620
Take overnight Trips	0.456	0.498	0.374	0.485	0.270	0.447	0.446	0.500

Table 4: Propensity Score Estimation Results from Probit Model

	Treatment Group 1		Treatment Group 2		Treatment Group 3	
	Coefficient	*p value*	Coefficient	*p value*	Coefficient	*p value*
Age	-0.831	0.211	-0.386	0.576	0.468	0.752
Age square	0.171**	0.011	0.080	0.280	0.081	0.567
Gender	0.002	0.994	0.218	0.376	-0.241	0.360
Education	0.176	0.222	0.085	0.582	0.365**	0.040
Age*Education	-0.060**	0.034	-0.046	0.146	-0.090***	0.008
Education*Gender	0.083	0.129	0.042	0.526	0.118*	0.085
Number of Children in Household	-0.158**	0.019	-0.125*	0.071	-0.198*	0.092
Rural	-0.436**	0.013	-0.291	0.213	-0.522**	0.018
Small Town	-0.082	0.553	0.199	0.225	-0.376*	0.052
Micropolitan	0.018	0.908	0.166	0.389	-0.199	0.312
Boat Activity	-0.226*	0.054	-0.159	0.292	-0.288*	0.042
Hunting	-0.251	0.318	-0.497	0.255	-0.144	0.620
Fishing	0.270**	0.023	0.271*	0.075	0.281*	0.067
Total Number of Trips	-0.001	0.928	-0.014*	0.065	0.008	0.197
Take overnight Trips	-0.031	0.794	-0.249	0.119	0.178	0.230
constant	-0.841	0.615	-1.291	0.449	-5.411	0.163
Model Statistics						
Number of observations	971		879		908	
Log pseudolikelihood	-353.617		-205.079		-213.674	
Wald chi2(15)	127.620		44.270		133.500	
Prob > chi2	0.000		0.000		0.000	
Pseudo R2	0.171		0.096		0.283	

Note: *** *p<0.01; *** ** *p<0.05; *** * *p<0.1.*

The propensity score equation has incorporated all the variables from our available information set that might be relevant for outcome equation as well as determining ones probability of being in the treatment group. Some interaction terms and polynomials are included in the propensity score estimation process so that matching based on those scores satisfies conditional independence assumption.

Table 5: Estimates of Average Treatment Effect on the Treated for Participation in Lake Recreation

	(I) Participation		(II) Difference in Participation	
Panel (a): Unemployed, Part-time Employed, and Retired People				
Matching Algorithm	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.059	0.055	0.033	0.060
Nearest Neighbor with replacement	0.091	0.075	0.052	0.076
Nearest 5 Neighbors with replacement	0.109*	0.060	0.101*	0.060
Radius Matching (caliper =0.5*SD)	0.083*	0.047	0.102**	0.051
Radius Matching (caliper =0.25*SD)	0.077	0.050	0.097*	0.053
Panel (b): Unemployed and Part-time Employed People				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.25***	0.088	0.20*	0.105
Nearest Neighbor with replacement	0.26***	0.102	0.26**	0.118
Nearest 5 Neighbors with replacement	0.16**	0.081	0.18*	0.096
Radius Matching (caliper =0.5*SD)	0.14**	0.064	0.130	0.081
Radius Matching (caliper =0.25*SD)	0.14**	0.067	0.110	0.084
Panel (c):Retired People				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	-0.045	0.073	0.022	0.077
Nearest Neighbor with replacement	-0.011	0.095	0.045	0.089
Nearest 5 Neighbors with replacement	0.029	0.074	0.037	0.074
Radius Matching (caliper =0.5*SD)	0.008	0.059	0.049	0.066
Radius Matching (caliper =0.25*SD)	0.014	0.063	0.060	0.069

Note: **** p<0.01; ** p<0.05; * p<0.1.Caliper is chosen as 1/2 of standard deviation of propensity score, and ¼ of standard deviation of propensity score. Standard errors reported are obtained from bootstrapping with 1000 replications.*

Table 6: Estimates of Average Treatment Effect on the Treated for Total Number of Lake-Trips

	(I) Total Trip		(II) Difference in Total Trip	
Panel (a):Unemployed, Part-time Employed, and Retired People				
Matching Algorithm	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.118	1.021	0.150	0.920
Nearest Neighbor with replacement	0.605	1.381	0.645	1.233
Nearest 5 Neighbors with replacement	0.958	1.061	-0.048	0.939
Radius Matching (caliper =0.5*SD)	0.716	0.782	0.433	0.731
Radius Matching (caliper =0.25*SD)	0.408	0.828	0.117	0.745
Panel (b): Unemployed and Part-time Employed				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	1.117	1.433	0.900	1.223
Nearest Neighbor with replacement	0.761	1.661	1.161	1.423
Nearest 5 Neighbors with replacement	0.191	1.287	0.324	1.136
Radius Matching (caliper =0.5*SD)	0.432	0.972	0.482	0.919
Radius Matching (caliper =0.25*SD)	0.567	1.000	0.318	0.938
Panel (c):Retired People				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	-0.775	1.388	0.584	1.355
Nearest Neighbor with replacement	-0.970	2.006	0.907	1.846
Nearest 5 Neighbors with replacement	-0.923	1.536	-0.065	1.416
Radius Matching (caliper =0.5*SD)	-0.496	1.183	-0.213	1.099
Radius Matching (caliper =0.25*SD)	-0.567	1.311	-0.063	1.193

Note: **** p<0.01; ** p<0.05; * p<0.1.Caliper is chosen as 1/2 of standard deviation of propensity score, and ¼ of standard deviation of propensity score. Standard errors reported are obtained from bootstrapping with 1000 replications.*

Table 7: Placebo Effect -Estimates of Average Treatment Effect on the Treated for Participation in Lake Recreation

	(I) Participation in Placebo recession year		(II) Difference in Participation between 2003(Placebo recession year) and 2005	
Panel (a): Unemployed, Part-time Employed, and Retired People				
Matching Algorithm	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.007	0.054	0.000	0.060
Nearest Neighbor with replacement	-0.005	0.076	0.037	0.077
Nearest 5 Neighbors with replacement	0.025	0.060	-0.035	0.065
Radius Matching (caliper =0.5*SD)	0.006	0.048	-0.020	0.053
Radius Matching (caliper =0.25*SD)	0.020	0.050	-0.043	0.056
Panel (b): Unemployed and Part-time Employed People				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.017	0.090	0.067	0.096
Nearest Neighbor with replacement	-0.017	0.106	0.100	0.114
Nearest 5 Neighbors with replacement	0.080	0.082	-0.094	0.088
Radius Matching (caliper =0.5*SD)	0.088	0.065	-0.075	0.069
Radius Matching (caliper =0.25*SD)	0.080	0.066	-0.051	0.072
Panel (c):Retired People				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.034	0.072	-0.023	0.079
Nearest Neighbor with replacement	0.045	0.092	0.023	0.097
Nearest 5 Neighbors with replacement	-0.016	0.073	0.002	0.082
Radius Matching (caliper =0.5*SD)	-0.058	0.059	0.014	0.066
Radius Matching (caliper =0.25*SD)	-0.051	0.063	0.016	0.070

Note: *** *p<0.01; ** p<0.05; * p<0.1.Caliper is chosen as 1/2 of standard deviation of propensity score, and ¼ of standard deviation of propensity score. Standard errors reported are obtained from bootstrapping with 1000 replications.*

Table 8: Placebo Effects-Estimates of Average Treatment Effect on the Treated for Total Number of Lake-Trips

	(II) Participation in Placebo recession year		(II)Difference in Participation between 2003 and 2005	
Panel (a):Unemployed, Part-time Employed, and Retired People				
Matching Algorithm	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.209	0.885	0.209	0.855
Nearest Neighbor with replacement	0.006	1.230	-0.326	1.276
Nearest 5 Neighbors with replacement	0.116	0.936	0.449	0.936
Radius Matching (caliper =0.5*SD)	-0.129	0.716	0.243	0.695
Radius Matching (caliper =0.25*SD)	0.076	0.764	0.171	0.741
Panel (b): Unemployed and Part-time Employed				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.100	1.233	0.267	1.133
Nearest Neighbor with replacement	-0.144	1.511	0.528	1.409
Nearest 5 Neighbors with replacement	0.154	1.086	-0.064	0.996
Radius Matching (caliper =0.5*SD)	0.487	0.772	-0.496	0.648
Radius Matching (caliper =0.25*SD)	0.494	0.816	-0.278	0.683
Panel (c):Retired People				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.670	1.282	0.648	1.264
Nearest Neighbor with replacement	0.736	1.736	-0.406	1.852
Nearest 5 Neighbors with replacement	-0.594	1.327	0.650	1.393
Radius Matching (caliper =0.5*SD)	-0.541	1.081	0.286	1.044
Radius Matching (caliper =0.25*SD)	-0.469	1.157	0.415	1.145

Note: **** p<0.01; ** p<0.05; * p<0.1.Caliper is chosen as 1/2 of standard deviation of propensity score, and ¼ of standard deviation of propensity score. Standard errors reported are obtained from bootstrapping with 1000 replication*

Table 9: Robustness Check of Estimates of Average Treatment Effect on the Treated for Participation in Lake Recreation (with different group of Covariates)

	Participation		Difference in Participation	
Panel (a):Unemployed and Retired People				
Matching Algorithm	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.046	0.056	0.059	0.058
Nearest Neighbor with replacement	0.065	0.064	0.075	0.058
Nearest 5 Neighbors with replacement	0.084	0.059	0.094*	0.056
Radius Matching (caliper =0.5*SD)	0.083*	0.050	0.104**	0.050
Radius Matching (caliper =0.25*SD)	0.089*	0.051	0.105**	0.050
Panel (b): Unemployed and Part-time Employed				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.129	0.087	0.177*	0.096
Nearest Neighbor with replacement	0.089	0.088	0.160*	0.095
Nearest 5 Neighbors with replacement	0.126	0.078	0.172**	0.088
Radius Matching (caliper =0.5*SD)	0.115*	0.069	0.163**	0.081
Radius Matching (caliper =0.25*SD)	0.118*	0.067	0.165**	0.080
Panel (c):Retired				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.065	0.080	0.048	0.073
Nearest Neighbor with replacement	0.061	0.072	0.029	0.071
Nearest 5 Neighbors with replacement	0.057	0.065	0.036	0.063
Radius Matching (caliper =0.5*SD)	0.069	0.066	0.024	0.065
Radius Matching (caliper =0.25*SD)	0.065	0.080	0.048	0.073

Note: **** p<0.01; ** p<0.05; * p<0.1.Caliper is chosen as 1/2 of standard deviation of propensity score, and ¼ of standard deviation of propensity score. Standard errors reported are obtained from bootstrapping with 1000 replications.*

Table 10: Robustness Check of Estimates of Average Treatment Effect on the Treated for Total Number of Trips (with different group of Covariates)

	Total Trip		Difference in Total Trip	
Panel (a):Unemployed and Retired People				
Matching Algorithm	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	-0.791	1.093	-0.098	1.020
Nearest Neighbor with replacement	-0.026	1.275	-0.017	1.067
Nearest 5 Neighbors with replacement	0.711	1.141	0.212	1.012
Radius Matching (caliper =0.5*SD)	0.265	0.982	0.398	0.883
Radius Matching (caliper =0.25*SD)	0.490	1.014	0.571	0.908
Panel (b): Unemployed and Part-time Employed				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	-1.855	1.639	1.597	1.455
Nearest Neighbor with replacement	-2.472	1.656	1.213	1.416
Nearest 5 Neighbors with replacement	-1.141	1.424	1.565	1.297
Radius Matching (caliper =0.5*SD)	-0.964	1.164	1.648	1.049
Radius Matching (caliper =0.25*SD)	-1.089	1.125	1.337	1.017
Panel (c):Retired				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.789	1.529	-0.544	1.385
Nearest Neighbor with replacement	1.944	1.603	-0.393	1.425
Nearest 5 Neighbors with replacement	1.969	1.476	-0.063	1.356
Radius Matching (caliper =0.5*SD)	1.248	1.421	-0.263	1.226
Radius Matching (caliper =0.25*SD)	1.605	1.470	-0.447	1.289

Note: **** p<0.01; ** p<0.05; * p<0.1.Caliper is chosen as 1/2 of standard deviation of propensity score, and ¼ of standard deviation of propensity score. Standard errors reported are obtained from bootstrapping with 1000 replications.*

Table 11: Robustness Check of Estimates of Average Treatment Effect on the Treated for Participation in Lake Recreation (Matching within Rural, Small Town, Micropolitan, and Metropolitan cell)

	Participation		Difference in Participation	
Panel (a):Unemployed and Retired People				
Matching Algorithm	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.071	0.048	0.096*	0.051
Nearest Neighbor with replacement	0.120*	0.071	0.114	0.070
Nearest 5 Neighbors with replacement	0.078	0.059	0.096*	0.059
Radius Matching (caliper =0.5*SD)	0.077	0.052	0.113**	0.053
Radius Matching (caliper =0.25*SD)	0.077	0.052	0.094*	0.053
Panel (b): Unemployed and Part-time Employed				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.164**	0.080	0.131	0.089
Nearest Neighbor with replacement	0.147	0.100	0.097	0.112
Nearest 5 Neighbors with replacement	0.129	0.080	0.142	0.091
Radius Matching (caliper =0.5*SD)	0.133*	0.071	0.116	0.083
Radius Matching (caliper =0.25*SD)	0.095	0.072	0.094	0.087
Panel (c):Retired				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.017	0.059	-0.023	0.061
Nearest Neighbor with replacement	0.025	0.090	0.044	0.079
Nearest 5 Neighbors with replacement	0.053	0.076	0.016	0.069
Radius Matching (caliper =0.5*SD)	0.058	0.070	0.020	0.067
Radius Matching (caliper =0.25*SD)	0.055	0.077	0.015	0.068

Note: *** *p<0.01; ** p<0.05; * p<0.1.Caliper is chosen as 1/2 of standard deviation of propensity score, and ¼ of standard deviation of propensity score. Standard errors reported are obtained from bootstrapping with 1000 replications.*

Table 12: Robustness Check of Estimates of Average Treatment Effect on the Treated for Total Number of Trips (Matching within Rural, Small Town, Micropolitan, and Metropolitan cell)

	Total Trip		Difference in Total Trip	
Panel (a):Unemployed and Retired People				
Matching Algorithm	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.750	0.938	0.750	0.863
Nearest Neighbor with replacement	0.795	1.365	0.660	1.327
Nearest 5 Neighbors with replacement	0.227	1.160	-0.023	1.004
Radius Matching (caliper =0.5*SD)	0.302	0.977	0.219	0.844
Radius Matching (caliper =0.25*SD)	0.148	0.994	-0.217	0.843
Panel (b): Unemployed and Part-time Employed				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.950	1.222	1.767	1.191
Nearest Neighbor with replacement	1.200	1.594	1.983	1.483
Nearest 5 Neighbors with replacement	0.146	1.239	0.617	1.188
Radius Matching (caliper =0.5*SD)	0.059	1.151	0.321	1.026
Radius Matching (caliper =0.25*SD)	-0.168	1.272	-0.245	1.110
Panel (c):Retired				
	Coefficient	Std. Err.	Coefficient	Std. Err.
Nearest Neighbor without replacement	0.659	1.379	-0.970	1.199
Nearest Neighbor with replacement	0.999	1.663	-0.838	1.502
Nearest 5 Neighbors with replacement	2.015	1.478	-0.114	1.368
Radius Matching (caliper =0.5*SD)	1.651	1.424	-0.333	1.256
Radius Matching (caliper =0.25*SD)	1.664	1.608	-0.818	1.320

Note: **** p<0.01; ** p<0.05; * p<0.1.Caliper is chosen as 1/2 of standard deviation of propensity score, and ¼ of standard deviation of propensity score. Standard errors reported are obtained from bootstrapping with 1000 replications.*

Table 13: Water Quality at the Destination Lakes, Employment Status, and Trip taking Behavior

Variables	Participation			Total Number of Trips		
	All (Unemployed and retired)	Unemployed Only	Retired Only	All (Unemployed and retired)	Unemployed Only	Retired Only
Panel(a): Include Fixed Effects						
Treatment* Recession (standard error)	0.069*** (0.003)	0.166*** (0.005)	0.002 (0.003)	0.143*** (0.048)	1.32*** (0.059)	-0.68*** (0.067)
Recession	+(sig)	+(sig)	+(sig)	-(sig)	(-sig)	(-sig)
Secchi Depth	+(insig)	-(insig)	+(insig)	+(insig)	+(insig)	+(insig)
Total Nitrogen	+(insig)	-(insig)	+(insig)	+(insig)	-(insig)	-(insig)
Total Phosphorus	-(insig)	+(insig)	-(insig)	-(insig)	-(insig)	-(insig)
Inorganic Suspended Solid	+(insig)	-(insig)	-(insig)	+(insig)	+(insig)	+(insig)
Volatile Suspended Solid	+(insig)	-(insig)	+(insig)	-(insig)	-(insig)	-(insig)
Chlorophyll	+(insig)	-(insig)	+(insig)	+(insig)	+(insig)	+(insig)
Individual FE	YES	YES	YES	YES	YES	YES
Panel(b): Include Demographics						
Treatment*Recession (standard error)	0.073*** (0.003)	0.18*** (0.005)	-0.003 (0.003)	0.299*** (0.047)	1.399*** (0.059)	-0.484*** (0.065)
Recession	+(sig)	+(sig)	+(sig)	-(sig)	(-sig)	(-sig)
Secchi Depth	- (insig)	- (insig)	- (insig)	- (insig)	- (insig)	- (insig)
Total Nitrogen	- (insig)	- (insig)	- (insig)	- (insig)	- (insig)	- (insig)
Total Phosphorus	+(insig)	+(insig)	+(insig)	-(insig)	-(insig)	+(insig)
Inorganic Suspended Solid	-(insig)	-(insig)	-(insig)	-(insig)	+(insig)	+(insig)
Volatile Suspended Solid	+(insig)	+(insig)	+(insig)	-(insig)	-(insig)	-(insig)
Chlorophyll	-(insig)	-(insig)	-(insig)	-(insig)	-(insig)	-(insig)
Age	-(sig)	-(sig)	-(sig)	-(sig)	-(sig)	-(sig)
Education	+(sig)	+(sig)	+(sig)	+(sig)	-(sig)	+(sig)
Child	-(sig)	-(sig)	-(sig)	-(sig)	-(sig)	-(sig)
Adults	+(sig)	+(sig)	+(sig)	+(sig)	+(sig)	+(sig)

Note: *** $p<0.01$; ** $p<0.05$; * $p<0.1$. *Demographics include age, education, children under age 2 and number of adults in the households. Since the magnitude of water quality indicators are extremely low in magnitude, we report their sign and statistical significance for convenience. All specifications estimate a recession indicator. In the fixed effect framework, treatment indicator is dropped. Although not reported here, we estimate models with three-way interaction among water quality measures, treatment indicator, and recession year. The estimate on the three-way interaction term never turned out to be significant but the interaction term between treatment indicator and recession year remain unaltered.*

Table 14: County level Unemployment and Participation in Lake Recreation

	I	II	III	IV	V
Panel (a): Sample used in DID matching estimator					
County Unemployment Rate	-0.059***	-0.0454***	-0.0168	-0.0380***	0.044
	[0.014]	[0.013]	[0.014]	[0.013]	[0.027]
Recession Year	-0.165***	-0.0895	-0.143*	-0.0117	0.114
	[0.056]	[0.056]	[0.059]	[0.069]	[0.113]
Recession*County Unemployment rate	0.0371**	0.0287**	0.0113	0.0139	-0.0550**
	[0.012]	[0.011]	[0.012]	[0.013]	[0.026]
Constant	0.975***	0.951***	0.867**	0.919***	0.628***
	[0.060]	[0.053]	[0.052]	[0.054]	[0.103]
Linear Trend		YES			
Quadratic Trend			YES		
County specific linear trend				YES	
County specific quadratic trend					YES
Sample Size	971	971	971	971	971
Panel (b): Balanced Sample					
	I	II	III	IV	V
Recession Year	-0.222	-0.136**	-0.182***	-0.115*	-0.011
	[0.047]	[0.048]	[0.048]	[0.052]	[0.077]
County Unemployment Rate	-0.061	-0.0428***	-0.020	-0.0479***	-0.008
	[0.010]	[0.010]	[0.012]	[0.010]	[0.018]
Recession*County Unemployment rate	0.045	0.0336***	0.0192*	0.0315**	-0.013
	[0.009]	[0.009]	[0.010]	[0.010]	[0.017]
Linear Trend		YES			
Quadratic Trend			YES		
county specific trend				YES	
county specific quadratic trend					YES
constant	0.956	0.926***	0.933***	0.947***	0.897***
	[0.042]	[-0.040]	[-0.039]	[-0.041]	[-0.056]
Sample Size	1494	1494	1494	1494	1494
Panel(c): Unbalanced Sample					
	I	II	III	IV	V
Recession Year	-0.221***	-0.146***	-0.198***	-0.138**	-0.105
	[0.038]	[0.038]	[0.038]	[0.042]	[0.072]
County Unemployment Rate	-0.055***	-0.042***	-0.017*	-0.045***	-0.011
	[0.008]	[0.008]	(0.009)	(0.009)	(0.015)
Recession*County Unemployment rate	0.042***	0.034***	0.019*	0.033***	0.002
	[0.008]	[0.007]	[0.008]	[0.008]	[0.015]
Linear Trend		YES			
Quadratic Trend			YES		
county specific trend				YES	
county specific quadratic trend					YES
Constant	0.931***	0.926***	0.947***	0.942***	0.930***
	[0.037]	[0.032]	[0.030]	[0.036]	[0.049]
Sample Size	3020	3020	3020	3020	3020

Note: *Standard errors are reported in bracket. ***, **, and * indicates significance at 1 percent, 5 percent and 10 percent level.*

APPENDIX A. APPENDIX TO CHAPTER 2

Table A1: Difference between treatment and control groups before matching

	Group 1(Unemployed and Retired)				Group 2(Unemployed)				Group 3(Retired)			
Variable	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Control	Mean Treated	%bias	P value
Age	5.16	4.43	93.90	0.00	4.75	4.43	39.30	0.00	4.43	5.45	145.80	0.00
Age square	27.28	20.21	95.80	0.00	23.22	20.21	40.90	0.00	20.21	30.05	146.50	0.00
Gender	1.35	1.25	23.80	0.01	1.43	1.25	39.10	0.00	1.25	1.30	13.00	0.23
Education	3.27	3.38	-10.80	0.21	3.17	3.38	-20.60	0.11	3.38	3.34	-4.40	0.68
Age*Education	16.52	14.87	28.90	0.00	14.70	14.87	-3.10	0.81	14.87	17.76	50.30	0.00
Education*Gender	4.31	4.11	9.40	0.27	4.43	4.11	14.70	0.24	4.11	4.23	5.60	0.61
Number of Children in Household	0.31	0.94	-63.30	0.00	0.49	0.94	-42.90	0.00	0.94	0.18	-79.20	0.00
Rural	0.10	0.15	-15.00	0.11	0.10	0.15	-17.60	0.21	0.15	0.11	-13.20	0.26
Small Town	0.21	0.21	0.50	0.95	0.30	0.21	20.80	0.09	0.21	0.15	-15.20	0.19
Micropolitan	0.15	0.13	4.60	0.59	0.17	0.13	11.70	0.35	0.13	0.13	-0.60	0.96
Boat Activity	0.43	0.55	-25.50	0.00	0.43	0.55	-24.90	0.06	0.55	0.42	-25.90	0.02
Hunting	0.04	0.06	-11.80	0.21	0.02	0.06	-25.10	0.12	0.06	0.05	-4.50	0.69
Fishing	0.50	0.51	-1.80	0.84	0.49	0.51	-4.00	0.76	0.51	0.51	-0.30	0.98
Total Number of Trips	7.07	7.35	-2.60	0.76	4.70	7.35	-29.90	0.04	7.35	8.70	11.20	0.25
Take overnight Trips	0.37	0.46	-16.60	0.06	0.27	0.46	-39.30	0.00	0.46	0.45	-2.00	0.85
Mean Bias	27.0				24.9				34.5			
Median Bias	15.00				24.9				13.00			

Table A2: Balancing for Matching with all Covariates for Treatment Group Consisting Unemployed and Retired (Group Two)

	Nearest Neighbor without replacement				Nearest Neighbor with replacement				Nearest 5 Neighbors with replacement			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	5.14	5.16	-1.70	0.88	5.14	5.20	-6.70	0.56	5.14	5.15	-1.10	0.93
Age square	27.08	27.16	-1.10	0.93	27.08	27.58	-6.70	0.57	27.08	27.15	-1.00	0.94
Gender	1.35	1.29	14.30	0.22	1.35	1.33	4.30	0.72	1.35	1.36	-1.90	0.87
Education	3.27	3.37	-9.40	0.43	3.27	3.33	-5.60	0.64	3.27	3.40	-11.60	0.33
Age*Education	16.55	17.11	-9.90	0.43	16.55	17.28	-12.70	0.31	16.55	17.26	-12.60	0.32
Education*Gender	4.31	4.22	4.70	0.71	4.31	4.35	-1.90	0.88	4.31	4.58	-12.70	0.32
No of Children in Household	0.31	0.31	0.70	0.94	0.31	0.27	4.60	0.57	0.31	0.31	0.30	0.97
Rural	0.10	0.11	-2.00	0.85	0.10	0.08	5.90	0.56	0.10	0.10	2.00	0.85
Small Town	0.22	0.22	-1.60	0.89	0.22	0.20	4.80	0.67	0.22	0.20	4.00	0.72
Micropolitan	0.15	0.14	1.90	0.87	0.15	0.16	-3.80	0.75	0.15	0.13	6.50	0.57
Boat Activity	0.43	0.50	-13.20	0.25	0.43	0.48	-9.20	0.42	0.43	0.42	1.60	0.89
Hunting	0.04	0.05	-2.90	0.78	0.04	0.04	0.00	1.00	0.04	0.06	-11.20	0.33
Fishing	0.51	0.50	2.60	0.82	0.51	0.43	15.70	0.17	0.51	0.46	9.70	0.40
Total Number of Trips	7.15	7.18	-0.30	0.98	7.15	7.19	-0.40	0.98	7.15	6.14	9.20	0.42
Take overnight Trips	0.38	0.35	5.30	0.64	0.38	0.34	8.00	0.48	0.38	0.34	8.40	0.45
Mean Bias				4.80				6.00				6.30
Median Bias				2.60				5.60				6.50

	Radius Matching (caliper =0.5*SD)				Radius Matching (caliper =0.25*SD)			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
	5.13	5.10	2.90	0.80	5.13	5.07	8.40	0.48
	26.90	26.69	2.80	0.82	26.96	26.36	8.10	0.51
	1.34	1.38	-8.10	0.51	1.34	1.36	-2.30	0.85
	3.29	3.44	-14.20	0.24	3.29	3.36	-6.70	0.58
	16.56	17.31	-13.20	0.30	16.61	16.75	-2.50	0.84
	4.29	4.68	-18.40	0.16	4.32	4.49	-8.20	0.52
	0.32	0.34	-2.40	0.78	0.32	0.40	-8.00	0.38
	0.11	0.10	0.70	0.95	0.11	0.11	-1.30	0.90
	0.22	0.19	6.70	0.56	0.22	0.20	5.70	0.62
	0.15	0.15	-1.50	0.90	0.15	0.15	-2.60	0.83
	0.44	0.43	2.10	0.86	0.44	0.45	-2.50	0.83
	0.04	0.06	-6.90	0.53	0.04	0.05	-5.70	0.60
	0.50	0.47	6.80	0.56	0.50	0.48	5.30	0.64
	5.13	6.85	2.70	0.82	5.13	6.90	2.40	0.83
	26.90	0.38	0.60	0.96	26.96	0.38	0.60	0.96
				6.00				6.00
				2.90				2.90

Table A3: Balancing for Matching with all Covariates for treatment group consisting unemployed (group two)

	Nearest Neighbor without replacement				Nearest Neighbor with replacement				Nearest 5 Neighbors with replacement			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	4.68	4.67	2.10	0.91	4.68	4.68	0.00	1.00	4.68	4.67	2.10	0.91
Age square	22.58	22.47	1.60	0.93	22.58	22.58	0.00	1.00	22.58	22.41	2.30	0.90
Gender	1.40	1.52	-25.00	0.20	1.40	1.52	-25.00	0.20	1.40	1.43	-6.40	0.74
Education	3.23	3.22	1.60	0.93	3.23	3.28	-4.90	0.77	3.23	3.18	5.30	0.78
Age*Education	14.83	14.83	0.00	1.00	14.83	15.30	-8.70	0.65	14.83	14.50	6.10	0.75
Education*Gender	4.45	4.82	-17.00	0.39	4.45	4.88	-20.10	0.29	4.45	4.45	0.00	1.00
No of Children in Household	0.52	0.50	1.60	0.92	0.52	0.48	3.20	0.84	0.52	0.49	2.70	0.86
Rural	0.10	0.10	0.00	1.00	0.10	0.10	0.00	1.00	0.10	0.13	-8.10	0.65
Small Town	0.28	0.22	15.30	0.40	0.28	0.22	15.30	0.40	0.28	0.22	15.00	0.41
Micropolitan	0.17	0.20	-9.20	0.64	0.17	0.23	-18.40	0.37	0.17	0.18	-2.80	0.89
Boat Activity	0.42	0.50	-16.70	0.36	0.42	0.45	-6.70	0.72	0.42	0.44	-4.30	0.81
Hunting	0.02	0.03	-8.50	0.56	0.02	0.03	-8.50	0.56	0.02	0.02	-3.40	0.80
Fishing	0.48	0.48	0.00	1.00	0.48	0.48	0.00	1.00	0.48	0.54	-12.2	0.51
Total Number of Trips	4.78	4.57	2.40	0.87	4.78	5.18	-4.50	0.77	4.78	4.92	-1.50	0.92
Take overnight Trips	0.28	0.33	-10.60	0.56	0.28	0.32	-7.00	0.69	0.28	0.30	-3.10	0.86
Mean Bias	7.40				8.20				5.00			
Median Bias	2.40				6.70				3.40			

	Radius Matching (caliper =0.5*SD)				Radius Matching (caliper =0.25*SD)			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	4.68	4.66	2.80	0.88	4.68	4.66	2.50	0.89
Age square	22.54	22.32	3.00	0.88	22.58	22.40	2.50	0.90
Gender	1.41	1.40	2.50	0.90	1.40	1.39	1.80	0.93
Education	3.24	3.25	-0.90	0.96	3.23	3.26	-2.20	0.91
Age*Education	15.09	14.91	3.20	0.86	14.83	15.01	-3.20	0.87
Education*Gender	4.53	4.43	4.70	0.81	4.45	4.44	0.30	0.99
No of Children in Household	0.53	0.52	0.20	0.99	0.52	0.55	-3.10	0.84
Rural	0.10	0.09	2.40	0.89	0.10	0.10	1.40	0.93
Small Town	0.27	0.28	-2.30	0.90	0.28	0.28	-0.10	1.00
Micropolitan	0.17	0.16	1.30	0.95	0.17	0.16	0.70	0.97
Boat Activity	0.41	0.45	-8.90	0.63	0.42	0.45	-6.10	0.74
Hunting	0.02	0.02	-3.10	0.82	0.02	0.03	-4.50	0.74
Fishing	0.47	0.50	-5.10	0.78	0.48	0.51	-5.00	0.79
Total Number of Trips	4.86	4.62	2.80	0.86	4.78	4.83	-0.60	0.97
Take overnight Trips	0.29	0.30	-3.40	0.85	0.28	0.31	-5.70	0.75
Mean Bias	3.10				2.60			
Median Bias	2.80				2.50			

Table A4: Balancing for Matching with all Covariates for Treatment Group Consisting Retired (group three)

	Nearest Neighbor without replacement				Nearest Neighbor with replacement				Nearest 5 Neighbors with replacement			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	5.42	5.37	6.50	0.61	5.42	5.37	6.50	0.61	5.42	5.38	5.10	0.69
Age square	29.71	29.15	8.40	0.54	29.71	29.15	8.40	0.54	29.71	29.29	6.30	0.66
Gender	1.29	1.31	-5.00	0.75	1.29	1.33	-7.50	0.63	1.29	1.27	4.90	0.75
Education	3.34	3.56	-21.00	0.17	3.34	3.47	-12.60	0.42	3.34	3.52	-17.5	0.27
Age*Education	17.78	18.80	-17.80	0.26	17.78	18.58	-14.10	0.36	17.78	18.34	-9.80	0.56
Education*Gender	4.19	4.56	-17.90	0.26	4.19	4.56	-17.90	0.26	4.19	4.37	-8.70	0.60
No of Children in Household	0.19	0.24	-4.70	0.61	0.19	0.25	-5.90	0.52	0.19	0.24	-4.90	0.61
Rural	0.11	0.06	16.60	0.18	0.11	0.06	16.60	0.18	0.11	0.10	3.40	0.80
Small Town	0.16	0.20	-11.70	0.44	0.16	0.18	-5.80	0.69	0.16	0.17	-4.10	0.78
Micropolitan	0.13	0.12	3.30	0.82	0.13	0.11	6.60	0.65	0.13	0.13	2.50	0.87
Boat Activity	0.44	0.42	4.50	0.76	0.44	0.42	4.50	0.76	0.44	0.43	1.80	0.90
Hunting	0.06	0.06	0.00	1.00	0.06	0.10	-18.90	0.27	0.06	0.05	1.60	0.91
Fishing	0.52	0.49	4.50	0.77	0.52	0.57	-11.20	0.46	0.52	0.52	0.00	1.00
Total Number of Trips	8.84	10.20	-11.30	0.50	8.84	10.72	-15.60	0.36	8.84	9.70	-7.10	0.67
Take overnight Trips	0.45	0.46	-2.30	0.88	0.45	0.49	-9.00	0.55	0.45	0.45	-0.50	0.98
Mean Bias	9				10.7				5.2			
Median Bias	6.5				9				4.9			

	Radius Matching (caliper =0.5*SD)				Radius Matching (caliper =0.25*SD)			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
	5.42	5.38	4.50	0.74	5.42	5.38	4.50	0.74
	29.71	29.39	4.80	0.74	29.71	29.39	4.80	0.74
	1.29	1.28	2.50	0.87	1.29	1.28	2.50	0.87
	3.34	3.49	-14.00	0.37	3.34	3.49	-14.00	0.37
	17.78	18.40	-10.90	0.50	17.78	18.40	-10.90	0.50
	4.19	4.39	-9.40	0.56	4.19	4.39	-9.40	0.56
	0.19	0.22	-3.00	0.75	0.19	0.22	-3.00	0.75
	0.11	0.12	-0.80	0.96	0.11	0.12	-0.80	0.96
	0.16	0.18	-5.90	0.69	0.16	0.18	-5.90	0.69
	0.13	0.12	2.90	0.84	0.13	0.12	2.90	0.84
	0.44	0.44	0.30	0.98	0.44	0.44	0.30	0.98
	0.06	0.06	-3.60	0.81	0.06	0.06	-3.60	0.81
	0.52	0.51	1.50	0.92	0.52	0.51	1.50	0.92
	8.84	9.28	-3.60	0.83	8.84	9.28	-3.60	0.83
	0.45	0.49	-7.80	0.61	0.45	0.49	-7.80	0.61
	5				6.1			
	3.6				4.5			

Table B1: Placebo Exercise-Difference between Placebo Treatment and Control groups Before Matching

	Group 1(Unemployed and Retired)				Group 2(Unemployed)				Group 3(Retired)			
Variable	Mean Treated	Mean Control	%bias	*p value*	Mean Treated	Mean Control	%bias	*p value*	Mean Control	Mean Treated	%bias	*p value*
Age	5.16	4.44	93.10	0.00	4.75	4.44	38.40	0.00	5.45	4.44	145.10	0.00
Age square	27.28	20.27	95.00	0.00	23.22	20.27	40.00	0.00	30.05	20.27	145.60	0.00
Gender	1.35	1.25	23.30	0.01	1.43	1.25	38.60	0.00	1.30	1.25	12.50	0.24
Education	3.27	3.39	-11.4	0.18	3.17	3.39	-21.3	0.10	3.34	3.39	-5.00	0.63
Age*Education	16.52	14.93	27.80	0.00	14.70	14.93	-4.20	0.75	17.76	14.93	49.20	0.00
Education*Gender	4.31	4.12	9.00	0.29	4.43	4.12	14.40	0.25	4.23	4.12	5.20	0.63
No of Children in Household	0.31	0.93	-63.0	0.00	0.49	0.93	-42.6	0.00	0.18	0.93	-78.90	0.00
Rural	0.10	0.15	-14.9	0.11	0.10	0.15	-17.5	0.22	0.11	0.15	-13.20	0.26
Small Town	0.21	0.21	0.50	0.95	0.30	0.21	20.80	0.09	0.15	0.21	-15.20	0.19
Micropolitan	0.15	0.13	5.90	0.49	0.17	0.13	13.00	0.29	0.13	0.13	0.70	0.95
Boat Activity	0.43	0.55	-24.9	0.01	0.43	0.55	-24.3	0.06	0.42	0.55	-25.30	0.02
Hunting	0.04	0.06	-9.40	0.32	0.02	0.06	-22.8	0.15	0.05	0.06	-2.00	0.86
Fishing	0.50	0.51	-0.50	0.96	0.49	0.51	-2.70	0.84	0.51	0.51	1.00	0.93
Total Number of Trips	7.07	6.82	2.40	0.77	4.70	6.82	-25.0	0.08	8.70	6.82	16.00	0.09
Take overnight Trips	0.37	0.45	-15.5	0.08	0.27	0.45	-38.2	0.01	0.45	0.45	-1.00	0.93
Mean Bias	26.4				24.3				34.4			
Median Bias	14.9				22.8				13.2			

Table B2: Placebo Exercise-Balancing for Matching with all Covariates for Treatment Group 1(Unemployed and Retired)

	Nearest Neighbor without replacement				Nearest Neighbor with replacement				Nearest 5 Neighbors with replacement			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	5.14	5.14	0.00	1.00	5.14	5.16	-1.70	0.88	5.14	5.14	5.7	0.94
Age square	27.08	27.04	0.50	0.97	27.08	27.18	-1.40	0.91	27.08	26.99	6.1	0.92
Gender	1.35	1.30	11.40	0.33	1.35	1.40	-10.00	0.41	1.35	1.39	-1.9	0.56
Education	3.27	3.42	-14.40	0.23	3.27	3.44	-15.60	0.19	3.27	3.40	5.2	0.30
Age*Education	16.55	17.28	-12.80	0.31	16.55	17.55	-17.50	0.15	16.55	17.15	8.2	0.40
Education*Gender	4.31	4.44	-5.90	0.63	4.31	4.86	-25.80	0.05	4.31	4.66	5.5	0.20
No of Children in Household	0.31	0.31	0.00	1.00	0.31	0.29	2.00	0.81	0.31	0.31	1.4	0.98
Rural	0.10	0.10	2.00	0.85	0.10	0.11	-2.00	0.85	0.10	0.11	-2	0.97
Small Town	0.22	0.22	-1.60	0.89	0.22	0.18	8.00	0.48	0.22	0.19	6	0.66
Micropolitan	0.15	0.16	-1.90	0.88	0.15	0.16	-3.80	0.75	0.15	0.14	-1.9	0.83
Boat Activity	0.43	0.48	-10.5	0.36	0.43	0.44	-1.30	0.91	0.43	0.42	-5.8	0.86
Hunting	0.04	0.05	-6.10	0.59	0.04	0.07	-12.1	0.31	0.04	0.06	-1.8	0.38
Fishing	0.51	0.45	11.70	0.31	0.51	0.44	14.40	0.21	0.51	0.48	-10	0.55
Total Number of Trips	7.15	6.73	3.90	0.73	7.15	7.52	-3.50	0.79	7.15	6.57	1.6	0.64
Take overnight Trips	0.38	0.39	-1.30	0.91	0.38	0.41	-5.30	0.64	0.38	0.36	-0.3	0.73
Mean Bias				5.6				8.3				5.7
Median Bias				3.9				5.3				5.1

	Radius Matching (caliper =0.5*SD)				Radius Matching (caliper =0.25*SD)			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
	5.14	5.09	7.50	0.53	5.13	5.10	3.30	0.78
	27.08	26.56	7.00	0.57	26.90	26.67	3.20	0.80
	1.35	1.35	-0.30	0.98	1.34	1.38	-8.70	0.48
	3.27	3.37	-9.30	0.43	3.29	3.43	-14.00	0.24
	16.55	16.89	-5.90	0.64	16.56	17.28	-12.60	0.32
	4.31	4.50	-8.80	0.48	4.29	4.69	-18.90	0.15
	0.31	0.39	-7.80	0.38	0.32	0.35	-2.60	0.77
	0.10	0.11	-2.40	0.82	0.11	0.11	-0.40	0.97
	0.22	0.20	3.20	0.78	0.22	0.19	6.70	0.56
	0.15	0.16	-3.00	0.80	0.15	0.15	-1.80	0.88
	0.43	0.45	-4.10	0.72	0.44	0.43	2.20	0.85
	0.04	0.05	-4.70	0.67	0.04	0.05	-5.80	0.61
	0.51	0.48	5.10	0.66	0.50	0.47	6.70	0.56
	7.15	7.10	0.50	0.97	7.14	6.93	2.00	0.87
	0.38	0.38	0.10	0.99	0.38	0.37	1.60	0.89
				6				4.6
				3.3				4.7

Table B3: Placebo Exercise-Balancing for Matching with all Covariates for Treatment Group Consisting Unemployed (Group 2)

	Nearest Neighbor without replacement				Nearest Neighbor with replacement				Nearest 5 Neighbors with replacement			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	4.68	4.67	2.10	0.91	4.68	4.68	0.00	1.00	4.68	4.64	5.70	0.76
Age square	22.58	22.47	1.60	0.93	22.58	22.62	-0.50	0.98	22.58	22.13	6.10	0.74
Gender	1.40	1.42	-3.60	0.85	1.40	1.37	7.10	0.71	1.40	1.41	-1.90	0.92
Education	3.23	3.12	11.50	0.52	3.23	3.15	8.20	0.65	3.23	3.18	5.20	0.78
Age*Education	14.83	14.60	4.30	0.82	14.83	14.80	0.60	0.97	14.83	14.39	8.20	0.67
Education*Gender	4.45	4.43	0.80	0.97	4.45	4.35	4.70	0.81	4.45	4.33	5.50	0.78
No of Children in Household	0.52	0.68	-16.0	0.35	0.52	0.70	-17.60	0.30	0.52	0.50	1.40	0.93
Rural	0.10	0.08	5.10	0.75	0.10	0.08	5.10	0.75	0.10	0.11	-2.00	0.91
Small Town	0.28	0.25	7.60	0.68	0.28	0.27	3.80	0.84	0.28	0.26	6.00	0.75
Micropolitan	0.17	0.23	-18.6	0.37	0.17	0.23	-18.60	0.37	0.17	0.17	-1.90	0.92
Boat Activity	0.42	0.48	-13.4	0.47	0.42	0.45	-6.70	0.72	0.42	0.45	-5.80	0.75
Hunting	0.02	0.00	8.80	0.32	0.02	0.00	8.80	0.32	0.02	0.02	-1.80	0.89
Fishing	0.48	0.50	-3.30	0.86	0.48	0.50	-3.30	0.86	0.48	0.53	-10.00	0.59
Total Number of Trips	4.78	4.42	4.30	0.79	4.78	4.40	4.50	0.79	4.78	4.65	1.60	0.92
Take overnight Trips	0.28	0.27	3.50	0.84	0.28	0.23	10.60	0.54	0.28	0.28	-0.30	0.99
Mean Bias	7				6.7				4.2			
Median Bias	4.3				5.1				5.2			

	Radius Matching (caliper =0.5*SD)				Radius Matching (caliper =0.25*SD)			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	4.68	4.65	3.20	0.87	4.68	4.66	3.10	0.87
Age square	22.54	22.29	3.40	0.86	22.58	22.36	3.10	0.87
Gender	1.41	1.40	1.60	0.93	1.40	1.40	0.50	0.98
Education	3.24	3.26	-1.80	0.92	3.23	3.25	-1.30	0.95
Age*Education	15.09	14.94	2.60	0.89	14.83	14.94	-2.00	0.92
Education*Gender	4.53	4.46	3.30	0.87	4.45	4.44	0.30	0.99
No of Children in Household	0.53	0.53	-0.40	0.98	0.52	0.54	-2.70	0.87
Rural	0.10	0.09	2.40	0.89	0.10	0.09	1.70	0.92
Small Town	0.27	0.28	-2.80	0.88	0.28	0.29	-0.50	0.98
Micropolitan	0.17	0.17	0.90	0.96	0.17	0.16	1.90	0.92
Boat Activity	0.41	0.45	-9.00	0.63	0.42	0.45	-5.80	0.75
Hunting	0.02	0.02	-3.20	0.82	0.02	0.02	-4.40	0.75
Fishing	0.47	0.50	-5.90	0.75	0.48	0.50	-4.00	0.83
Total Number of Trips	4.86	4.61	3.00	0.86	4.78	4.78	0.00	1.00
Take overnight Trips	0.29	0.30	-3.50	0.85	0.28	0.31	-5.70	0.75
Mean Bias	2.5				3.1			
Median Bias	2				3			

Table B4: Placebo Exercise-Balancing for Matching with all Covariates for Treatment Group Consisting Unemployed (Group Two)

	Nearest Neighbor without replacement				Nearest Neighbor with replacement				Nearest 5 Neighbors with replacement			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	5.41	5.41	0.00	1.00	5.41	5.43	-3.30	0.80	5.41	5.39	3.20	0.81
Age square	29.64	29.55	1.40	0.92	29.64	29.80	-2.40	0.86	29.64	29.39	3.70	0.80
Gender	1.30	1.28	2.50	0.87	1.30	1.31	-2.50	0.87	1.30	1.26	8.10	0.60
Education	3.33	3.47	-12.80	0.40	3.33	3.43	-9.60	0.52	3.33	3.51	-17.0	0.29
Age*Education	17.71	18.42	-12.40	0.42	17.71	18.53	-14.4	0.31	17.71	18.47	-13.2	0.42
Education*Gender	4.19	4.38	-8.80	0.59	4.19	4.45	-12.6	0.42	4.19	4.41	-10.3	0.54
No of Children in Household	0.19	0.24	-4.80	0.64	0.19	0.24	-4.80	0.66	0.19	0.21	-1.30	0.89
Rural	0.11	0.11	0.00	1.00	0.11	0.14	-6.70	0.65	0.11	0.10	3.80	0.79
Small Town	0.16	0.14	5.90	0.67	0.16	0.13	8.80	0.52	0.16	0.18	-5.30	0.72
Micropolitan	0.14	0.15	-3.40	0.83	0.14	0.14	0.00	1.00	0.14	0.12	4.70	0.75
Boat Activity	0.43	0.41	4.60	0.76	0.43	0.36	13.70	0.36	0.43	0.43	-0.40	0.98
Hunting	0.06	0.03	9.80	0.47	0.06	0.01	19.60	0.10	0.06	0.06	-2.20	0.89
Fishing	0.51	0.48	6.80	0.65	0.51	0.44	13.60	0.37	0.51	0.49	3.70	0.81
Total Number of Trips	8.44	7.19	10.70	0.49	8.44	8.22	1.90	0.91	8.44	8.52	-0.60	0.97
Take overnight Trips	0.44	0.41	6.80	0.65	0.44	0.39	11.40	0.45	0.44	0.46	-4.00	0.79
Mean Bias	6.00				8.40				5.40			
Median Bias	5.90				8.80				3.80			

	Radius Matching (caliper =0.5*SD)				Radius Matching (caliper =0.25*SD)			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
	5.41	5.33	10.70	0.45	5.41	5.38	4.10	0.76
	29.64	28.93	10.50	0.49	29.64	29.34	4.40	0.77
	1.30	1.27	5.50	0.72	1.30	1.27	5.30	0.73
	3.33	3.49	-14.90	0.34	3.33	3.51	-16.6	0.29
	17.71	18.22	-8.90	0.59	17.71	18.46	-13.1	0.42
	4.19	4.34	-6.90	0.67	4.19	4.38	-8.80	0.59
	0.19	0.26	-7.40	0.47	0.19	0.22	-3.30	0.73
	0.11	0.11	-0.40	0.98	0.11	0.10	2.80	0.84
	0.16	0.18	-4.50	0.76	0.16	0.18	-4.20	0.78
	0.14	0.13	2.90	0.85	0.14	0.13	1.80	0.91
	0.43	0.46	-6.30	0.68	0.43	0.44	-1.60	0.91
	0.06	0.06	0.30	0.98	0.06	0.06	-2.50	0.87
	0.51	0.49	4.30	0.78	0.51	0.50	2.10	0.89
	8.44	8.77	-2.80	0.87	8.44	8.55	-0.90	0.96
	0.44	0.46	-4.30	0.78	0.44	0.47	-4.90	0.74
	6				5.1			
	5.5				4.1			

Table C1: Balancing for Matching with all Covariates for Treatment Group Consisting Unemployed and Retired (Group One)

	Nearest Neighbor without replacement				Nearest Neighbor with replacement				Nearest 5 Neighbors with replacement				Radius Matching (caliper =0.5*SD)				Radius Matching (caliper =0.25*SD)			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	5.14	5.16	-1.70	0.88	5.14	5.16	-2.50	0.83	5.14	5.16	-2.50	0.83	5.14	5.09	6.90	0.56	5.14	5.12	3.50	0.77
Age square	27.08	27.17	-1.20	0.92	27.08	27.24	-2.20	0.85	27.08	27.22	-2.00	0.87	27.08	26.60	6.50	0.60	27.08	26.83	3.40	0.78
Gender	1.35	1.30	11.50	0.33	1.35	1.35	1.40	0.91	1.35	1.34	2.90	0.81	1.35	1.34	3.80	0.75	1.35	1.36	-0.90	0.94
Education	3.27	3.54	-25.0	0.04	3.27	3.37	-9.40	0.42	3.27	3.43	-14.9	0.20	3.27	3.35	-7.10	0.55	3.27	3.42	-14.1	0.23
Age*Education	16.55	17.87	-23.2	0.06	16.55	17.12	-10.1	0.38	16.55	17.60	-18.4	0.13	16.55	16.75	-3.50	0.78	16.55	17.26	-12.5	0.32
Education*Gender	4.31	4.52	-9.60	0.43	4.31	4.47	-7.50	0.53	4.31	4.54	-10.8	0.37	4.31	4.40	-3.90	0.75	4.31	4.57	-11.9	0.34
No of Children in Household	0.31	0.38	-6.60	0.45	0.31	0.34	-2.60	0.76	0.31	0.29	2.50	0.76	0.31	0.41	-9.60	0.29	0.31	0.34	-2.80	0.74
Rural	0.10	0.12	-3.90	0.72	0.10	0.11	-2.00	0.85	0.10	0.09	4.20	0.68	0.10	0.11	-2.70	0.80	0.10	0.09	5.50	0.59
Small Town	0.22	0.23	-3.20	0.78	0.22	0.20	3.20	0.78	0.22	0.20	3.80	0.74	0.22	0.20	2.60	0.82	0.22	0.23	-2.60	0.83
Micropolitan	0.15	0.14	3.80	0.75	0.15	0.13	5.60	0.62	0.15	0.12	7.40	0.52	0.15	0.15	1.20	0.92	0.15	0.14	2.00	0.86
Mean Bias	9.00				4.70				6.90				48.00				5.90			
Median Bias	5.20				2.90				4.00				3.90				3.40			

Table C2: Balancing for Matching with all Covariates for Treatment Group Consisting Unemployed (Group Two)

	Nearest Neighbor without replacement				Nearest Neighbor with replacement				Nearest 5 Neighbors with replacement			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	4.726	4.839	-14.20	0.447	4.726	4.839	-14.20	0.447	4.726	4.833	-13.40	0.456
Age square	23.016	24.065	-14.20	0.458	23.016	24.065	-14.20	0.458	23.016	23.928	-12.40	0.506
Gender	1.419	1.403	3.500	0.857	1.419	1.403	3.500	0.857	1.419	1.353	14.30	0.450
Education	3.210	3.097	11.10	0.527	3.210	3.097	11.10	0.527	3.210	3.081	12.70	0.486
Age*Education	14.839	14.871	-0.600	0.972	14.839	14.871	-0.60	0.972	14.839	14.526	5.80	0.757
Education*Gender	4.468	4.274	9.000	0.624	4.468	4.274	9.00	0.624	4.468	4.004	21.60	0.253
No of Children in Household	0.500	0.500	0.000	1.000	0.500	0.500	0.00	1.00	0.500	0.457	4.20	0.771
Rural	0.097	0.129	-9.800	0.574	0.097	0.129	-9.80	0.574	0.097	0.103	-1.80	0.911
Small Town	0.290	0.274	3.700	0.843	0.290	0.274	3.70	0.843	0.290	0.297	-1.60	0.934
Micropolitan	0.177	0.145	8.900	0.629	0.177	0.145	8.90	0.629	0.177	0.133	12.40	0.496
Mean Bias	7.50				7.50				10.00			
Median Bias	9.00				9.00				12.40			

	Radius Matching (caliper =0.5*SD)				Radius Matching (caliper =0.25*SD)			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	4.726	4.710	1.900	0.918	4.726	4.748	-2.800	0.884
Age square	23.016	22.862	2.100	0.913	23.016	23.245	-3.100	0.872
Gender	1.419	1.405	3.200	0.868	1.419	1.402	3.700	0.845
Education	3.210	3.202	0.700	0.968	3.210	3.214	-0.400	0.981
Age*Education	14.839	14.760	1.500	0.936	14.839	15.029	-3.500	0.843
Education*Gender	4.468	4.368	4.600	0.808	4.468	4.426	1.900	0.919
No of Children in Household	0.500	0.538	-3.700	0.812	0.500	0.496	0.300	0.982
Rural	0.097	0.101	-1.400	0.934	0.097	0.093	1.200	0.941
Small Town	0.290	0.280	2.300	0.904	0.290	0.292	-0.400	0.984
Micropolitan	0.177	0.172	1.600	0.934	0.177	0.170	2.200	0.909
Mean Bias	2.3				2.0			
Median Bias	2.0				2.1			

Table C3: Balancing for Matching with all Covariates for Treatment Group Consisting Retired (Group Three)

	Nearest Neighbor without replacement				Nearest Neighbor with replacement				Nearest 5 Neighbors with replacement			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	5.422	5.422	0.000	1.000	5.422	5.422	0.000	1.000	5.422	5.423	-0.100	0.994
Age square	29.778	29.778	0.000	1.000	29.778	29.778	0.000	1.000	29.778	29.777	0.000	0.999
Gender	1.300	1.244	12.400	0.405	1.300	1.256	9.900	0.508	1.300	1.244	12.500	0.401
Education	3.344	3.578	-21.80	0.161	3.344	3.489	-13.50	0.385	3.344	3.500	-14.60	0.364
Age*Education	17.844	19.200	-23.50	0.132	17.844	18.733	-15.40	0.307	17.844	18.725	-15.30	0.347
Education*Gender	4.233	4.378	-7.000	0.636	4.233	4.300	-3.200	0.825	4.233	4.294	-2.900	0.847
No of Children in Household	0.189	0.200	-1.200	0.903	0.189	0.189	0.000	1.000	0.189	0.175	1.500	0.875
Rural	0.111	0.111	0.000	1.000	0.111	0.133	-6.600	0.651	0.111	0.141	-9.000	0.544
Small Town	0.156	0.156	0.000	1.000	0.156	0.111	11.500	0.383	0.156	0.125	8.000	0.552
Micropolitan	0.133	0.133	0.000	1.000	0.133	0.111	6.600	0.651	0.133	0.115	5.500	0.709
Mean Bias	6.600				6.700				6.900			
Median Bias	0.600				6.600				6.000			

	Radius Matching (caliper =0.5*SD)				Radius Matching (caliper =0.25*SD)			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	5.422	5.380	6.100	0.661	5.422	5.433	-1.60	0.904
Age square	29.778	29.391	5.800	0.698	29.778	29.898	-1.80	0.900
Gender	1.300	1.273	6.200	0.685	1.300	1.271	6.40	0.671
Education	3.344	3.429	-7.900	0.608	3.344	3.430	-8.00	0.608
Age*Education	17.844	18.242	-6.900	0.665	17.844	18.495	-11.30	0.478
Education*Gender	4.233	4.285	-2.500	0.873	4.233	4.272	-1.90	0.904
No of Children in Household	0.189	0.246	-6.000	0.543	0.189	0.222	-3.50	0.718
Rural	0.111	0.112	-0.200	0.990	0.111	0.121	-2.80	0.844
Small Town	0.156	0.176	-5.300	0.714	0.156	0.162	-1.70	0.906
Micropolitan	0.133	0.142	-2.500	0.869	0.133	0.130	1.10	0.943
Mean Bias	4.900				4.000			
Median Bias	5.900				2.300			

Table D1: Balancing of Covariates for Matching within Rural, Urban, Micropolitan and Metropolitan Cell (Group 1: Unemployed and Retired)

	Nearest Neighbor without replacement				Nearest Neighbor with replacement				Nearest 5 Neighbors with replacement			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	5.16	5.12	5.00	0.66	5.16	5.16	0.00	1.00	5.16	5.16	-0.10	1.00
Age square	27.28	26.77	6.80	0.56	27.28	27.27	0.20	0.99	27.28	27.27	0.10	0.99
Gender	1.35	1.30	12.70	0.28	1.35	1.34	4.20	0.72	1.35	1.35	1.60	0.89
Education	3.27	3.48	-19.70	0.09	3.27	3.32	-4.90	0.68	3.27	3.50	-22.1	0.06
Age*Education	16.52	17.47	-16.60	0.18	16.52	16.50	0.30	0.98	16.52	17.91	-24.4	0.05
Education*Gender	4.31	4.37	-3.10	0.81	4.31	4.22	4.30	0.73	4.31	4.67	-17.3	0.18
No of Children in Household	0.31	0.31	0.00	1.00	0.31	0.30	0.60	0.94	0.31	0.34	-2.60	0.76
Boat Activity	0.43	0.51	-16.90	0.14	0.43	0.43	-1.30	0.91	0.43	0.45	-4.90	0.67
Hunting	0.04	0.03	2.90	0.76	0.04	0.01	11.60	0.15	0.04	0.06	-8.70	0.43
Fishing	0.50	0.49	2.60	0.82	0.50	0.39	21.90	0.05	0.50	0.48	4.30	0.71
Total Number of Trips	7.07	7.07	0.00	1.00	7.07	6.94	1.20	0.91	7.07	6.82	2.30	0.84
Take overnight Trips	0.37	0.39	-2.60	0.82	0.37	0.31	13.10	0.23	0.37	0.39	-2.30	0.84
Mean Bias	7.40				5.30				7.60			
Median Bias	4.00				2.80				3.50			

	Radius Matching (caliper =0.5*SD)				Radius Matching (caliper =0.25*SD)			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	5.11	5.06	5.70	0.64	5.08	5.07	1.50	0.90
Age square	26.71	26.34	5.10	0.68	26.39	26.31	1.10	0.93
Gender	1.33	1.31	5.90	0.62	1.33	1.35	-3.30	0.79
Education	3.29	3.39	-9.00	0.45	3.33	3.46	-11.9	0.33
Age*Education	16.53	16.95	-7.30	0.56	16.65	17.35	-12.3	0.33
Education*Gender	4.28	4.37	-4.50	0.72	4.33	4.60	-12.8	0.33
No of Children in Household	0.33	0.39	-6.80	0.46	0.34	0.37	-3.00	0.75
Boat Activity	0.44	0.46	-3.50	0.77	0.44	0.44	-0.20	0.99
Hunting	0.04	0.05	-6.00	0.59	0.04	0.05	-5.40	0.64
Fishing	0.50	0.48	3.80	0.75	0.50	0.47	6.50	0.59
Total Number of Trips	7.22	7.13	0.80	0.95	7.15	6.79	3.30	0.77
Take overnight Trips	0.39	0.39	-1.20	0.92	0.38	0.38	0.60	0.96
Mean Bias	5.00				5.20			
Median Bias	5.40				3.30			

Table D2: Balancing of Covariates for Matching within Rural, Urban, Micropolitan and Metropolitan Cell (Group Two: Unemployed)

	Nearest Neighbor without replacement				Nearest Neighbor with replacement				Nearest 5 Neighbors with replacement			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	4.700	4.733	-4.200	0.823	4.700	4.733	-4.200	0.818	4.700	4.710	-1.200	0.949
Age square	22.767	23.033	-3.600	0.849	22.767	22.967	-2.700	0.884	22.767	22.794	-0.400	0.984
Gender	1.417	1.417	0.000	1.000	1.417	1.400	3.600	0.854	1.417	1.400	3.500	0.858
Education	3.233	3.217	1.600	0.928	3.233	3.150	8.200	0.649	3.233	3.208	2.500	0.893
Age*Education	15.133	15.017	2.200	0.906	15.133	14.750	7.100	0.696	15.133	14.740	7.300	0.694
Education*Gender	4.550	4.400	7.000	0.710	4.550	4.250	13.900	0.447	4.550	4.323	10.600	0.595
No of Children in Household	0.517	0.450	6.400	0.676	0.517	0.417	9.600	0.531	0.517	0.450	6.400	0.675
Boat Activity	0.417	0.383	6.700	0.712	0.417	0.383	6.700	0.712	0.417	0.431	-2.900	0.873
Hunting	0.017	0.000	8.500	0.319	0.017	0.000	8.500	0.319	0.017	0.020	-1.700	0.893
Fishing	0.483	0.417	13.300	0.467	0.483	0.417	13.300	0.467	0.483	0.482	0.200	0.990
Total Number of Trips	4.783	5.600	-9.200	0.599	4.783	5.567	-8.800	0.617	4.783	5.254	-5.300	0.738
Take overnight Trips	0.283	0.300	-3.500	0.842	0.283	0.267	3.500	0.840	0.283	0.317	-7.000	0.693
Mean Bias	5.500				7.500				4.100			
Median Bias	5.300				7.700				3.200			

Radius Matching (caliper =0.5*SD)				Radius Matching (caliper =0.25*SD)			
Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
4.655	4.672	-2.100	0.913	4.700	4.694	0.800	0.968
22.310	22.466	-2.100	0.912	22.767	22.705	0.800	0.966
1.397	1.393	0.700	0.972	1.417	1.378	8.200	0.671
3.241	3.232	0.900	0.960	3.233	3.210	2.300	0.899
15.034	14.891	2.700	0.885	15.133	14.759	6.900	0.701
4.500	4.392	5.000	0.795	4.550	4.301	11.600	0.542
0.534	0.551	-1.600	0.920	0.517	0.567	-4.800	0.763
0.431	0.485	-10.800	0.566	0.417	0.465	-9.700	0.596
0.017	0.020	-1.400	0.912	0.017	0.026	-4.700	0.730
0.500	0.496	0.900	0.963	0.483	0.480	0.600	0.974
4.828	4.853	-0.300	0.985	4.783	5.045	-2.900	0.850
0.293	0.310	-3.600	0.844	0.283	0.320	-7.700	0.666
2.700				5.100			
1.800				4.700			

Table D3: Balancing of Covariates for Matching within Region-Rural, Urban, Micropolitan and Metropolitan (Group Three: Retired)

	Nearest Neighbor without replacement				Nearest Neighbor with replacement				Nearest 5 Neighbors with replacement			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	5.44	5.32	17.40	0.19	5.44	5.42	3.20	0.81	5.44	5.42	3.50	0.80
Age square	29.99	28.64	20.10	0.16	29.99	29.73	3.90	0.79	29.99	29.70	4.40	0.76
Gender	1.30	1.31	-2.50	0.87	1.30	1.33	-7.40	0.63	1.30	1.31	-3.10	0.84
Education	3.35	3.70	-32.90	0.03	3.35	3.44	-8.20	0.59	3.35	3.52	-16.00	0.30
Age*Education	17.96	19.59	-28.40	0.06	17.96	18.42	-8.00	0.59	17.96	18.88	-16.10	0.28
Education*Gender	4.23	4.54	-14.80	0.36	4.23	4.58	-16.90	0.28	4.23	4.57	-16.20	0.30
No of Children in Household	0.19	0.37	-19.60	0.07	0.19	0.18	1.20	0.89	0.19	0.15	3.50	0.67
Boat Activity	0.43	0.51	-16.70	0.26	0.43	0.55	-24.30	0.10	0.43	0.48	-10.00	0.50
Hunting	0.05	0.06	-3.60	0.81	0.05	0.13	-31.10	0.09	0.05	0.07	-7.70	0.62
Fishing	0.52	0.39	25.80	0.08	0.52	0.46	10.70	0.47	0.52	0.38	27.20	0.07
Total Number of Trips	8.79	7.16	13.50	0.37	8.79	6.95	15.30	0.28	8.79	6.66	17.70	0.23
Take overnight Trips	0.45	0.44	1.90	0.90	0.45	0.38	14.00	0.34	0.45	0.39	11.90	0.42
Mean Bias	16.40				12.00				11.40			
Median Bias	17.10				9.50				11.00			

	Radius Matching (caliper =0.5*SD)				Radius Matching (caliper =0.25*SD)			
	Mean Treated	Mean Control	%bias	P value	Mean Treated	Mean Control	%bias	P value
Age	5.43	5.36	11.20	0.44	5.41	5.39	3.50	0.81
Age square	29.92	29.17	11.20	0.47	29.71	29.47	3.70	0.81
Gender	1.29	1.30	-2.20	0.89	1.29	1.30	-3.10	0.84
Education	3.37	3.37	-0.40	0.98	3.36	3.42	-6.30	0.69
Age*Education	18.02	17.82	3.50	0.82	17.89	18.25	-6.30	0.68
Education*Gender	4.23	4.26	-1.50	0.92	4.21	4.31	-4.90	0.75
No of Children in Household	0.19	0.25	-6.00	0.54	0.20	0.21	-1.40	0.88
Boat Activity	0.43	0.51	-16.40	0.28	0.45	0.53	-16.60	0.28
Hunting	0.06	0.07	-5.60	0.71	0.06	0.08	-8.50	0.60
Fishing	0.51	0.43	17.10	0.25	0.53	0.40	25.00	0.10
Total Number of Trips	8.77	6.78	16.50	0.27	9.06	6.57	20.60	0.17
Take overnight Trips	0.46	0.38	14.40	0.33	0.47	0.40	14.00	0.36
Mean Bias	8.80				9.50			
Median Bias	8.60				6.30			

Table E1: Validity of Common Support Condition- Number of Unmatched and Matched Treatments

	Unemployed and Retired (Group 1)			Unemployed (Group 2)			Retired (Group 3)		
	(a) Matching with all Covariates included								
	Unmatched	Matched	Total	Unmatched	Matched	Total	Unmatched	Matched	Total
Nearest Neighbor without replacement	2	153	155	3	60	63	3	89	92
Nearest Neighbor with replacement	2	153	155	3	60	63	3	89	92
Nearest 5 Neighbors with replacement	2	153	155	3	60	63	3	89	92
Radius Matching (caliper =0.5*SD)	4	151	155	3	60	63	3	89	92
Radius Matching (caliper =0.25*SD)	5	150	155	4	59	63	3	89	92
	(b) Matching with a subset of Covariates included								
	Unmatched	matched	Total	Unmatched	matched	Total	Unmatched	matched	Total
Nearest Neighbor without replacement	2	153	155	1	62	63	2	90	92
Nearest Neighbor with replacement	2	153	155	1	62	63	2	90	92
Nearest 5 Neighbors with replacement	2	153	155	1	62	63	2	90	92
Radius Matching (caliper =0.5*SD)	2	153	155	1	62	63	2	90	92
Radius Matching (caliper =0.25*SD)	2	153	155	1	62	63	2	90	92
	(c) Matching within Rural, Urban, Micropolitan and Metropolitan Cell								
	Unmatched	matched	Total	Unmatched	matched	Total	Unmatched	matched	Total
Nearest Neighbor without replacement	0	155	155	3	60	63	1	91	92
Nearest Neighbor with replacement	0	155	155	3	60	63	1	91	92
Nearest 5 Neighbors with replacement	0	155	155	3	60	63	1	91	92
Radius Matching (caliper =0.5*SD)	8	147	155	4	59	63	2	90	92
Radius Matching (caliper =0.25*SD)	13	142	155	6	57	63	5	87	92

CHAPTER 3. DOES JOB LOSS DURING RECESSION AFFECT RECREATIONEXPENDITURE? AN INVESTIGATION INTO PANEL STUDY OF INCOME DYNAMICS DATA

1. Introduction

The 2008–2009 recession affected individual economic well-being through job loss, stock market crashes, and falling real estate prices, which generated low consumer confidence. Job loss almost always causes household income to shrink, and in response to reduced income, spending on normal goods is supposed to fall. However, unemployment also reduces the opportunity cost of time, and, for some household members allows more time to spend on time-intensive activities, such as recreation. This opposing effects of lower income and cheaper time motivates the research question of how households experiencing a job loss during a recession change expenditures on time-intensive trips and recreation.

The association between recessionary job loss and recreation expenditure is a relatively untapped area. This paper is one of the few to investigate the impact of unemployment and retirement during a recession on household recreation, and is the first in exploiting a national longitudinal dataset to address this research question. To the best of our knowledge, this is also the first endeavor in framing the recession as a quasi-experimental setting and applying treatment effect framework to study households' counterfactual behaviors. A number of studies examined the impact of great recession on physical and mental health **(**Ruhm 2000; 2005; Dehejia and Lleras-Muney 2004; Currie and Schwandt 2014; Currie and Tekin 2014**)**, where, in most cases the authors captured the intensity of recession through variation in aggregate unemployment over time and across states, metropolitan statistical areas, or county. Only recently have studies begun to address for individual heterogeneity **(**McInerney, Mellor, and Nicholas 2013; Marcus 2013;

Currie, Duque, and Garfinkel 2015**).** Although outdoor recreation is an important part of Americans' daily lives and the industry is comparable to that of the automobile industry in size (OIF 2010; Loomis and Keske 2012), similar studies focusing on recreation are not common.

Our economics literature search suggests that Loomis and Keske (2012) is the only study that explored the relationship between recession and recreation. Exploiting two intercept surveys conducted in 2006 and 2009 on Quandary Peak, they did not find any significant impact of recession on total visits, travel expenditure, and willingness to pay for visits. However, since the respondent groups studied before and after the recession are different in their study, it is not clear whether the survey respondents experienced any employment or wealth shock during the recessionary period. In a companion paper where we utilized "Iowa Lakes Project" panel data—a rich dataset containing detailed information on household demographics, employment, and recreation expenditure both before and after the recession—we found that Iowans who became unemployed or retired during the recession did not reduce outdoor recreation at lake sites during the recession of 2009. Since the study was based on data restricted to Iowa and frequency of trips, this study will utilize a nationally representative Panel Study of Income Dynamics (PSID) dataset to examine the question in a broader context. Instead of trip frequency and local recreation, we focus on total household expenditure on recreation and trips. Facing a job loss, a household might choose a *stay-at-home* option (Egan, Herriges, and Kling 2009) and substitute their expensive recreation for relatively cheaper options including local outdoor recreation. This hypothesis would be better captured through examining households' recreation expenditures.

The PSID data contains detailed household-level information on broad expenditure items including trips exclusively for recreation purposes, employment status, income, wealth, and a rich set of socioeconomic factors. Being longitudinal data, it also allows observation of the same households' recreation expenditures and employment statuses both before and during the recession. We utilize four rounds of PSID surveys spanning the period 2004–2010. Although, officially, the recession ends in June of 2009, for our empirical design, we consider both 2008 and 2010 as recessionary periods since aggregate unemployment was high throughout 2010 and 2011. Note that aggregate national unemployment rate was less than 5% at the beginning of recession, reached at its peak at 10% in October 2009, and was above 9% until the third quarter of 2011 (US Bureau of Labor Statistics (BLS). 2015).

Following other literature studying the relationship between recession and wellbeing, we started the investigation by a household fixed-effect model that attempts to explain variation in household recreation expenditure by state unemployment rate, which controls for household-level heterogeneity. However, since state-level unemployment captures overall economic condition, we cannot disentangle the impact of job loss at the household level. To learn the true impact of job losses during the recession on recreation expenditure, we need to know the counterfactual behavior—how much households would spend on recreation had they not lost their jobs. One challenge in this regard is the potential selection issue with job loss. Although recession affects everyone, not all households are affected in a similar fashion. We adopted Rubin's (1983) potential outcome framework to address the selection issue and derive counterfactual outcome.

Studies focusing on the 2008–2009 recession documented that unmarried, black, and Hispanic individuals with low education levels were disadvantaged during the recession and suffered more from unemployment and poor health (Hoynes, Miller, and Schaller 2012; Currie, Duque, and Garfinkel 2015). We address the selection issues related to job loss during the recession by using propensity score matching (PSM) method to identify a household that is similar to the treatment but did not suffer job loss during the recession. Our information set is not exhaustive—there is a possibility that PSM cannot account for all observable characteristics that determine employment loss during recession. There can still be both time-variant and time-invariant *unobservable* household characteristics that PSM cannot control for. However, since PSID provides information on households' pre-recession recreation expenditures, a difference-in-difference approach would take care of time-invariant confounders. The difference-in-difference approach, in contrast to the household fixed-effect model with state unemployment rate, can disentangle the individual impact of unemployment.

In the empirical analysis, we define the treatment status based on both household head's and spouse's employment status in a pre-recession and recession year. We consider both unemployment and retirement during the recession as evidence of job loss to form a combined treatment group. We then split the treatment group recognizing that the retiree might be different from the unemployed group. We conduct several recommended tests and placebo exercises to test the validity of the key underlying assumptions behind matching framework.

Empirical results reveal that aggregate unemployment provides a different picture on the impact of recessionary job loss on recreation expenditure. Findings based on household-level unemployment show that households losing jobs during a recession did not change recreation expenditure at the intensive margin. However, estimates at the extensive margin shows that unemployment during the recession led some households to completely stop recreating.

2. Panel Study of Income Dynamics (PSID) Data

We use PSID data to investigate the impact of unemployment during recession on trip expenditure. PSID, a longitudinal study conducted by the University of Michigan, began interviewing 18,000 individuals from 5,000 families in 1968. Annual interview of this nationally representative sample and descendants from the original families continued until 1997. After 1997, the survey was conducted on a biennial basis. The survey collected detail household information including employment, education, income, wealth, health, and expenditures, along with many other socioeconomic factors. Eventually, the surveys incorporated questions to address scientific and policy needs. The interviews always collected detail information on the household head, and spouses as well, if applicable.

To answer the research question, we draw on the data from PSID survey rounds of 2005, 2007, 2009, and 2011. We begin with 2005 because that year is the first time PSID included questions on recreation trip expenditure.[32] In any survey year, PSID collects information on consumption expenditure incurred or employment status from the previous calendar year (e.g., recreation expenditure enumerated in 2009 survey was incurred by the household in 2008). In this paper, we state the expenditure or employment status by the year it took place.

The 19 calendar-month-long recession in the US, according the national Bureau of Economic Research, started in December of 2007 and ended in June of 2009. Based on those dates, the 2009 survey round exactly matches with the recession year, and represents the treatment year in our research. Although it officially ended in June of 2009, the recession was still widespread after the official end date, as the economy recovered slowly. Because of the slow recovery, households did not regain confidence to spend at pre-recession levels, which leads us to consider 2010 a recession year.

Drawing on samples from the survey rounds, we include households that provided information to conduct the analysis following empirical strategy and methods proposed in the

[32] PSID asked the respondents "*How much did you (and your family living there) spend altogether in 2008 on trips and vacations, including transportation, accommodations, and recreational expenses on trips?*"

previous section. In the fixed-effect framework, where the fixed effect is assumed on the households, the data requirement is minimal. Since we only need recreation trip expenditure and state identification, we include all households that provided complete information on these two items, allowing us to utilize the maximum number of survey households. Our tabulation of the data reveals that 11,387 households appear at least once with complete information on recreation trips in survey rounds from 2005 to 2011. Out of these responses, 9,123 households appear at least twice and 5,603 households appear across all four rounds with complete information on recreation expenditure. Additionally, 809 households, 8.38% of the original sample, reported a different state of residence across survey rounds. While they might be different from those not changing states, we keep them in the sample. We match these 11,387 households with the average yearly unemployment rate in their state of residence and utilize them in the fixed-effect regression analysis. The state-level yearly average unemployment rate is drawn from the Local Area Unemployment Statistics provided by the Bureau of Labor Statistics.

Figure 1 plots state-level unemployment rates, participation in recreation, and total expenditure in recreation trips for all households that provided trip expenditure information at least once. We observe an opposite movement of participation and trip expenditure against state-level unemployment rate. Figure 2, a replication of Figure 1 with households that appear in our fixed-effect exercise, depicts that both participation and average expenditure on trips across years are higher compared to the corresponding averages of those households utilized in Figure 1. In Figure 2, the pattern of movement of both participation and trip expenditure and state-level unemployment rate remains the same as in Figure 1, except in 2006, where we observe a rise in participation with a fall in the state-level unemployment rate. The pattern in Figure 1 remains unchanged if we focus only on the balanced sample comprising 5,587 households.

For the matching and difference-in-difference exercises, we include all households that provided complete information on employment status and trip-expenditure in both of the pre-recession year and recession year, but provided complete socioeconomic factors only in the pre-recession year. The total number of treatment and control households are presented in Table 1. Between 2006 and 2008, out of 4,627 households with complete information, 83.66% revealed that both spouses were employed in both 2006 and 2008, forming the control. Of the respondent households, 16.34% reported at least one spouse who became unemployed or retired during the recession year 2008, forming treatment group one. Treatment group one is split to form treatment group two, consisting only of those who became unemployed (11.89%), and treatment group three, consisting only of retirees (4.5%). Following a similar approach, we have a sample of 4,129 households from the years 2006 and 2010, of which 80.55% form the control group, 19.5% form treatment group one, 11.12% form treatment group two, and 8.5% from treatment group three. Finally, between the years 2008 and 2010, 14.22% household respondents out of 4,422 became either unemployed or retired (treatment group one), 9.7% became unemployed (treatment group two), and 4.55% retired (treatment group three).[33]

In the matching framework, we consider 2006 as the baseline year in the analysis of 2006 vs. 2008 and 2006 vs. 2010; while in the analysis of 2008 and 2010, we consider 2008 as the baseline year. The summary statistics of the covariate used in the baseline years of 2006 and 2008 are presented in Table 2. In our sample, household heads are, on average, 42.5 years old, and the average spousal age is 25.5 years. The sample consists of 62% white, 30% black, and

[33] Note that treatment group two (unemployed) and three (retired) do not sum to treatment group one (unemployed and retired). In a few household cases, spouses experienced both retirement and unemployment. For example, a husband became unemployed while a wife retired from a fulltime employment position in the pre-recession year. Since the percentage of such households in this sample are small, instead of discarding them, we include them in our analysis.

8% from other races, 56% are married, the average family size is 2.74, and 14% of households have children under the age of two. The schooling profile indicates that 11% of respondents did not complete high school, 28% are high school graduates, and 28% have some college education. In our sample, approximately 76% of households reside in urban areas and 3% reside in rural areas. Around 65% of the households own a house and 93% own a vehicle. Average household wealth is a little over $200,000 but there is huge variation across households. The standard deviation of household wealth is $12.44 million USD. Most households are employed in manufacturing, retail trade, and services sectors. The construction industry employed 7% of the households, and around 10% were employed in agriculture and public administration, two of the relatively less-affected industries during the 2009 recession. A majority of the households come from states in the southern US. The overall health status of the sample seems good—89% self-reported good health, 65% are involved in physical activities, and only 20% are smokers.

The averages of the covariates are almost the same in the baseline year of 2008, as presented in Table 2, except those on household wealth. Since, in 2008, households were already affected by the recession, we expect the average household wealth to fall. Accordingly, we observe that average household wealth in 2008 is approximately 7% lower than in 2006, and standard deviation of household wealth increases by 1.43 times in the recession year.

3. **Empirical Design and Strategy**

We investigate the impact of employment shock during a recession on recreation expenditure both at the extensive and intensive margin. The extensive margin results will reveal if more households stop or start recreating during the recession, while the intensive margin results will show the magnitude of change. We adopt two different empirical approaches. First, we apply a linear fixed-effect model with fixed effects assumed at the household level. We exploit variation in state-level unemployment rate to capture the recession and estimate how changes in state-level unemployment rate affect household recreation expenditure. Although state-level unemployment rate reflects the overall macroeconomic conditions, it will not inform us how changes in employment status alters average recreation behavior. To investigate the role of changes in individual employment status on recreation, we apply the treatment effect framework (Rosenbaum and Rubin 1983), where we consider recession as a quasi-experiment and define treatment and control status based on household members' exposure to employment shock during the recession.[34]

3.1 Linear fixed-effect model

One advantage of using state unemployment rate in studying household recreation expenditure is that it is less likely to be endogenous to household decision making on various consumption expenditures when compared to household members' labor market status. Several studies investigating the relationship between individual health behavior and recession have utilized group variables such as state-level unemployment rate to represent business cycles (Ruhm 2000; 2005; Dehejia and Lleras-Muney 2004; Currie and Schwandt 2014). In our setting for household-level analysis, state-level unemployment rate varies across states and years, but is invariant for households within the state in a particular year. We estimate the following reduced form specification:

$$Participation_{jst} = \delta_1 Unemployment_{st} + \delta_2 Recession_t + \delta_3\, Unemployment_{st} * Recession_t + \gamma_j + \gamma_{st} + \mu_t + \epsilon_{jst}. \quad (1)$$

[34] The empirical framework adopted here follows the same strategy applied in our companion paper "*Is Outdoor Recreation Recession-Proof? An investigation on Lake Recreation Behavior During 2009 Recession*" in chapter one, and this section heavily draws from there.

$$TripExpenditure_{jst} = \beta_1 Unemployment_{st} + \beta_2 Recession_t \ + \beta_3 Unemployment_{st} * Recession_t + \gamma_j + \gamma_{st} + \mu_t + \vartheta_{jst}. \quad (2)$$

where $Participation_{jst}$ is a binary variable indicating whether household "*j*" in state "*s*" takes any recreation trip in year "*t*" or not, $Unemployment_{st}$ stands for unemployment rate in state "s" in year "*t*," $Recession_t$ is an indicator variable assuming a value of 1 if year "*t*" is a recession year and 0 otherwise, γ_j are household-specific fixed effects which take care of time-invariant demographics such as race, gender, education, preference for recreation or work, risk attitudes etc., δ_t are year fixed effects to address time-varying, but household-level invariant, factors such as national and global economic condition, cost of living, commodity prices (including gas prices, etc.), γ_{st} are state-specific time trends, and $\in_{jst}$ and ϑ_{jst} are error terms that are assumed to be uncorrelated with the explanatory variables in the above specifications. Note that including both recession effects and year dummies will drop the recession year while estimating the individual year fixed effects. Standard errors are clustered within the household to allow the errors to be correlated within the household across years.

Fixed effects model can control for household-level time-invariant characteristics that might be correlated with the probability of employment shock exposure during a recession, or living in a state with high unemployment, and with recreation behavior. We are interested in the sign of the parameters δ_3 and β_3, the coefficients on the interaction term between state-level unemployment and a recession indicator. Under the household fixed-effect framework, the identification of δ_3 and β_3 comes from within-state variation in unemployment rate across years. If state-level unemployment exerts a notable effect on average household recreation expenditure in a recession year compared to a typical year, it will be is reflected through δ_3 and β_3.

3.2 Treatment effect framework

We model the impact of employment change during recession on recreation expenditure in a non-experimental design. Our treatment group includes those who experience a recessionary change in employment. In contrast to an experimental setting exposure to recession is non-random—there are selections on who is affected during a recession and to what extent. Recent studies find evidence that during a recession the intensity of losses is high among young and less-educated workers, minorities, unmarried, black, and Hispanic populations; men were more affected than women and their recovery was faster (Hoynes, Miller, and Schaller 2012; Currie, Duque, and Garfinkel *forthcoming*). Selection due to such non-random treatment assignment might hide the true causal effect of a change in employment status during a recession on recreation expenditure.

We want to learn how recessionary job loss alters affected households' recreation expenditure patterns. This requires knowing the counterfactual expenditure of the households that lost jobs during the recession. This behavior is not observable since we can observe only one outcome in one state. Propensity Score Matching is a widely used method (Rosenbaum Rubin 1983; Dehejia and Wahba 2002; Jalan and Ravallion 2003; Greenstone 2004; List et al. 2003; Imbens and Woolridge 2009; Ferret and Subervie 2013) that derives the missing counterfactual outcome in a non-experimental setting under certain assumptions and conditioning on observables. In the first step of a two-step procedure, the method estimates a propensity score (one's probability of being in the treatment) conditioned on observed covariates. Based on the estimated metric in the second step, it matches the treatment observations with similar control to estimates the impact of treatment on outcome of interest.

The study periods for this analysis are 2004–2010, out of which the pre-recession years are 2004 and 2006, and the recession years are 2008 and 2010. Officially the recession was over in June 2009 but its affect was still in full swing in 2010. The economy was sluggish to recover and job creation was insignificant, which leads us to consider both 2008 and 2010 as recession years. Our empirical design involves the comparison of recreation expenditure between the treatment and control across a pre-recession and the recession year. We draw an analysis across each of the three pairs of years: (*a*) 2006 vs. 2008, (*b*) 2006 vs. 2010, and (*c*) 2008 vs. 2010. In our household-level setting, treatment status T_j for household "*j*" is defined as

$$T_j = \begin{cases} 1 \ \ if \text{ i was fulltime employed in } t-1 \text{ but Unemployedd or Retired in year } t \\ 0 \qquad\qquad\qquad\qquad if \text{ i was fulltime employed in year } t-1 \text{ and } t \end{cases}.$$

The treatment group includes all households where at least one spouse became unemployed or retired during the recession but was employed before the recession, while the control group contains those households where at least one spouse had been employed full time and employment status of the other spouse remained unchanged across the pre-recession and recession year. Recognizing the possible differences between unemployed and retirees, we split the combined treatment group into three separate groups, including and excluding the retirees. The outcome variable for household "*j*" in year "*t*," following specifications one and two, assume a binary indicator $Participation$ and a continuous variable $TripExpenditure_{jt}$ to represent recreation expenditure at the extensive and intensive margin, respectively.

The next step involves constructing the propensity score based on observable covariates, conditioned on which the treatment and control group will be matched so that the two groups look similar like an experimental setting. There is no clear set of standards on variables to include in the propensity score equation while constructing the propensity score measure, so our strategy is to include a vector of covariates that make the treatment and control groups similar. The literature suggests incorporating all important and necessary variables from the pre-treatment period that might influence outcome and treatment variables (Heckman, Ichimura, and Todd 1997; Smith and Todd 2005; Caliendo and Kopeinig 2008). Accordingly, economic theory, previous research, and institutional setting can help to characterize the covariates.

Incorporating the treatment status as a dependent variable and the covariate $\boldsymbol{X}$ in a probit model, we estimate the probability of being unemployed or retired during the recession and obtain the propensity score $P(\boldsymbol{X})$**.** To identify the true impact of job loss during the recession, the following three assumptions are critical for the counterfactual framework described above:

Identification Assumption 1: Conditional Independence Assumption

Conditional on the observables, there does not exist any selection effects (i.e., no difference in potential recreation expenditure between the treatment and control absent the treatment), which can be expressed as:

$$TripExpenditure_{j,t}^{0}\, , Participation_{j,t}^{0} \perp T|P(X). \qquad (3)$$

Identification Assumption 2: Stable Unit Treatment Value Assumption (SUTVA)

Treatment households do not affect the recreation behavior of control households. This assumption is to rule out interaction across households and any general equilibrium effect.

Identification Assumption 3: Common Support Assumption

In the sample, treatment households have corresponding control households. For households with respondents that have become unemployed or retired, there must exist households with respondents that are still employed, but are otherwise similar with respect to other characteristics. This is also known as overlapping condition.

$P(T = 1|X) > 0, \forall\, j \epsilon \{i: T_i = 0\}.$

The conditional independence assumption, stated in equation (3), implies the following mean independence condition, which provides the counterfactual observation for the treatment.

$E[Participation^0_{j,t}|P(X), T = 1] = E[Participation^0_{j,t}|P(X), T = 0]$, and

$E[TripExpenditure^0_{j,t}|P(X), T = 1] = E[TripExpenditure^0_{j,t}|P(X), T = 0]$.

We estimate the impact of a change in employment status during the recession on recreation adopting the following average treatment effect on the treated (ATT) estimators:

$$\widehat{ATT}_{extensive} = E\left[Participation^1_{j,t} - Participation^0_{j,t}|\, T = 1]\right]$$

$$=E[Participation^1_{j,t}|\, T = 1] - E[Participation^0_{j,t}|P(X), T = 1]$$

$$= E[Participation^1_{j,t}|\, T = 1] - E[Participation^0_{j,t}|P(X), T = 0]^{35}. \quad (4)$$

For trip expenditure, the estimator is

$$\widehat{ATT}_{intensive} = E[TripExpenditure^1_{j,t}|\, T = 1] - E[TripExpenditure^0_{j,t}|P(X), T = 0]. \quad (5)$$

We apply three different matching algorithms to ensure that results are not driven by any particular procedure, and to facilitate comparison across procedures.[36] While matching is conducted we discard all treatments that lie outside the common support to ensure that *identification assumption two* is satisfied. The *conditional independence assumption* is also testable drawing a comparison of covariates' means across the treatment and matched control groups. Standard errors are estimated following Abadie and Imbens (2008) for nearest-neighbor matching algorithms, and bootstrapped procedure for radius matching estimator.

[35] Note that we replace the counterfactual participation of the treatment households by participation of similar households from the control group.

[36] In nearest-neighbor matching, for each treatment, we pick the control with the closest propensity score with a replacement. Nearest-five-neighbors matching picks the five controls with the closest propensity score. Radius matching: for each exposed individual, we pick all the controls whose propensity score lies within a radius distance of ½ of standard deviation of the estimated propensity score.

Selection on Unobservables**:** If the selection into change in employment status is due to unobservables rather than observable factors, the PSM estimators will not reveal the true impact of recession. In such cases, difference-in-difference (DID) matching estimators is highly recommended (Heckman, Ichimura, and Todd 1997; Heckman, et al. 1998; Smith and Todd 2005; Abadie 2005; Imbens and Woolrdige 2009). For example, poor mental health condition is common during a recession (Currie and Tekin 2014) and can affect employability as well as willingness to take recreation trips. Similarly, if wages go down below one's reservation wage during the recession, s/he might choose a voluntary unemployment and spend time on vacation and trips. If we can observe the households both before and during the recession, we can net-out such factors by a simple first differencing with respect to the pre-treatment period. Combining propensity score matching with DID makes the treatment and control similar both in terms of observables as well as unobservables. The panel setting of our data allows us to implement this technique to control for all potential time-invariant unobservable factors that might be associated both with recreation expenditure and change in employment status.

The estimation procedure for the DID matching estimator is the same as the propensity score matching estimators stated above, except that we will conduct the matching on differences of recreation expenditure and differences of participation across pre-recession periods. However, the DID matching estimators must satisfy one additional identifying assumption relative to the matching estimators: the parallel trend.

Parallel Trend Assumption: In the absence of recessionary unemployment, the average change in recreation expenditure across periods is the same for the treatment and control households.

$$E[Participation^0_{j,t-1} - Participation^0_{j,t-2} | P(\boldsymbol{X}), T = 1]$$
$$= E[Participation^0_{j,t-1} - Participation^0_{j,t-2} | P(\boldsymbol{X}), T = 0]. \quad (6)$$

Similarly, for total recreation expenditures,

$$E[TripExpenditure^0_{j,t-1} - TripExpenditure^0_{j,t-2}|P(\boldsymbol{X}), T = 1]$$

$$= E[TripExpenditure^0_{j,t-1} - TripExpenditure^0_{j,t-2}|P(\boldsymbol{X}), T = 0]. \quad (7)$$

The common trend assumption is testable if multiple pre-treatment period's data is available. We have recreation expenditure data available for two pre-recession periods, 2004 and 2006. We conduct a placebo exercise across 2004 and 2006 to test if the parallel trend conditions holds for the analysis across 2006 and 2008, which will reveal if DID matching estimates are unbiased. We also conduct a placebo exercise for the matching estimates to test if the treatment and control groups are similar in terms of pre-treatment outcome.

We adopt three matching estimators to check the robustness of the estimates of impact of recessionary job loss on recreation spending during the recession. In addition, we apply doubly robust (DR) estimators (Imbens and Woolridge 2009; Woolrdige 2010) as an additional robustness check. The DR estimator exploits the same specification for both propensity score estimation and outcome equation. While estimating the outcome equation, DR weights by the inverse probability obtained from the propensity score estimation stage. The argument for DR estimator is that if one of the two specifications is wrong, estimates are still consistent. This seems appealing when we do not know much about the treatment assignment equation.

4. Results and Interpretation

4.1 State unemployment and recreation expenditure

The estimates from regression specifications one and two are presented in Tables 3 and 4. We report estimates from five different specifications, where specifications gradually increase control by including various combinations of household, state, year fixed effects, and trends. Specification one does not include household fixed effect to accommodate state fixed effects.

Specifications four and five are the most restricted specifications since they control for either or both of year fixed effect and state-specific linear trend to control for time-varying factors.

Panel (a) in Table 3 reports results for participation assuming 2008 as a recession year, panel (b) reports results assuming 2010 as a recession year, while panels (c) and (d) report the estimates assuming 2004 and 2006 as placebo recession years. The coefficient of state unemployment rate is negative and statistically insignificant in all cases, except when we consider 2010 as the recession year. The statistical significance does not persist once we control for time-varying factors by including year fixed effects and state trends. The coefficient on recession is statistically insignificant in most of the cases. The coefficient on the interaction term between state unemployment and recession, δ_3, consistently fails to exhibit statistical significance, except in two specifications under the assumption of 2010 as a recession year. Overall, results suggest that higher state unemployment during the recession, assumed different from a normal year, does not exhibit any statistically significant relationship with recreation participation.

In contrast to participation, recreation expenditure at the intensive margin (Table 4) exhibits different patterns based on the specified recession year. When we consider 2008 the recession year, in three out of five specifications, including specification four, which includes year fixed effect, state unemployment rate exhibits a negative, statistically significant relationship with recreation expenditure. If 2008 is considered the recession year, the recession indicator is always negative and statistically significant, but switches sign, and is not consistently significant when 2010 is considered as the recession year. The unemployment and recession coefficients are never significant under placebo recession years. The coefficient on the interaction term between unemployment and recession indicator, β_3, is always positive and statistically significant if 2008 is assumed as the recession year. However, if 2010 is assumed the recession year, β_3 exhibits a negative and statistically significant relationship in specification two with quadratic trend, and specification five, which controls for year fixed effect and state linear trend. Under the assumption that 2004 or 2006 is the placebo recession year, β_3 is consistently found to be statistically insignificant across all specifications.

After controlling for the level effect of state-level unemployment, recession, household fixed effects, and various trends, the consistent finding of $\beta_3 > 0$ suggests that state unemployment in 2008 had a differentiated effect on recreation expenditure compared to state unemployment in any other year. Since the entire calendar year 2008 was part of an officially announced recession year, there is some evidence that recessionary high state unemployment increases recreation expenditure.

4.2 Job loss within household during recession and recreation expenditure

We present the ATT estimates here (i.e., the impact of household members' job loss on recreation expenditure). Before reporting these results, we report the estimates from propensity score estimation and evaluate the quality of matching. For convenience, we will name the ATT estimates from propensity score matching estimators as "matching estimates" and difference-in-difference matching estimates as "DID estimates."

Propensity score estimation

Table 5 presents propensity score estimates. Across each pair of years, we ran a separate *probit* regression for each of the three treatment groups: (*a*) combining all who lost jobs during a recession year, (*b*) including only those who were unemployed, and (*c*) including the retired. The objective of these estimations is to obtain propensity score metrics, and therefore, we are not interpreting the estimates in greater detail. Overall, the estimates suggest that being Hispanic, black, unmarried, less educated, not owning a house and car, smoking, and self-reporting good health are positively associated with unemployment during a recession. In addition, working in the agricultural sector and living in the south are found to be negatively associated with unemployment. Retirement status is negatively associated with linear component of age, larger family size, living in the south, and not self-reporting good health. On the other hand, being black, less educated, a smoker, and working in public administration increases the likelihood of unemployment during a recession.

Quality of matching

The matching quality is assessed based on whether assumptions 1–3, as laid out in the previous section, are satisfied. One implication of the conditional mean independence assumption is that the treatment and matched control groups look similar in terms of observables. The standardized mean difference of the covariates between the treatment and matched control group indicates whether this similarity holds.[37] We consider a standardized difference of means of 20 as large (Rosenbaum and Rubin 1983; Imbens and Woolridge 2009). Overall, matching has been successful. Although there were significant differences across treatment and control groups with respect to observables before matching, the difference goes away in the matched sample. In a few nearest-neighbor matching cases, around 5% of the covariates exhibited standardized mean difference above 20. We find no such violation across nearest-five-neighbor matching and radius matching estimators. The results are reported in Table A1-A9 in the *appendix* B.

[37] The formula for standardized difference in means is $\frac{Mean_{treated} - Mean_{control}}{\sqrt{0.5(Variance_{\ treated} + Variance_{control})}}$.

The overlapping condition, as stated in assumption 2, is satisfied as well. While matching is conducted, we exclude all treatments that do not find similar control. Table A10 in the *appendix* B presents the results. Across the matching algorithms exercised, less than 2% of treatments fail to find a matching counterfactual. Finally, SUTVA is less likely to be violated in our framework. Since our empirical design models household's recreation expenditure, where decisions are made at the household setting and incorporates both spouse's employment while constructing a treatment and control status, such spillover effect is less likely to take place. All these are crucial to interpret the treatment effect in a causal manner. The parallel trend results for DID matching are discussed later.

2006 as pre-recession and 2008 as recession year

Table 6 reports the estimates on how households that experienced a job loss during the 2008 recession changed recreation participation and expenditures during 2008. Panel (a) in Table 6 shows that treatment households recreated less during the recession. The matching estimators reveal that recreation participation falls in the range of 7.8 to 11.6 percentage points. DID estimates are smaller in magnitude and lie in the range of 6.2 to 10.4 percentage points. All estimates exhibit statistical significance. Panels (b) and (c) suggests that the decline in recreation participation is mainly driven by the unemployed and not retired households. The DID estimates reveal that unemployed households' participation in recreation in 2008 fell by 8.2 to 9.2 percentage points. We also report the recreation participation in post-recession year 2010.

For a subsample of the sample utilized in 2006 and 2008, recreation information in 2010 is available. We use this information to observe how the treatment groups from 2008 change participation in recreation in 2010 compared to the control group of 2008. All matching estimates as well as DID estimates reveal that treatment households' participation is significantly lower even in the post-recession year. The estimates are larger in magnitude compared to those obtained for recreation in the treatment year 2008, and the decline in participation consistently comes from the unemployed group. Retired households in 2008 would not have changed recreation participation in 2010 had they been employed full time during the recession of 2008. However, for the retired households in 2008, when we take differences in participation between 2008 and 2010, all of the DID estimates suggest decreased participation in recreation in post-treatment years.

Note that the level and DID estimates obtained for the post-treatment years cannot be interpreted as ATT since many households in the sample experience a change in employment status in 2010. This suggests that the treatment and control status have also changed for some households between 2008 and 2010. In this context, the decline in recreation participation in the post-recession years by the households with respondents that were unemployed in 2008 cannot be attributed to recessionary unemployment.

Table 6 also presents how the households with employed respondents in 2006 that lost jobs in 2008 changed recreation expenditures during the recession and post-recession years. The matching estimates suggests that treatment households experienced a decrease in recreation expenditure in the recession year of 2008 compared to the control households. The magnitude of the ATT estimates indicate that household respondents who became unemployed during the recession spent \$249–\$344 less on recreation expenditure compared to what they would have spent had they not been unemployed. However, the DID estimates do not support this fall in recreation expenditure caused by recessionary unemployment, all the three estimators fail to exhibit statistical significance. None of the ATT estimates for the retired groups exhibit statistical significance. The retired households, as the matching estimates as well as DID estimates suggest, did not change recreation expenditure in 2008 due to recessionary retirement.

There is some evidence that household respondents who became unemployed during the recession spent less on recreation expenditure during the post-treatment year of 2010. Both matching and DID estimates for the unemployed group suggest this. However, due to the change in employment status between 2008 and 2010, we will not attribute this fall in recreation expenditure to recessionary unemployment in 2008.

2006 as pre-recession and 2010 as recession year

Table 7 reports how households respondents that were employed in 2006 but became unemployed or retired in 2010 changed participation in recreation and recreation spending during the recession year of 2010. The matching estimates suggest that treatment households participated less in 2010. The unemployed households' participation shrank by 14.6 to 17.3 percentage points while that for the retired shrank by around 7 percentage points. The DID estimates also suggest that the unemployed households' participation in 2010 fell compared to what it would have been were they not exposed to recessionary unemployment. For retired households, DID estimates support recessionary retirement causing reduced recreation participation in 2010. Matching estimates suggest both treatment groups participated less in 2008 compared to their employed counterparts, but DID estimates do not support this evidence.

At the intensive margin, the matching estimates suggest that treatment household respondents that became unemployed reduced recreation expenditure in the treatment year 2010, but none of the DID estimates provide support in favor of that finding. Retired households are spending the same in 2008 as their counterfactual spending. We did not notice any significant difference in mean spending in 2008, relative to spending absent the treatment, for the unemployed or retired group.

2008 as pre-recession and 2010 as recession year

Table 8 presents findings on recreation participation and expenditure for the treatment household respondents employed in 2008 that lost jobs during 2010.[38] The matching estimates suggest that both participation and spending in recreation in 2010 drop for treatment group two, but DID estimates do not consistently agree with this—only one out of three DID estimates supports a fall in spending.

Retirees do not exhibit any change in recreation behavior in 2010 at the extensive or intensive margin. Unemployed households, as matching estimates reveal, exhibited significantly less participation in recreation in 2008 compared to their corresponding controls. In contrast to participation, there is no strong evidence that respondents who became unemployed during the recession year 2010 spent less in 2008.

Placebo exercises

Table 9 presents the results from the placebo exercise implemented as laid out in Section 3. For the analysis between 2006 and 2008, the exercise incorporating 2004 as a placebo recession year suggests that the treatment and control group was similar in terms of their recreation participation and expenditure in 2004. Both for participation and expenditure, the DID matching estimates between 2004 and 2006 confirm that the parallel trend assumption is

[38] Treatment households in this sample did not lose jobs in 2008, but became unemployed or retired in 2010, after the recession.

satisfied. The parallel trend assumption is satisfied as well for the DID estimates obtained from the analysis across "2006 and 2010," and "2008 and 2010."[39]

Across the comparison pairs "2006 and 2010" and "2008 and 2010," matching estimates suggest that the treatment group consisting of unemployed households exhibited a significant difference in participation compared to their corresponding controls in the placebo year, which is concerning as the PSM method attempts to make the two groups identical even in terms of the baseline outcome. One indication of such difference is the possible existence of unobserved confounders that can make the two groups different. This suggests that the DID matching estimates, which control for household-specific fixed effects, will be more reliable. Therefore, we are more confident in interpreting the difference-in-difference results as a causal effect of recessionary job loss on recreation expenditure. Note that for the retirement group as well as for recreation at the intensive margin, difference in outcome in the placebo year is not evident.

Robustness

The DR estimates will indicate: (*a*) if the negative impact of recessionary unemployment on recreation participation during the recession is robust, and (*b*) the placebo estimates obtained for the unemployed treatment exhibit any statistical significance under DR method. Table 10 reports the DR estimates for the treatment household respondents that became unemployed during a recession year. The DR estimates exhibit similar patterns as the matching and DID matching estimates reported previously. However, the magnitude of the DR estimates for participation in recreation during a recession year is larger compared to the matching estimates. The magnitude of the estimates is comparable with matching estimates obtained for expenditure and DID

[39] In a comparison pair of years for each of the treatment groups, at most one out of three DID estimators exhibit statistical significance. Nine percent (5 out of 54) of the DID matching estimates violated parallel trend assumption.

matching estimates for participation and expenditure. The DR estimates suggest that our finding that households' unemployment during a recession causes a drop in household participation in recreation spending is robust. Moreover, the DR estimates on placebo exercises support that the parallel trend assumption is satisfied. However, the treatment and control group in "2006 vs. 2010" and "2008 vs. 2010" still exhibit significant difference in terms of placebo year participation in recreation, which imply the plausibility of confounding *unobservables*.

5. **Conclusion**

This study examines how the recession of 2008–2009 affected household recreation expenditure. We started with a household-specific fixed-effect mode that controls household-level heterogeneity while explaining variation in recreation participation and spending by variation in yearly state unemployment rate. The results consistently reveal that the recessionary state unemployment rate does not matter for participation in recreation. Although we did not observe any consistent relationship of state unemployment with recreation spending, recessionary unemployment in 2008 appears to have exerted a positive influence on recreation spending. We are cautious interpreting this as an effect of unemployment, as state unemployment is a catch-all for overall macroeconomic conditions. Recognizing that state-level unemployment rate captures overall unemployment scenarios, uncertainty, consumer confidence, aggregate productivity loss, we applied treatment-effect framework to study the effects of unemployment on recreation spending at the household level. Several PSM estimators and DID matching estimators showed that average household respondents that became unemployed during the recession participated less in recreation during the recession. However, at the intensive margin, we do not find consistent and clear evidence of a drop in recreation spending during the recession year. The retired households showed a mixed picture: they did not exhibit any change in participation in 2008 but there is some evidence that participation dropped in 2010.

Overall, we do not observe any consistent impact of recessionary unemployment on recreation spending. However, this study suggests that the aggregate unemployment statistics and micro-unemployment indicate a different effect on recreation expenditure, the former at the intensive margin and the latter at the extensive margin. The findings imply the significance of exploiting unemployment at the household level in studying the impact of recessionary events on wellbeing. As a caveat, we recognize that unemployment is only one channel through which households are affected during the recession. Accordingly, our estimates cannot be interpreted as the impact of recession, but rather the impact of recessionary job loss.

6. **References**

Abadie, A. 2005. "Semiparametric Difference-in-Difference Estimators." *Review of Economic Studies* 72(2005): 1–19.

Abadie, A., and G.W. Imbens. 2008. "On the Failure of the Bootstrap for Matching Estimators." *Econometrica* 76(2008): 1537–1557.

Black, D., N. Kolesnikova, S. Sanders, and L. Taylor. 2013. "Are Children Normal?" *Review of Economic and Statistics* 95(1): 21–33.

Caliendo, M., and S. Kopeinig. 2011."Some Practical Guidance for the Implementation of Propensity Score Matching." *Journal of Economic Surveys* 22(1): 31–72.

Chabé-Ferret, S. and J. Subervie. 2013. "How Much Green for the Buck? Estimating Additional and Windfall Effects of French Agro-environmental Schemes by DID-matching." *Journal of Environmental Economics and Management* 65(2013): 12–27.

Currie, J. and H. Schwandt. 2014. "Short- and Long-term Effects of Unemployment on Fertility." *Proceedings of the National Academy of Sciences* 111(41): 14734–14739.

Currie, J., V. Duque, V., and I. Garfinkel. Forthcoming. "The Great Recession and Mother's Health." *Economic Journal*.

Currie, J., and E. Tekin. 2014. "Is there a Link between Foreclosure and Health?" *American Economics Journals: Economic Policy* January (2014).

Dehejia, R., and Adriana Lleras-Muney. 2004. "Booms, Busts, and Babies' Health." *Quarterly Journal of Economics* 119(2004): 1091–1130.

Dehejia R.H., and S. Wahba. 2002. "Propensity Score-matching Methods for Non-experimental Causal Studies." *Review of Economics and Statistics* 84(2002): 151–161.

Egan, Kevin J., J.A. Herriges, C.L. Kling, and J.A. Downing. 2009. "Valuing Water Quality as a Function of Water Quality Measures." *American Journal of Agricultural Economics* 91(2009): 106–123.

Greenstone, M. 2004. "Did the Clean Air Act Cause the Remarkable Decline in Sulfur Dioxide Concentrations?" *Journal of Environmental Economics and Management* 47(2004): 585–611.

Heckman, J., H. Ichimura, and P. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies* 64(4): 605–654.

———. 1998. "Characterizing Selection Bias Using Experimental Data." *Econometrica* 66(5): 1017–1098.

Hoynes, Hilary, Douglas L. Miller, and Jessamyn Schaller. 2012. "Who Suffers during Recessions?" *Journal of Economic Perspectives* 26(3): 27–48.

Imbens, G.W., and J.M. Woolridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47(1): 5–86.

Jalan, J., and M. Ravallion. 2003. "Does Piped Water Reduce Diarrhea for Children in Rural

India?" *Journal of Econometrics* 112(2003): 153–173.

List, John A., Daniel Millimet, Per Fredriksson, and Warren McHone. 2003. "Effects of Environmental Regulations on Manufacturing Plant Births: Evidence from a Propensity Score Matching Estimator," *Review of Economics and Statistics* 85(2003): 944–952.

Loomis, J., and C. Keske. 2012. "Did the Great Recession Reduce Visitors Spending and Willingness to Pay for Nature-based Recreation? Evidence from 2006 and 2009." *Contemporary Economic Policy* 30(2012): 238–246.

Marcus, J. 2013. "The Effect of Unemployment on the Mental Health of Spouses—Evidence from Plant Closures in Germany." *Journal of Health Economics* 32: 546–558.

McInerney, M., J.M. Mellor, and L.H. Nicholas. 2013. "Recession Depression: Mental Health Effects of the 2008 Stock Market Crash." *Journal of Health Economics* 32: 1090–1104.

Outdoor Industry Foundation (OIF). 2010. *Outdoor Recreation Participation Report*. Available at: http://www.outdoorindustry.org/images/researchfiles/TOF_ResearchParticipation2010.pdf?121. (*last accessed on July20th, 2015*).

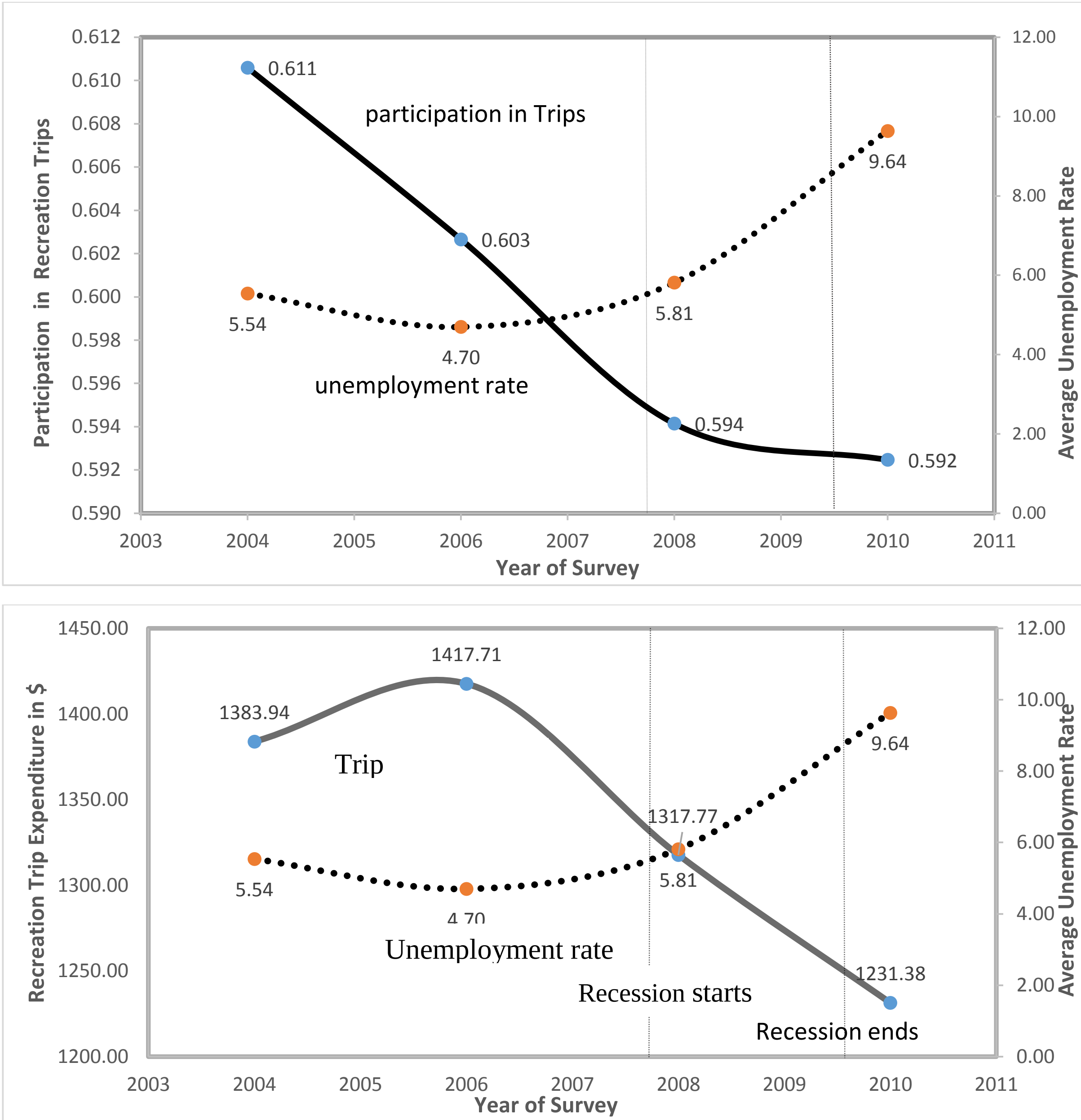
Rosenbaum, P.R., and D.B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1983): 41–55.

Ruhm, C.J. 2000. "Are Recessions Good for Your Health?" *Quarterly Journal of Economics* 115 (2000): 617–650.

———. 2000. "Healthy Living in Hard Times." *Journal of Health Economics* 24(2005): 241–364.

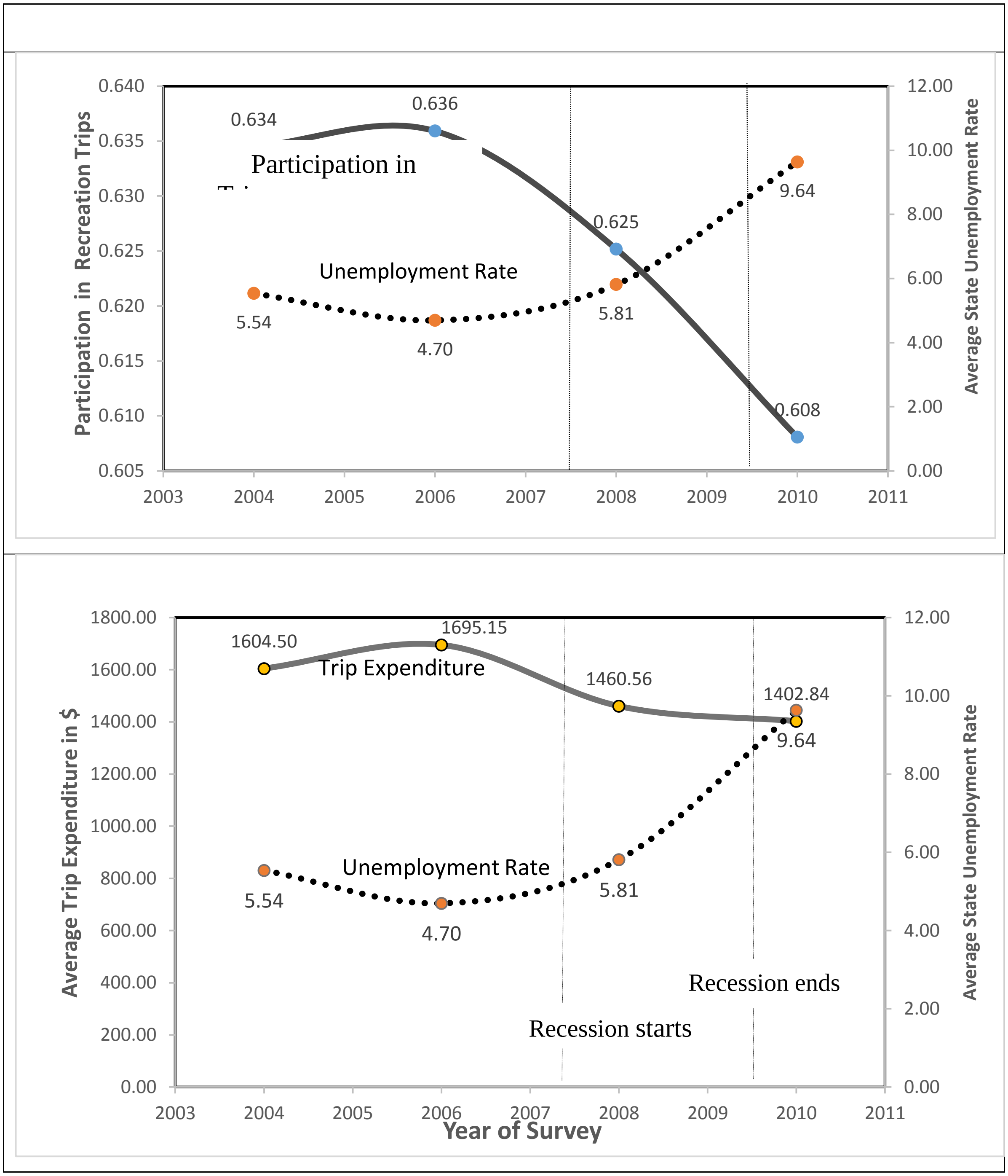
Smith, J., and P. Todd. 2005. "Does Matching Overcome Lalonde's Critique of Non-experimental Estimators?" *Journal of Econometrics* 125(2005): 305–353.

Wooldridge, J. 2010. *Econometric Analysis of Cross Section and Panel Data*, 2nd edition. MIT Press: Cambridge.



Note: We include household respondents that appear at least once with trip expenditure information in the PSID survey rounds from 2005 to 2011.

Figure 1. Average State Unemployment, Participation in Recreation, and Expenditures in Recreation Trips (in US$) 2004–2010



Note: We include household respondents that appear at least twice with trip expenditure information in the PSID survey rounds from 2005 to 2011. The households we analyze here corresponds to the sample used in the fixed-effect model

Figure 2. Average State Unemployment, Participation in Recreation Trips, and Total Expenditure in US$ on Recreation Trips from 2004 to 2010

Table 1. Treatment and Control Based on Employment Status across Years

	Control Group (Employed)	Treatment Group 1 (Unemployed and Retired)	Treatment Group 2 (Unemployed)	Treatment Group 3 (Retired)
		Year 2006 and 2008		
Both spouses are employed in 2006	Both spouses are employed in 2008 3871	Either spouse is unemployed or retired in 2008 756	Either spouse is unemployed in 2008 550	Either spouse is retired in 2008 207
		Year 2006 and 2010		
Both spouses are employed in 2006	Both spouses are employed in 2010 3325	Either spouse is unemployed or retired in 2010 804	Either spouse is unemployed in 2010 459	Either spouse is retired in 2010 349
		Year 2008 and 2010		
Both spouses are employed in 2008	Both spouses are employed in 2010 3793	Either spouse is unemployed or retired in 2010 629	Either spouse is unemployed in 2010 429	Either spouse is retired in 2010 201

Table 2. Summary Statistics of Covariates Used in the Matching Exercises

		Baseline Year 2006				Baseline Year 2008			
	Variable	Mean	std. dev	Min	Max	Mean	std. dev	Min	Max
	Age of Head	42.48	12.83	18.00	82.00	42.33	12.93	17.00	84.00
	Age of Wife	25.48	22.68	0.00	82.00	24.58	22.84	0.00	83.00
	Gender of Head	0.73	0.44	0.00	1.00	0.72	0.45	0.00	1.00
	Hispanic	0.08	0.26	0.00	1.00	0.08	0.27	0.00	1.00
Demographics	Black	0.30	0.46	0.00	1.00	0.32	0.47	0.00	1.00
	Marital Status	0.56	0.50	0.00	1.00	0.53	0.50	0.00	1.00
	Family Size	2.74	1.40	1.00	9.00	2.70	1.43	1.00	12.00
	Child Under Age 2	0.14	0.35	0.00	1.00	0.13	0.34	0.00	1.00
	Less Than High School	0.11	0.31	0.00	1.00	0.11	0.32	0.00	1.00
Education of	High School Graduate	0.28	0.45	0.00	1.00	0.26	0.44	0.00	1.00
Households	Some College	0.28	0.45	0.00	1.00	0.27	0.44	0.00	1.00
Housing,	Wealth in 2007	208250	1245408	-336269	49800000	194148	1790732	-1974000	100000000
Wealth, and	Own House	0.65	0.48	0.00	1.00	0.60	0.49	0.00	1.00
Vehicle	Own Private Vehicle	0.93	0.25	0.00	1.00	0.92	0.28	0.00	1.00
	Manufacturing	0.13	0.34	0.00	1.00	0.13	0.34	0.00	1.00
	Agriculture	0.04	0.20	0.00	1.00	0.04	0.19	0.00	1.00
Industry	Construction	0.07	0.25	0.00	1.00	0.07	0.25	0.00	1.00
	Services	0.34	0.47	0.00	1.00	0.34	0.48	0.00	1.00
	Retail Trade	0.13	0.34	0.00	1.00	0.13	0.34	0.00	1.00
	Public Administration	0.06	0.24	0.00	1.00	0.06	0.24	0.00	1.00
	North East	0.15	0.35	0.00	1.00	0.14	0.35	0.00	1.00
Geographic	Urban	0.21	0.41	0.00	1.00	0.20	0.40	0.00	1.00
Region	Metro	0.76	0.43	0.00	1.00	0.77	0.42	0.00	1.00
	South	0.42	0.49	0.00	1.00	0.43	0.49	0.00	1.00
	West	0.19	0.39	0.00	1.00	0.19	0.39	0.00	1.00
Health and	Health Status of HD	0.89	0.32	0.00	1.00	0.89	0.31	0.00	1.00
Lifestyle	Smoker	0.20	0.40	0.00	1.00	0.20	0.40	0.00	1.00
	Physical Activity of HD	0.65	0.48	0.00	1.00	0.65	0.48	0.00	1.00
	Total Number of Observations	4627				4422			

Table 3. State Unemployment and Household Participation in Recreation

	2008 is Recession Year					**2010 is Recession Year**				
	I	II	III	IV	V	I	II	III	IV	V
State Unemployment	**-0.004***	-0.002	-0.001	-0.003	-0.001	**-0.008***	**-0.007***	**-0.009***	-0.006	-0.008
	[0.002]	[0.002]	[0.004]	[0.002]	[0.004]	[0.004]	[0.004]	[0.006]	[0.004]	[0.006]
Recession*Unemployment	-0.002	-0.004	-0.004	-0.004	-0.004	0.006	**0.008***	**0.009***	0.007	0.008
	[0.004]	[0.005]	[0.005]	[0.005]	[0.005]	[0.004]	[0.004]	[0.005]	[0.004]	[0.005]
Recession	0.01	0.023	0.024	0.023	0.029	-0.04	-0.056	-0.079	-0.052	-0.057
	[0.027]	[0.030]	[0.030]	[0.030]	[0.032]	[0.034]	[0.035]	[0.049]	[0.035]	[0.036]
Linear Trend	*YES*	*YES*	*YES*	*No*	*No*	*YES*	*YES*	*YES*	*No*	*No*
Quadratic Trend	*No*	*No*	*YES*	*No*	*No*	*No*	*No*	*YES*	*No*	*No*
State Specific Linear Trend	*No*	*No*	*No*	*YES*	*YES*	*No*	*No*	*No*	*YES*	*YES*
State Fixed Effect	*No*	*No*	*No*	*No*	*YES*	*No*	*No*	*No*	*No*	*YES*
Adjusted R-square	*0.031*	*0.001*	*0.001*	*0.003*	*0.003*	*0.024*	*0.004*	*0.004*	*0.009*	*0.009*
	2004 is Placebo Recession Year					2006 is Placebo Recession Year				
	I	II	III	IV	V	I	II	III	IV	V
State Unemployment	-0.002	-0.002	-0.001	-0.002	-0.001	-0.003	-0.002	-0.002	-0.002	-0.001
	[0.003]	[0.003]	[0.004]	[0.003]	[0.004]	[0.002]	[0.002]	[0.004]	[0.002]	[0.004]
Recession*Unemployment	0.053	0.068	0.07	0.06	0.054	0.001	-0.001	-0.001	-0.007	-0.003
	[0.051]	[0.044]	[0.046]	[0.044]	[0.055]	[0.032]	[0.029]	[0.029]	[0.029]	[0.031]
Recession	-0.011	**-0.013***	**-0.013***	-0.012	-0.012	0.001	0	0	0.002	0.002
	[0.009]	[0.008]	[0.008]	[0.008]	[0.008]	[0.007]	[0.006]	[0.006]	[0.006]	[0.006]
Linear Trend	*YES*	*YES*	*No*	*No*	*No*	*YES*	*YES*	*No*	*No*	*No*
Quadratic Trend	*No*	*No*	*YES*	*No*	*No*	*No*	*No*	*YES*	*No*	*No*
State Specific Linear Trend	*No*	*No*	*No*	*YES*	*YES*	*No*	*No*	*No*	*YES*	*YES*
State Fixed Effect	*YES*	*No*	*No*	*No*	*No*	*No*	*No*	*No*	*No*	*No*
Year Fixed Effect	*No*	*No*	*No*	*No*	*YES*	*No*	*No*	*No*	*No*	*YES*
Adjusted R-square	*0.024*	*0.004*	*0.004*	*0.009*	*0.009*	*0.024*	*0.004*	*0.004*	*0.008*	*0.008*
Number of Households	*11387*	*9086*	*9086*	*9086*	*9086*	*11387*	*9086*	*9086*	*9086*	*9086*

Note: Standard errors are reported in the bracket immediately under the estimates. Significance level can be read as * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. *Specification I does not include any household fixed effect to accommodate state fixed effects.*

Table 4. State Unemployment and Household Recreation Expenditure in $

	2008 is the Recession Year					2010 is the Recession Year				
	I	II	III	IV	V	I	II	III	IV	V
State Unemployment	-47.57***	-45.80**	-5.80	-47.12**	-3.54	-35.79	-47.05	45.35	-50.11	54.22
	[15.833]	[19.911]	[35.954]	[20.584]	[39.368]	[26.494]	[35.829]	[51.063]	[37.562]	[56.592]
Recession*Unemployment	92.02***	82.85**	77.10*	87.18**	83.03**	-11.22	6.29	-52.68*	5.96	-58.60*
	[27.977]	[41.576]	[40.063]	[42.058]	[40.885]	[27.042]	[30.566]	[31.773]	[30.629]	[32.836]
Recession	-674.32***	-629.73***	-622.89***	-657.37***	-511.22**	211.95	108.44	993.10***	122.38	394.76*
	[187.834]	[241.072]	[238.717]	[243.276]	[218.922]	[227.561]	[240.707]	[345.535]	[241.530]	[214.243]
Linear Trend	*YES*	*YES*	*YES*	*No*	*No*	*YES*	*YES*	*YES*	*No*	*No*
Quadratic Trend	*No*	*No*	*YES*	*No*	*No*	*No*	*No*	*YES*	*No*	*No*
State Linear Trend	*No*	*No*	*No*	*YES*	*YES*	*No*	*No*	*No*	*YES*	*YES*
State Fixed Effect	*No*	*No*	*No*	*No*	*YES*	*No*	*No*	*No*	*No*	*YES*
Adjusted R-square	0.022	0.002	0.002	0.005	0.005	0.022	0.001	0.002	0.004	0.005
	2004 is Placebo Recession Year					2006 is Placebo Recession Year				
	I	II	III	IV	V	I	II	III	IV	V
State Unemployment	24.58	29.27	-0.06	31.81	1.05	1.79	6.24	1.67	6.38	2.44
	[25.794]	[28.268]	[36.935]	[29.876]	[40.183]	[15.609]	[16.968]	[37.497]	[17.471]	[40.605]
Recession*Unemployment	35.05	6.56	7.81	10.76	10.39	-23.79	-31.83	-31.11	-35.79	-35.35
	[49.920]	[48.442]	[48.497]	[48.575]	[48.547]	[31.797]	[33.340]	[34.586]	[32.674]	[33.726]
Recession	-492.04	-342.38	-562.84	-378.80	-6.07	274.03	317.537*	318.843*	338.651*	326.71
	[330.436]	[335.002]	[368.173]	[343.770]	[432.097]	[168.048]	[186.759]	[184.479]	[183.626]	[233.522]
Linear Trend	*YES*	*YES*	*No*	*No*	*No*	*YES*	*YES*	*No*	*No*	*No*
Quadratic Trend	*No*	*No*	*YES*	*No*	*No*	*No*	*No*	*YES*	*No*	*No*
State Linear Trend	*No*	*No*	*No*	*YES*	*YES*	*No*	*No*	*No*	*YES*	*YES*
State Fixed Effect	*YES*	*No*	*No*	*No*	*No*	*YES*	*No*	*No*	*No*	*No*
Year Fixed Effect	*No*	*No*	*No*	*No*	*YES*	*No*	*No*	*No*	*No*	*YES*
Adjusted R-square	0.022	0.002	0.002	0.005	0.005	0.022	0.002	0.002	0.005	0.005
Number of Households	11387	9086	9086	9086	9086	11387	9086	9086	9086	9086

Note: Standard errors are reported in the bracket immediately under the estimates. Significance level can be read as * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.
Specification I does not include any household fixed effect to accommodate state fixed effects.

Table 5. Estimation of Propensity Score Equation

	2008 and 2010			*2006 and 2010*			*2006 and 2008*		
	Unemployed & retired	Retired	Unemployed	Unemployed & retired	Retired	Unemployed	Unemployed & retired	Retired	Unemployed
Age of Head	-0.169***	-0.395***	-0.110*	-0.181***	-0.38***	-0.076	-0.141***	-0.381**	-0.049
	[0.048]	[0.151]	[0.061]	[0.051]	[0.116]	[0.066]	[0.048]	[0.150]	[0.059]
(Age of Head)^2	0.003***	0.009***	0.002	0.003***	0.009***	0.001	0.003**	0.009***	0.001
	[0.001]	[0.003]	[0.001]	[0.001]	[0.002]	[0.002]	[0.001]	[0.003]	[0.001]
(Age of Head)^3	-0.000*	-0.000***	0	-0.000*	-0.000***	0	0	-0.00***	0
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Age of Wife	0.012	-0.052	-0.001	0.013	-0.033	-0.005	0.022*	-0.025	0.026*
	[0.013]	[0.047]	[0.017]	[0.014]	[0.034]	[0.019]	[0.013]	[0.041]	[0.015]
(Age of Wife)^2	0	0.002	0.001	0	0.002	0.001	-0.001	0.001	-0.001
	[0.001]	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]
(Age of Wife)^3	0	0	0	0	0	0	0	0	0
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Gender of Head	0.143*	0.19	0.137	0.148*	0.193	0.142	0.140*	-0.011	0.176**
	[0.085]	[0.183]	[0.091]	[0.084]	[0.150]	[0.093]	[0.081]	[0.187]	[0.086]
Hispanic	0.158	-0.156	0.241**	0.122	-0.242	0.195*	0.197**	0.029	0.246**
	[0.100]	[0.244]	[0.106]	[0.096]	[0.201]	[0.105]	[0.090]	[0.205]	[0.096]
Black	0.367***	0.242*	0.402***	0.381***	0.238**	0.441***	0.325***	0.205*	0.336***
	[0.065]	[0.134]	[0.070]	[0.063]	[0.108]	[0.071]	[0.061]	[0.120]	[0.066]
Marital Status	-0.132	0.241	-0.16	-0.366***	-0.182	-0.359***	-0.156	0.352	-0.191*
	[0.106]	[0.312]	[0.110]	[0.102]	[0.223]	[0.109]	[0.098]	[0.290]	[0.102]
Family Size	0.011	-0.169***	0.033	-0.001	-0.161***	0.033	0.016	-0.123**	0.040*
	[0.023]	[0.062]	[0.025]	[0.023]	[0.049]	[0.025]	[0.022]	[0.053]	[0.023]
Child Under Age of 2	0.028	-0.174	0.02	0.005	-0.075	-0.037	-0.138*	0.672***	-0.216**
	[0.087]	[0.461]	[0.089]	[0.089]	[0.328]	[0.092]	[0.083]	[0.245]	[0.086]
Urban	-0.127	-0.15	-0.084	-0.004	-0.046	0.063	0.035	0.024	-0.008
	[0.154]	[0.279]	[0.178]	[0.154]	[0.235]	[0.191]	[0.148]	[0.289]	[0.160]
Metro	-0.201	-0.266	-0.119	-0.163	-0.211	-0.042	-0.091	0.054	-0.128
	[0.149]	[0.271]	[0.172]	[0.150]	[0.228]	[0.187]	[0.145]	[0.284]	[0.156]
Less Than High School	0.265***	0.311*	0.240**	0.296***	0.001	0.378***	0.306***	-0.11	0.373***
	[0.087]	[0.167]	[0.097]	[0.083]	[0.141]	[0.094]	[0.080]	[0.169]	[0.086]
High School Graduate	0.140**	-0.003	0.182**	0.059	-0.147	0.145*	0.125*	0.156	0.131*
	[0.067]	[0.124]	[0.076]	[0.066]	[0.104]	[0.078]	[0.065]	[0.117]	[0.072]
Some College	-0.056	-0.069	-0.038	-0.041	-0.014	-0.051	0.044	0.169	0.032
	[0.070]	[0.125]	[0.079]	[0.068]	[0.104]	[0.082]	[0.066]	[0.116]	[0.074]
Wealth in 2007	0	0	0	0	0	-0.000**	0	0	0
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]

Table 5 Continued.									
	2008 and 2010			2006 and 2010			2006 and 2008		
	Unemployed & retired	Retired	Unemployed	Unemployed & retired	Retired	Unemployed	Unemployed & retired	Retired	Unemployed
Own House	-0.159**	0.16	-0.214***	-0.088	0.093	-0.112	-0.172***	0.09	-0.219***
	[0.064]	[0.143]	[0.069]	[0.062]	[0.117]	[0.068]	[0.059]	[0.133]	[0.063]
Own Private Vehicle	-0.328***	-0.007	-0.372***	-0.147*	0.082	-0.187**	-0.390***	0.126	-0.434***
	[0.085]	[0.220]	[0.088]	[0.088]	[0.191]	[0.094]	[0.083]	[0.242]	[0.086]
Manufacturing Sector	0.022	0.014	0.018	0.028	-0.003	0.042	-0.07	0.026	-0.06
	[0.084]	[0.149]	[0.095]	[0.082]	[0.130]	[0.096]	[0.080]	[0.151]	[0.087]
Agriculture	-0.363**	-0.221	-0.371**	-0.275**	-0.257	-0.188	-0.335**	-0.238	-0.292**
	[0.149]	[0.234]	[0.181]	[0.140]	[0.220]	[0.162]	[0.133]	[0.256]	[0.144]
Construction	-0.019	-0.108	-0.001	0.138	0.059	0.155	0.089	0.135	0.077
	[0.108]	[0.201]	[0.122]	[0.103]	[0.166]	[0.119]	[0.097]	[0.185]	[0.107]
Services	-0.041	-0.054	-0.046	-0.01	0.003	-0.052	-0.127**	0.125	-0.203***
	[0.067]	[0.137]	[0.073]	[0.066]	[0.115]	[0.075]	[0.064]	[0.132]	[0.069]
etail Trade	-0.101	-0.14	-0.085	-0.051	-0.006	-0.094	0	0.323**	-0.059
	[0.088]	[0.174]	[0.096]	[0.086]	[0.143]	[0.099]	[0.079]	[0.154]	[0.086]
Public Administration	0.034	0.328*	-0.152	0.195*	0.343**	0.07	-0.142	0.121	-0.201
	[0.111]	[0.178]	[0.139]	[0.105]	[0.151]	[0.131]	[0.108]	[0.176]	[0.127]
North East	-0.159*	-0.184	-0.152	-0.159*	-0.131	-0.153	-0.349***	-0.320**	-0.326***
	[0.085]	[0.143]	[0.099]	[0.082]	[0.125]	[0.099]	[0.082]	[0.151]	[0.091]
South	-0.148**	-0.330***	-0.085	-0.155**	-0.182*	-0.142**	-0.193***	-0.083	-0.207***
	[0.065]	[0.121]	[0.073]	[0.062]	[0.102]	[0.072]	[0.060]	[0.114]	[0.065]
West	-0.153*	-0.193	-0.102	-0.017	-0.087	0.069	-0.086	0.038	-0.128
	[0.080]	[0.140]	[0.092]	[0.075]	[0.119]	[0.088]	[0.073]	[0.132]	[0.081]
Health Status of Head	-0.044	-0.287**	0.087	-0.143**	-0.287***	-0.077	-0.253***	-0.34***	-0.183**
	[0.078]	[0.123]	[0.095]	[0.071]	[0.104]	[0.086]	[0.067]	[0.110]	[0.077]
Smoker	0.228***	0.068	0.262***	0.308***	0.276***	0.301***	0.289***	0.356***	0.244***
	[0.061]	[0.132]	[0.065]	[0.058]	[0.101]	[0.065]	[0.055]	[0.110]	[0.060]
Physical Activity of Head	-0.088*	-0.072	-0.075	-0.099*	-0.209***	-0.029	-0.082	-0.143	-0.064
	[0.053]	[0.097]	[0.060]	[0.052]	[0.081]	[0.061]	[0.050]	[0.091]	[0.055]
Constant	1.757**	2.53	0.538	1.790**	2.601	-0.025	1.545**	1.185	0.195
	[0.693]	[2.540]	[0.843]	[0.721]	[1.814]	[0.890]	[0.680]	[2.530]	[0.802]
N	4422	3994	4222	4129	3674	3784	4627	4078	4421
Log Likelihood	-1635.707	-473.88	-1269.84	-1773.5	-671.088	-1281	-1871.104	-519.764	-1511.28
Pseudo R^2	0.096	0.405	0.085	0.129	0.418	0.084	0.092	0.365	0.09
chi2	345.915	645.623	235.093	524.138	964.637	233.91	378.106	597.768	298.607

Note: Standard errors are reported in the bracket immediately under the estimates. Significance level can be read as * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

Table 6. Comparison of Recreation Participation and Expenditure across 2006(base) and 2008(treatment)

Treatment Group	Matching Method	Participation		Difference in Participation			Trip Expenditure		Difference in Trip Expenditure		
		2008	2010	2006 & 2008	2006 & 2010	2008 & 2010	2008	2010	2006 & 2008	2006 & 2010	2008 & 2010
		Treatment year	*Post treatment year*	*Treatment year*	*Post treatment years*		*Treatment year*	*post treatment year*	*Treatment year*	*Post treatment years*	
Panel a											
Employed in 2006 but Retired or Unemployed in 2008	NN1	-0.116***	-0.114***	-0.104**	-0.102**	0.002	-393.86**	-274.30	-246.403	-280.76	-96.968
		[0.037]	[0.038]	[0.041]	[0.045]	[0.047]	[167.872]	[232.036]	[175.081]	[216.41]	[229.718]
	NN5	-0.078***	-0.121***	-0.064**	-0.11***	-0.043	-151.334	-185.75	-2.253	-138.07	-52.473
		[0.024]	[0.024]	[0.027]	[0.027]	[0.027]	[106.348]	[114.534]	[117.240]	[116.33]	[103.518]
	Radius Caliper	-0.098***	-0.117***	-0.062***	-0.09***	-0.039	-312.43***	-273.91**	-109.68	-122.85	51.453
		[0.019]	0.022	[0.023]	0.025	0.026	[104.000]	134.667	[101.359]	120.74	110.468
Panel b											
Employed in 2006 but Unemployed in 2008	NN1	-0.102**	-0.122***	-0.092*	-0.112**	-0.02	-249.211*	-402.3***	-1.719	-24.20	-31.368
		[0.045]	[0.044]	[0.053]	[0.051]	[0.057]	[140.068]	[147.352]	[181.820]	[166.36]	[169.035]
	NN5	-0.118***	-0.127***	-0.082**	-0.091**	-0.009	-343.70***	-530.4***	-108.178	-387***	-242.73*
		[0.028]	[0.029]	[0.033]	[0.036]	[0.034]	[96.406]	[112.233]	[119.131]	[114.01]	[134.086]
	Radius Caliper	-0.135***	-0.157***	-0.084***	-0.12***	-0.03	-333.44***	-533.6***	-118.908	-275.9**	-101.835
		[0.023]	[0.026]	[0.028]	[0.033]	[0.030]	[92.227]	[116.711]	[106.272]	[112.92]	[123.271]
Panel c											
Employed in 2006 but Retired in 2008	NN1	0.112*	-0.039	0.151**	0	-0.151***	312.854	423.066	215.051	-35.134	-220.753
		[0.063]	[0.059]	[0.065]	[0.062]	[0.057]	[357.765]	[361.443]	[299.687]	[363.06]	[305.304]
	NN5	0.054	-0.054	0.045	-0.063	-0.107**	-90.168	-101.649	415.905	103.7	110.234
		[0.041]	[0.040]	[0.044]	[0.040]	[0.047]	[276.097]	[318.800]	[289.615]	[289.15]	[235.404]
	Radius Caliper	0.023	-0.057	0.02	-0.046	-0.091*	-205.009	145.888	21.868	221.025	320.572
		[0.036]	[0.040]	[0.041]	[0.044]	[0.048]	[309.553]	[379.661]	[255.212]	[342.59]	[234.424]

Note: Standard errors are reported in the bracket immediately under the estimates. Significance level can be read as * p≤0.05, ** p≤0.01, *** p≤0.001. *Standard errors for nearest neighbor matching estimates are Abadie-Imbens standard errors while those for radius matching are obtained from 1000 bootstrapped sample.*

Table 7. Comparison of Recreation Participation and Expenditure across 2006(base) and 2010(treatment)

Treatment Group	Matching Method	Participation		Difference in Participation			Trip Expenditure		Difference in Trip Expenditure		
		2008	2010	2006 & 2008	2006 & 2010	2008 & 2010	2008	2010	2006 & 2008	2006 & 2010	2008 & 2010
		Non treatment years	Treatment year	Non treatment years	Treatment year	Non treatment years	Non treatment year	Treatment year	Non treatment years	Treatment year	Non treatment years
		Panel a									
Employed in 2006 but Retired or Unemployed in 2010	NN1	-0.109***	-0.120***	-0.072*	-0.060*	-0.01	-13.209	-271.585	-9.943	-123.797	1.325
		[0.031]	[0.032]	[0.037]	[0.036]	[0.035]	[182.087]	[313.230]	[202.073]	[272.61]	[174.806]
	NN5	-0.089***	-0.117***	-0.046*	-0.07***	-0.042*	-3.235	-96.387	189.439	128.256	-86.947
		[0.022]	[0.022]	[0.025]	[0.025]	[0.025]	[149.986]	[141.165]	[183.368]	[147.59]	[145.703]
	Radius Caliper	-0.063***	-0.121***	-0.022	-0.072***	-0.054**	-83.539	-89.699	109.718	103.558	-6.161
		[0.022]	[0.020]	[0.024]	[0.024]	[0.026]	[156.971]	[137.559]	[187.467]	[164.29]	[154.932]
		Panel b									
Employed in 2006 but Unemployed in 2010	NN1	-0.046	-0.173***	0.023	-0.088*	-0.06	-161.41	-659.52***	259.90	-73.83	-304.45
		[0.044]	[0.044]	[0.048]	[0.050]	[0.047]	[219.713]	[246.897]	[269.864]	[202.62]	[228.686]
	NN5	-0.087***	-0.146***	-0.006	-0.086***	-0.074**	32.67	-367.69***	216.03	-177.38	-449.4**
		[0.029]	[0.027]	[0.033]	[0.033]	[0.033]	[180.701]	[107.460]	[186.374]	[123.75]	[192.373]
	Radius Caliper	-0.101***	-0.150***	-0.027	-0.087***	-0.065**	-142.80	-359.56***	182.27	-115.769	-306.17
		[0.028]	[0.025]	[0.033]	[0.032]	[0.032]	[203.247]	[81.731]	[217.296]	[98.530]	[209.108]
		Panel c									
Employed in 2006 but Retired in 2010	NN1	-0.026	-0.066	-0.044	-0.092*	-0.038	74.574	-19.897	113.429	-109.034	142.508
		[0.041]	[0.046]	[0.038]	[0.051]	[0.051]	[291.852]	[359.524]	[389.236]	[300.02]	[290.977]
	NN5	-0.076**	-0.072**	-0.056*	-0.060*	0.013	21.927	36.169	214.249	208.863	127.538
		[0.030]	[0.036]	[0.033]	[0.036]	[0.041]	[201.244]	[326.022]	[231.044]	[294.97]	[212.952]
	Radius Caliper	-0.063**	-0.075**	-0.027	-0.047	-0.016	-117.587	94.48	106.039	335.911	183.825
		[0.032]	[0.032]	[0.035]	[0.036]	[0.038]	[248.241]	[291.276]	[317.993]	[331.03]	[267.298]

Note: Standard errors are reported in the bracket immediately under the estimates. Significance level can be read as * p≤0.05, ** p≤0.01, *** p≤0.001. *Standard errors for nearest neighbor matching estimates are Abadie-Imbens standard errors while those for radius matching are obtained from 1000 bootstrapped sample.*

Table 8. Comparison of Recreation Participation and Expenditure across 2008(base) and 2010(treatment)

Treatment Group	Matching Method	Participation		Trip Expenditure		Trip Expenditure		Difference in Trip Expenditure	
		2008	2010	2006 & 2010	2008 & 2010	2008	2010	2006 & 2010	2008 & 2010
		Base year	Treatment year	Not treatment years	Treatment year	Base year	Treatment year	Not treatment years	Treatment year
Employed in 2008 but Retired or Unemployed in 2010	NN1	-0.059	-0.081**	-0.013	-0.022	263.879	-62.591	-25.955	-326.469
		[0.037]	[0.038]	[0.050]	[0.042]	[252.256]	[192.955]	[274.738]	[254.523]
	NN5	-0.035	-0.096***	-0.036	-0.060**	242.295	-129.698	115.822	-371.993*
		[0.023]	[0.023]	[0.028]	[0.027]	[229.072]	[135.423]	[169.486]	[211.079]
	Radius Caliper	-0.053**	-0.093***	-0.03	-0.04	131.356	-137.135	80.9	-268.49
		[0.023]	[0.024]	[0.028]	[0.027]	[273.182]	[129.464]	[151.515]	[238.321]
Employed in 2008 but Unemployed in 2010	NN1	-0.126***	-0.110**	-0.068	0.016	-108.218	-132.798	-143.353	-24.58
		[0.042]	[0.045]	[0.049]	[0.043]	[223.843]	[135.466]	[150.989]	[217.642]
	NN5	-0.081***	-0.130***	-0.062*	-0.049	-90.305	-362.187***	-136.231	-271.881
		[0.028]	[0.028]	[0.037]	[0.034]	[180.206]	[90.922]	[121.096]	[179.079]
	Radius Caliper	-0.065**	-0.141***	-0.025	-0.077**	-34.534	-422.120***	-127.352	-387.585*
		[0.027]	[0.026]	[0.034]	[0.034]	[210.916]	[98.402]	[119.372]	[220.930]
Employed in 2008 but Retired in 2010	NN1	-0.015	-0.02	0.051	-0.005	448.427	-177.295	443.974	-625.722
		[0.064]	[0.056]	[0.061]	[0.081]	[712.435]	[643.834]	[391.704]	[705.144]
	NN5	-0.064*	-0.019	-0.057	0.045	503.345	251.691	313.659	-251.655
		[0.033]	[0.039]	[0.044]	[0.046]	[602.634]	[353.388]	[406.265]	[509.426]
	Radius Caliper	-0.044	-0.028	-0.041	0.016	384.959	167.094	334.378	-217.865
		[0.038]	[0.042]	[0.043]	[0.045]	[648.075]	[386.700]	[400.686]	[574.775]

Note: Standard errors are reported in the bracket immediately under the estimates. Significance level can be read as * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. *Standard errors for nearest neighbor matching estimates are Abadie-Imbens standard errors while those for radius matching are obtained from 1000 bootstrapped sample.*

Table 9. Estimates from Placebo Exercises & Testing the Parallel Trend Assumption

Treatment Group	Matching Method	Participation			Difference in Participation			Trip Expenditure			Difference in Trip Expenditure		
	Analysis across	2006 vs. 2008	2006 vs. 2010	2008 vs. 2010	2006 vs. 2008	2006 vs. 2010	2008 vs. 2010	2006 vs. 2008	2006 vs. 2010	2008 vs. 2010	2006 vs. 2008	2006 vs. 2010	2008 vs. 2010
		Placebo year is 2004	Placebo year is 2004	Placebo year is 2006	Placebo across 2004 & 2006	Placebo across 2004 & 2006	Placebo across 2006 & 2008	Placebo year is 2004	Placebo year is 2004	Placebo year is 2006	Placebo across 2004 & 2006	Placebo across 2004 & 2006	Placebo across 2006 & 2008
Treatment including both Retired and Unemployed	NN1	-0.032	**-0.12*****	-0.032	-0.02	**-0.085****	0.019	**-1019.6****	-212.065	1.956	**-859.30***	-324.591	477.962
		[0.035]	[0.032]	[0.039]	[0.040]	[0.039]	[0.038]	[501.612]	[223.359]	[185.336]	[478.191]	[236.536]	[304.564]
	NN5	-0.018	**-0.08*****	**-0.05****	-0.004	-0.018	-0.004	-95.325	-59.514	-186.014	-43.597	187.148	391.764
		[0.024]	[0.023]	[0.023]	[0.028]	[0.026]	[0.026]	[123.973]	[147.509]	[145.783]	[143.412]	[180.740]	[274.870]
	Radius	-0.020	**-0.07*****	**-0.06*****	0.001	-0.023	0.01	-133.366	-173.239	**-218.03***	17.686	20.018	349.39
		0.021	[0.022]	[0.022]	0.024	[0.025]	[0.024]	109.852	[183.196]	[128.813]	134.611	[208.091]	[285.644]
Treatment including Unemployed	NN1	-0.055	**-0.16*****	-0.074	-0.045	**-0.123****	0.053	-232.767	**-365.97***	10.936	311.052	-86.22	230.856
		[0.045]	[0.045]	[0.046]	[0.053]	[0.063]	[0.045]	[186.776]	[215.249]	[132.397]	[207.892]	[185.560]	[246.478]
	NN5	-0.037	**-0.10*****	**-0.08*****	-0.001	-0.033	0.027	-142.247	-43.53	-165.394	193.506	**295.78***	110.505
		[0.028]	[0.028]	[0.031]	[0.033]	[0.035]	[0.034]	[121.692]	[115.657]	[112.151]	[128.789]	[177.674]	[230.110]
	Radius	**-0.046***	**-0.11*****	**-0.12*****	-0.009	-0.039	0.051	-123.571	-168.99	**-294.7*****	134.167	156.08	260.233
		[0.026]	[0.028]	[0.028]	[0.031]	[0.033]	[0.032]	[111.301]	[116.727]	[100.972]	[121.376]	[135.871]	[228.023]
Treatment including Retired	NN1	0.006	0.018	-0.056	0.045	-0.036	0.026	-632.172	129.129	-36.075	-585.273	-128.28	739.115
		[0.058]	[0.051]	[0.055]	[0.052]	[0.058]	[0.058]	[476.269]	[288.921]	[475.525]	[613.049]	[362.778]	[680.088]
	NN5	0.022	-0.041	0.043	0.013	-0.02	**-0.109*****	-162.047	-303.398	-69.817	69.527	9.152	572.478
		[0.037]	[0.030]	[0.034]	[0.043]	[0.033]	[0.039]	[217.074]	[331.281]	[322.621]	[268.748]	[393.729]	[628.406]
	Radius	0.023	-0.044	0.013	0.034	-0.008	-0.057	-202.873	-176.508	-167.284	-127.736	47.119	552.243
		[0.039]	[0.031]	[0.036]	[0.045]	[0.034]	[0.039]	[243.004]	[300.359]	[360.570]	[364.328]	[363.824]	[674.970]

Note: Standard errors are reported in the bracket immediately under the estimates. Significance level can be read as * p≤0.05, ** p≤0.01, *** p≤0.001. *Standard errors for nearest neighbor matching estimates are Abadie-Imbens standard errors while those for radius matching are obtained from 1000 bootstrapped sample.*

Table 10. Doubly Robust Estimators for Unemployed Treatment Across all Comparison Years

	Participation		Difference in Participation		Trip Expenditure		Difference in Trip Expenditure	
Analysis Across Base Year 2006 and Treatment Year 2008								
	2004 (Placebo)	*2008 (Treatment)*	*2004 & 2006 (placebo)*	*2006 & 2008 (Base vs. treatment)*	*2004 (Placebo)*	*2008 (Treatment)*	*2004 & 2006 (placebo)*	*2006 & 2008 (Base vs. treatment)*
Employed in 2006 but Retired or Unemployed in 2008	-0.119	-0.638***	0.018	-0.093***	-236.96**	-277.953**	-40.655	-96.134
	[0.176]	[0.138]	[0.034]	[0.028]	[118.531]	[116.707]	[139.877]	[109.749]
Analysis Across Base Year 2006 and Treatment Year 2010								
	2004 (Placebo)	*2008 (Treatment)*	*2004 & 2006 (placebo)*	*2006 & 2008 (Base vs. Treatment)*	*2004 (Placebo)*	*2008 (Treatment)*	*2004 & 2006 (placebo)*	*2006 & 2008 (Base vs. Treatment)*
Employed in 2006 but Retired or Unemployed in 2010	-0.452**	-0.778***	-0.024	-0.092***	-33.06	-342.557***	162.47	-123.50
	[0.226]	[0.166]	[0.037]	[0.030]	126.47	[89.347]	[162.382]	[106.130]
Analysis Across Base Year 2008 and Treatment Year 2010								
	2006 (Placebo)	2010 (Treatment)	2006 & 2008 (placebo)	2008 & 2010 (Base vs. Treatment)	2006 (Placebo)	2010 (Treatment)	2006 & 2008 (placebo)	2008 & 2010 (Base vs. Treatment)
Employed in 2008 but Retired or Unemployed in 2010	-0.509**	-0.695***	0.051	-0.078**	-146.165	-352.134***	46.444	-262.701
	[0.221]	[0.156]	[0.034]	[0.032]	[120.898]	[82.439]	[215.896]	[161.423]

Note: Standard errors are reported in the bracket immediately under the estimates. Significance level can be read as * p≤0.05, ** p≤0.01, *** p≤0.001. *Standard errors for nearest neighbor matching estimates are Abadie-Imbens standard errors while those for radius matching are obtained from 1000 bootstrapped sample. Estimates from the placebo exercises under "differences in participation" and "differences in trip expenditure" heading will test if the parallel trend assumption is satisfied*

APPENDIX B. APPENDIX TO CHAPTER 3

Table A1: Balancing of Covariates, 2006 and 2008, Treatments are Unemployed and Retired Group

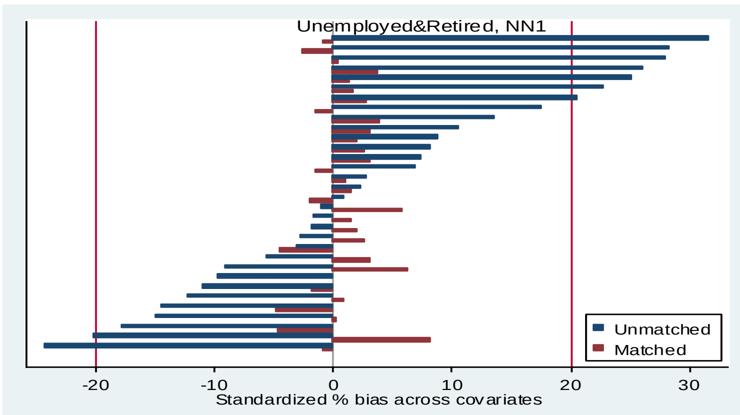
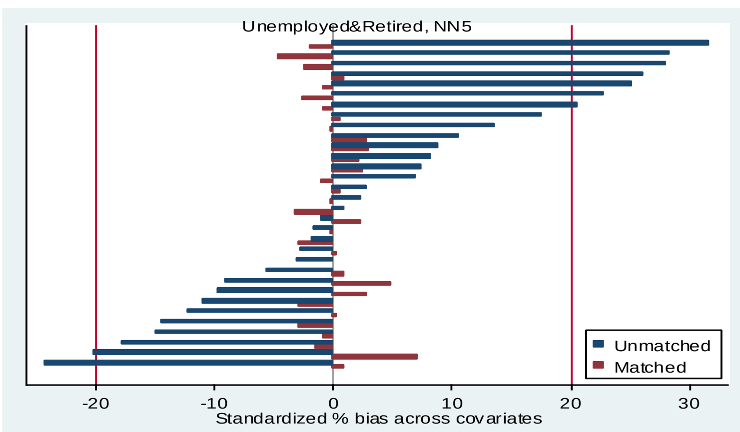
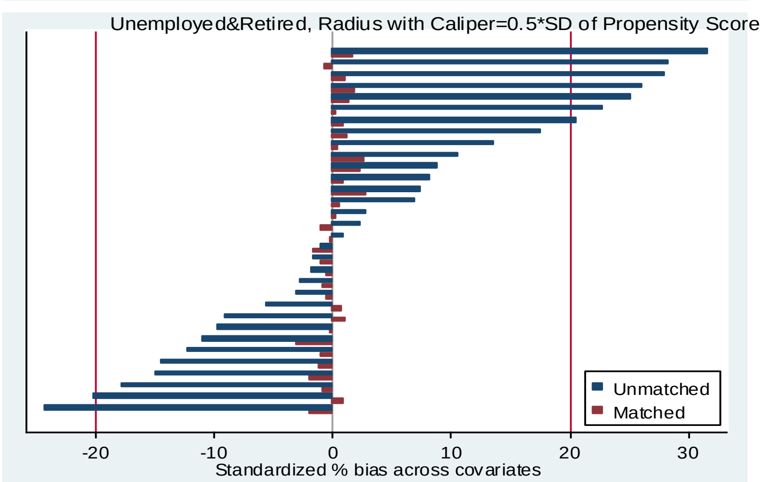
Treatment Year is 2008 and Baseline Year is 2006	Mean Bias Unmatched	Mean Bias Matched	Median Bias Unmatched	Median Bias Matched
	12.9	2.6	10.7	2
	12.9	1.9	10.7	1.8
	12.9	1.2	10.7	1

Table A2: Balancing of Covariates, 2006 and 2008, Treatments are Unemployed Group

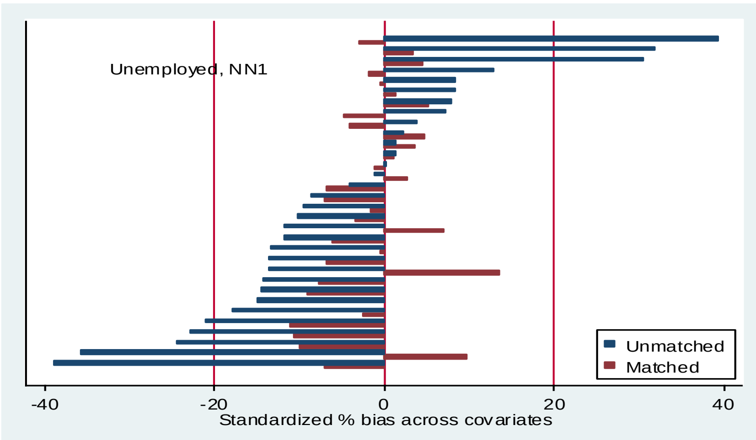
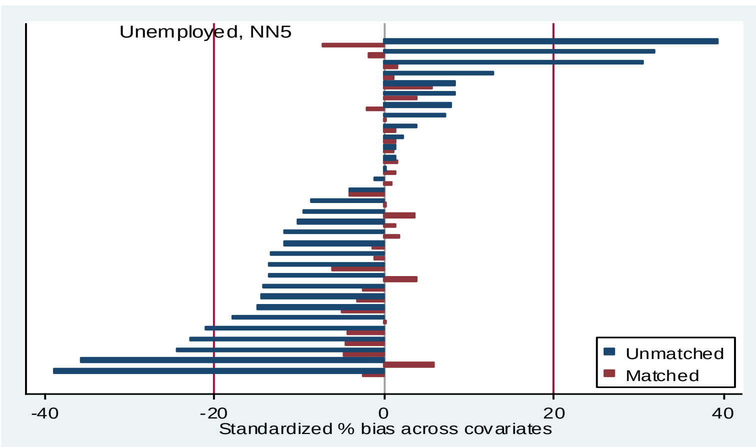
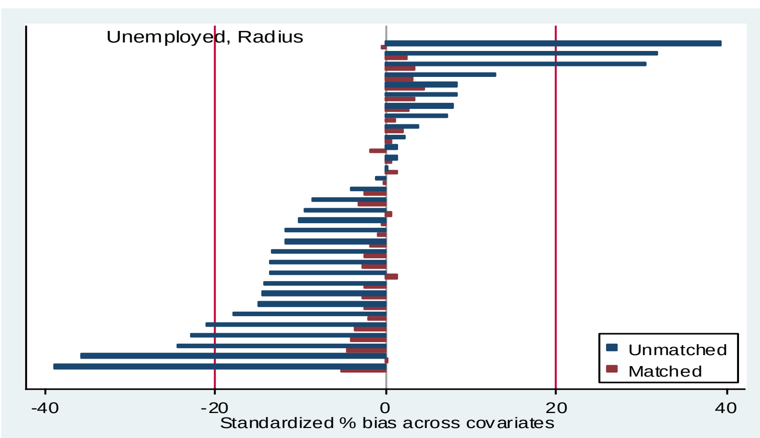
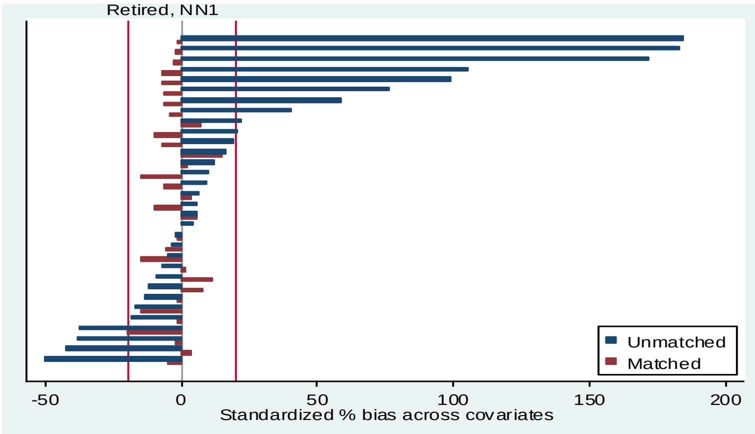
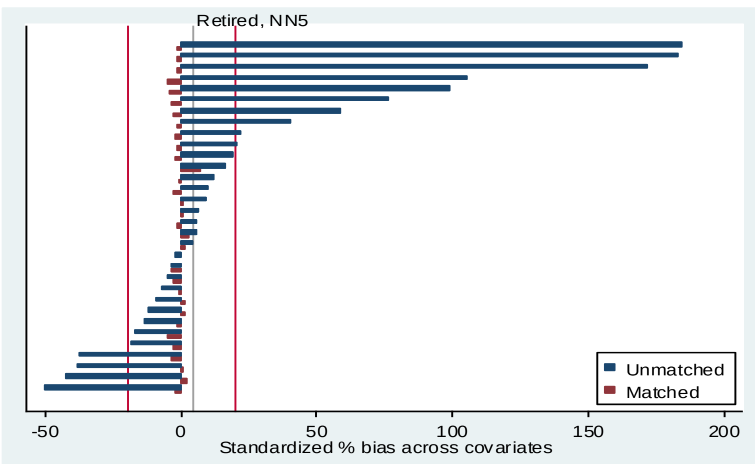
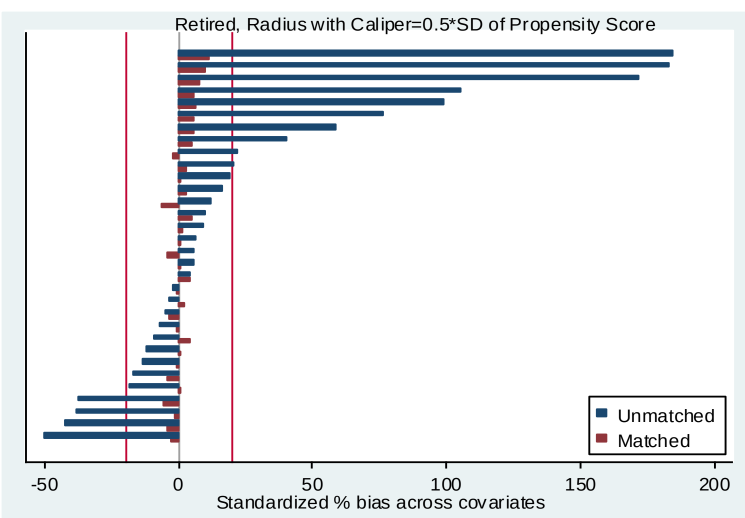
Treatment Year is 2008 and Baseline Year is 2006	Mean Bias		Median Bias	
	Unmatched	Matched	Unmatched	Matched
	14.3	5.1	12.4	4.7
	14.3	2.7	12.4	1.9
	14.3	2.1	12.4	2

Table A3: Balancing of Covariates, 2006 and 2008, Treatments are Retired Group

Treatment Year is 2008 and Baseline Year is 2006







	Mean Bias		Median Bias	
	Unmatched	Matched	Unmatched	Matched
Retired, NN1	40.9	6.5	17.7	5.9
Retired, NN5	40.9	2.2	17.7	1.8
Retired, Radius with Caliper=0.5*SD of Propensity Score	40.9	3.7	17.7	3.7

Table A4: Balancing of Covariates, 2006 and 2010, Treatments are Unemployed and Retired Group

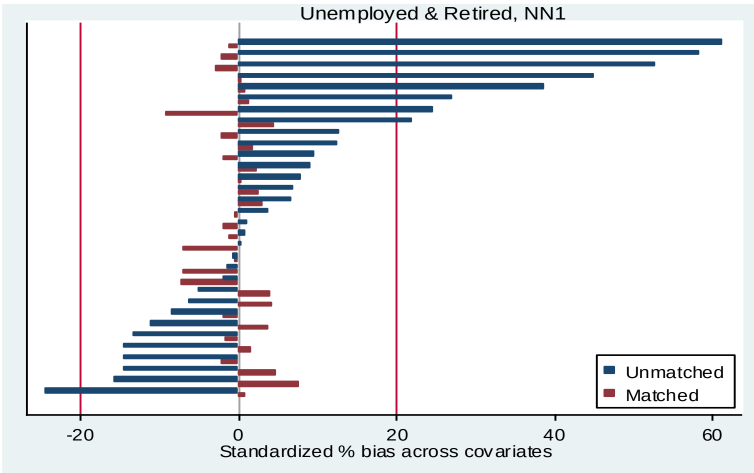
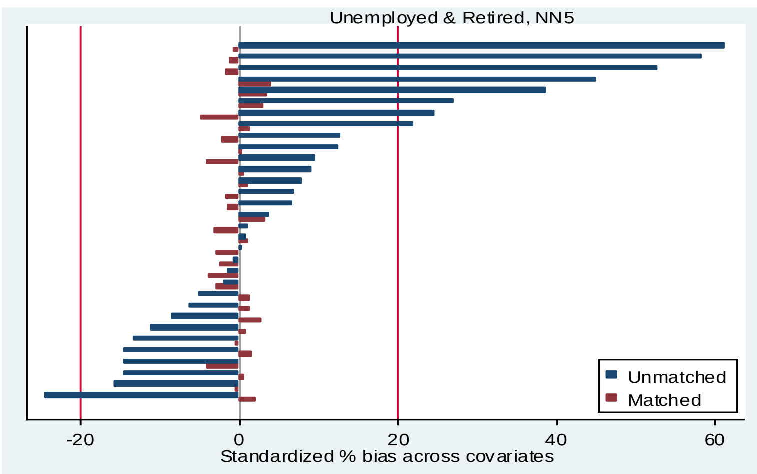
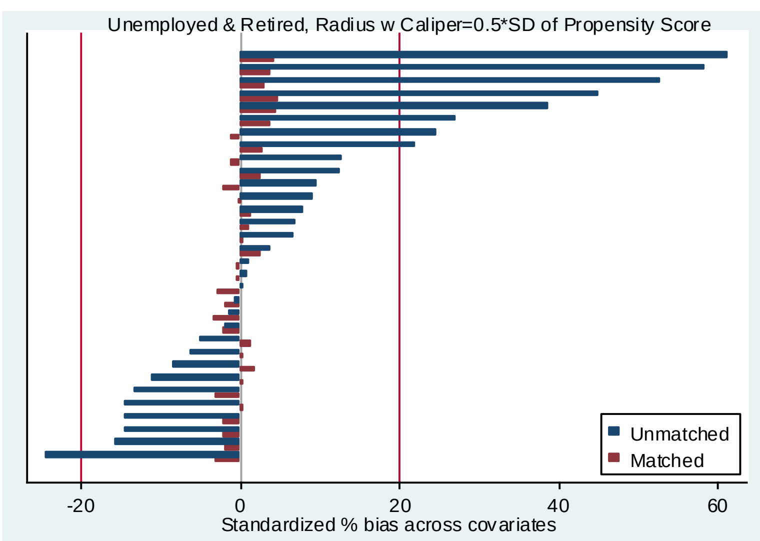
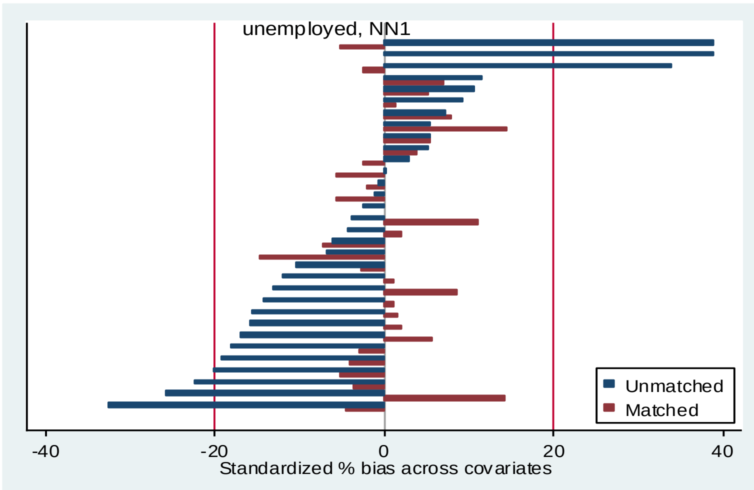
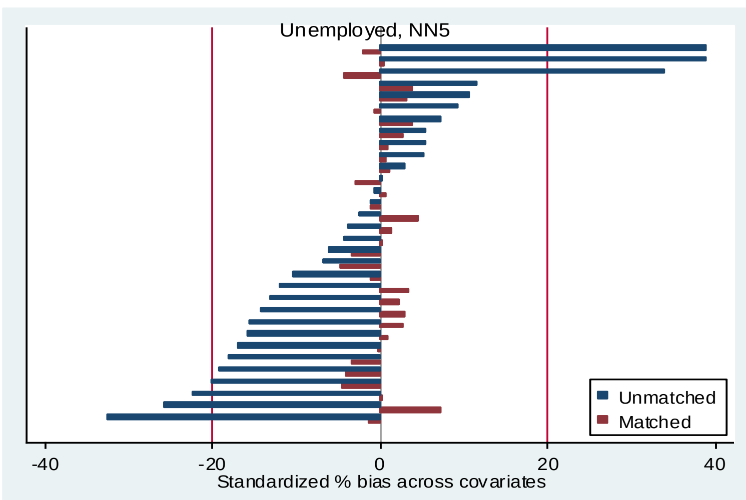
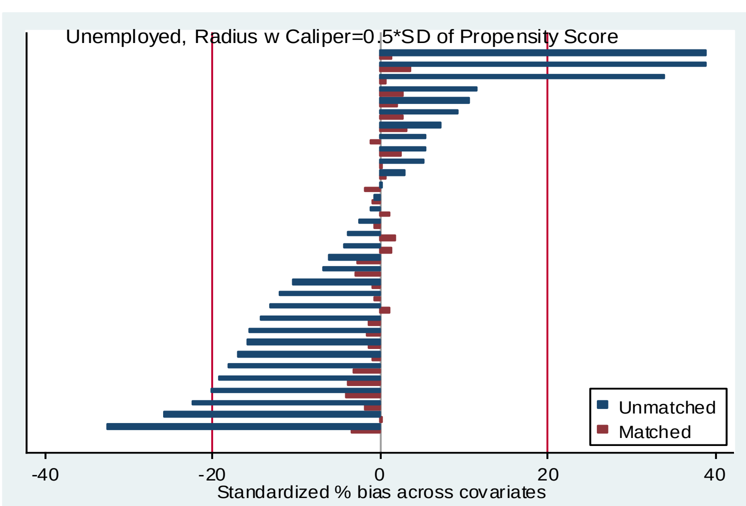
Treatment Year is 2010 and Baseline Year is 2006	Mean Bias		Median Bias	
	Unmatched	Matched	Unmatched	Matched
	16.6	2.9	11.7	2.1
	16.6	2.1	11.7	1.7
	16.6	2.1	11.7	2.1

Table A5: Balancing of Covariates, 2006 and 2010, Treatments are Unemployed Group

Treatment Year is 2010 and Baseline Year is 2006



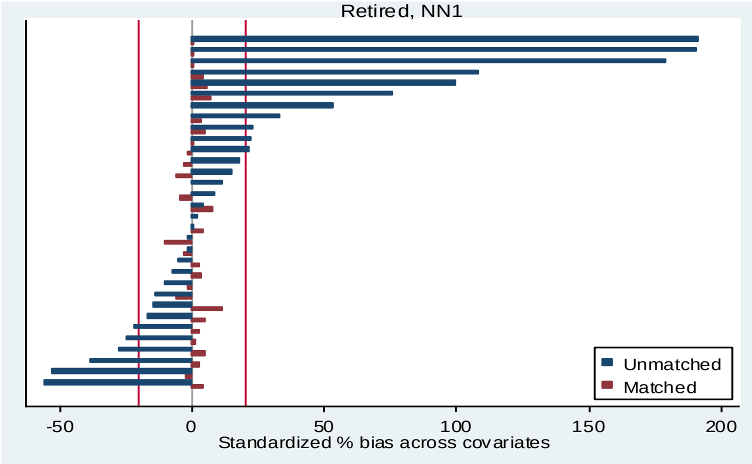


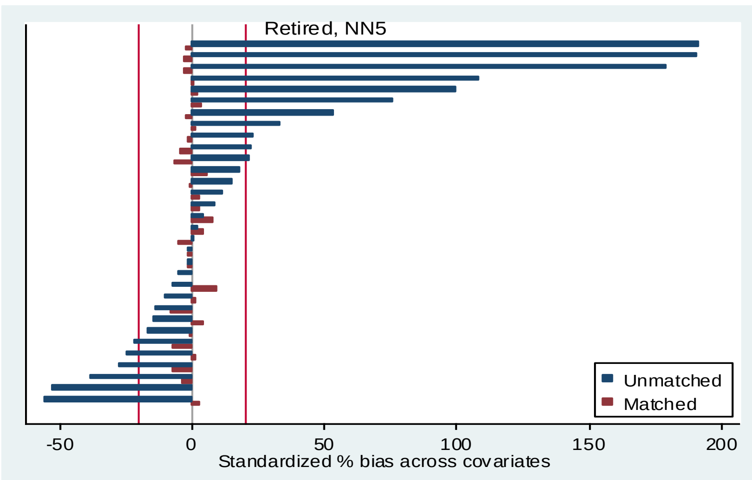


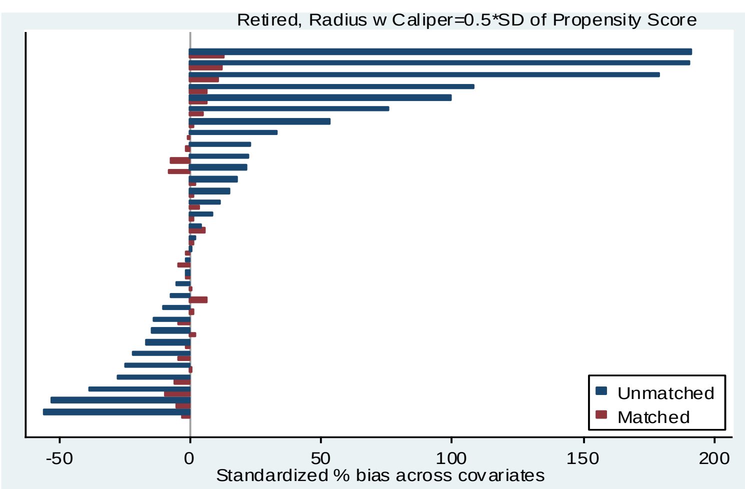
Mean Bias		Median Bias	
Unmatched	Matched	Unmatched	Matched
13.5	5.1	11.2	4.3
13.5	2.5	11.2	2.5
13.5	1.9	11.2	1.6

Table A6: Balancing of Covariates, 2006 and 2010, Treatments are Retired Group

Treatment Year is 2010 and Baseline Year is 2006



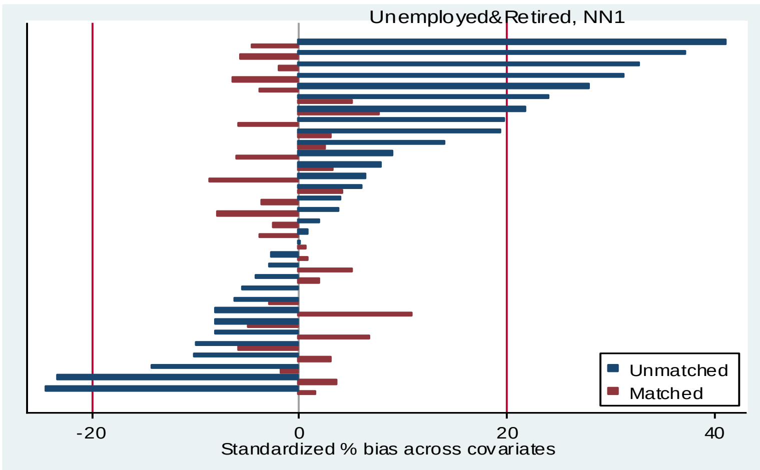




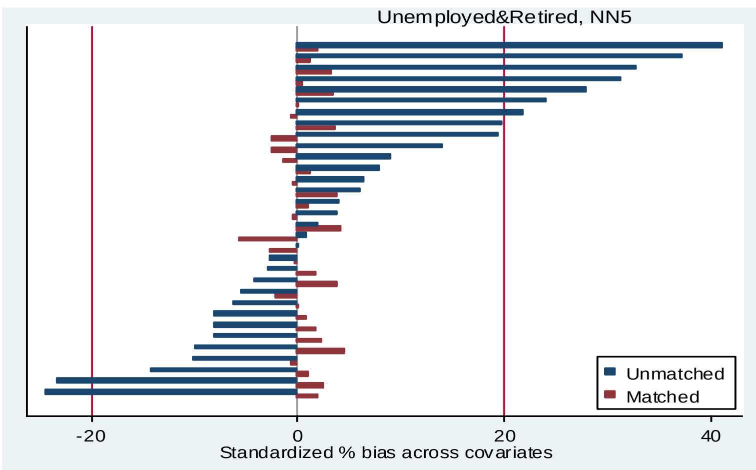
	Mean Bias		Median Bias	
	Unmatched	Matched	Unmatched	Matched
Retired, NN1	42.3	3.7	21.7	3.6
Retired, NN5	42.3	3.4	21.7	3
Retired, Radius w Caliper=0.5*SD of Propensity Score	42.3	4.4	21.7	4

Table A7: Balancing of Covariates, 2008 and 2010, Treatments are Unemployed and Retired Group

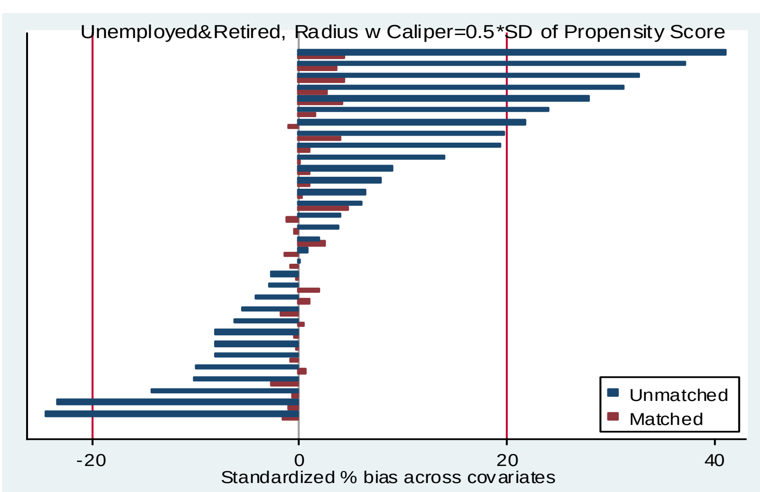
Treatment Year is 2010 and Baseline Year is 2008	Mean Bias		Median Bias	
	Unmatched	Matched	Unmatched	Matched



	Mean Bias Unmatched	Mean Bias Matched	Median Bias Unmatched	Median Bias Matched
NN1	13.6	4.2	8.5	3.8



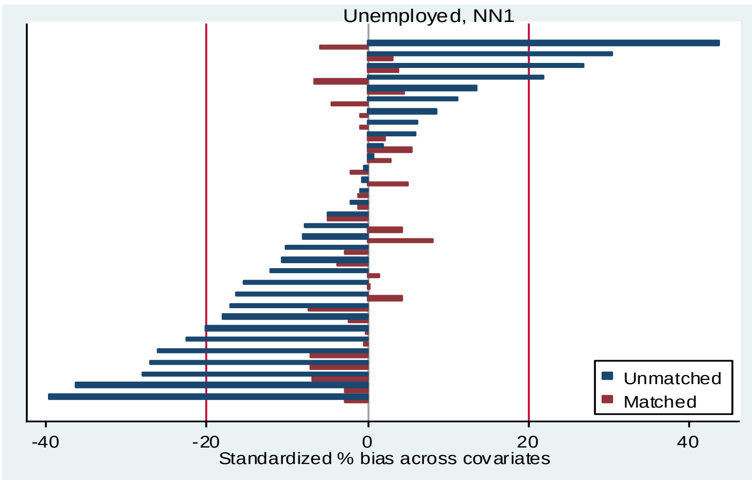
	Mean Bias Unmatched	Mean Bias Matched	Median Bias Unmatched	Median Bias Matched
NN5	13.6	2	8.5	1.9

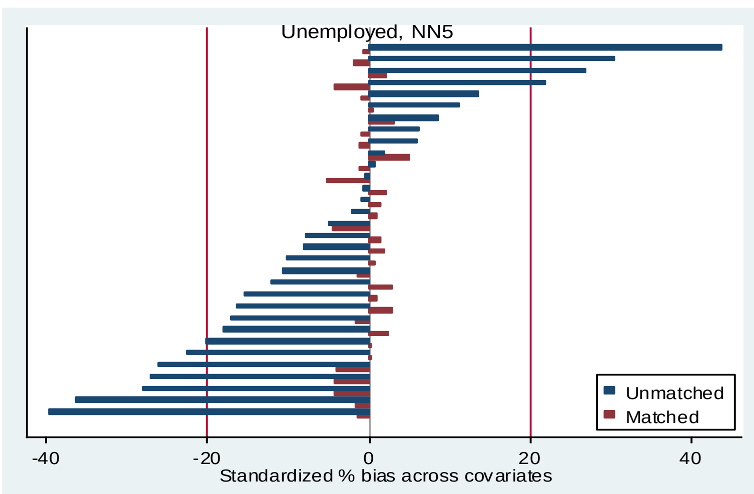


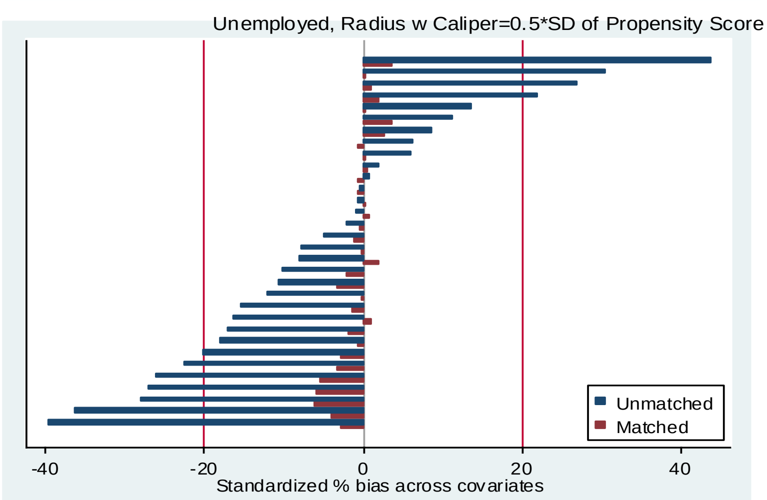
	Mean Bias Unmatched	Mean Bias Matched	Median Bias Unmatched	Median Bias Matched
Radius	13.6	1.7	8.5	1.1

Table A8: Balancing of Covariates, 2008 and 2010, Treatments are Unemployed Group

Treatment Year is 2010 and Baseline Year is 2008







	Mean Bias		Median Bias	
	Unmatched	Matched	Unmatched	Matched
	15.5	3.7	12.9	3.5
	15.5	2.2	12.9	1.6
	15.5	1.9	12.9	1.3

Table A9: Balancing of Covariates, 2008 and 2010, Treatments are Retired Group

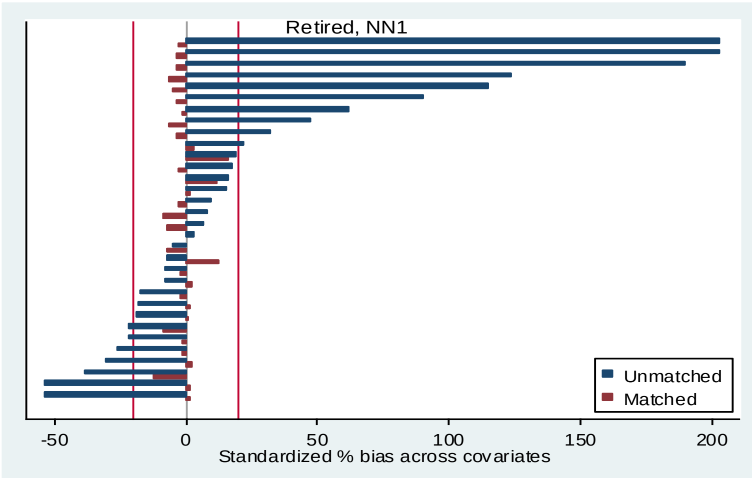
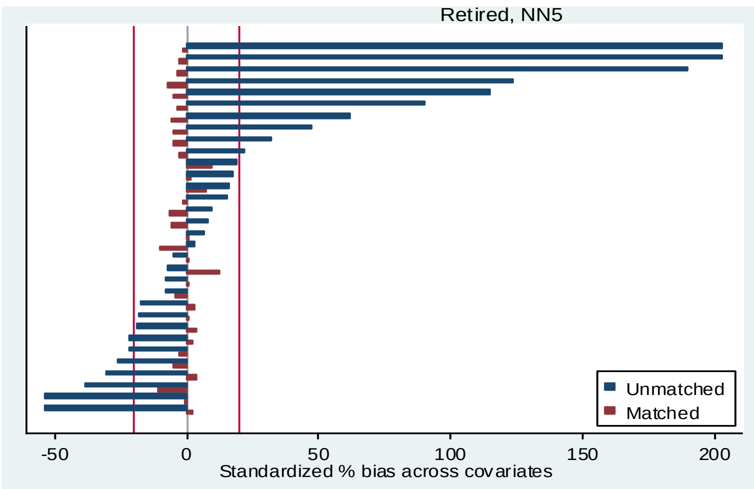
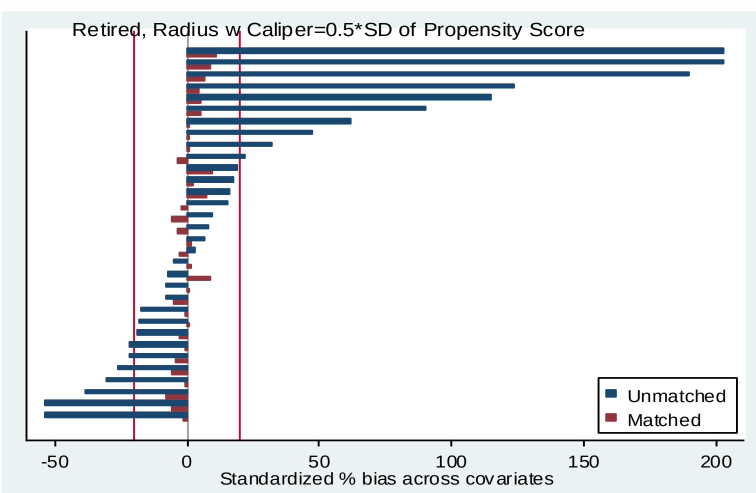
Treatment Year is 2010 and Baseline Year is 2008	Mean Bias Unmatched	Mean Bias Matched	Median Bias Unmatched	Median Bias Matched
	47.2	4.7	21.6	3
	47.2	4.3	21.6	3.7
	47.2	4	21.6	3.2

Table A10: Validity of Common Support Condition, Number of Matched and Unmatched Treatments.

	Unemployed and Retired (group 1)			Unemployed (Group 2)			Retired (Group 3)		
	(a) Analysis on 2006 and 2008								
	Unmatched	matched	Total	Unmatched	matched	Total	Unmatched	matched	Total
Nearest Neighbor with replacement	0	756	756	0	550	550	6	201	207
Nearest 5 Neighbors with replacement	0	756	756	0	550	550	6	201	207
Radius Matching (caliper =0.5*SD)	0	756	756	0	550	550	6	201	207
	(b) Analysis on 2006 and 2010								
	Unmatched	matched	Total	Unmatched	matched	Total	Unmatched	matched	Total
Nearest Neighbor with replacement	2	457	459	1	803	804	1	348	349
Nearest 5 Neighbors with replacement	2	457	459	1	803	804	1	348	349
Radius Matching (caliper =0.5*SD)	2	457	459	1	803	804	1	348	349
	(c) Analysis on 2008 and 2010								
	Unmatched	matched	Total	Unmatched	matched	Total	Unmatched	matched	Total
Nearest Neighbor with replacement	1	428	429	0	629	629	1	200	201
Nearest 5 Neighbors with replacement	1	428	429	0	629	629	1	200	201
Radius Matching (caliper =0.5*SD)	1	428	429	0	629	629	1	200	201

CHAPTER 4. HOW LIFE EXPECTANCY AT BIRTH AFFECTS SCHOOLING INVESTMENT AND LIFETIME EARNINGS: EVIDENCE FROM CROSS-COUNTRY HOUSEHOLD SURVEYS

1. Introduction

Improvements in nutrition, education, transportation, sanitation and knowledge of diseases have had a dramatic effect on life expectancy worldwide. Vallin and Meslé (2009) estimated that the various improvements in knowledge and technology increased potential life expectancy at birth by 0.33 years per year from 1885 to 1960 and by 0.2 years per year thereafter. A woman born in 2000 was expected to live 33 years longer than a woman born in 1885.[40] This 63% increase in expected length of life was also accompanied by improved health and enhanced physical ability to work which should have profound effects on life-cycle investments of time in the acquisition of skills, the application of skills to the labor market, and the ability to consume leisure.

Ben-Porath (1967) and Becker (1993) showed that increased length of life would have an unambiguous effect toward increased investment in human capital. Both formulations limited agents to choices of working versus acquiring additional human capital, ignoring the possibility that individuals would consume more leisure rather than spending more time working or learnings. Heckman (1976) extended the model to allow both time and financial investments into human capital production and to allow agents to choose labor supply and the consumption of goods and leisure over the life cycle. Still, his model predicts that increased life expectancy at birth would cause individuals to increase their lifetime human capital production.

[40] Vallin and Meslé (2009) focused on women's life expectancy to avoid the effects of war and the higher probability of violent or accidental deaths in their analysis of vital statistics. Potential life expectancy is based on the highest country life expectancy in each of the years they evaluated between 1750 and 2000.

The link between health and human capital investment has been examined intensively in previous research and at different points in the life-cycle. Shocks to fetal or infant health such as maternal malnutrition (Field et al. 2009; Maluccio et al. 2009; Almond and Majumder 2011); low birth-weight (Behrman and Rosenzweig 2004); excessive or insufficient rainfall during the first year of childhood (Maccini and Yang 2009; Shah and Steinberg 2014); exposure to diseases (Almond 2006; Currie and Vogl 2013); or famine during early childhood (Almond and Currie 2011; Gorgens et al. 2012) have all been shown to lower educational investments or returns. Exogenous shocks from exposure to environmental pollution or hazards (Chay and Greenstone 2003; Foster et al. 2009; Jayachandran 2009; and Almond et al. 2009); or exposure to violence or civil war (Akresh et al. 2011; Camacho 2008; Blattman and Annan 2010; Leon 2012; and Yuksel 2014) have also reduced educational attainment and labor market earnings.

While the link between these shocks might be due to health outcomes or to related income shocks, other studies have been able to isolate the effects of health on human capital investments. Exploiting variation in the timing and intensity of hookworm and malaria eradication in the American South and across developing countries, Bleakley (2007, 2010a) and Lucas (2010) demonstrated improved education and earnings outcomes for populations with the earliest compared to later exposure to the public health intervention. Miguel and Kremer (2004) found that children who received treatment for intestinal worms were absent 25% less than students in schools that were randomly assigned to receive the treatment 2-3 years later. Follow-up surveys reveal that compared to the children in the control schools, children under treatment worked 13% more hours and earned 20-29% more (Karlan and Appel 2011). Bhalotra and Venkataramani (2012) showed that the availability of antibiotics to combat pneumonia at the time of infancy increased education and earnings when the infants reached adulthood.

A few studies have examined the role of life expectancy on education. Jayachandran and Lleras-Muney (2009) used improvements in maternal health to instrument for presumed endogenous life expectancy of their children, and found that an additional year of expected life increased schooling by 0.11 years. Oster (2013) used information on when individuals learned that they had the fatal Huntington's disease as the life cycle shock, and found that schooling increased by 0.17 years for every additional year of expected life.

Studies that examine the effect of increased life expectancy on human capital investment have had more mixed results. Acemoglu and Johnson (2006, 2007) found no effect of increasing life expectancy due to improved control of infectious disease on schooling. They argued that because cohort size increases with improved length of life, returns to human capital may fall due to rising labor supply outpacing any growth in demand for skills. Bloom, Canning and Fink (2014) found the Acemoglu and Johnson result reverses when controls for initial health are added. Hazan (2009) argued that the Ben Porath model required an increase in lifetime labor supply for the gain in life expectancy to increase investment in human capital. He found that American men born between 1840 and 1970 actually reduced lifetime labor supply, from which he concluded that life expectancy has either a negligible or possibly even a negative effect on investments in education. His subsequent analysis (Hazan 2012) found no correlation between life expectancy at age 5 and schooling.[41] However, other analyses of similar country-level data still retain the positive correlation between life expectancy and schooling (Cervellati and Sunde 2013; Hansen 2013; and Cohen and Leker 2014).

[41]Following Soares (2005), Hazan included "post demographic transition countries" which are basically group of countries which exhibited life expectancy at birth above 50 in 1960. He preferred life expectancy at 5 or 10 instead of that at birth since the later displayed widespread variability due to high infant mortality.

This paper makes several improvements over the previous cross-country studies of life expectancy on human capital investments. First, it is based on a much larger and broader set of 111 developed and developing countries. The analysis is conducted at the individual cohort level so that there is a one-to-one correspondence between year of birth in a country and the corresponding life expectancy at birth. Exact measures of years of schooling by cohort were generated from household surveys for each country in place of the noisier approximations based on school enrollment data, extrapolated estimates for missing values, and *ex post* adjustments for mortality that were used in previous studies. Estimates are reported separately for men and women and for urban and rural residents to establish the robustness of the findings. Estimates of parental life expectancy at birth are incorporated into the analysis to examine evidence of the intergenerational transmission of human capital from parent to child. Finally, we incorporate estimate of the impact of life expectancy on both lifetime years of schooling and lifetime earnings. We find that an additional year of life expected at time of birth increases years of schooling by about 0.12 years and increases earnings by about 1%. The implied Wald estimate of returns to schooling are 9.9% for men, 4.3% for women, 10.2% for urban residents and 2.9% for rural residents. These cross-country results are very consistent with the findings based on individual data.

The next section applies Heckman's (1976) model of life-cycle earnings, learnings and consumption to the question of how increased life expectancy at birth will affect lifetime schooling and earnings. The next section is on how we utilize these implications to derive the reduced form specifications for our empirical exercises. Section four elaborates on the data sources while section five specifies the empirical model. Section six reports findings and presents some robustness tests. The final section discusses, interprets the findings, and draws some concluding remarks.

2. Theoretical Framework

Heckman (1976) developed a life cycle model of earnings, learnings, and consumption by merging the theory of labor supply with that of human capital production. His model relaxes several assumptions of the Ben-Porath (1967) model that are important to our analysis, including that labor supply decisions are made endogenous, that budgets can be used to consume leisure and invest in human capital as well as to purchase market goods and services, and that initial endowments of assets and human capital will alter the entire trajectory of consumption, investment and labor supply. This study uses the Heckman's framework to motivate the analysis of how life expectancy at birth will alter lifetime human capital investment and earnings.

At each instant the individual is endowed with 1 unit of time, which s/he allocates among leisure $L(t)$, investment in human capital $I(t)$, and work $(1 - L(t) - I(t))$. Human capital $H(t)$, augments individual time in the production of additional human capital, in income generation and in leisure consumption. Human capital is accumulated at the rate

$$\dot{H(t)} = F[I(t)H(t), D(t)] - \sigma H(t). \tag{1}$$

$$H(0) = H_0, \tag{2}$$

where $I(t)$ is the time allocated to human capital production, and $D(t)$ is the input of market goods into human capital production in period *t*. F is a concave production function. No human capital is produced unless time is allocated to it, *i.e.,* $I(t) > 0$. Human capital depreciates at the rate of σ in every period.[42]

[42] Although human capital might exhibit differentiated rate of depreciation at older ages, we adopt Heckman's assumption that σ is constant through the life. This will not change the implication of the model for our setting

Consumers' income in period t comes from two sources: interest earnings from assets accumulated and wage earnings conditional on accumulated human capital. The market price for a unit of human capital is R. Labor income in period t can at most be $RH(t)$if the individual devotes no time to human capital production or leisure. Income can be allocated to direct investment in education goods ($D(t)$), and consumption of durable and nondurable goods ($X(t)$) which are priced at P. Given an endowment of initial assets A_0 and the human capital endowmentH_0, the individual will accumulate wealth $A(t)$ according to

$$\dot{A(t)} = rA(t) + RH(t)[1 - I(t) - L(t)] - PD(t) - PX(t), \tag{3}$$

$$A(0) = A_0. \tag{4}$$

In equation (3), r is the risk-free rate of return on accumulated assets. The individual instantaneous utility function takes the form

$$U[X(t), L(t)H(t)],$$

where utility is concave in its arguments $X(t)$ and $L(t)H(t)$. Note that $L(t)$ is leisure in natural units of time while $H(t)L(t)$ is the human capital augmented leisure. The individual seeks to maximize lifetime utility as

$$\int_0^T e^{-\rho t} U[X(t), L(t)H(t)]dt. \tag{5}$$

This is maximized subject to the constraints (1)-(4).[43] The current value Hamiltonian of the above problem is

[43] We ignore the bequest motive for simplicity. Bequests do not make much sense in a model without hierarchical families.

$J(t){:}\, e^{-\rho t}U[X(t), L(t)H(t)] + \lambda(t)\{ rA(t) + RH(t)[1 - I(t) - L(t)] - PD(t) - PX(t)\} +$

$\mu(t)\{F[I(t)H(t), D(t)] - \sigma H(t)\},$

(6)[44]

where $\lambda(t)$ is the shadow value of an additional unit of wealth and $\mu(t)$ is the shadow value of an additional unit of human capital in period *t*. The first order conditions are

$J(t)_{X(t)}{:}\, U_1(t) = \lambda(t)e^{\rho t}P.$ (7)

$J(t)_{L(t)}{:}\, U_2(t)H(t) = \lambda(t)e^{\rho t}RH(t).$ (8)

$J(t)_{I(t)}{:}\, \mu(t)F_1(t)H(t) = \lambda(t)RH(t).$

(9)

$J(t)_{D(t)}{:}\, \mu(t)F_2(t) = \lambda(t)P.$ (10)

$J(t)_{A(t)}{:}\, \dot{\lambda}(t) = -r\lambda(t).$ (11)

Equation (11) is a first order differential equation in $\lambda(t)$. The solution for $\lambda(t)$ is

$\lambda(t) = \lambda(0)e^{-rt}.$ (12)

The last first order condition is

$J(t)_{H(t)}{:}\, \dot{\mu}(t) = \sigma\mu(t) - R\lambda(t).$ (13)

The terminal condition for human capital is

$\mu(T) = 0,$ (14)

while the assumption of non-satiation $\lambda(T) > 0$ together with the "no Ponzi scheme" condition $A(T) \geq 0$ implies that the terminal condition for wealth is

$\lambda(T)A(T) = 0.$ (15)

---

[44] For simplification, we assume that there is no income tax in the model. Heckman assumed a tax rate of (1-α) so that household could keep only a fraction (α) of the income.

In equation (12), $\lambda(0)$ is the shadow value or marginal utility of wealth at the beginning of life. Therefore, this is also the period 0 shadow value of lifetime earnings. Since resources are finite and an assumption of non-satiation holds, $\lambda(0)$ must be positive.

To simplify further analysis, we define the shadow value of human capital in terms of wealth, which is the ratio of the shadow value of human capital to that of assets: $g(t) = \frac{\mu(t)}{\lambda(t)}$.

Utilizing this along with equation (14), the first order differential equation for $g(t)$ is

$$\dot{g(t)} = (\sigma + r)g(t) - R.$$

Utilizing the terminal condition for shadow value of human capital, as stated in equation (14), we can derive the solution for $g(t)$:

$$g(t) = \frac{R}{(\sigma+r)}\left[1 - e^{(\sigma+r)(t-T)}\right]. \tag{16}$$

Since two basic assumptions of the model are strict concavity and differentiability of the utility and production functions, we can invert equations (7), (8), and (12) to obtain the ($\lambda(0)$) constant demand function for consumption good $X(t)$ and effective leisure $L(t)H(t)$

$$X(t) = X\left[\lambda(0)e^{(\rho-r)t}P,\ R\lambda(0)e^{(\rho-r)t}\right], \tag{17}$$

$$H(t)L(t) = HL\left[\lambda(0)e^{(\rho-r)t}P,\ R\lambda(0)e^{(\rho-r)t}\right]. \tag{18}$$

Similarly, the demand functions for the two human capital investment inputs can be obtained by inverting equations (9) and (10).

$$D(t) = D\left[\frac{P}{g(t)}, \frac{R}{g(t)}\right]. \tag{19}$$

$$I(t)H(t) = IH\left[\frac{P}{g(t)}, \frac{R}{g(t)}\right]. \tag{20}$$

Because $g(t)$ gets smaller as $t \rightarrow$T, the price of purchased educational inputs and the opportunity cost of time devoted to human capital production increase as the individual ages. As a result, time invested in human capital production, $I(t)$, decreases as an individual ages and approaches zero at *T*. Before time *T*, human capital production takes place in every period to offset human capital depreciation.

The stock of human capital at any period *t*, *H(t)*, is the depreciation-weighted accumulated investment in human capital till period *t* plus the depreciated initial stock. Human capital stock at period *t* is specified as following

$$H(t) = \int_0^t e^{\sigma(\tau-t)} F[I(\tau)H(\tau), D(\tau)]d\tau + H(0)e^{-\sigma t}. \tag{21}$$

So, the lifetime human capital stock is the accumulated human capital over a lifetime *T*

$$H(T) = \int_0^T e^{\sigma(\tau-T)} F[I(\tau)H(\tau), D(\tau)]d\tau + H(0)e^{-\sigma T}. \tag{22}$$

The value of human capital is equal to the earnings generated from selling this human capital in the market in each period, net of its explicit and implicit production cost. The shadow value of human capital in terms of wealth, *g(t)*, can be used to convert human capital into wealth. The value of lifetime accumulated human capital stock evaluated at the initial period is

$$V = g(0)H(0) + \int_0^T e^{-rt}\{g(t)F[I(t)H(t), D(t)] - PD(t) - RIH(t)\}\, dt. \tag{23}$$

The term inside the integral is the lifetime net earnings from human capital investment.

2.1 Comparative dynamics from increase in life expectancy at birth T

This section derives the effect of changes in life expectancy, *T*, on human capital investment decisions, lifetime human capital production, and lifetime earnings from human capital investment.

Proposition 1: The shadow value of human capital in terms of wealth, $g(t)$, increases when life expectancy, *T*, increases.[45] This happens for all $t \in [0, T]$ as $\frac{\partial g(t)}{\partial T} = Re^{(\sigma+r)(t-T)} > 0$.

[45] All proofs to these propositions are presented in the *appendix* C.

As $g(t)$ increases, the effective cost of the inputs into human capital production fall, as expressed in equations (18-19), that leads us to proposition 2.

Proposition 2**:** As life expectancy at birth, *T*, increases, purchases of educational-investment goods, *D(t)*, and effective time investment, *I(t)H(t)* increase in every period of life $t < T$.

Proposition 3: As life expectancy at birth, *T*, increases, the human capital stock *H*(*t*) accumulated by time *t* increases in every period $t \in [0, T]$ as does the total human capital *H*(*T*) accumulated over the lifetime. This is a direct consequence of the increased use of *D(t)* and *I(t)H(t)* in every period *t* as governed by the production function $F[I(t)H(t), D(t)]$.[46]

Proposition 4: As life expectancy at birth, *T*, increases, lifetime labor income increases.

Proposition 5: As life expectancy at birth, *T*, increases marginal utility of lifetime wealth at the beginning of life, $\lambda(0)$, decreases.

Proposition 6: As life expectancy at birth, *T*, increases, consumption of leisure in human capital adjusted efficiency units, $HL(t)$, increases. However, measured hours of leisure, $L(t)$, may increase or decrease.

The proof in *appendix* C shows that an increase in life expectancy at birth has two opposing effects on measured units of leisure. Effective leisure becomes cheaper due to a fall in the marginal utility of wealth, but at the same time, the opportunity cost of hours spent in leisure increases due to rising human capital investments. It is possible that lifetime leisure will rise or fall as *T* increases, contrary to the assertion made by Hazan (2009).

[46] We are assuming that per unit value of human capital is not bid downward due to the outward shift of the supply of skilled workers. As will demonstrate, the estimated impact of increased life expectancy on human capital investment is sufficiently small in magnitude that the positive effects of human capital on income have dominated the downward pressure from increased supply.

3. Reduced Form and Econometric Specification

The model predicts that in every period of life, increased life expectancy at the start of life will increase accumulated human capital and will raise lifetime earnings. Simultaneously solving the first-order conditions (7-15) results in reduced-form solutions of the lifetime sequences of expected paths of goods consumption, investments in human capital, leisure consumption, and the planned accumulations of assets and human capital in every period, conditional on available information on the exogenous variables at time 0. The exogenous variables include the rates of interest (r) and human capital depreciation (σ), the price of human capital inputs (P), the rental rate of human capital (R), the endowments of human capital and assets (H_0, A_0), and life expectancy at birth (T). At the time of birth, the individual can set the optimal trajectory of the human capital stock at every point in the life cycle based on information available at that time Ω_0 :

$$E[H(t)|\Omega_0] = f^H\big(r, \sigma, P, R, H_0, A_0, T, g(t)\big)\ \forall t \in [0, T]. \tag{24}$$

The investment trajectory for time and goods inputs into human capital investment in each period is set by the expected paths of the shadow values of assets and human capital

$$g(t) = g(0)e^{(\sigma+r)t} + \frac{R}{(\sigma+r)}\left[1 - e^{(\sigma+r)t}\right] = \frac{\mu(0)}{\lambda(0)}e^{(\sigma+r)t} + \frac{R}{(\sigma+r)}\left[1 - e^{(\sigma+r)t}\right]. \tag{25}$$

A change in T increases the projected lifetime wealth at the beginning of life, which causes the marginal utility of wealth at birth, $\lambda(0)$, to fall. That increases $g(0)$, the value of human capital relative to wealth at the start of life. The increase in $g(0)$increases all the subsequent values of $g(t)$ by equation (25).

The planned sequence of $g(t)$ is based on information available at time 0. Unanticipated shocks to the exogenous variables in equation (24) will cause the individual to re-optimize. As a result, the sequence of $g(t)$ will evolve. Critically, however, the new information will be orthogonal to the information set Ω_0. As a result, changes to the sequence of $g(t)$ will be uncorrelated with Ω_0. For example, suppose at time *t'* the individual finds out that life expectancy has changed from *T* to *T'*. The individual will re-optimize including new values of the $g(t)$ sequence from *t'* through the end of life at *T'*. However, $E[\mathscr{g}(t) - g(t)|\Omega_0] = 0 \; \forall \, t > t'$ where $\mathscr{g}(t)$ represents the re-optimized sequence of relative shadow values of human capital to assets. That means that changes made to planned sequences of human capital investments, labor supply and lifetime consumption paths from the plans made at time 0 will be uncorrelated with the values of the exogenous variables at time 0 including the value of life expectancy at birth.[47]

This has important implications for estimating lifetime human capital investments and earnings as a function of life expectancy at birth. Suppose that the planned human capital stock at time *t* conditional on initial information is $H(t|T)$ and the updated plan after changes in information on life expectancy is $\mathcal{H}(t|T')$. A survey will reveal information on actually completed human capital investments $\mathcal{H}(t)$, but $E[\mathcal{H}(t|T') - \mathcal{H}(t)|\Omega_0] = 0$ and so the projection of observed $\mathcal{H}(t)$ on *T* will yield the effect of life expectancy at birth on planned human capital investments at birth. On the other hand, $E[(\mathcal{H}(t|T') - H(t)|\, \Omega_0, T'] \neq 0$ and so a regression of $\mathcal{H}(t)$ on T' will not generate the unbiased effect of life expectancy on planned human capital investments. In particular, if individual decisions made after birth due to new information on any of the exogenous variables result in changes in life expectancy, the observed human capital outcomes $\mathcal{H}(t)$ and the observed life expectancy T'will be jointly determined. A

[47] Some recent studies, for example Hazan (2012), have proposed that life expectancy at five instead of life expectancy at birth is more suitable to explain human capital investment decision due to selection problem with respect to who survives infancy or early childhood. In practice, by the time a child reaches age five, parents or government or both have made significant investment in the child, which makes life expectancy at age five higher compared to what it was at birth, and thus makes life expectancy at five an endogenous variable.

similar argument suggests that to derive the effect of life expectancy on lifetime earnings, one should also regress observed earnings on life expectancy at birth and not life expectancy at later ages.

4. Data

We require data with considerable variation in life expectancy at birth and information on lifetime human capital investment and earnings. We exploit the World Bank's *International Income Distribution Database (I2D2)* for that purpose. I2D2 is a harmonized collection of household surveys conducted in 111 countries. A list of the countries and survey years is presented in table B1 in the *appendix* C. The database includes countries from all regions and income groups. Of our 111 countries, 30 are developed countries, 11 from Asia and Pacific, 17 from Central Asia and Eastern Europe, 23 from Latin America, 4 from the Middle East and North Africa, and 26 from Sub Saharan Africa and South Asia.

From each survey, we keep only those individuals who have complete information on both education and wages. Since our interest lies in how life expectancy at birth affects human capital accumulation and earnings, we focus only on those individuals who are working. We focus on the prime working age groups between 25 and 60 years who report positive incomes. Our focus on individuals over age 25 limits the probability that individuals are still in school. [48] The upper age threshold of 60 years was selected to avoid selection issues related to retirements and rising mortality in some of the countries.

Because life expectancy at birth sets the trajectory for lifetime human capital investment and earnings, we cannot aggregate across individuals with different life expectancies. Therefore, we define each cohort in each country as the unit of observation. Our earliest available survey is

[48] Both Barro and Lee (2010), and Cohen and Soto (2007) assume that years of schooling are fixed by age 25.

in 1970 while the latest is in 2012. To fit our age range of 25-60, that means we include 77 birth cohorts born from 1911 to 1987.

There are multiple surveys for many of the countries, and so we have repeated observations for many cohorts. However, completed schooling will be the same for the same cohort across surveys. We opted to use the earliest available survey for each country to limit mortality bias in the estimated completed schooling and then the most recent survey to capture the completed schooling for the youngest cohorts in the country. In total, we used 188 surveys across 111 countries to create 4670 cohort observations covering almost 4 million individual observations. We further disaggregate the cohorts by gender, and if possible, by urban and rural residence. For each birth cohort, we computed average years of schooling, average earnings, and incidence of marriage. To compare lifetime earnings across countries, we require a common unit of time. Across the 188 surveys, wages are measured per hour, day, week, month, quarter and year. However, the surveys are internally consistent, and so a survey-specific dummy variable will standardize time units. The survey specific dummy variable also will standardize the currency units and so we do not have to rely on exchange rates. The survey-specific dummy variable will also control for country-specific effects. In effect, the source of identification will be variation in schooling and earnings across cohorts within surveys.

Our data on life expectancy at birth by country were compiled from 1950 on from the United Nation's Population database.[49] For earlier birth cohorts, Gap Minder constructs a measure of life expectancy at birth for almost 200 countries back to 1900. Figure 1 shows the pattern of life expectancy at birth by birth cohort starting in 1910. Worldwide life expectancy has risen from 38 to 72 years over the 90-year period. Over that same period, average years of schooling rose from

[49] The UN maintains a rich database on various socio-economic indicators http://data.un.org/Default.aspx.

6.8 to 12.4 years. As shown in Figure 2, these patterns are common across regions and income groups.

As life expectancy increases, the fraction of the birth cohort that enters working age increases. If workers of different ages are not perfect substitutes for one another, members of unusually large working-age cohorts will face depressed earnings (Welch, 1979). We use the number in the cohort relative to the total population as our measure of the relative cohort supply. Figure 3 shows the path of average wages across birth cohorts after netting out the survey fixed effect.[50] Starting with the oldest cohorts, average earnings rise over birth cohorts until the mid-1950s when the average earnings begin to decline. The reversal is due to the declining age of the more recent birth cohorts, illustrating that we will need to control for position in the life cycle to remove the effects of age on lifetime earnings. As we demonstrate in the next section, use of quadratic terms in age of the cohort or using cohort-specific fixed effects serve to correct for the age effect on earnings.

5. Empirical Specification

The theory suggests that the reduced-form equation for completed schooling and earnings will depend on conditions known at the time of birth plus changes to those variables conditioned on information orthogonal to those variables known at birth. We specify these equations for completed years of schooling S_{jct} and log earnings $\ln(Y_{jct})$ for cohort *j*, country *c*, and survey year *t* by

$$S_{jct} = \gamma_0 + \gamma_L LE_{jc} + \gamma_R RUR_{jct} + \gamma_M MALE_{jct} + a_j + \theta_{Sct} + \varepsilon_{jct}, \tag{26}$$

[50] The survey fixed effect would take care of the problems we described above, and therefore, would give us an estimate of average wages net of fixed factors such as currency, time unit used for measurement of wages, country specific fixed effect and inflation. The wages are presented on a logarithmic scale.

$$\ln(Y_{jct}) = a_0 + \beta_L LE_{jc} + \beta_R RUR_{jct} + \beta_M MALE_{jct} + \sum_{p=1}^{J} \beta_p \boldsymbol{Z}_{jct} + a_j + \theta_{Yct} + \omega_{jct}. \qquad (27)$$

The focus on years of schooling is a matter of convenience in that we know that human capital investment will rise in every period *t* as life expectancy LE_{jc} for cohort *j* rises, but schooling is the most readily observable and consistent measure of human capital investment across countries and time. As it is also a form of human capital investment that is fixed at relatively young ages, we can assume that for birth cohorts aged 25 and over, their years of schooling are fixed for the rest of their lives. In equation (26) and (27), a_j includes cohort-specific effects that are known at birth and common across countries; θ_{Sct} and θ_{Yct} are survey-specific fixed effects that also incorporate country-specific effects that are common across cohorts within the country. $MALE_{jct}$ and RUR_{jct} are respectively average proportion male and average proportion rural in cohort *j*, country *c*, and survey year *t*. The log earnings equation (27) shares many of the same features as (26). Unique elements in $\boldsymbol{Z}$ include the cohort specific marriage incidence rate and size of the cohort within a country-survey year. Cohort-specific fixed effects a_j will correct for position in the life cycle. Alternatively, we can conserve on parameters and specify the lifetime log earnings function as

$$\ln(Y_{jct}) = a_0' + \beta_L' LE_{jc} + \beta_A' AGE_{jct} + \beta_{AA}' (AGE_{jct})^2 + \beta_R' RUR_{jct} + \beta_M' MALE_{jct} + \sum_{p=1}^{J} \beta_p \boldsymbol{Z}_{jct} + \theta_{Yct}' + \omega_{jct}', \qquad (28)$$

where the quadratic terms in the age of the cohort control for position in the life cycle. We can compute the returns to schooling applying the Wald estimator: $\frac{d\ln Y}{dS} = \frac{\frac{\partial \ln Y}{\partial LE}}{\frac{\partial S}{\partial LE}} = \frac{\beta_L}{\gamma_L}$. This estimate uses life expectancy at birth as an instrument for completed years of schooling by cohort. We can compare this estimate with the traditional estimate using the *Mincerian* earnings function specification

$$\ln(Y_{jct}) = \varphi_0 + \varphi_A(AGE_{jct}) + \varphi_{AA}(AGE_{jct})^2 + \varphi_S(SchoolYears_{jct}) + \varphi_R RUR_{jct} + \varphi_M MALE_{jct} + \varphi_{Yct} + \vartheta_{jct}, \quad (29)$$

which yields biased estimates of the returns to schooling due to presumed endogeneity of the schooling choice (Card, 1999).

Observations are weighted to reflect the cell share of the total population in the country. We further weighted the data by the square root of the cell-size to correct for differences in measurement error variance between thin and thick cell samples.[51] Finally, we cluster all errors at the country level to correct for likely correlated errors across cohorts within a country.

6. Results

6.1 Life expectancy at birth and education

Table 1 reports the estimates obtained from regression specification (26). We start with the simplest bivariate specification relating life expectancy at birth and lifetime schooling. All of the specifications include a survey fixed effect which controls for country specific fixed factors, cyclical factors related to the timing of the survey, and survey type: household expenditure survey vs. labor force survey. We also control for common birth cohort-specific effects across countries.[52] In specification IV, we include cohort dummies, where cohort is defined by the year of birth, to control time varying factors across countries. [53]

[51] Cell size is the total number of observations belonging to a specific cohort in a survey. We divide the cell size by how many times each cohort appears and use that to construct weight to be applied in the regression. In our sample, out of 111 countries, for 77 countries we add younger birth-year cohorts from the most recent survey.
[52] Cohorts born during the Great Depression or during World War II might experience common shocks to schooling availability. Similarly, there were several United Nations programs and activities to improve health and education across the countries implying that cohorts born after 1960s in the low income countries might have been exposed to similar global campaigns for education.
[53] Later, in the robustness section, we include cohort dummies defining cohorts by five-year birth range.

The coefficient of life expectancy at birth, γ_L in equation (26), is always positive and statistically significant. The effect ranges from 0.094 to 0.12 years of schooling per year of gain in life expectancy. The specifications imply that individuals add one year of schooling for every 8.5 years of gain in life expectancy, suggesting that the34 years of increased life expectancy between the 1910 and 1987 birth cohorts, as revealed in Figure 1, would have increased time in school by 3.98 years. The coefficient of parental life expectancy in specification IV suggests that there is an intergenerational channel through which parental health affect education of their children. The effect of parental life expectancy is 22% of the own life expectancy effect.

The effect of life expectancy gain at birth might have different effects on different groups. We investigate this by estimating equation (26) for four groups separately: (i) Male (ii) Female (iii) Urban, and (iv) Rural. Table 2 presents separate estimates of the life expectancy effect on schooling for each of these groups. Across all specifications, we observe a larger effect of life expectancy gain at birth on women's schooling compared to men. There is little difference in the life expectancy effect on schooling between rural and urban residents. For all four groups, the parental life expectancy does not seem to influence schooling decision for their children.

6.2 Life expectancy at birth and earnings

Table 3 reports estimates obtained from regression specification (27) and (28). Similar to the inverted u-shaped plot of log of wages across birth cohorts in figure three, the negative life expectancy coefficient obtained from a simple regression of log of wages against life expectancy at birth, as reported in the specification one in table three, implies the role of age factor in determining lifetime earnings. Once we control for lifecycle effects by including ages or cohort-fixed effects, and other potential confounders, the coefficient of life expectancy at birth reveres sign-the effect of life expectancy at birth on log of wages turn out to be positive and statistically significant. Increasing life expectancy at birth by one year increases lifetime earnings by roughly 1%. In all of the specifications, we observe that log of earnings increases with age at a decreasing rate. Consistent with the findings in the literature, we find that married and urban people earn relatively more compared to their unmarried and rural counterparts. Larger cohort-size lowers cohort earnings, consistent with presumption that unusually large cohorts receive depressed earnings.[54] As with schooling, there is an intergenerational gain from parental life expectancy that is statistically significant in all of the specifications with cohort fixed effects. The parental life expectancy effect is 60-88% as large as the own life expectancy effect.

In table 4, we report separate estimates by gender and region. The returns to life expectancy are somewhat larger for men than women and for urban compared to rural residents. However, for male, the life expectancy coefficient turns out to be statistically significant only in specification III where we control for lifecycle position by birth-year fixed effect. For rural residents, life expectancy at birth fail to show statistical significance in any of the specifications. Parental life expectancy retains a small positive effect on the earnings of their children, but the effect is statistically significant for men, women and rural residents, and only in the specification with birth-year fixed effect.

6.3 Return to schooling

The *Mincerian* earnings function (29) generates a measure of the returns to education. This can be compared to the Wald estimator as laid out in the previous section.[55] Estimates from the *Mincerian* earnings function are reported in table 5 while the comparisons of these estimates with those from Wald estimators are drawn in table 6. Additional schooling increases lifetime

---

[54] Cohort-specific dummies would absorb this effect if one specific cohort experiences a surge in population across the world.

[55] The Wald estimator for $Return\ to\ Education = \frac{return\ to\ lifetime\ earnings\ from\ 1\ year\ gain\ in\ life\ expectancy\ at\ birth}{return\ to\ schooling\ from\ 1\ year\ \ gain\ in\ life\ expectancy\ at\ birth}$.

earnings irrespective of the specification used. For the pooled sample, the coefficient of school-years show that one additional year of schooling increases lifetime-earnings by 11.8% in the simple specification I, which shrinks to 9.4% once we utilize a broad set of controls.[56] The estimates on age terms, percentage male, percentage urban-rural, percentage married, and cohort-size exhibit right sign pattern and statistical significance. The estimates of the earnings function by gender and urban-rural group do not show any notable variation across groups. Specification II reveals that one additional year of schooling increases lifetime earnings for male, female, urban and rural group by 10.6%, 7.3%, 9.5%, and 8.8% respectively. Table 6 compares the return to education estimates obtained from specification *II* of the *Mincerian* earnings function with the *Wald* estimates.

For the Pooled sample, the Wald estimate of returns to schooling is obtained by dividing the life expectancy coefficients from specification IV in table 3 by the life expectancy coefficient from specification IV in table 1. Similarly, for male, female, urban, and rural group we divide the group-specific life expectancy coefficient obtained from specification II in table 4 by the life expectancy coefficient from specification IV in table 2. The standard errors for the *Wald* estimates are obtained by five hundred bootstrap replications with the corresponding sample. In the pooled sample, the *Wald* estimates are 16% lower compared to the *Mincerian* estimates (9.40% vs. 7.90%). The *Wald* estimate is slightly higher for urban residents (10.21% vs. 9.50%). In contrast, the *Wald* estimates are lower for males (9.90% vs. 10.60%), females (4.30% vs. 7.30%) and rural residents (2.90% vs. 8.8%). For both *Mincerian* and *Wald* estimates, the return to schooling is always higher for males than females and higher for urban than rural residents.

[56] In all specifications reported in table five, we have used survey fixed effects to facilitate comparison across countries and time periods.

6.4 Robustness

We reexamined our findings with several differences in estimation methods and samples. We re-estimated the models (i) without weights, (ii) with a different sample consisting only of one survey per country, (iii) using an alternative definition of parents' life expectancy[57], (iv) including cohort fixed effect with an alternative definition of cohort[58] (v) with a sample consisting only of young age groups, and (v) using a higher order age variable to wipe out all age effects while estimating the life expectancy effects on earnings.[59] The robustness checks generate similar estimates of life expectancy effects on schooling and earnings. In most cases, the sign and statistical significance of the life expectancy coefficient is positive and statistically significant. The results for all robustness checks are briefly discussed and reported in the *appendix C.*

The final check of robustness involves addition of some exogenous variables that vary by country and cohort at the time birth, and substitution of life expectancy at birth by life expectancy measures at higher ages. The theory predicts that the effect of life expectancy at birth is exogenous to any random shock realized in a later period in life. To demonstrate the validity of this, we incorporate average temperature and average precipitation that was observed for a birth-cohort at the time of birth in its country of origin. Note that similar to life expectancy at birth,

---

[57] Previously, parents' life expectancy was constructed by taking a 25 year lag of life expectancy at birth. The youngest cohort in our survey was born in 1987. In the 1980s, in many regions, especially in Sub Saharan Africa and South Asia, mother's age at first child birth was less than 20. For example, in Niger half of the women gave birth by age 18 (Source: http://www.unicef.org/pon95/fami0009.html). Alternative measure of parents life expectancy assumed a 15 years lagged value of own life expectancy.

[58] While constructing the five-year birth cohorts, we collapse all individuals aged 25-59 into different five year birth range except the first and last cohort. In total, we have defined 13 cohorts based on 5-year birth groups. Since the number of observations before 1930 is too thin, we group them into one cohort. Similarly all individuals, who were born during 1985-87, were collapsed to form the last cohort.

[59] In addition to those attempted for schooling, we try one additional robustness check for lifetime earnings. Following Card (1999), and Murphy and Welch (1990), we add higher order age terms to sponge out all age effects.

average temperatures and average precipitation also vary by birth-year and country. The results and discussion on this are presented in the *appendix* C. In total, the estimates reveal that the statistically significant positive association of life expectancy at birth with schooling and earnings is not altered after inclusion of these exogenous variables.

Concerns related to high infant and child mortality rate led some recent papers to argue for appropriateness of use of life expectancy at birth. Hazan (2012) suggests that life expectancy beyond early childhood is more appropriate to capture its true effect on human capital investment decision since parents make schooling decision for their children at age five or later.[60] However, by the time a child reaches age five or ten, the parents have made substantive investments in the child's health based on new information on the child's survival prospects. That makes the life expectancy at ages 5 or 10 endogenous responses to child survival, making their use inappropriate as explanatory variables for other parental investments in their children. However, results using life expectancy at higher ages, still have positive and significant effects on years of schooling and lifetime earnings, as reported in *appendi*x C. Note that this contrast to Hazan's (2012) finding that life expectancy at higher ages do not exhibit any statistically significant association with schooling years in a cross-country panel.

7. Extension

We extend the analysis utilizing individual level observations instead of cohort level mean observations as reported above. Note that the theoretical model suggests that life expectancy at birth is exogenous in determining human capital investment and lifetime earnings, an individual level analysis will confirm if country-cohort specific unobservable is contaminating the cohort-

[60] The argument is based on the observation that cross-country life expectancy at birth exhibits more variation compared to life expectancy at five due to high infant and child mortality.

mean based empirical results. We estimate three equivalent specifications of equation 26, 28, and 29 using individual level data.

$$S_{ict} = \gamma_0 + \gamma_L LE_{jc} + \gamma_R RUR_{ict} + \gamma_M MALE_{ict} + a_j + \theta_{Sct} + \varepsilon_{ict}, \tag{26 a}$$

$$\ln(Y_{ict}) = a_0' + \beta_L' LE_{jc} + \beta_A' AGE_{ict} + \beta_{AA}' (AGE_{ict})^2 + \beta_R' RUR_{ict} + \beta_M' MALE_{ict} + \sum_{p=1}^{J} \beta_p \boldsymbol{Z}_{ict} + \theta_{Yct}' + \omega_{ict}, \tag{28 a}$$

$$\ln(Y_{jct}) = \varphi_0 + \varphi_A(AGE_{ict}) + \varphi_{AA}(AGE_{ict})^2 + \varphi_S(SchoolYears_{ict}) + \varphi_R RUR_{ict} + \varphi_M MALE_{ict} + \varphi_{Yct} + \vartheta_{ict}. \tag{29 a}$$

In our setting, we cannot observe individual life expectancy at birth; however, cohort life expectancy at birth, an average measure of individual life expectancies at birth across individuals within a cohort ($LE_{jc} = \frac{\sum_{i=1}^{N} LE_{ijc}}{N}$), is exogenous to an individual's completed years of schooling or lifetime earnings. Group mean is often used as an instrument to resolve *endogeneity* issue in individual level empirical analysis[61]. In the above specifications, $LE_{ijc} = LE_{jc} + \mu_{ij}$, which states that individual *i*'s life expectancy at birth in the country "*c*" deviates from cohort *j*'s mean life expectancy by μ_{ij}, which by construction orthogonal to mean. Since μ_{ij} will be contained in the error term, the following conditions hold -

$Cov(\varepsilon_{ict}, LE_{jc}) = 0$, $Cov(\omega_{ict}', LE_{jc}) = 0$, and $Cov(\vartheta_{ict}, LE_{jc}) = 0$.

In contrast to cohort mean level analysis, to save time and space, here, we estimate one specification for each of the pooled sample, and male-female, rural-urban subsample[62]. As table 7 reveals, the estimates conform to those obtained from the cohort-mean level analysis, life expectancy at birth exhibits a positive and statistically significant association both with

[61] Royalty (2000) has used state tax rate as an instrument for marginal tax rate in explaining employees' health insurance eligibility. Similarly, a series of studies following Ruhm (2000) exploited variation in state or county level unemployment rate while explaining individual health behavior during a recession.
[62] We choose specification IV from table 1 for schooling, and specification IV from table 3 for earnings

completed years of schooling and lifetime earnings. The life expectancy effects on schooling regression reveal similar pattern to those obtained from cohort-mean level analysis as reported in table 1 and 2. However, except for the urban sample, the life expectancy effects on earnings are always larger compared to similar estimates obtained from cohort-mean level analysis. In the pooled sample, one year gain in life expectancy at birth leads one to complete 0.115 years more completed years of schooling, and increases lifetime earnings by 1.3%. The largest effect on earnings is observed for the male sample-1.5% raise in lifetime earnings from each additional year of gain in life expectancy at birth.

Following the procedure followed in table 7, table 8 reports the implied *Wald* estimate of return to schooling and draw a comparison with the estimates obtained for *Mincerian* return to schooling. The *Mincerian* return to education estimates lie around 10%. In contrast, the indirect least square estimates exhibit more variation across groups-for male the return to schooling is 15% while that for the female is 7.8%. Except for the female sample, in all other cases *Wal*d estimates of return to schooling outweighs corresponding *Mincerian* estimates. The *Mincerian* return to schooling for the pooled sample is 15.5% smaller compared to the *Wald* estimate.

In contrast to the cohort-mean level analysis, the findings from individual empirical analysis consistently suggest that parent’s life expectancy exhibit a statistically significant positive influence both on human capital investment and lifetime earnings in the pooled sample as well as across male, urban and rural groups. Parents’ life expectancy effects on years in school stand around 19-26% of own life expectancy effect while for earnings the impacts are even larger- parents’ life expectancy effects are 38.5% to 54.5% of cohorts’ own life expectancy effect. There is strong evidence of intergenerational transfer to rising life expectancy at birth. Such high transmission across generations is not uncommon considering the recent findings by

Lindahl *et al.* (2015) that intergenerational persistence in human capital and lifetime earnings persists across several generations.

7.1 Selection due to participation in the labor force

Our analysis for lifetime earnings includes those who are in the labor force at the time of survey. If life expectancy at birth affects an individuals' labor force participation decision, our estimates will be subject to bias and the direction of bias is uncertain. Note that Hazan (2009) observed that gain in life expectancy at birth was associated with decrease in labor force participation for the US male born between 1840 and 1970. We investigate the labor force selection issue with the individual data in two ways: (i) first, including a birth-year specific correction measure for selection in each survey, which is constructed as the proportion of individuals participating in the labor force within the birth year cohort in a survey, (ii) two-step Heckman selection correction for an individual's labor force participation status. The selection equation includes household size (number of household members) and interviewee's relationship with the household head to fulfill the exclusion criterion.[63] We examine the selection issue for each of the pool, male, female, urban rural sample separately.

The results are reported in table 9 and 10. After controlling for the proportion of the birth-year cohort in the labor force, neither for the pooled sample nor for any of the subsample, does the effects of gain in life expectancy at birth on lifetime earnings show a different pattern from that obtained without addressing the selection issue as reported in table 8. However, estimates obtained following Heckman's two-step procedure, as reported in table 10, where we utilize

[63] In the first stage we estimate a fixed effect logit model of labor force participation decision including age, gender, urban/rural, marital status, life expectancy at birth, parents' life expectancy, household size, relationship with head, and survey/country fixed effects. Based on the parameters extracted from the logit model, we calculate linear predicted probability, which is then converted into normal densities to construct the inverse mills ratio and to use in the second stage for selection correction.

information on individual labor force status, the coefficients of life expectancy at birth in the second stage earnings estimation are positive, statistically significant, and except for the female group, are larger in magnitude compared to those obtained without any selection correction.[64] The life expectancy effects on lifetime earnings for the pool, male, and rural subsample are respectively 17%, 20%, and 15% larger in magnitude compared to those obtained from the corresponding sample without correcting for selection. For urban population, the selection correction does not alter the life expectancy effects on lifetime earnings. Household size and relationship with head, the identification variables in the selection equation, always turn out to be statistically significant predictor of one's labor force participation decision.

The life expectancy coefficient in selection equation is positive and statistically significant except for the rural sample though smaller in magnitude. Since we control for the life cycle position in the selection specification, this positive association suggests that gain in life expectancy at birth influences labor force participation decision marginally at any stage of the lifecycle. In the earnings specification, the negative and statistically significant inverse mills ratios suggest possibly a negative selection-individuals who we observe with wage information are with lower wages compared to the counterfactuals. But this does not seem consistent for the female group. After controlling for selection, the life expectancy coefficients for female has shrunk by 17%. Possibly, we female population with complete wage information are drawn from a relatively high income group. Overall, the effects of gain in life expectancy at birth on lifetime earnings without correcting for selection terms from the cohort-based analysis are conservative.

[64] Replicating the analysis in table 10 on the same sample but without including the selection correction term produces life expectancy coefficients of 0.13, 0.15, 0.12, 0.12 and 0.10 for the pooled, male, female, rural, and urban sample.

8. Interpretation and Conclusion

We find that gain in life expectancy at birth increases investment in human capital and lifetime earnings. An individual spent 0.12 years in school out of each additional year of gain in life expectancy at birth. This is comparable to the estimates of 0.11 years in Sri Lanka (Jayachandran and Lleras-Muney 2009), and 0.17 years in a cross-country study (Hansen 2013). Global average life expectancy at birth increased 29.74 years between 1922 and 1987 birth cohorts. That implies a 3.56 years of increase in terms of completed schooling. In our sample, actual years of schooling increased by 4.76 years for the same cohorts. Our estimates suggest that increased life expectancy at birth explains 75% of the increase in average schooling worldwide. To put it into the United States context, life expectancy at birth rose 28 years from the 1880 to the 1980 birth cohorts. Actual years of schooling increased 6.5 years. The implied increase in schooling by our model is 3.36 years. Therefore, increased life expectancy at birth explains 51.6% of the increased years of schooling in the United States over the 100-year period.

Each additional year of life expectancy at birth leads to a 1% gain in lifetime earnings. Global per capita GDP has increased by 380% between 1900 and 2000. Our analysis implies that the gain in life expectancy alone could explain approximately one-third of this gain in per capita GDP. Utilizing the estimates obtained from the simple regression specification of life expectancy effects on earnings (specification II in table 3), we conduct a simulation exercise to observe how lifetime earnings of urban-male respond to improvement in life expectancy at birth across birth cohorts. Figure 4 suggests that due to gain in life expectancy at birth, lifetime earnings of cohorts born in 1987 increased by 26% to 28% compared to that of the cohort born in1922.

The estimates of return to education from Wald estimator and *Mincerian* earnings function are consistent. However, both of the cohort-mean based analysis and individual observation based analysis reveal that the Wald estimate of male returns to schooling exceeds female returns to schooling. Similar evidence for rural-urban returns is not consistent across cohort-mean and individual level analysis. Selection due to participation in the labor force is not driving the life expectancy effects on earnings upward, the cohort mean level estimate of 1% return of life expectancy gain to lifetime earnings is rather conservative. Finally, a series of robustness checks give confidence in reinforcing the role of health on human capital accumulation and wellbeing.

9. References

Acemoglu, D., and S. Johnson 2007. “Disease and Development: The Effect of Life Expectancy on Economic Growth.” *Journal of Political Economy* 115(6): 925–985.

Acemoglu, D., and S. Johnson. 2006. “Disease and development: the effect of life expectancy on economic growth.” *NBER working paper*, No. 12269.

Akbulut-Yuksel, M. 2014. “Children of War: The Long-Run Effects of Large Scale Physical Destruction and Warfare on Children.” *Journal of Human Resources* 49(3):634-662.

Akresh, Richard, and Damien de Walque. 2011. “Armed Conflict and Schooling: Evidence from the 1994 Rwandan Genocide.” *World Bank Policy Research Working Paper* No 4606.

Almond, Douglas. 2006. “Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 U.S. population.” *Journal of Political Economy* 114 (4):672-712.

Almond, Douglas, L. Edlund L, and M. Palme. 2009. “Chernobyl’s subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in Sweden.” *Quarterly Journal of Economics* 124:1729–72.

Almond, Douglas, and Janet Currie. 2011. “Killing Me Softly: The Fetal Origins Hypothesis.” *Journal of Economic Perspectives* 25(3): 153-72.

Almond, Douglas, and B. Mazumder. 2011. "Health Capital and the Prenatal Environment: The Effect of Ramadan Observance during Pregnancy." *American Economic Journal: Applied Economics* 3(4): 56-85.

Barro, Robert and J-W Lee. 2013. “A new data set of educational attainment in the world, 1950–2010.” *Journal of Development Economics* 104:184–198.

Becker, Gary S. 1993. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, 3rd edition*. Chicago: University of Chicago Press.

Behrman, J. R., and M. R. Rosenzweig. 2004. “Returns to Birthweight.” *Review of. Economics and. Statistics* 86(2):586–601.

Ben-Porath, Y. 1967. “The Production of Human Capital and the Life Cycle of Earnings.” *Journal of Political Economy* 75: 352–365.

Bhalotra, Sonia, and Atheendar Venkataramani. 2012. “Shadows of the Captain of the Men of Death: Early Life Health Interventions, Human Capital Investments, and Institutions.” *unpublished manuscript*, University of Bristol.

Blattman, C., and J. Annan. 2010. “The Consequences of Child Soldiering.” *Review of Economics and Statistics* 92(4):882-98.

Bleakley, Hoyt. 2007. “Disease and Development: Evidence from Hookworm Eradication in the American South.” *Quarterly Journal of Economics* 122(1): 73–11.

Bleakley, Hoyt. 2010. “Malaria eradication in the Americas: a retrospective analysis of childhood exposure.” *American Economic Journal: Applied. Economics* 2:1–45.

Bloom, David, D. Canning, and G. Fink. 2014. “Disease and development revisited.” *Journal of Political Economy, forthcoming*.

Card, David. 1999. “The causal effect of education on earnings.” In O. Ashenfelter & D. Card (ed.), *Handbook of Labor Economics*, edition 1, volume 3, chapter 30, pages 1801-1863 Elsevier.

Cervellati, M., and S. Uwe. 2013. "Life expectancy, schooling, and lifetime labor supply: Theory and evidence revisited." *Econometrica* 81: 2055–2086.

Chay, K.Y., and M. Greenstone. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." *Quarterly Journal of Economics* 118(3):1121-67.

Camacho, A. 2008. "Stress and Birth Weight: Evidence from Terrorist Attacks." *American Economic Review, Paper and. Proceedings*. 98(2):511-15.

Cohen, Daniel, and Marcelo, Soto. 2007. "Growth and human capital: Good data, good results." *Journal of Economic Growth* 12: 51–76.

Cohen, Daniel, and Laura, Leker. 2014. "Health and Education: Another Look with the Proper Data." Available at *http://www.parisschoolofeconomics.eu/docs/cohen-daniel/cohen-leker-health-and-education-2014.pdf* (*last accessed on July 17th, 2015*).

Currie, Janet, and Tim Vogl. 2013. "Early-life health and adult circumstance in developing countries." *Annual Review of Economics* 5:7.1-7.36.

Dell, Melissa, B.F. Jones, and B. A. Olken. 2014. "What Do We Learn from the Weather? The New Climate-Economy Literature." *Journal of Economic Literature* 52(3):740-798.

Field, Erica, O. Robles, and M. Torero. 2009. "Iodine deficiency and schooling attainment in Tanzania." *American Economic Journal: Applied Economics* 1(4):140–69.

Finis, Welch.1979. "Effect of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust." *Journal of Political Economy* 87(5):S65-S97.

Foster, Andrew, Emilio Gutierrez, and Naresh Kumar. 2009. "Voluntary Compliance, Pollution Levels, and Infant Mortality in Mexico." *American Economic Review* 99(2): 191-97.

Gørgens, Tue, X. Meng, and R. Vaithianathan. 2012. "Stunting and selection effects of famine: A case study of the Great Chinese Famine." *Journal of Development Economics*, 97 (1): 99–111.

Halla, Martin, and M. Zweimüller. 2014. “Parental Response to Early Human Capital Shocks: Evidence from the Chernobyl Accident.” *Discussion Paper 7968*, Institute for the Study of Labor.

Hansen, Casper W. 2013. “Life expectancy and human capital: evidence from the international epidemiological transition.” *Journal of Health Economics* 32 (6):1142–1152.

Hazan. M. 2009. “Longevity and lifetime labor supply: evidence and implications.” E*conometrica* 77(6):1829–1863.

Hazan. M. 2012. "Life Expectancy and Schooling: New Insights from Cross-Country Data." *Journal of Population Economics* 25 (4):1237-1248.

Heckman, James J. 1976. “A life-cycle model of earnings, learnings, and consumption.” *Journal of Political Economy* 84: S9–S44.

Jayachandran, Seema, and Adriana Lleras-Muney. 2009. “Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines.” *Quarterly Journal Economics* 124(1):349–398.

Jayachandran, Seema. 2009. *“*Air quality and early-life mortality evidence from Indonesia’s wildfires.” *Journal of Human Resources* 44:916–54.

Karlan, Dean, and J. Appel. 2011. *More than Good Intentions: How a New Economics is Helping to Solve Global Poverty.* New York: Dutton.

León, G. 2012. “Civil Conflict and Human Capital Accumulation: Long Term Consequences of Political Violence in Perú.” *Journal of Human Resources* 47(4): 991-1023.

Lucas, Adrienne M. 2010. “Malaria eradication and educational attainment: Evidence from Paraguay and Sri Lanka.” *American Economic Journal: Applied Economics* 2(2): 46-71.

Maccini, Sharon, and D. Yang. 2009. “Under the weather: health, schooling, and economic consequences of early-life rainfall.” *American Economic Review*. 99:1006–26.

Maluccio, John A., J. Hoddinott, J. R. Behrman, R. Martorell, A.R. Quisumbing, and A. D. Stein. 2009. “The Impact of Improving Nutrition during Early Childhood on Education among Guatemalan Adults.” *Economic Journal* 119(537): 734–63.

Lindahl, Mikael, M. Palme, S.S. Massih, and A. Sjögren. 2015. “Long-Term Intergenerational Persistence of Human Capital: An Empirical Analysis of Four Generations.” *Journal of Human Resources 50:1-33.*

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

Murphy, Kevin M., and F. Welch.1990. "Empirical age-earnings profiles." *Journal of Labor Economics* 8:202-229.

Oster, Emly, I. Shoulson, and R. Dorsey. 2013. “Limited life expectancy, human capital and health investments.” *American Economic Review* 103(5):1977–2002.

Royalty, A.B. 2000. “Tax Preferences for Fringe Benefits and the Health Insurance Offered by Employers.” *Journal of Public Economics* 75(2):209-227.

Ruhm, C.J. 2000. “Are Recessions Good for Your Health?” *Quarterly Journal of Economics* 115 (2000): 617–650.

Shah, Manisha, *and* B. M. Steinberg. 2014*. “*Drought of opportunities: contemporaneous and long term impacts of rainfall shocks on human capital.*” Working Paper No. 19140, NBER.*

Soares, Rodrigo R. 2005. “Mortality reductions, educational attainment, and fertility choice.” *American Economic Review* 95(3):580–601.

University of East Anglia Climatic Research Unit (CRU); [Ian Harris]. Climatic Research Unit (CRU) Year-by-Year Variation of Selected Climate Variables by Country(CY), [Internet]. NCAS British Atmospheric Data Centre, 2013*, 25th August, 2014*. Available at http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__DE_56531370-2613-11e3-9fca-00163e251233

Vallin, Jacques, and F. Meslé. 2009. “The Segmented Trend Line of Highest Life Expectancies.” *Population and Development Review* 35(1): 159-187.

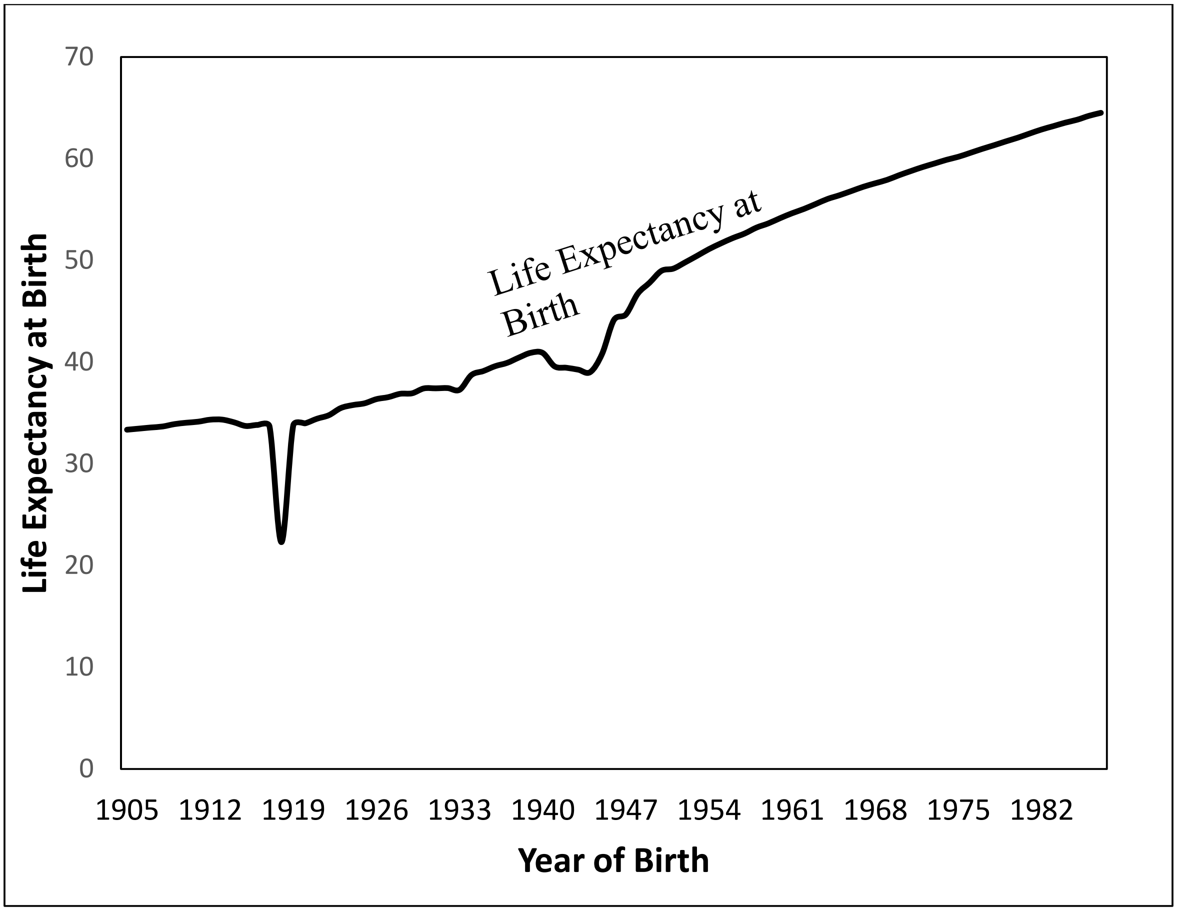


Figure 1: Life Expectancy at Birth across Cohorts, World Average

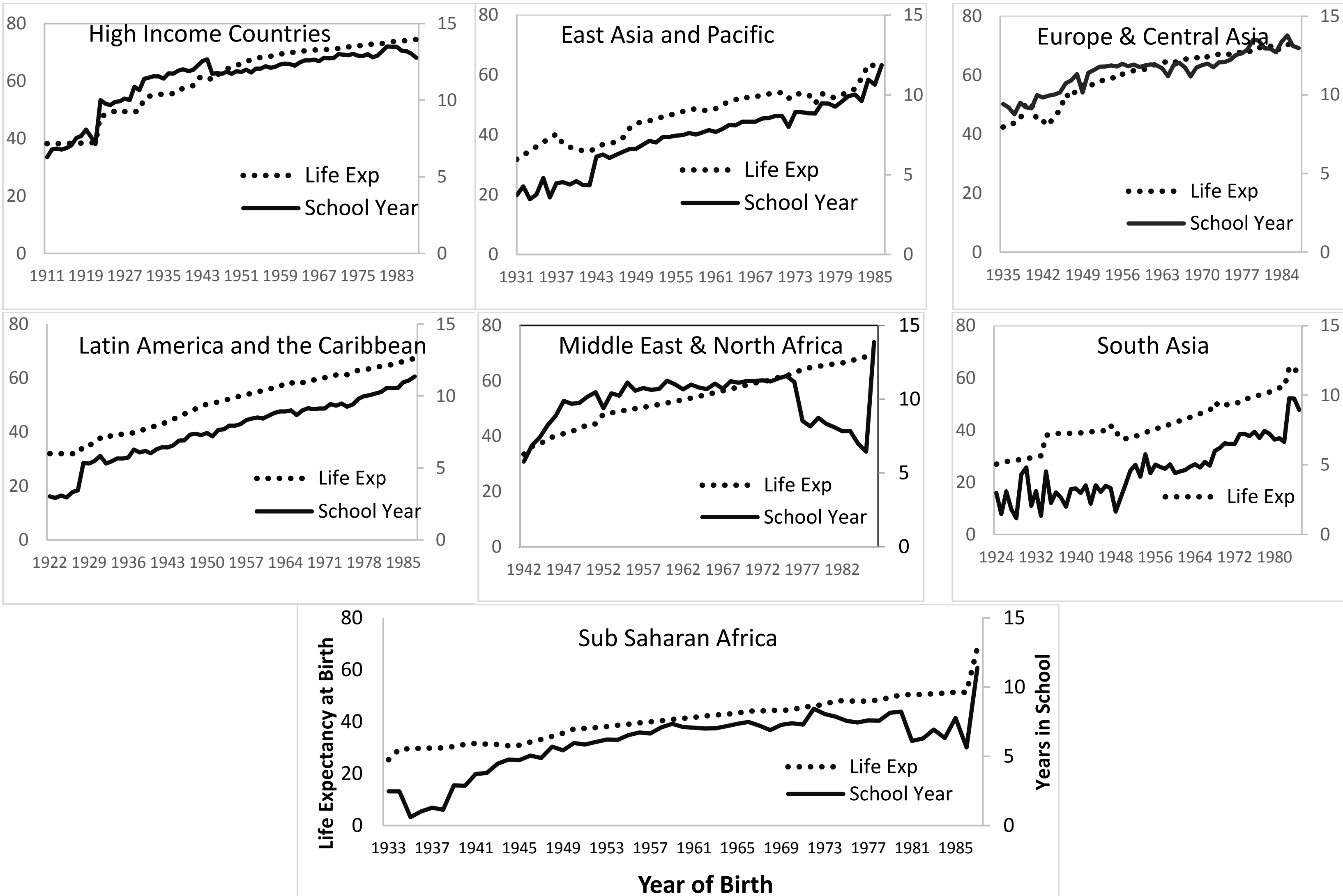
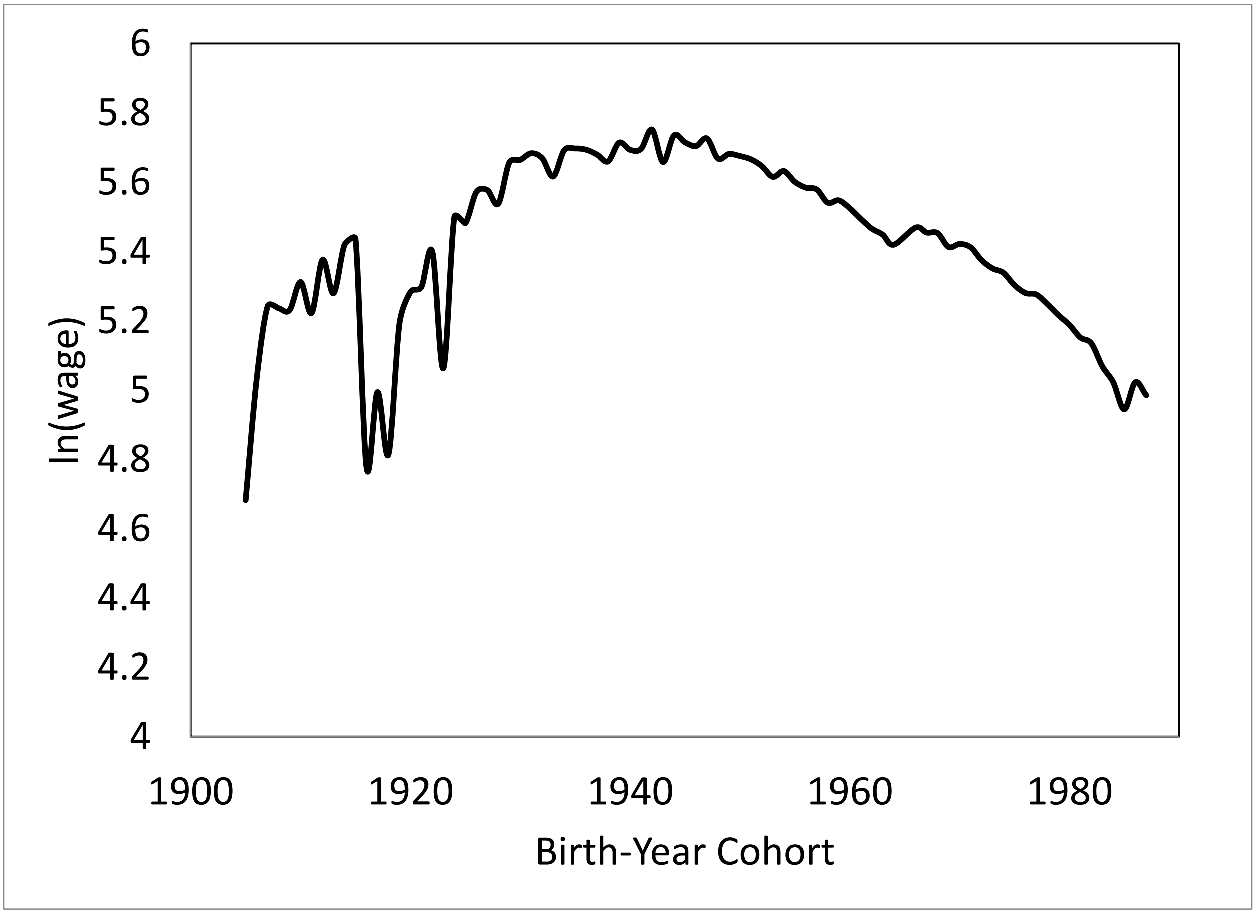
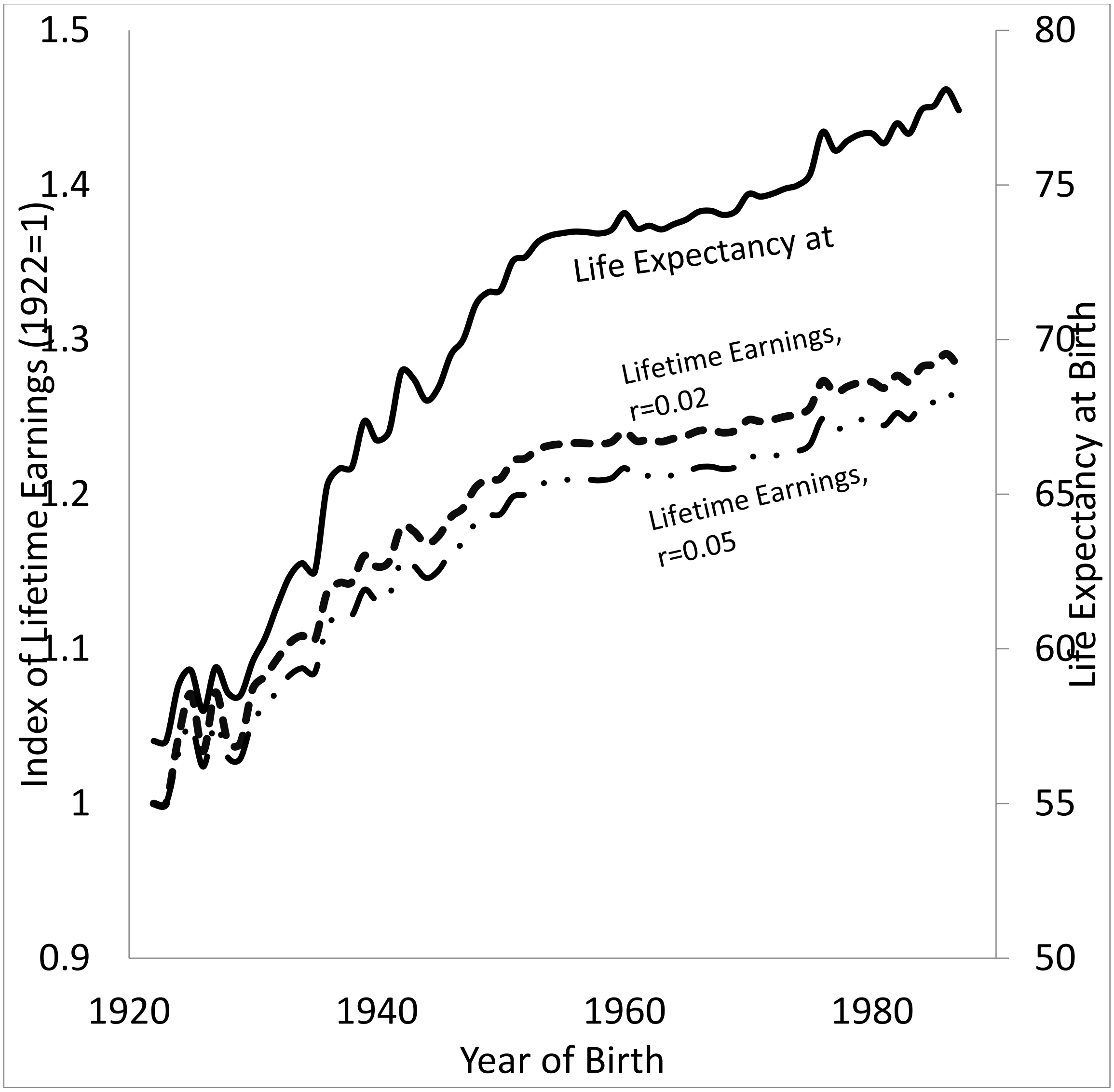


Figure 2: Life Expectancy and Schooling by Birth Cohort across Regions



Note: The birth-year specific wages are the coefficients of Birth-Year dummies in the regression of log (wages) on birth-year dummies and survey dummies, where the survey dummies are taking care of across survey differences in exchange rates, inflation, unit of wages, differences in survey instruments.

Figure 3: Log(wage) by 1905-1987 Birth-year Cohorts, World Averages



Note: We assume 1922 as the base year. We plot the implied present value of lifetime earnings (adjusted for the base year) from the specification II in table 3 against the Birth Year. The lifetime earnings estimates are assumed for male residing in urban areas. The life expectancy at birth numbers are the maximum life expectancy enjoyed by a cohort across the countries, which is to capture what an average person would expect to enjoy staying on the frontier of health technology at the time of birth. While calculating the net present value of log of lifetime earnings, we try two different discount rates 2% and 5%. The period in the figure ranges from 1922 to 1987 as prior to 1922, the information on urban/rural residence is missing.

Figure 4: World Average Life Expectancy at Birth and Implied Lifetime Earnings Index

Table 1: Life Expectancy at Birth and Education

	I	II	III	IV
Life Expectancy at Birth	*0.120****	*0.097****	*0.094****	*0.117****
	[0.015]	*[0.014]*	*[0.008]*	*[0.026]*
% Urban		6.009***	6.043***	6.737***
		[1.687]	[1.748]	[1.057]
% Male		-1.212	-1.342	-0.858*
		[1.047]	[0.922]	[0.516]
Parents Life Expectancy			0.004	0.026*
			[0.019]	[0.014]
Birth Year FE				YES
Survey FE				
Constant	3.476***	1.574	1.625	-1.748
	[0.916]	[2.298]	[2.455]	[2.256]
N	4670	4185	3861	3861
Adjusted R-square	0.987	0.985	0.985	0.987

Note: Standard errors in brackets. Significance level can be read as * p<0.05, ** p<0.01, *** p<0.001.

Table 2: Heterogeneity in Effects of Life Expectancy at Birth on Education

Panel a

	MALE				FEMALE			
	I	II	III	IV	I	II	III	IV
Life Expectancy at Birth	***0.098****	***0.082****	***0.084****	***0.107****	***0.157****	***0.141****	***0.135****	***0.157****
	[0.015]	***[0.013]***	***[0.009]***	***[0.028]***	***[0.018]***	***[0.018]***	***[0.014]***	***[0.032]***
% Urban		5.679***	5.766***	5.958***		4.249***	4.351***	4.737***
		[1.860]	[1.911]	[1.155]		[0.845]	[0.927]	[0.645]
% Male								
Parents Life Expectancy			-0.005	0.02			0.007	0.019
			[0.022]	[0.017]			[0.018]	[0.015]
Birth Year FE				YES				YES
Constant	4.420***	1.801	1.888	-1.172	1.622	-0.419	-0.52	-3.001
	[0.905]	[1.713]	[1.949]	[2.288]	[1.148]	[1.491]	[1.755]	[2.623]
N	4690	4204	3878	3878	4622	4149	3822	3822
Adjusted R-square	0.984	0.98	0.98	0.983	0.983	0.981	0.982	0.982

Panel b

	URBAN				RURAL			
Life Expectancy at Birth	***0.112****	***0.103****	***0.097****	***0.108****	***0.129****	***0.119****	***0.114****	***0.118****
	[0.018]	***[0.018]***	***[0.011]***	***[0.028]***	***[0.017]***	***[0.011]***	***[0.010]***	***[0.029]***
% Urban								
% Male		-0.25	-0.291	-0.106		-0.56	-0.68	-0.448
		[0.807]	[0.707]	[0.441]		[0.548]	[0.576]	[0.467]
Parents Life Expectancy			0.007	0.02			0.01	0.012
			[0.016]	[0.013]			[0.019]	[0.018]
Birth Year FE				YES				YES
Constant	4.539***	5.461***	5.492***	3.861*	2.195**	2.567***	2.444**	1.785
	[1.104]	[1.433]	[1.513]	[2.167]	[0.996]	[0.884]	[1.202]	[2.510]
N	4684	4200	3874	3874	4446	3959	3657	3657
Adjusted R-square	0.982	0.963	0.963	0.965	0.992	0.987	0.988	0.988

Note: Standard errors in brackets. Significance level can be read as * p<0.05, ** p<0.01, *** p<0.001.

Table 3: Effect of Life Expectancy at Birth on Earnings

	I	II	III	IV	V	VI
Life Expectancy at Birth	***-0.017****** ***[0.005]***	***0.010**** ***[0.006]***	***0.011**** ***[0.006]***	***0.009**** ***[0.005]***	***0.013***** ***[0.005]***	***0.009***** ***[0.004]***
Age		0.091*** [0.015]	0.094*** [0.015]	0.077*** [0.013]		
Age square*(1/100)		-0.10*** [0.000]	-0.10*** [0.000]	-0.10*** [0.000]		
% Urban		1.303*** [0.421]	1.396*** [0.423]	1.112*** [0.196]	1.348*** [0.323]	1.087*** [0.186]
% Male		0.409*** [0.111]	0.412*** [0.116]	0.206* [0.118]	0.154 [0.130]	-0.059 [0.145]
Parents Life Expectancy			0.004 [0.003]	0.003 [0.003]	0.008** [0.003]	0.008** [0.003]
% Married				0.471*** [0.153]		0.610*** [0.124]
Cohort Size				-8.577*** [1.534]		-7.733*** [1.544]
Birth Year FE					YES	YES
Constant	8.490*** [0.331]	3.738*** [0.746]	3.341*** [0.819]	4.206*** [0.671]	5.278*** [0.481]	5.637*** [0.415]
N	4670	4185	3861	3861	3861	3861
Adjusted R-square	0.996	0.998	0.998	0.998	0.998	0.998

Note: Standard errors in brackets. Significance level can be read as * p<0.05, ** p<0.01, *** p<0.001.

Table 4: Heterogeneity in Effect of Life Expectancy at Birth on Earnings

	MALE			FEMALE		
	I	II	III	I	II	III
Age	0.103***	0.084***		0.069***	0.054***	
	[0.018]	[0.018]		[0.011]	[0.007]	
Age square*(1/100)	-0.10***	-0.10***		-0.10***	-0.10***	
	[0.000]	[0.000]		[0.000]	[0.000]	
% Urban	1.252***	0.947***	0.892***	0.907***	0.933***	0.949***
	[0.475]	[0.201]	[0.154]	[0.121]	[0.130]	[0.151]
% Male						
Life Expectancy at Birth	***0.011***	***0.010***	***0.012*****	***0.009*****	***0.007*****	***0.009******
	[0.008]	***[0.007]***	***[0.006]***	***[0.004]***	***[0.003]***	***[0.003]***
Parents Life Expectancy		0.002	0.008*		0.001	0.004*
		[0.004]	[0.004]		[0.003]	[0.002]
% Married		0.472**	0.578***		0.429***	0.434***
		[0.180]	[0.144]		[0.125]	[0.074]
Cohort Size		-12.837***	-12.065***		-9.243	-10.247
		[0.643]	[1.059]		[7.188]	[7.816]
Birth Year FE			YES			YES
Constant	3.796***	4.350***	5.671***	4.596***	4.905***	5.708***
	[0.893]	[0.980]	[0.562]	[0.370]	[0.334]	[0.324]
N	4204	3878	3878	4149	3822	3822
adj. R-square	0.997	0.998	0.998	0.997	0.998	0.998

	URBAN			RURAL		
	I	II	III	I	II	III
Age	0.096***	0.083***		0.069***	0.054***	
	[0.012]	[0.010]		[0.013]	[0.015]	
Age square*(1/100)	-0.10***	-0.10***		-0.10***	-0.10***	
	[0.000]	[0.000]		[0.000]	[0.000]	
% Urban						
% Male	0.495***	0.377***	0.103	0.193	0.124	0.074
	[0.091]	[0.074]	[0.091]	[0.140]	[0.133]	[0.147]
Life Expectancy at Birth	***0.012******	***0.011******	***0.010******	***0.002***	***0.003***	***0.007***
	[0.004]	***[0.003]***	***[0.003]***	***[0.008]***	***[0.009]***	***[0.007]***
Parents Life Expectancy		0.002	0.005		0.007	0.011**
		[0.002]	[0.003]		[0.006]	[0.005]
% Married		0.294***	0.501***		0.534**	0.592***
		[0.089]	[0.095]		[0.237]	[0.124]
Cohort Size		-7.169*	-5.264		-7.316***	-6.624***
		[4.065]	[3.542]		[1.637]	[2.369]
Constant	4.337***	4.653***	6.396***	5.706***	5.380***	5.988***
	[0.543]	[0.473]	[0.336]	[0.900]	[1.086]	[0.642]
N	4200	3874	3874	3959	3657	3657
adj. R-square	0.998	0.999	0.998	0.996	0.996	0.996

Note: Standard errors in brackets. Significance level can be read as * p<0.05, ** p<0.01, *** p<0.00

Table 5: *Mincerian* Earnings Functions and Return to Schooling

	POOL		Male		Female		URBAN		RURAL	
	I	II	I	II	I	II	I	II	I	II
Age	0.081***	0.072***	0.088***	0.077***	0.069***	0.054***	0.087***	0.081***	0.073***	0.049***
	[0.007]	[0.006]	[0.008]	[0.008]	[0.005]	[0.005]	[0.007]	[0.005]	[0.009]	[0.009]
Age-square	-0.10***	-0.10***	-0.10***	-0.10***	-0.10***	-0.10***	-0. 10***	-0.10***	-0.10***	-0.000***
*(1/100)	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Years of	0.121***	0.094***	0.137***	0.106***	0.093***	0.073***	0.102***	0.095***	0.096***	0.088***
Schooling	[0.025]	[0.013]	[0.028]	[0.020]	[0.012]	[0.008]	[0.018]	[0.010]	[0.022]	[0.024]
%Urban		0.528***		0.406**		0.604***				
		[0.172]		[0.171]		[0.093]				
%Male		0.363***						0.392***		0.244**
		[0.116]						[0.099]		[0.108]
% Married		0.329***		0.308***		0.341***		0.172*		0.406**
		[0.110]		[0.097]		[0.127]		[0.103]		[0.182]
Cohort Size		-5.78***		-7.35***		-6.91		-4.073*		-4.967***
		[1.208]		[1.141]		[6.369]		[2.291]		[1.486]
Constant	4.291***	4.243***	4.135***	4.416***	4.665***	4.748***	4.302***	4.352***	4.746***	5.110***
	[0.337]	[0.307]	[0.373]	[0.335]	[0.211]	[0.126]	[0.343]	[0.230]	[0.348]	[0.386]
Survey FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	4670	4185	4690	4204	4622	4149	4687	4201	4446	3959
Adjusted R-square	0.998	0.998	0.998	0.998	0.997	0.998	0.998	0.999	0.996	0.996

Note: Standard errors in brackets. Significance level can be read as *p<0.05, ** p<0.01, ***p<0.001.

Table 6: Comparison of Returns to Schooling from Wald estimator and *Mincerian* functions

	POOL	MALE	FEMALE	URBAN	RURAL
Mincerian Earnings Function [std. error]	9.40%*** [0.013]	10.60%*** [0.020]	7.30%*** [0.008]	9.50%*** [0.010]	8.80%*** [0.024]
Indirect Least Square Estimates [std. error]	7.90%*** [0.018]	9.90%*** [0.024]	4.30%*** [0.011]	10.21%*** [0.017]	2.90% [0.026]

Note: The *Mincerian* return to schooling estimates are taken from specification II for each group in table 5 while the Indirect Least Square estimates for pooled sample is derived by dividing the coefficients of life expectancy at birth from specification VII in table 3 by the coefficient of life expectancy at birth in specification IV in table 1. Similarly, for each gender and Rural-Urban group, the ILS estimates are obtained by dividing the life expectancy coefficients from specification II in table 4 by life expectancy coefficients from specification II in table 2. Standard errors for indirect least square estimates are obtained by 500 bootstrap replications.

Table 7: Life Expectancy at Birth Effects on Schooling and Earning, Individual Level analysis

	POOL		MALE		FEMALE		RURAL		URBAN	
	Schooling	*Earning*	*Schooling*	*Earning*	*Schooling*	*Earning*	*Schooling*	*Earning*	*Schooling*	*Earning*
Urban	-1.621*** [0.087]	-0.424*** [0.017]	-1.812*** [0.087]		-1.253*** [0.074]			-0.483*** [0.016]		-0.396*** [0.017]
Male/Female	0.247*** [0.032]	-0.366*** [0.010]		-0.358*** [0.012]		-0.355*** [0.008]	0.0932* [0.041]		0.287*** [0.031]	
Life Expectancy at Birth	***0.115*** [0.008]***	***0.013*** [0.001]***	***0.103*** [0.008]***	***0.015*** [0.002]***	***0.150*** [0.011]***	***0.012*** [0.002]***	***0.120*** [0.009]***	***0.012*** [0.002]***	***0.0996*** [0.007]***	***0.011*** [0.001]***
Parents Life Expectancy	0.0219*** [0.006]	0.005*** [0.001]	0.0270*** [0.007]	0.008*** [0.002]	0.008 [0.006]	0.002 [0.001]	0.0233** [0.007]	0.006*** [0.002]	0.0233*** [0.006]	0.006*** [0.001]
Age		0.077*** [0.004]		0.087*** [0.005]		0.060*** [0.005]		0.064*** [0.003]		0.081*** [0.005]
Age square*(1/100)		-0.10*** [0.000]		-0.10*** [0.000]		-0.10*** [0.000]		-0.10*** [0.000]		-0.10*** [0.000]
Marital Status		-0.107*** [0.008]		-0.245*** [0.008]		0.024 [0.015]		-0.100*** [0.008]		-0.113*** [0.009]
Cohort Size		-2.483*** [0.286]		-2.666*** [0.357]		-1.270*** [0.375]		-1.771*** [0.335]		-2.497*** [0.279]
Constant	3.315*** [0.645]	4.925*** [0.217]	3.749*** [0.640]	4.474*** [0.279]	2.112* [0.835]	4.979*** [0.245]	1.297 [0.796]	5.062*** [0.218]	4.242*** [0.569]	4.838*** [0.225]
Birth-Year FE	YES		YES		YES		YES		YES	
Survey Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	2725213	2543741	1589055	1473199	1136158	1070542	872016	832947	1853197	1710794
adj. R-square	0.444	0.898	0.442	0.916	0.451	0.869	0.589	0.888	0.340	0.904
F	36.730	1371.40	39.200	1040.50	49.770	350.70	39.120	292.50	30.190	468.70

Note: Standard errors in brackets. Significance level can be read as *p<0.05, ** p<0.01, *** p<0.001.

Table 8: Return to Schooling from *Mincerian* Earnings Functions and Indirect Least Square Estimates, Individual Level analysis.

	POOL		MALE		FEMALE		RURAL		URBAN	
	Mincerian Estimate	*Indirect Least Square*	*Mincerian Estimate*	*Indirect Least Square*	*Mincerian Estimate*	*Indirect Least Square*	*Mincerian Estimate*	*Indirect Least Square*	*Mincerian Estimate*	*Indirect Least Square*
Years of Schooling	***0.097****** ***[0.001]***	***0.112***	***0.089****** ***[0.001]***	***0.150***	***0.107****** ***[0.001]***	***0.078***	***0.086****** ***[0.001]***	***0.102***	***0.100****** ***[0.001]***	***0.105***
Age	0.062*** [0.003]		0.068*** [0.004]		0.052*** [0.004]		0.052*** [0.003]		0.066*** [0.004]	
Age square*(1/100)	-0.10*** [0.000]		-0.10*** [0.000]		-0.10*** [0.000]		-0.10*** [0.000]		-0.10*** [0.000]	
URBAN	-0.222*** [0.004]		-0.210*** [0.006]		-0.231*** [0.004]					
MALE/FEMALE	-0.450*** [0.014]						-0.493*** [0.016]		-0.427*** [0.015]	
Marital Status	-0.061*** [0.006]		-0.204*** [0.006]		0.073*** [0.012]		-0.067*** [0.007]		-0.062*** [0.007]	
Cohort Size	-0.326 [0.203]		-0.417 [0.260]		0.083 [0.294]		-0.331 [0.274]		-0.348* [0.197]	
Constant	5.102*** [0.075]		5.164*** [0.096]		4.573*** [0.093]		5.444*** [0.059]		4.865*** [0.084]	
Survey Fixed Effect	YES		YES		YES		YES		YES	
N	2572536		1491901		1080635		842749		1729787	
adj. R-square	0.912		0.928		0.888		0.901		0.919	
F	3299.304		2002.804		3300.83		1117.629		2029.129	

Note: Standard errors in brackets. Significance level can be read as *p<0.05, ** p<0.01, *** p<0.001. Indirect least square estimates for pool, male-female, and urban-rural groups are obtained by dividing the coefficients of life expectancy at birth from earning specification by the coefficient of life expectancy at birth from schooling specification in table 7.

Table 9: Life Expectancy at Birth Effects on Lifetime Earnings with Correction for Selection including Cohort Size in Labor Force

	POOL	MALE	FEMALE	RURAL	URBAN
Life Expectancy at Birth	***0.013****	***0.015****	***0.012****	***0.013****	***0.011****
	[0.001]	***[0.002]***	***[0.001]***	***[0.001]***	***[0.001]***
age	0.077***	0.091***	0.061***	0.062***	0.083***
	[0.005]	[0.006]	[0.007]	[0.005]	[0.006]
Age square*(1/100)	-0.10***	-0.10***	-0.10***	-0.10***	-0.10***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Rural	-0.371***	-0.363***	-0.362***		
	[0.010]	[0.012]	[0.008]		
Male	0.425***			0.485***	0.396***
	[0.016]			[0.015]	[0.017]
Marital Status	-0.109***	-0.248***	0.022	-0.103***	-0.116***
	[0.009]	[0.008]	[0.015]	[0.009]	[0.010]
Cohort Size	-2.05	0.22	-1.641	-3.361*	-0.671
	[1.489]	[1.697]	[1.990]	[1.944]	[1.400]
Parents Life Expectancy	0.005***	0.008***	0.002	0.006***	0.006***
	[0.001]	[0.002]	[0.001]	[0.002]	[0.001]
Correction term for selection (Proportion of the Birth-year cohort in Labor Force)	-0.743	-4.248*	0.204	2.243	-2.622
	[2.000]	[2.324]	[2.600]	[2.664]	[1.851]
Constant	4.498***	4.459***	4.986***	4.606***	4.428***
	[0.214]	[0.273]	[0.258]	[0.221]	[0.220]
N	2585968	1498455	1087513	849833	1736135
Adjusted R square	0.896	0.915	0.868	0.886	0.903

Note: Standard errors in brackets. Significance level can be read as *p<0.05, ** p<0.01, *** p<0.001.

Table 10: Life Expectancy at Birth Effects on Lifetime Earnings with Correction for selection due to participation in Labor Force

	POOL		MALE		FEMALE		RURAL		URBAN	
	Selection Equation	Earnings Equation	Selection Equation	Earnings Equation	Selection Equation	Earnings Equation	Selection Equation	Earnings Equation	Selection Equation	Earnings Equation
Household Size	*-0.039**** *[0.001]*		*-0.011**** *[0.001]*		*-0.075**** *[0.001]*		*-0.022**** *[0.001]*		*-0.055**** *[0.001]*	
Respondent is Household Head	*0.784**** *[0.003]*		*0.785**** *[0.007]*		*0.396**** *[0.005]*		*0.730**** *[0.006]*		*0.801**** *[0.004]*	
Life Expectancy at Birth	***0.002***** ***[0.000]***	***0.015***** ***[0.001]***	***0.002***** ***[0.001]***	***0.018***** ***[0.002]***	***0.006***** ***[0.000]***	***0.010***** ***[0.002]***	***-0.006***** ***[0.001]***	***0.014***** ***[0.002]***	***0.015***** ***[0.001]***	***0.010***** ***[0.001]***
Age	0.202*** [0.001]	0.040*** [0.004]	0.206*** [0.002]	0.037*** [0.005]	0.210*** [0.001]	-0.038*** [0.007]	0.198*** [0.002]	0.042*** [0.004]	0.206*** [0.002]	0.038*** [0.005]
Age square	-0.003*** [0.000]	-0.000*** [0.000]	-0.003*** [0.000]	-0.000* [0.000]	-0.003*** [0.000]	0.001*** [0.000]	-0.003*** [0.000]	-0.000*** [0.000]	-0.003*** [0.000]	-0.000*** [0.000]
Rural	0.107*** [0.003]	-0.382*** [0.009]	-0.062*** [0.006]	-0.351*** [0.012]	0.177*** [0.003]	-0.428*** [0.006]				
Male	1.688*** [0.003]	0.090*** [0.019]					1.844*** [0.006]	0.248*** [0.018]	1.629*** [0.004]	0.039* [0.020]
Marital status, if single	0.127*** [0.005]	-0.174*** [0.008]	-0.672*** [0.011]	-0.097*** [0.008]	0.520*** [0.007]	-0.303*** [0.021]	-0.01 [0.009]	-0.125*** [0.008]	0.183*** [0.007]	-0.201*** [0.010]
Cohort Size		-2.226*** [0.317]		-2.436*** [0.389]		-0.65 [0.424]		-1.725*** [0.343]		-2.043*** [0.327]
Parents Life Expectancy	0.008*** [0.000]	0.004*** [0.002]	-0.002*** [0.001]	0.008*** [0.002]	0.012*** [0.000]	-0.002* [0.001]	0.012*** [0.001]	0.005** [0.002]	-0.001*** [0.000]	0.007*** [0.002]
Inverse Mills Ratio		-0.867*** [0.050]		-1.566*** [0.115]		-1.602*** [0.097]		-0.537*** [0.046]		-0.986*** [0.056]
Constant	-3.395*** [0.038]	5.519*** [0.224]	-0.558*** [0.084]	5.219*** [0.275]	-4.105*** [0.045]	7.990*** [0.306]	-3.130*** [0.058]	5.161*** [0.229]	-3.700*** [0.050]	5.673*** [0.234]
Number of Obs.	4453457	2478294	2149295	1434525	2304162	1043769	1831123	822189	2622334	1656105
Adjusted. R square		0.90		0.92		0.87		0.89		0.91
F-stat		1301.789		916.29		724.047		220.543		330.477

Note: Standard errors in brackets. Significance level can be read as *p<0.05, ** p<0.01, *** p<0.001. From the selection equation, we calculate the linear predicted probabilities probability, and then convert these into normal densities to calculate inverse mills ratio. The group-specific inverse-mills ratios are used in the earnings specifications, which is estimated only on those who are in the labor force, to correct for the selection.

APPENDIX C. APPENDIX TO CHAPTER 4

Proofs of Propositions Presented in section Three

Proposition 1: shadow value of human capital in terms of wealth, $g(t)$, increases when life expectancy, *T*, increases.
Proof: differentiation of equation (15) w.r.t. *T* yields

$$\frac{dg(t)}{dT} = \frac{d}{dT}\left(\frac{R}{(\sigma+r)}\left[1 - e^{(\sigma+r)(t-T)}\right]\right)$$

$$= \frac{R}{(\sigma+r)}\left[-(\sigma+r)(-e^{(\sigma+r)(t-T)})\right]$$

$$= R\left[(e^{(\sigma+r)(t-T)})\right].$$

Since market rental rate of human capital $R > 0, \frac{dg(t)}{dT} = R\left[(e^{(\sigma+r)(t-T)})\right] > 0.$

Proposition 2**:** If life expectancy at birth, *T*, increases, purchases of educational-investment goods, *D(t)*, and effective time investment, *I(t)H(t)* would increase in every period of life.

Proof: from equation (19)-(20), we see that price of *D(t)* is $\frac{P}{g(t)}$, and price of *I(t)H(t)* is $\frac{P}{g(t)}$. *P* and *R* are assumed to be constant over lifetime. Therefore, $g(t)$ determines the movement of prices in the demand functions for investment goods. Proposition 1 shows that $\frac{dg(t)}{dT} > 0$. So, in response to a rise in *T*, prices of both *D(t)* and *I(t)H(t)* would decrease as well. Since both of these inputs are assumed normal, own price decrease would increase purchase of both inputs $\forall\, t \in [0, T]$ *i.e.*, $\frac{dD(t)}{dT} > 0$ and $\frac{dI(t)H(t)}{dT} > 0$. We explicitly show the case for $D(t)$.

$$D(t) = D\left[\frac{P}{g(t)}, \frac{R}{g(t)}\right] \forall\, t \in [0, T]$$

$$\frac{dD(t)}{dT} = -\frac{P}{[g(t)]^2} * D_1 * \frac{dg(t)}{dT} - \frac{R}{[g(t)]^2} * D_2 * \frac{dg(t)}{dT}$$

$$\frac{dD(t)}{dT} = -\left[\frac{P}{g(t)} * D_1 + \frac{R}{g(t)} * D_2\right] * \frac{\frac{dg(t)}{dT}}{g(t)}.$$

Strong concavity and twice differentiability of production function implies that

$\frac{P}{g(t)} * D_1 + \frac{R}{g(t)} * D_2 < 0$, and from proposition 1, $\frac{dg(t)}{dT} > 0$. These together imply that $\frac{dD(t)}{dT} > 0$.

Proposition 3: If life expectancy, *T*, increases, total human capital stock accumulated over lifetime increases as well.
Proof: From equation 21, human capital stock at any time *t* is

$$H(t) = \int_0^t e^{\sigma(\tau-t)} F[I(\tau)H(\tau), D(\tau)]d\tau + H(0)e^{-\sigma t}.$$

$$\frac{d}{dT}[H(t)] = \frac{d}{dT}\left[\int_0^t e^{\sigma(\tau-t)}F[I(\tau)H(\tau), D(\tau)]d\tau + H(0)e^{-\sigma t}\right]$$
$$= \int_0^t e^{\sigma(\tau-t)}\frac{d}{dT}\{F[I(\tau)H(\tau), D(\tau)]\}d\tau + \frac{d}{dT}(H(0)e^{-\sigma t})$$
$$= \int_0^t e^{\sigma(\tau-t)}\left\{F_1\frac{\delta I(\tau)H(\tau)}{\delta T} + F_2\frac{\delta D(\tau)}{\delta T}\right\}d\tau.$$

Proposition 2 demonstrates that $\frac{dD(t)}{dT} > 0$, and $\frac{dI(t)H(t)}{dT} > 0$. Again, $F[I(\tau)H(\tau), D(\tau)]$ is increasing in both of its argument if $I(\tau) > 0$. So, the term in braces inside the integral is positive, *i.e.*, $\left\{F_1\frac{\delta I(\tau)H(\tau)}{\delta T} + F_2\frac{\delta D(\tau)}{\delta T}\right\} > 0$. These together imply that $\frac{d}{dT}[H(t)] > 0$.

Since in response to a rise in life expectancy at birth, human capital increases $\forall\, t \in [0, T]$, lifetime human capital stock, $H(T) = \int_0^T\left[e^{\sigma(\tau-t)}F[I(\tau)H(\tau), D(\tau)]d\tau + H(0)e^{-\sigma T}\right]dt$, will increase as well. Further, since lifetime accumulation of human capital increases, the value of the stock would change as well in response to a rise in life expectancy. Proposition 4 below explains this.

Proposition 4: If life expectancy *T* increases, lifetime labor income increases.

Proof: the present value of lifetime labor earnings, as stated in equation 23, is
$V = g(0)H(0) + \int_0^T e^{-rt}\{g(t)F[I(t)H(t), D(t)] - PD(t) - RIH(t)\}\, dt,$

where the term inside the integral is the time *0* present value of net profits from human capital accumulation as of time *t*. Differentiating w.r.t *T* yields

$$\frac{dV}{dT} = H(0)\frac{d}{dT}\left[\frac{R}{(\sigma+r)}\left[1 - e^{(\sigma+r)(-T)}\right]\right] + \frac{d}{dT}\left[\int_0^T e^{-rt}\{g(t)F[I(t)H(t), D(t)] - PD(t) - RIH(t)\}\, dt\right].$$

Applying the Leibniz Rule on the above yields

$$\frac{dV}{dT} = H(0)\left[R\left[e^{(\sigma+r)(-T)}\right]\right] + \left[\int_0^T e^{-rt}\frac{d}{dT}\{g(t)F[I(t)H(t), D(t)] - PD(t) - RIH(t)\}\, dt\right] + e^{-rT}[g(T)F[I(T)H(T), D(T)] - PD(T) - RIH(T)],$$

$$\frac{dV}{dT} = H(0)\left[R\left[e^{(\sigma+r)(-T)}\right]\right] + \left[\int_0^T e^{-rt}\left\{\left[g(t)F_1\frac{\delta I(\tau)H(\tau)}{\delta T} + g(t)F_2\frac{\delta D(\tau)}{\delta T}\right] + F[I(t)H(t), D(t)]\frac{dg(t)}{dT} - P\frac{dD(t)}{dT} - R\frac{dH(t)}{dT}\right\}dt\right] + e^{-rT}[g(T)F[I(T)H(T), D(T)] - PD(T) - RIH(T)].$$

From the first order conditions (9), we substitute $g(t)F_1 = R$, and from (10), $g(t)F_2 = P$ into the above equation to yield

$\frac{dV}{dT} = H(0)\left[R\left[e^{(\sigma+r)(-T)}\right]\right] + \left[\int_0^T e^{-rt}\left\{R\frac{\delta I(\tau)H(\tau)}{\delta T} + P\frac{\delta D(\tau)}{\delta T} + F[I(t)H(t), D(t)]\frac{dg(t)}{dT} - P\frac{\delta D(\tau)}{\delta T} - R\frac{\delta I(\tau)H(\tau)}{\delta T}\right\}dt\right] + e^{-rT}[g(T)F[I(T)H(T), D(T)] - PD(T) - RIH(T)]$,

$\frac{dV}{dT} = H(0)\left[R\left[e^{(\sigma+r)(-T)}\right]\right] + \left[\int_0^T e^{-rt}\left\{F[I(t)H(t), D(t)]\frac{dg(t)}{dT}\right\}dt\right] + e^{-rT}[g(T)F[I(T)H(T), D(T)] - PD(T) - RIH(T)]$.

In the RHS of the above equation, clearly the first term $H(0)\left[R\left[e^{(\sigma+r)(-T)}\right]\right] > 0$. Since $\frac{dg(t)}{dT} > 0$ from proposition 1, the middle term $\left[\int_0^T e^{-rt}\left\{F[I(t)H(t), D(t)]\frac{dg(t)}{dT}\right\}dt\right]$>0. Finally, the last term $e^{-rT}[g(T)F[I(T)H(T), D(T)] - PD(T) - RIH(T)]$ is the present value of net profit from human capital investment in the last period *T*. Although, Heckman assumes that human capital investment might be taken even at a financial loss because of nonmarket benefit of education, in the current setting, condition 14 states that shadow value of human capital is *0* in the last period *T*. It implies that an individual at her last stage of life would not invest in human capital since she will not survive in periods after T to reap the benefits of the investment. Accordingly, in the last period *T*, $D(T) = 0$, and, therefore, the last term is 0. It suggests that lifetime labor income from human capital investment is positive, i.e., $\frac{dV}{dT} > 0$.

Proposition 5: If *T* increases, marginal utility of lifetime wealth, $\lambda(0)$, decreases.
Proof*:* Since $\lambda(0)$ is the marginal utility of wealth or shadow value of lifetime wealth as of period 0, and utility function follows concavity, it would decrease if lifetime wealth increases. Throughout the lifetime, wealth comes from two sources-labor income from exploiting human capital in the labor market and initial asset. Proposition 4 shows that lifetime labor income increases in response to gain in *T*. However, *T* does not affect initial endowment of assets $A(0)$. These together imply that an increase in *T* would increase lifetime wealth, which, in turn, suggests that $\lambda(0)$ falls when life expectancy at birth increases.

Proposition 6: If life expectancy *T* increases, consumption of leisure in human capital adjusted efficiency units, i.e., effective leisure $HL(t)$ increases. However, measured hours of leisure,$L(t)$, responds in an ambiguous manner.

Proof: Since leisure is by assumption a normal good and an increase in life expectancy at birth (*T*) increases lifetime income, value of leisure should increase$\forall\ t \in [0, T]$. From equation 17, $H(t)L(t) = HL\left[\lambda(0)e^{(\rho-r)t}P,\ R\lambda(0)e^{(\rho-r)t}\right]$.

Since proposition 5 shows that $\frac{d\lambda(0)}{dT} < 0$, when T increases effective leisure becomes cheaper through reduced value for $\lambda(0)$. It implies that in response to gain in T, for an individual value of leisure increases at all ages. However, the direction of change in consumption of leisure in natural units of time, $L(t)$, is not quite clear. For exposition, note that effective leisure can be expressed as

$$InL(t) = InHL(t) - InH(t)$$

$$\frac{d}{dT}[InL(t)] = \frac{d}{dT}[InHL(t)] - \frac{d}{dT}[InH(t)],$$

$$= \frac{d[InHL(t)]}{d\lambda(0)} * \frac{d\lambda(0)}{dT} - \frac{d[InH(t)]}{dT}.$$

Since leisure is a normal good, $\frac{d[InHL(t)]}{d\lambda(0)} * \frac{d\lambda(0)}{dT} > 0$, and in proposition 3, we have already shown that the second term $\frac{d[InH(t)]}{dT} > 0$. Therefore, we cannot sign $\frac{d}{dT}[InL(t)]$. If $\frac{d[InHL(t)]}{d\lambda(0)} * \frac{d\lambda(0)}{dT} > \frac{d[InH(t)]}{dT}$, $L(t)$ increases in response to a rise in T and *vice versa*.

Table B1: List of Countries and Years of Surveys

Country	Survey Year	Country	Survey Year
Afghanistan	2007	Latvia	2004, 2012
Albania	2003	Moldavia	2002, 2005
Argentina	2012	Maldives	1998, 2004
Austria	2004, 2012	Mexico	1989, 2012
Azerbaijan	1995	Macedonia	2003, 2005
Belgium	2004, 2011	Malta	2009, 2012
Burkina Faso	1994, 2009	Mongolia	2002, 2011
Bulgaria	2003, 2012	Mozambique	2002
Bosnia-Herzegovina	2001, 2004	Mauritius	1999, 2012
Belarus	1998	Malawi	2004, 2010
Belize	1993, 1999	Namibia	1993
Bolivia	1992, 2012	Niger	1995, 2011
Brazil	1981, 2012	Nigeria	1993
Canada	1981, 2001	Nicaragua	1993, 2009
Switzerland	2011	Holland	2005, 2012
Chile	1987, 2011	Norway	2004, 2012
China	2002	Pakistan	2010
Cameroon	2001	Panama	1989, 2012
Colombia	2001, 2012	Peru	1997, 2012
Costa Rica	1989, 2009	Philippines	2003, 2011

Table B1 continued.

Country	**Survey Year**	**Country**	**Survey Year**
Cyprus	2005, 2012	Poland	2005, 2012
Czech Republic	2005, 2012	Puerto Rica	1970, 2005
Germany	2005, 2012	Portugal	2004, 2012
Denmark	2004, 2012	Paraguay	1990, 2011
Dominican Republic	1996, 2011	Romania	1994, 2012
Ecuador	1994, 2012	Russia	1994, 2009
Spain	2004, 2012	Senegal	2011
Estonia	2004, 2012	Solomon Islands	2005
Ethiopia	2005	Sierra Leone	2003, 2011
Finland	2004, 2012	El Salvador	1991, 2009
France	2004, 2012	Serbia	2008
Micronesia, Fed. States.	2000	Sao Tome and Principe	2000, 2010
Gabon	2005	Surinam	1999
United Kingdom	2005, 2012	Slovakia	2003, 2012
Greece	2004, 2012	Slovenia	2005, 2012
Guatemala	2000, 2011	Sweden	2004, 2012
Guyana	1992	Swaziland	2000
Honduras	1991, 2011	Chad	2003
Croatia	2004, 2012	Togo	2006
Haiti	2001	Thailand	1990, 2009
Hungary	2004, 2012	Tajikistan	2003
Indonesia	1998, 2010	Turkmenistan	1998
India	1983, 2007	East Timor	2001, 2007
Ireland	2004, 2009	Tunisia	2001
Iceland	2004, 2012	Turkey	2002
Italy	2004, 2012	Tanzania	2000
Jamaica	1990, 2002	Uganda	1992
Jordan	2002, 2010	Ukraine	2000, 2005
Kenya	2005	Uruguay	1989, 2012
Kyrgyzstan	1997	USA	1990, 2010
Cambodia	1997, 2008	Venezuela	1989, 2006
Lao PDR	1997, 2008	Vietnam	2002, 2010
Lebanon	2011	West Bank and Gaza	1998, 2008
Sri Lanka	1993, 2009	Zaire	2005
Lithuania	2005, 2012	Zambia	1998, 2010
Luxembourg	2004, 2012		

Total 188 surveys from 111countries spanning the years 1970-2012.

Robustness

Table C1 reports the first set of robustness results for schooling. In general, estimates of life expectancy effect on schooling from all of the specifications for robustness show similar pattern as we observed before: the coefficients of life expectancy at birth always turn out to be positive and statistically significant at less than one percent level. The estimates lie in the range of 0.04-0.111. Life expectancy effect on schooling is smaller in magnitude both in the un-weighted case, and when a young group sample is used. The estimates are similar in magnitude when we use only one survey per country (the latest possible survey), or we replace parents' life expectancy by a 15-year lagged value of life expectancy at birth. Similar to previously reported estimates, if life expectancy at birth increases by one year, the birth-year cohort will spend 0.1 years more time in school.

For lifetime earnings, table C2 reports that the positive and statistically significant relationship between life expectancy at birth and earnings is robust across sample consisting only young-age group, controls through higher order age terms, alternative definition of parents' life expectancy, and alternative assumption on cohort fixed effect. In contrast to that for schooling, we fail to notice any statistical significance for un-weighted regression and sample consisting only one survey per country. Incorporating higher order age terms in our specification, we obtain a 0.8% increase in income from an additional year of gain in life expectancy at birth, which is similar to what we observe in our main specification (specification IV in table 3). Including cohort fixed effect by defining cohort at five-year birth range shrinks life expectancy effect on lifetime earnings, as reported in column IV. Parents' life expectancy if constructed by taking a 15-year lag does not affect the life expectancy effect on earnings. Next, we restrict our sample on young working group by including only those who are in age group 25-45. The coefficient of life expectancy at birth now shrinks further in magnitude (1.6% per year of added life expectancy) and still remain statistically significant. Although we do not report in table C2, the coefficient of life expectancy at birth turns out to be positive and statistically significant if we alternatively define the young working group to be those in the age ranges of 25-50 or 25-40.

Life Expectancy at birth or something else?

The literature finds that weather shocks impact well-being through multifaceted channels including reduced labor productivity, agricultural output shock, mortality due to disease outbreak, and political instability instigating civil war (Dell et al. 2014; Maccini and Yang 2009). As a control for other factors prevailing at the time of birth, we incorporate country-cohort specific average temperature and average precipitation. The weather attributes that prevailed at the time of birth appear to be exogenous. This robustness check will give us an indication of whether the positive and statistically significant positive association of life expectancy at birth with schooling and lifetime earnings is truly an exogenous impact of life expectancy at birth, or it is actually due to any cohort and country specific omitted or unobserved factors which influence both health and human capital.[65]

The time series on country averages on yearly temperature and precipitation is compiled from CRY-CY dataset, produced and maintained by Climate Research Unit at the University of East Anglia, UK. [66] The CRU CY dataset maintains information on monthly, seasonal or annual

---

[65] For illustration, while one was in the womb, if there was a severe flood in his/her locality, which caused food scarcity and high infant mortality in that area, then any life expectancy effects we observe in our model would actually be the true effect of weather shocks.

[66] The underlying dataset behind the construction of CRU-CY dataset is CRU TS dataset. The construction is described as "*The original data (CRU TS 3.21) took the form of a value for each month and each box on a 0.5 degree latitude/longitude grid. CRU assigned each box to a single country. For each country CRU calculated the weighted mean of the values from its constituent grid boxes for each month in turn. Each grid box was weighted by surface area, using the cosine of the latitude. The seasonal and annual values are the means of their constituent months. The CRU TS dataset prioritizes completeness, and has no missing data over land. Where observations are unavailable, the 1961-90 monthly climatic mean is used as a substitute. In data sparse regions of the world, this can lead to repeated values, and this can show up in derived products such as CRU CY.*"

spatial averages on ten climate variables including temperature and precipitation. We utilize annual averages of these two variables- temperature and precipitation. The first two columns in table C3 report the results from the earnings specifications with weather variables while the last two columns report those for years in school.

The estimates show that the inclusion of birth-year-and-country specific weather variables do not alter the impact of life expectancy at birth on either schooling or lifetime earnings compared to what we observed above. Controlling for any possible weather shock at the time of birth, we observe that one additional year gain in life expectancy increases investment in schooling by 0.11 year and lifetime earnings by 0.8%. We do not observe any such independent effects of temperature and precipitation on either schooling or earnings in our complete specification. However, specifications excluding life expectancy variables in column II and IV reveal that any possible temperature shock at the time of birth is associated with lifetime earnings but not schooling, while high precipitation at the time of birth lead one to spend more time in school.

Life expectancy at higher ages

Some recent papers question the appropriateness of use of life expectancy at birth emphasizing on concerns related to high infant and child mortality rate. We check the strength of our findings to life expectancy at ages beyond infancy by incorporating life expectancy at age five and ten in place of that at birth in our empirical analysis. However, since life expectancy at age 5 and 10 are not available before 1950 for most of the countries in our sample, we use a truncated sample of what we have used so far. Data on life expectancy at age 5 and 10 is available from world population prospects (2012) published by Population division, Department of Economic and Social Affairs, United Nations.[67] Note that this data is available at 5-year range, for example, those who were born between 1950 and 1955 in Brazil share the same life expectancy at five or ten. So, in this empirical exercise life expectancy at age five or ten varies by five-year-cohort within a country, whereas previously it varied by birth-year. Other variables would remain the same, and will vary birth-year. To consistently assign life expectancy measures at higher ages, for a birth-year cohort we assign a five year forwarded value as measures of life expectancy at five, and ten-year forwarded value as measure of life expectancy at ten.

The columns I-III in table C4 reports the results for schooling while columns IV-VI for lifetime earnings. To facilitate a comparison of the coefficients of life expectancy at birth with life expectancy at five and ten, we estimate one specifications with life expectancy at birth. The impact of life expectancy at age five or ten on time spent in school is similar to what we observe for life expectancy at birth. An additional year of life expectancy at birth increases school-years by 0.11 year while an additional year of gain in life expectancy at age five and ten increases years in school respectively by 0.115 and 0.10 year. Overall we observe estimates in similar magnitude compared to our estimates in table 1 and 2. Similarly, the impact of life expectancy at higher ages, at five and ten, on lifetime earnings is positive and statistically significant in this truncated sample. An additional year of life expectancy gain at age five or ten increases lifetime earnings by 1.1% and 1.3%, which is close to the 0.9% effect that we have observed for life expectancy at birth.

[67] Various region, gender and age specific life expectancy data is available at http://esa.un.org/wpp/Excel-Data/mortality.htm (*last accessed on November 13th, 2014*)

Table C5 reports the results from a replication of table C4 with individual level data.[68] In this table, both for schooling and earnings, as we move from life expectancy at birth to higher ages, the coefficient of life expectancy at birth increases in magnitude. An additional year of gain in life expectancy at age 10 increase school years by 0.129 years, which is 16% larger compared to the effect of life expectancy at birth. For earnings, the life expectancy at age 10 exhibits a 61% larger effect compared to similar effect from life expectancy at birth. One interesting finding from this robustness exercise is that parents' life expectancy turnout to be positive and statistically significant in almost all of the specifications with life expectancy at higher ages for both of schooling and earnings. It implies that parents with higher life expectancy were expectedly healthier and possibly more educated and rich that is transmitted to their children either through better health or more investment into children.

[68]To facilitate a comparison of the coefficients of life expectancy at birth with life expectancy at five and ten, we also replicate table C5 exclusively for the sample for which life expectancy at age of five and ten are available. The estimates from the balanced and unbalanced samples are close.

Table C1: Life Expectancy at Birth and Schooling- Robustness Check with Different Weighting, Surveys and Sample Groups

	I	II	III	IV	V
	No Weight	Single Survey(Oldest)	Alternative Definition of Parents Life Expectancy	Alternative Definition of Cohort Fixed Effect	Sample Consisting only Young Group (Age 25 to 50)
% Urban	3.811***	6.028***	6.477***	6.715***	6.237***
	[0.654]	[1.415]	[1.234]	[1.220]	[1.464]
% Male	0.991*	-1.336**	-0.896*	-0.932*	-0.182
	[0.526]	[0.643]	[0.476]	[0.495]	[0.483]
Life Expectancy at Birth	***0.039****	***0.104*****	***0.107*****	***0.112*****	***0.111*****
	[0.017]	***[0.023]***	***[0.021]***	***[0.023]***	***[0.023]***
Parents Life Expectancy (lag of 25 years)	-0.008	-0.008		0.023**	0.031***
	[0.010]	[0.009]		[0.011]	[0.010]
Parents Life Expectancy (lag of 15 years)			0.026***		
			[0.010]		
Cohort FE	YES	YES	YES	YES	YES
Constant	4.896***	1.566	-1.064	-1.198	-1.681
	[1.111]	[1.841]	[1.617]	[1.736]	[1.777]
N	3861	3022	3977	3861	3090
Adjusted R-square	0.945	0.986	0.987	0.987	0.988

Note: Standard errors in brackets. Significance level can be read as * p<0.05, ** p<0.01, *** p<0.001. Young Age group specific analysis is robust to age group 25-45 and 25-40. In column IV, instead of birth-year level, cohort fixed effect is defined at five year birth range.

Table C2: Life Expectancy at Birth and Earnings: Robustness Check with Different Weighting, Specification and Survey Selection

	I	II	III	IV	V	VI
	No Weight	Single Survey(Oldest)	Higher Order Age Variable	Alternative Definition of Cohort Fixed Effect	Alternative Definition of Parents Life Expectancy	Sample Consisting only Young Group (Age 25 to 45)
Age	0.068***	0.072***	-0.094		0.079***	0.089***
	[0.007]	[0.012]	[0.168]		[0.014]	[0.009]
Age Square	-0.001***	-0.001***	0.006		-0.001***	-0.001***
	[0.000]	[0.000]	[0.006]		[0.000]	[0.000]
% Urban	0.883***	1.006***	1.109***	1.043***	1.118***	0.953***
	[0.101]	[0.206]	[0.193]	[0.221]	[0.203]	[0.166]
% Male	0.480**	-0.007	0.210*	-0.11	0.210*	0.363**
	[0.186]	[0.145]	[0.121]	[0.155]	[0.111]	[0.140]
Life Expectancy at Birth	***0.004***	***0.005***	***0.008****	***0.005****	***0.008****	***0.016******
	[0.003]	***[0.005]***	***[0.005]***	***[0.003]***	***[0.004]***	***[0.006]***
Parents Life Expectancy	0.00	-0.001	0.003	0.004		0.008**
(lag of 25 years)	[0.003]	[0.003]	[0.003]	[0.003		[0.004]
Parents Life Expectancy					0.005	
(lag of 15 years)					[0.004]	
% Married	0.249***	0.491***	0.491***	0.779***	0.446***	0.305**
	[0.063]	[0.169]	[0.179]	[0.200]	[0.152]	[0.138]
Cohort Size	-5.843***	-10.320***	-8.636***	-7.792***	-8.684***	-7.948***
	[2.105]	[1.624]	[1.513]	[1.785]	[1.352]	[1.619]
Age Cube			0			
			[0.000]			
Age^4			-0.094			
			[0.168]			
Constant	4.943***	5.184***	5.881***	6.072***	4.087***	3.241***
	[0.477]	[0.565]	[1.747]	[0.316]	[0.700]	[0.712]
Cohort FE				YES		
Survey FE	YES	YES	YES		YES	YES
Number of Observations	3861	3022	3861	3861	3977	2638
Adjusted R square	0.995	0.998	0.998	0.998	0.998	0.999

Note: Standard errors in brackets. Significance level can be read as * p<0.05, ** p<0.01, *** p<0.001.Specification VI is robust to age group 25-40 and 25-50. In column IV, position in the life cycle is controlled by including cohort fixed effect while defining cohort at 5-year birth range.

Table C3: Life Expectancy at Birth, Earnings and Years in School: Robustness Check including Temperature and Precipitation

	Log Wage		Years in School	
	I	II	III	IV
Age	0.077***	0.071***		
	[0.013]	[0.011]		
Age Square*100	-0.10***	-0.10***		
	[0.000]	[0.000]		
% Urban	1.107***	1.082***	6.772***	6.737***
	[0.198]	[0.207]	[1.233]	[1.355]
% Male	0.212*	0.035	-0.981**	-3.303***
	[0.122]	[0.166]	[0.506]	[1.223]
Life Expectancy at Birth	***0.008****		***0.112******	
	[0.005]		***[0.023]***	
Average Precipitation at the time of Birth	-0.007	-0.01	0.028	0.071***
	[0.005]	[0.006]	[0.036]	[0.023]
Average Temperature at the time of Birth	0.004	0.007**	-0.003	0.006
	[0.004]	[0.003]	[0.010]	[0.012]
Parent's Life expectancy	0.003	0.002	0.024**	0.013
	[0.003]	[0.003]	[0.011]	[0.012]
% Married	0.477***	0.549***		
	[0.156]	[0.164]		
Cohort Size	-8.680***	-8.503***		
	[1.494]	[1.616]		
Constant	4.256***	5.133***	-1.624	6.885***
	[0.623]	[0.314]	[1.605]	[1.412]
Cohort FE	YES	YES	YES	YES
Survey FE	YES	YES	YES	YES
N	3751	3751	3751	3751
Adjusted R-square	0.998	0.998	0.987	0.982

Note: Standard errors in brackets. Significance level can be read as * p<0.05, ** p<0.01, *** p<0.001.

Table C4: Life Expectancy, Earnings and Education: Robustness Check with Life Expectancy at Higher Ages.

	Years in School			**Log of Earnings**		
	I	II	III	IV	V	VI
% Urban	6.715***	5.633***	5.777***	1.112***	0.834***	0.902***
	[1.220]	[1.143]	[1.323]	[0.196]	[0.146]	[0.150]
% Male	-0.932*	-1.136*	-1.444*	0.206*	0.296***	0.261**
	[0.495]	[0.573]	[0.756]	[0.118]	[0.098]	[0.106]
Life Expectancy at Birth	***0.112****			***0.009****		
	[0.023]			***[0.005]***		
Life Expectancy at Age 5		***0.115*****			***0.011******	
		[0.045]			***[0.004]***	
Life Expectancy at Age 10			***0.101*****			***0.013******
			[0.044]			***[0.004]***
Parent's Life expectancy	0.023**	0.048***	0.039**	0.003	0.007**	0.005*
	[0.011]	[0.018]	[0.018]	[0.003]	[0.003]	[0.003]
Age				0.077***	0.085***	0.082***
				[0.013]	[0.012]	[0.013]
Age square				-0.10***	-0.10***	-0.10***
				[0.000]	[0.000]	[0.000]
% Married				0.471***	0.298***	0.369***
				[0.153]	[0.103]	[0.110]
Cohort Size				-8.577***	-6.136***	-6.368***
				[1.534]	[1.719]	[1.916]
Cohort Fixed Effect	YES	YES	YES			
Survey Fixed Effect	YES	YES	YES	YES	YES	YES
Constant	-1.198	-1.834	-0.091	4.206***	3.856***	3.814***
	[1.736]	[3.447]	[3.117]	[0.671]	[0.610]	[0.650]
N	3861	3341	3546	3861	3341	3546
Adjusted R-square	0.987	0.984	0.984	0.998	0.999	0.999

Note: Standard errors in brackets. Significance level can be read as * p<0.05, ** p<0.01, *** p<0.001.Although not reported here, life expectancy at age 15 results in similar estimate for both of schooling and earnings as those of life expectancy at age 5 or 10.

Table C5: Life Expectancy, Earnings and Education - Individual Level Data: Robustness Check with Life Expectancy at Higher Ages

	Years in School			Log of Earnings		
	I	II	III	I	II	III
Age				0.077***	0.079***	0.081***
				[0.004]	[0.006]	[0.006]
Age Square				-0.001***	-0.001***	-0.001***
				[0.000]	[0.000]	[0.000]
% Urban	-1.621***	-1.612***	-1.625***	-0.366***	-0.361***	-0.365***
	[0.087]	[0.209]	[0.202]	[0.010]	[0.022]	[0.022]
% Male	0.247***	0.329***	0.296***	-0.424***	-0.387***	-0.402***
	[0.032]	[0.069]	[0.068]	[0.017]	[0.037]	[0.036]
Life Expectancy at Birth	***0.113****			***0.013****		
	[0.008]			***[0.001]***		
Life Expectancy at Age 5		***0.126****			***0.015****	
		[0.025]			***[0.003]***	
Life Expectancy at Age 10			***0.129****			***0.021****
			[0.026]			***[0.003]***
Parent's Life expectancy	0.020***	0.048***	0.041***	0.005***	0.008***	0.007***
	[0.007]	[0.010]	[0.010]	[0.001]	[0.002]	[0.001]
Marital Status				-0.107***	-0.118***	-0.115***
				[0.008]	[0.018]	[0.018]
Cohort Size				-2.483***	-1.729***	-1.770***
				[0.286]	[0.388]	[0.379]
Cohort Fixed Effect	YES	YES	YES			
Survey Fixed Effect	YES	YES	YES	YES	YES	YES
Constant	3.718***	1.221	1.810	4.925***	4.564***	4.259***
	[0.635]	[1.775]	[1.765]	[0.217]	[0.286]	[0.279]
N	2726329	2345389	2478221	2543741	2222008	2333455
Adjusted R-Square	0.444	0.418	0.427	0.898	0.903	0.901
	101.021	25.791	35.514	1371.380	312.059	355.925

Note: Standard errors in brackets. Significance level can be read as * p<0.05, ** p<0.01, *** p<0.001.Although not reported here, life expectancy at age 15 results in similar estimate for both of schooling and earnings as those of life expectancy at age 5 or 10.

CHAPTER 5. LIFE EXPECTANCY AT BIRTH AND LIFETIME HUMAN CAPITAL INVESTMENT

1. Introduction

The maximum life expectancy at birth has increased worldwide from 57 years for the 1922 birth cohort to 77 for the 1987 birth cohort. As shown in Figure 1, this growth in life expectancy would raise expected discounted lifetime earnings by about 27%. Lifetime earnings can increase because individuals work more years over their lifetime, because the improved health increases the hours they can work per day or per year, or because they invest more in human capital which raises their productivity per hour of work. This study examines the last link, using data on life expectancy at birth and lifetime years of schooling for birth cohorts from 1905 to 1988. The data spans 147 countries and 919 surveys. Results are presented separately for men and women, and for urban and rural residents. The results show overwhelming support for a consistent and stable positive relationship between years of schooling and life expectancy at birth.

Despite a long-held theoretical presumption that increased life-expectancy will increase lifetime human capital investment (Ben Porath 1967; Becker1993; Heckman 1976), recent studies examining the link between life expectancy and years of schooling has found mixed evidence. In a cross-country long panel, Acemoglu and Johnson (2006, 2007) failed to observe any significant association between life expectancy, schooling, and per capita income that heightened the debate on the impact of life expectancy gain on human capital investment. Later, Lorentzen, McMillan and Wacziarg (2008) and Hazan (2006, 2009) drew similar conclusion. The empirical findings from all these four studies that gains in life expectancy at birth were not translated into longer time spent in school actually cast doubt on the Ben Porath mechanism. In contrast to the macro studies, the findings from micro studies more consistently support that improvement in health has a positive impact on human capital investment, except few of the recent studies (Bleakley 2010; Cutler 2010). Note that the micro studies link health status at different stages of lifecycle (e.g. in-utero health, early childhood health) instead of life expectancy at birth with individual school-enrollment information. The mixed evidences from recent empirical research motivate us to explore the issue rigorously in a broader setting.

The theoretical motivation comes from Bleakley (2007; 2010b) that incorporates health related productivity components to explain human capital investment in different economic setting. The empirical framework is built on the hypothesis that individuals take expected length of life at the time of birth as a measure of future health to plan out human capital investment path. Our empirical design follows a macro pattern of cross-country analysis but links micro level schooling information with cohorts' life expectancy at birth.

This paper contributes in establishing the causal link from life expectancy at birth to human capital investment. The study covers a large number of developing countries from sub-Saharan Africa and South Asia, and adopts a birth-cohort framework to exploit within country, across birth-cohort variation in life expectancy at birth and years of schooling. In contrast to other studies (Acemoglu and Johnson 2006; Hazan 2012; Hansen 2013; Cohen and Leker 2014), the measure of schooling, completed years of schooling, is extracted from household-surveys rather than noisier school-enrollment data based on population census. The matching of birth-year specific completed years of schooling with life expectancy at birth avoids the noisy aggregation of health and schooling decisions made by different birth cohorts. This is the first study in incorporating parents' life expectancy while controlling for the intergenerational link. The extension with individual level schooling data is innovative in examining the robustness of causal link. Finally, the regional and survey-by-survey analysis will be unique in facilitating a comparison of the effect of life-expectancy at birth on schooling across income regions and countries.

The next two sections review the recent studies on life expectancy at birth and schooling to adapt a theoretical framework that will guide our empirical framework. Section four reviews the data and variable construction. Section five and six reports the findings from empirical exercises and some extensions. The final section compares the result with other studies and draw the implications for policy.

2. Relevant Literature

We briefly discuss some of the recent studies that found contrary evidence to this prediction from the Ben Porath (1967) model. While examining the consequence of diseases for economic development based on the prediction from neoclassical growth models, Acemoglu and Johnson (2006, 2007) exploited cross-country variation in mortality reduction from global epidemiological intervention in the 1940s, and reported that increased life expectancy did not affect average years of schooling.[69] Bloom, Canning, and Fink (2014) suggested that controlling for initial health is critical since countries with poor initial health endowment experienced most improvement from mortality reduction but did not exhibit a similar increase in completed years of schooling. In another cross-country study, Lorentzen, McMillan and Wacziarg (2008) failed to observe any statistically significant impact of life expectancy at birth on secondary schooling.

Both of these two studies examined the effect of life expectancy on the stock of schooling years in a country (Cohen and Leker 2014). Note that life expectancy at birth is a decisive factor

[69] The explanation lies in the higher fertility in response to an improved health, at least in the initial stage of development. Unexpectedly large cohorts will face depressed earnings from a crowding of more individuals into the market, particularly if different birth cohorts are not perfect substitutes in production. To the extent that the potential for crowding in is expected, individuals in cohorts with rising life expectancy will moderate their human capital investment decisions to reflect anticipated depreciated earnings per unit of human capital.

for schooling decision of younger cohorts and not relevant for people who have already completed their schooling. In our birth-cohort framework, we avoid such aggregation by relating life expectancy at birth in period *t* with completed years of schooling of only cohort born in *t*.[70]

In another influential study, Hazan (2009) included retirement into the *Ben Porath* model, and derived a necessary condition that increased life expectancy has to increase lifetime labor supply to raise lifetime earnings because increased schooling is not a rational decision otherwise. As a supporting evidence, he noted that American cohorts born between 1840 and 1970 did not exhibit an increase in labor supply, from which he concluded that increases in life expectancy do not cause a rise in human capital investment through schooling. Hazan (2012) reinforces this view by studying cross-country data for the later part of the twentieth century, in which he finds positive and significant association between life expectancy at birth and schooling while such correlation goes away once life expectancy at age 5 or 10 replace the life expectancy at birth. Note that although life expectancy at birth exhibits higher variation across countries due to high rate of child and infant mortality, life expectancies at higher ages are more likely to be endogenous since parents make investment into children health. If such investments are not controlled in an empirical analysis, the life expectancy effects on schooling will be biased. In this study, we address this issue by choosing observations at the birth cohort level.

According to the *Ben-Porath* framework, the longer time spent in school in response to a higher life expectancy at birth can increase lifetime earnings because the individuals enter the labor market with an increased stock of human capital. The resulting gain in lifetime earnings from increased human capital stock can offset the reduced earnings from decreased labor supply, which suggests that increased labor supply does not need to be a necessary condition in *Ben*

[70] In contrast to our birth-cohort strategy, Hansen (2013) adopted a school-cohort based strategy and their schooling measures are not constructed from completed years of schooling.

Porath model. Cervellati and Sunde (2013) revisited Hazan's model and data. They argued that Hazan adopted a special case of *Ben Porath* model, and the fall in expected lifetime labor supply by successive cohorts of US men born between 1840 and 1930 was actually due to the fact that the cohorts spent more time in school and delayed entry into the labor market. The decreased lifetime labor supply indicates to the possibility of increased leisure as well. The increased lifetime earnings in response to a rise in life expectancy at birth implies that individuals in successive cohorts can afford more leisure. Heckman (1976) theoretically demonstrated that increased human capital investment, decreased labor supply, and higher consumption of leisure is consistent in an extension of Ben-Porath model with endogenous leisure.

In a micro-empirical study, Bleakley (2010a) compared the cohort's schooling across sites with different malarial intensity both before and after the malaria eradication campaign in several South American countries and Mexico. The impact on schooling is mixed, in fact, in one country he noted that cohorts born after the malaria eradication obtained less schooling. In another study, Cutler et al. (2010) reported that there was no statistically significant relationship between schooling and health improvement among post malaria eradication cohorts in India. Several arguments exist that would explain such an inconclusive findings on schooling (Costa 2014; Bleakley *et al*. 2014). Better nutrition increases the number of healthy workers who are more productive in unskilled works as well as better performers in school as they learn faster. Now, whether improved health would have a positive impact on education actually depends on whether the economy is brain or brawn based. In the later type, demand for physical labor is high which implies that marginal cost of time spent in school to be higher compared to a brain based economy. Consequently, improved health in such economy might not lead to higher average schooling in the population (Pitt *et al.* 2012).

There is a view that significant association between economic outcome and life expectancy at birth in cross-country regression based studies is merely a correlation rather than causal effect because of some omitted or unobserved decisive factors. Some studies have argued that better educated individuals with higher income would be more knowledgeable to spend more in preventive health care and medical expenditure and health thus might be associated with income and education in the other direction. However, studies could not reach a conclusion on this. For example, Clark and Royer (2013) did not find any positive impact of British education reform on health although it reported such changes increased wages and educational attainment.

3. Theoretical Motivation and Empirical Framework

We illustrate the various ways that an improvement in health at the time of birth can affect years of schooling following the framework advanced by Bleakley (2007, 2010b). Let the expected health at the time of birth for individual *i* be summarized by life expectancy at birth (l_{i0}). Improved health at the time of birth will alter the expected length of productive life which increases potential lifetime earnings. If health and time in school (S_{it}) are complementary inputs in the production of health, improved health will also increase the human capital that can be produced per year of schooling. The expected lifetime benefits from additional time in school at age *t* can be summarized by the marginal benefit equation $B(l_{i0}, S_{it}, P_{it}, q, \varepsilon_{it})$ where P_{it} is a vector of parental inputs and *q* an index of school quality that are inputs into the human capital production process. The unobserved term ε_{it} represents individual-specific productivity in producing human capital that are uncorrelated with health, parental or schooling inputs. The function *B* can be viewed as the impact of an additional year of schooling at time *t* on lifetime earnings or utility.

It is optimal to continue investing time in school until $B = C$, the marginal cost of an additional year of schooling. The cost of education depends on monetary costs of schooling (p_{it}), and the opportunity cost of time spent in school equal to the wage the child could earn given past investments in human capital $W(l_{i0}, P_{i\_t}, S_{i\_t}, Y_{it})$ where the notation *_t* reflects parental and school time investments before age *t*. The index Y_{it} reflects the state of the labor market for workers with skill that are close substitutes for *i* with larger values indicating stronger demand for similarly skilled workers. The opportunity cost of schooling is rising in all past accumulations of human capital, and the marginal cost of schooling $C(p_{it}, W(l_{i0}, P_{i_t}, S_{i_t}, Y_{it}))$ is rising in both direct and opportunity cost of schooling. We further assume that the marginal benefit from schooling is subject to diminishing returns ($\frac{\partial B}{\partial S_{it}} < 0$). Because the opportunity cost of schooling is rising in years of schooling, $\frac{\partial C}{\partial S_{i\_t}} > 0$.

We can illustrate how changes in expected health at the time of birth will alter expected time spent in formal schooling using Figure two. Consider two health states, one with a good draw and the other with a bad draw on life expectancy at birth. Note that at the time of birth, all planned parental inputs are conditioned on the parents' endowment at the time of birth, P_{i0}, and all subsequent parental inputs will be endogenous. Similarly, all planned trajectories for the direct and opportunity costs of schooling will be based on information at the time of birth. Therefore, all other factors affecting the marginal benefit and marginal cost of schooling are the same across the two health states. The parents will plan for the child to remain in school as long as the marginal benefit exceeds the cost. The good health state raises the marginal benefit per year of schooling because of the complementarity between health and productivity in school, but also because the child will have a longer potential time to productively exploit human capital. At the same time, the good health state has a higher opportunity cost of an additional year of schooling because of the faster accumulation of human capital. As illustrated, expected time in school increases because the increased marginal benefit rose more than the marginal cost. But the opposite could have happened, in which case the child would spend less time in school in the good health state.

In either case, the present value of lifetime earnings, given by the area under the marginal benefit curve, rises as a result of the increase in life expectancy at birth. However, the greater share of the benefits from improved child health will come from greater efficiency in the production of human capital per year of schooling (illustrated by the change in the height of the marginal benefit curves shaded by diagonal lines), and only a modest share of the increased lifetime income will come from the induced increase in years of schooling (illustrated by the cross-hatched area between $S*$ and $S**$.[71] Consequently, any response of years of schooling to increased life expectancy at birth will understate the induced increase in human capital resulting from the improved health.

This discussion suggests that life expectancy at birth could raise or lower years of schooling and that the effect of higher life expectancy at birth can be confounded with other factors as the cohort ages. Equating marginal benefit and marginal cost of schooling yields the relationship

$$S_{it} = f(l_{i0}, q, P_{it}, p_{it}, P_{i\_t}, S_{i\_t}, Y_{it}, \varepsilon_{it}) \text{ ,}$$

but as suggested by our previous discussion, $P_{it}, p_{it}, P_{i_t}, S_{i\_t}$ and Y_{it} will all be endogenously determined by information obtained as the child ages. Some of the reduced form effect of life expectancy at birth will be found through these other factors whose values will depend in part on life expectancy at birth and in part on new information revealed over time. To make this point

[71] This point was made by Bleakley (2007, 2010b).

more precise, consider the projection of the cost of schooling at time *t* on information available at the time of birth.

$$p_t = E(p_t|\Omega_0) + \xi_t$$

Innovations in the cost of schooling will be uncorrelated with information known at the time of birth. This will be true for the other factors P_{it}, P_{i_t}, S_{i_t} and Y_{it} as well. For this reason, we propose to measure the effect of life expectancy on completed schooling using only information known at the time of birth. Note that even later innovations in life expectancy can endogenously reflect investments by the parents and will generate biased inference regarding the effect of life expectancy on schooling.

Define the relevant sample as individuals of an age such that they have completed their schooling. For individual *i* in cohort *j* and country *c*, consider the specification

$$S_{ijc} = \gamma_1 LE_{jc0} + \gamma_2 LE_{jcP} + \gamma_M M_{ijc} + \gamma_U U_{ijc} + \alpha_c + a_j + \alpha_Y + \alpha_{jc} + \varepsilon_{ijc}. \quad (1)$$

The dependent variables in the above equation, S_{ijc} is the completed years of schooling for individual *i* in birth cohort *j* and country *c*. The key exogenous variable LE_{jc0} is the average life expectancy at birth for individuals in cohort *j* and country *c*. The coefficient γ_1 will provide the change in completed years of schooling for every one year increase in life expectancy. The other key independent variable is LE_{jcP}, taken as the life expectancy at birth for the parents of individuals in cohort *j* and country *c*. We use the life expectancy for birth cohorts 25 years prior as the parents' life expectancy at birth.[72] We know that increases in the parents' life expectancy at birth will increase their lifetime earnings, whether from more schooling, more human capital accumulated per year of schooling, or more years of productive work, and so we should find that

[72] We also experimented with life expectancy at birth 20 and 30 years prior as our measure of the parents' health endowment. In practice, life expectancy at birth 20, 25, and 30 years prior were highly correlated.

some of that increased parental wealth is transferred to their children in the form of greater human capital investments (γ_2 >0).[73] We also control for the fraction of the birth cohort that is male in the survey and the fraction that reside in urban areas.

The error terms include α_c, a country-specific fixed effect that holds constant the level of economic development and other political, social and economic attributes that are common across birth cohorts; α_Y, a fixed effect for the year of the survey that controls for any economic, political or health shocks that are common across states; and a_j, a fixed effect for the year of birth that is controls for health innovations and pandemics as well as other factors that would affect a birth cohort across countries. The error term ε_{ijc} represents the purely random factors that affect years of completed schooling.

The remaining variation that we use to identify our life expectancy effects is due to variation across cohorts within a country. The possible bias in our estimate is due to α_{jc}, a shock to completed schooling that is specific to birth cohort *j* within the country. Our estimate of γ_1 will be biased if this shock is correlated with changes in life expectancy for the cohort, as might be the case if a country always introduces improvements in public health with improvements in school quality.

We apply this model to two units of observation. Our most comprehensive data set aggregates completed schooling decisions to the birth cohort level within a country. For a subset of these countries, we also have data on individual completed years of schooling. The latter data set allows us additional controls for the possible bias related to the country-cohort specific fixed effect α_{jc} as we will discuss below. We cluster the standard error, ε_{ijc}, at the survey level to correct for correlated errors across birth cohorts *j* within the country *c*. We weight the

[73] See Becker and Tomes (1979, 1986).

observations to reflect the cell share of the total population in the country. We further weighted the data by the square root of the cell-size to correct for differences in measurement error variance between thin and thick cell samples.

4. Data

This study uses the World Bank's *International Income Distribution Database (I2D2),* a harmonized collection of 919 household surveys from 147 countries. A list of the countries and total number of surveys from each country is presented in table A1 in the *appendix* D. The surveys were conducted between 1960 and 2012 with 78% of the surveys collected on or after 2000. The database includes countries from all regions and income groups. Of the 147 countries, 32 are from industrialized nations, 16 from Asia and the Pacific, 20 from Central Asia and Eastern Europe, 23 from Latin America, 10 from the Middle East and North Africa, 8 from South Asia, and 38 from Sub Saharan Africa. From each survey we keep observations that have completed information on years of schooling. We include individuals in the age range 25 to 60. We use the 25 year age cut-off because individuals are likely to have completed their schooling by that age. The upper bound of age 60 is chosen to avoid the selection issues related to mortality.

Our observations are aggregated to birth-year cohorts from each survey in each country. This allowed us to access the full set of data, as many of the data sets are privileged and not open to use by non-Bank researchers. The 919 surveys totaling 44.6 million individuals in the age range 25-60 were placed in one of 3583 country survey-birth year cohorts. There were up to 87 birth-year groups per country with birth years ranging from 1901 to 1987. We further subdivided our birth cohorts by urban versus rural residence and by gender. Our analysis requires information on each birth-cohort's average completed years of schooling, proportion living in urban or rural residence, and gender.[74] All of these variables are harmonized and consistent across surveys. We also compiled information on the surveyed population versus the total population for each birth cohort in order to construct sample weights used in our regression analysis.

Our key independent variable "Life Expectancy at Birth" is compiled from United Nation's Population database and "Gap Minder".[75] "Gap Minder" constructs a measure of life expectancy at birth for almost 200 countries back from 1900 by compiling pre-1950 data on mortality rates from the Human Mortality Database and the United Nations Population Division's *World Population Prospects*.[76] We also utilize life expectancy at age 5 and 10. Because life expectancy at older ages will reflect parental investments in their children's health and human capital in response to updated information on the cohort's health, these measures are inherently endogenous, but are used to compare our findings to previously published estimates. Life expectancy at ages 5 and 10 is gathered from world population prospects (2012) published by the Population Division of the United Nations Department of Economic and Social Affairs.[77] The data are reported for 5-year birth cohorts rather than specific birth cohorts.

Figure 3 illustrates how life expectancy evolves globally across cohorts in our sample. The regional scatterplot in Figure 4 indicates that all of regions have experienced similar rising trend of life expectancy at birth across cohorts. The difference across regions in life expectancy at birth has converged overtime except for Sub Saharan Africa. In Latin America and the Caribbean, life

---

[74] We will not know where an individual was at the time of birth and so we will have some mismatch between urban and rural residence during the survey versus birth-place.

[75] The UN maintains a rich database on various socio-economic indicators http://data.un.org/Default.aspx.

[76] In the case where no estimates are available, they rely on simple model of interpolation and extrapolation to reach an approximate measure. Although "Gap Minder" admits that quality of life expectancy at birth data would vary across countries for the period before 1950, our extensive search suggests that this is the best available information covering such a wide set of countries for a long period before 1950.

[77] Various region, gender and age specific life expectancy data is available at http://esa.un.org/wpp/Excel-Data/mortality.htm

expectancy at birth was only 68.6% of that in the rich and developed countries, and the ratio has reached to 90.5% by early nineties. In the similar period, Sub Saharan Africa could narrow the gap in life expectancy at birth with the rich countries only by 24%.[78] Besides the evolution of life expectancy at birth, figure 2 and 3 also show how global and regional mean years of schooling evolve across birth cohorts. Average years of schooling across all birth cohorts follow similar pattern as life expectancy at birth does. The regional scatterplot indicates that all regions have experienced a rising trend of schooling across cohorts.

In the survey specific analysis, we utilize all 919 surveys. That means that the same cohort may show up multiple times across surveys. Since cohort-specific schooling does not change within a country, the repeated cohort observations are redundant and would overweight repeated cohorts. To correct this, we only include one observation for each country-birth cohort in our cross-country analysis. We used information from the oldest survey from each country. If multiple surveys are available in a country, we used the most recent survey to add in the birth cohorts that did not appear in the first survey.[79]

5. Results

In this section, first we report the survey specific estimates. In a following subsection we analyze the results obtained from pooling the surveys in a cross-country analysis.

5.1 Survey by survey estimates

We have estimated the life expectancy effects on completed school years in each of the 919 surveys. A notable proportion of these surveys cover developing countries, approximately

[78] Due to AIDS epidemic, the convergence for Sub Saharan Africa has been disrupted.
[79] For example, Germany has two surveys in our survey-pool, one in 2005 and the other in 2012. The youngest cohort in the former survey was born in 1980 while in the later survey the youngest was born in 1987. Since the cohorts who were born between 1980 and 1987 were under 25 during the survey of 2005, we only include these new cohorts from the second survey for Germany in our sample.

twenty percent of the total surveys originated from South Asia and Sub Saharan Africa. Table 1 reports the number of surveys we have utilized from each region, and a descriptive statistics on the estimates obtained from the survey specific exercises. We estimate specification (1) for each of these surveys-The identification comes from within survey, across cohort variation in life expectancy at birth. Note that in few of the surveys we had to exclude the control on urban-rural location because either the information is missing or because only the urban population was surveyed.

The impact of life expectancy at birth on completed years in school is quite consistent across surveys and regions. The survey specific estimates from 95% of the surveys reveal that life expectancy at birth has a positive and statistically significant impact on years of schooling. Only in 2.2% of the surveys, life expectancy coefficients turn out to be negative. A simple average of life expectancy coefficients across these surveys shows that for each additional year of gain in life expectancy at birth, individuals spent approximate 0.155 years in school. Figure 5 presents region specific kernel distributions of the coefficients of life expectancy at birth obtained from survey-specific regressions. The region specific median value of the coefficient, as indicated in the graph, reveals that highest median life expectancy effect is observed in Latin America, while the lowest in Central Asia and Europe.[80] The life expectancy estimates exhibit wide variation across surveys in Africa, East Asia and Pacific, and Industrialized countries.

5.2 Estimates from the regression on pooled surveys

As reported in table 2, life expectancy at birth is found to be positively associated with schooling. The coefficients of life expectancy at birth, γ_1, imply that a one year increase in life expectancy at birth increases years in school in the range of 0.13 to 0.15 years.

[80] For comparison, the distribution of all survey-specific estimates shows a median life expectancy effect of 0.148.

From our discussion in section 3 we realize that there might exist cohort-specific fixed factors that influence cohort's schooling investment decision. For example, cohorts born during 1930-1934 or before who reach school-going age during World War II might experience different environment than those who were born after that war.[81] Considering this, we control for such cohort specific fixed effects in two different ways. In specification IV, the cohort dummies is incorporated by defining cohorts in five years birth range while in specification V cohort is defined by birth year. [82] The inclusion of cohort specific fixed effects in specification IV and V does not change the coefficient of life expectancy at birth. If one lives 7.5 years longer, s/he would spend an additional year in school. Only in the specification with birth-year fixed effect, parent's life expectancy exhibits a positive, but small effect which is marginally significant.

5.3 Heterogeneity across groups

The effect of gain in life expectancy at birth might differ across groups. We investigate this by male, female, urban and rural groups separately. Table 3 presents the group-specific results. The estimates, in general, reveal similar pattern as those observed in the pooled sample. Increases in life expectancy at birth increase schooling more for rural than urban birth cohorts. Increased life expectancy increases years of schooling more for female than male birth cohorts. Another year of life expectancy at birth adds 0.15 years of schooling for women and 0.11 years of schooling for men. In specification V, where we control for birth-year fixed effect, a gain of 10 years in life expectancy at birth will lead a female cohort to take 0.43 years of more schooling

[81] Again, there were several United Nations programs and activities to improve health and education across the developing countries. So, cohorts born after 1960s in those countries might have enjoyed favorable environment for schooling.

[82] While constructing the five-year birth cohorts, we collapse all individuals aged 25-60 into different five year birth cohorts except the first and the last. In total, we have defined 13 cohorts based on 5-year birth groups. Since the number of observations before 1930 is too thin, we group them into one cohort. Similarly all individuals, who were born during 1985-87, were collapsed to form the last cohort.

compared to the male cohort. In specifications with birth-cohort fixed effects, parents' life expectancy turn out to be positive and statistically significant for the rural and female group.

Life expectancy at birth might affect the schooling decision of birth-cohorts differently across regions. Therefore, we extend the empirical exercise by seven regions based on the World Bank classification of countries based on income and region.[83] The results reported in table 4 reveal that although life expectancy effects are consistently positive and statistically significant, they vary in magnitude across regions. The coefficient of life expectancy at birth shrinks for all regions once we control for birth-cohort specific fixed effects. The estimates suggest that compared to other regions cohorts in the Middle East and North Africa spent longer time in school in response to a rise in life expectancy at birth. Consistent with what we observe in survey specific estimates, life expectancy effect is smaller in East Europe and Central Asia. However, the life expectancy effect in South Asia is not significant after including birth-year specific fixed effect. Parents' life expectancy does not exhibit consistency across specifications and regions.

5.4 Life expectancy at higher ages

There exist concerns related to appropriateness of using life expectancy at birth in explaining human capital investment decision. Hazan (2012) argued that life expectancy at birth exhibits more variation across countries and cohorts due to high infant and child mortality, based on that he suggests that life expectancy at age five will be more appropriate to capture its true effect on human capital investment decision. However, we must recognize that in consistent with the predictions of the model laid out in section two and three, life expectancy at age 5 or 10 is

[83] World Bank classifies the developing economies into six regions: "East Asia and Pacific", Europe and Central Asia", "Latin America and Caribbean", "Middle East and North Africa", "South Asia" and "Sub-Saharan Africa". We added to this the pool of industrialized countries into "High Income Countries". For World Bank classification please see http://data.worldbank.org/about/country-and-lending-groups

endogenous because by the time one reaches age five or ten, parents have made substantive investment into his/her health. To check the robustness of our findings to life expectancy at higher ages, we incorporate life expectancy at age five, ten, and fifteen in our empirical analysis. Since life expectancy at exact ages is not available before 1950 for many countries, we use a truncated sample. In addition, since the data is available at 5-year range, life expectancy measures at higher ages vary by five-year birth cohorts within a country.[84]

Table 5 reports that the effect of life expectancy at higher ages on time spent in school is consistently positive and statistically significant that contrasts Hazan's (2012) findings. A one year of gain in life expectancy at age five, ten, and fifteen increases time in school by 0.185 year, 0.17 year, and 0.129 year respectively. Parents' life expectancy turns out to be positive and statistically significant in specifications with life expectancy at birth, but not consistently at higher ages. Note that the value of parental life expectancy falls, at least in precision, as we measure life expectancy at higher ages, which probably indicate that parents' endowment is not as crucial as they are in early childhood.

5.5 Robustness checks

This section incorporates several country-cohort specific measures to check robustness of the effects of life expectancy at birth on completed years of schooling. We extend the paper by incorporating cohort-country specific weather and polity variables. To investigate the quality of institutions and political regimes at the time of one's birth, we utilize polity measure that ranks countries by their strength of democratic institutions. We use Polity IV data, which assigns a

[84] Moreover, since we are using birth-year cohort specific data and life expectancy at five and ten are available by five-year range, for a birth-year cohort in a country we assign a five year forwarded value of life expectancy at five, and ten year forwarded value of life expectancy at ten.

polity score to 167 countries which as of 2013 has a population of more than 500,000.[85] Although the data goes back to 1800 for some countries, for many countries the polity constructs start after their independence. For a few countries we impute the missing polity information by the polity score of their origin country prior to the split, for example, all of the Post-Soviet states and states formed after the dissolution of former Yugoslavia and Czechoslovakia.[86] We also exploit information from Nunn and Puga (2014) on whether a cohort in a country was born under colonial regime, and if yes, their colonial origin. The time series data on country averages on yearly temperature and precipitation is obtained from CRY-CY dataset, produced and maintained by Climate Research Unit at University of East Anglia, UK.[87] Table 6 reports the results from robustness checks. In specification I, we incorporate two weather measures while in specification II and III we include a quality of governance measure, polity score that cohorts experienced at the time of birth in a country. Since the polity variable is missing for many birth cohorts, we start the estimation with a sample consisting observations on polity score. In specification III, we divide the sample into two groups on the basis of having polity data to check if missing polity sample exhibit a differentiated effect of life expectancy at birth on schooling across groups.[88] We

---

[85] The data and documentation is available at http://www.systemicpeace.org/inscrdata.html (*accessed on October 8th, 2014*). The polity scale varies from "*strongly autocratic*" coded as -10 to "*strongly democratic*" coded as 10.
[86] Belize, though included in our sample does not have any polity data. In some cases, we could not use the available polity data since two countries have been consolidated into one, and the surveys do not identify respondents by the origin. For example, in surveys from Germany, we could not utilize cohorts born after 1945 since the surveys do not identify individuals born between 1946 and 1987 by place of birth, i.e., whether one was born in West or the Eastern part. We exclude cohorts born before 1976 in Vietnam, and all cohorts born in Yemen for similar reason.
[87] The CRU-CY dataset is constructed from the CRU TS dataset. The original data (CRU TS 3.21) took the form of a value for each month and each box on a 0.5 degree latitude/longitude grid. CRU assigned each box to a single country. For each country CRU calculated the weighted mean of the values from its constituent grid boxes for each month in turn. Each grid box was weighted by surface area, using the cosine of the latitude. The seasonal and annual values are the means of their constituent months. The CRU TS dataset prioritizes completeness, and has no missing data over land. Where observations are unavailable, the 1961-90 monthly climatic mean is used as a substitute. In data sparse regions of the world, this can lead to repeated values, and this can show up in derived products such as CRU CY."
[88] As most of the developing countries were European colonies in early twentieth century and polity data is missing for cohorts born during the colonial regime in those countries, we divide the sample into four groups based on non-missing polity data and colonial exposure of countries: (i) cohorts from countries that were never colonies but lack polity data at the time of birth (ii) cohorts from countries that were colonies and lack polity data at the time of

investigate in speciation IV if colonial exposure affects life expectancy effects on schooling in systematic way on the basis of polity data availability.

In table 6, under specification (I) and (II), we observe that inclusion of weather or polity measures does not alter the impact of life expectancy at birth on completed years of schooling. Average yearly precipitation at the time of birth has no effect on schooling while average temperature at the time of birth is positively associated with years in school. However, the estimates on polity variable itself and its interaction with life expectancy reveal that quality of government at the time of birth does not matter for schooling in our sample. However, are the cohorts with polity data different from that without polity data? In specification III, the dummy for having polity data turns out to be positive and statistically significant, which implies that average schooling is higher for the group having polity data. But the life expectancy at birth effect on years in school does not differ between the two groups in a notable manner.[89]

Next, specification IV, where we incorporate the colonial exposure of a country, reveals that effect of gain in life expectancy at birth does not vary in a statistically significant manner across the four groups formed on the basis of availability of polity data and colonial exposure. The historic colonial origin and exposure does not matter for life expectancy effect on cohort's decision on time spent in school. The interaction of polity with life expectancy at birth do not vary on the basis of colonial exposure - higher polity score does not cause higher life expectancy effect on schooling in countries with colonial exposure compared to those who were never colonies. In all three specifications with information included on quality of governance, the

birth(iii) cohorts from countries that were never colonies and have polity data at the time of birth(iv)cohorts from countries that were never colonies and have polity data at the time of birth.

[89] Although we do not report in table 6**,** estimates from a simple modification of specification I with a colony dummy and interaction of colony with life expectancy at birth reveal that life expectancy effect on years in school is positive and statistically significant but does not exhibit any statistically significant difference at 10 percent level across groups of countries who were never colonies and who had colonial exposure.

finding of positive and statistically significant coefficient on parents' life expectancy suggests the possibility of transmission of intergenerational health.

5.6 Extension

We extend the analysis utilizing individual level observations instead of cohort level means from 66% (173) of the original surveys used for cross-country analysis. Note that the theoretical model suggests that life expectancy at birth is exogenous in determining human capital investment, an individual level analysis will confirm if country-cohort specific unobservable is contaminating the cohort-mean based empirical results. We estimate the equivalent specifications of equation 1 using individual level data.

In our setting, we cannot observe individual life expectancy at birth; however, cohort life expectancy at birth, an average measure of individual life expectancies at birth across individuals within a cohort ($LE_{jc} = \frac{\sum_{i=1}^{N} LE_{ijc}}{N}$), is exogenous to an individual's completed years of schooling or lifetime earnings. Group mean is often used as an instrument to resolve *endogeneity* issue in individual level empirical analysis[90]. In the above specifications, $LE_{ijc} = LE_{jc} + \mu_{ij}$, which states that individual *i*'s life expectancy at birth in the country *c* deviates from cohort *j*'s mean life expectancy by μ_{ij}, which is by construction orthogonal to mean. Since μ_{ij} will be contained in the error term, the condition, $Cov(\varepsilon_{ict}, LE_{jc}) = 0$, must hold. In contrast to cohort mean level analysis, to save time and space, here, we estimate one specification for each of the pooled, male, female, rural, and urban subsamples.[91] As table 7 reveals, the estimates conform to those obtained from the cohort-mean level analysis, life expectancy at birth exhibits a positive and

[90] Royalty (2000) has used state tax rate as an instrument for marginal tax rate in explaining employees' health insurance eligibility. Similarly, a series of studies following Ruhm (2000) exploited variation in state or county level unemployment rate while explaining individual health behavior during a recession.
[91] We choose specification V from table 2, and specification III & VI from table 3.

statistically significant association with completed years of schooling. The pattern is similar to those obtained from cohort-mean level analysis reported in table 1 and 2. In the pooled sample, one year gain in life expectancy at birth leads one to complete 0.11 years more completed years of schooling. In general, the life expectancy coefficients are smaller compared to the cohort-mean level analysis.

In contrast to the cohort-mean level analysis, the findings from individual empirical analysis consistently suggest that parent's life expectancy exhibit a statistically significant positive influence both on human capital investment in the pooled sample as well as across male, female, urban and rural groups. Parents' life expectancy effects on years in school are 20-34% of an individuals' own life expectancy effect. This evidence implies the possibility of intergenerational transfer.

6. Discussion & Conclusion

This study covers a wide group of countries, extensive time range, and exploits the across-cohort variation within a country to identify the impact of life expectancy at birth on human capital accumulation. We find that gain in life expectancy at birth increases investment in human capital. An individual spent 11% to 15% time in school out of every one additional year of gain in life expectancy at birth. One out of every 6.5 to 9 years of additional life years is translated into time spent in school. This is comparable to the estimates of 0.11 years in Sri Lanka (Jayachandran and Lleras-Muney 2009), and 0.17 years in a cross-section of countries (Hansen, 2013).

In our sample, life expectancy at birth and completed years of schooling increase by 31 years and 5 years respectively for the younger cohort compared to the older birth cohort. Our estimates imply that gain in life expectancy at birth explains at least 70% of this rise in schooling years. To put this in the U.S. perspective, life expectancy at birth in the U.S. rose by 28 years from 1880 to 1980 birth cohorts and years of schooling rose by about 6.5 years. Our estimates suggest that rising life expectancy at birth in the U.S. explains 4 out of the 6.5 years of increased schooling.

We complement our main findings by several robustness checks. To address concern with the appropriateness of using life expectancy at birth in explaining human capital investment, we took attempt to estimate life expectancy effects at higher ages. Although our estimates exhibit robustness against life expectancy at age five or ten, we claim that life expectancy at age five or ten are endogenous, and the interpretation of coefficients of life expectancy at higher ages instead of that at birth is difficult. There is evidence that parents make compensatory and complementary investment in children to offset early life health shock. For example, in Austria, Halla and Zweimuller (2013) find that parents of children exposed to Chernobyl Accident made compensatory investment through choosing smaller family size and decreasing participation in the labor force. Further, within a birth-cohort, if only the kids with better health survive beyond infancy, the issue with mortality selection arises. Since life expectancy at higher ages are correlated with one's probability of being alive at higher ages, and this probability is correlated with ones' performance in school but part of unobservable, the coefficient of life expectancy at higher ages would be contaminated due to mortality selection bias.

The robustness of life expectancy at birth effects to the inclusion of birth-cohort and country specific weather and quality of governance information, gives us confidence in the impact of life expectancy at birth on completed years of schooling. The consistent estimates on the effects of life expectancy at birth on schooling reinforce the importance of investment in infant and child health for long run human development. The finding suggests that health could be an important avenue for developing countries to catch up the developed nation.

7. References

Acemoglu, D., and S. Johnson. 2007. "Disease and Development: The Effect of Life Expectancy on Economic Growth." *Journal of Political Economy* 115(6): 925–985.

Acemoglu, D., and S. Johnson. 2006. "Disease and development: the effect of life expectancy on economic growth." *NBER working paper* no 12269.

Ben-Porath, Y. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75: 352–365.

Becker, Gary S., and Nigel Tomes.1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy* 87(1979):1153-1189.

Becker, Gary S., and Nigel Tomes. 1986. "Human capital and the rise and fall of families." *Journal of Labor Economics* 4(1986):S1-S39.

Bleakley, Hoyt. 2007. "Disease and Development: Evidence from Hookworm Eradication in the American South." *Quarterly Journal of Economics* 122(1): 73–11.

Bleakley, Hoyt. 2010a. "Malaria eradication in the Americas: a retrospective analysis of childhood exposure." *American Economic Journal: Applied. Economics* 2:1–45.

Bleakley, Hoyt. 2010b. "Health, Human Capital, and Development." *Annual Reviews of Economics* 2:283-310.

Bleakley, Hoyt, D. Costa, and A. Lleras-Muney. 2014. "Health, Education and Income in the United States, 1820-2000." In Leah P. Boustan, Carola Frydman, and Robert A. Margo, eds. *Human Capital in History: The American Record.* University of Chicago Press.

Bloom, David, D. Canning, and G. Fink. 2014. "Disease and development revisited." *Journal of Political Economy* 122(6):1355–66.

Clark, Damon, and H. Royer. 2013. "The Effect of Education on Adult Mortality and Health: Evidence from Britain." *American Economic Review* 103(6): 2087-2120.

Cervellati, Matteo, and U. Sunde. 2013. “Life expectancy, schooling, and lifetime labor supply: Theory and evidence revisited.” *Econometrica*.81: 2055–2086.

Costa, Dora. *Forthcoming.* “Health and the Economy in the United States, from 1750 to the Present.” *Journal of Economic Literature*.

Cohen, Daniel, and Laura, Leker. 2014. “Health and Education: Another Look with the Proper Data.” Available at *http://www.parisschoolofeconomics.eu/docs/cohen-daniel/cohen-leker-health-and-education-2014.pdf* (*last accessed on July 17th, 2015*).

Cutler, David, W. Fung, M. Kremer, M. Singhal, and T. Vogl. 2010. “Early-life malaria exposure and adult outcomes: evidence from malaria eradication in India.” *American Economic Journal: Applied. Economics* 2:72–94.

Halla, Martin, and M. Zweimüller. 2014. “Parental Response to Early Human Capital Shocks: Evidence from the Chernobyl Accident.” *Discussion Paper 7968,* Institute for the Study of Labor (IZA).

Hansen, C. W. 2013. “Life expectancy and human capital: evidence from the international epidemiological transition.” *Journal of Health Economics* 32 (6):1142–1152.

Hazan. M. 2009. “Longevity and lifetime labor supply: evidence and implications.” E*conometrica* 77(6):1829–1863.

Hazan. M. 2012. "Life Expectancy and Schooling: New Insights from Cross-Country Data." *Journal of Population Economics* 25 (4):1237-1248.

Heckman, J. J. 1976. “A life-cycle model of earnings, learning, and consumption.” *Journal of Political Economy* 84: S9–S44.

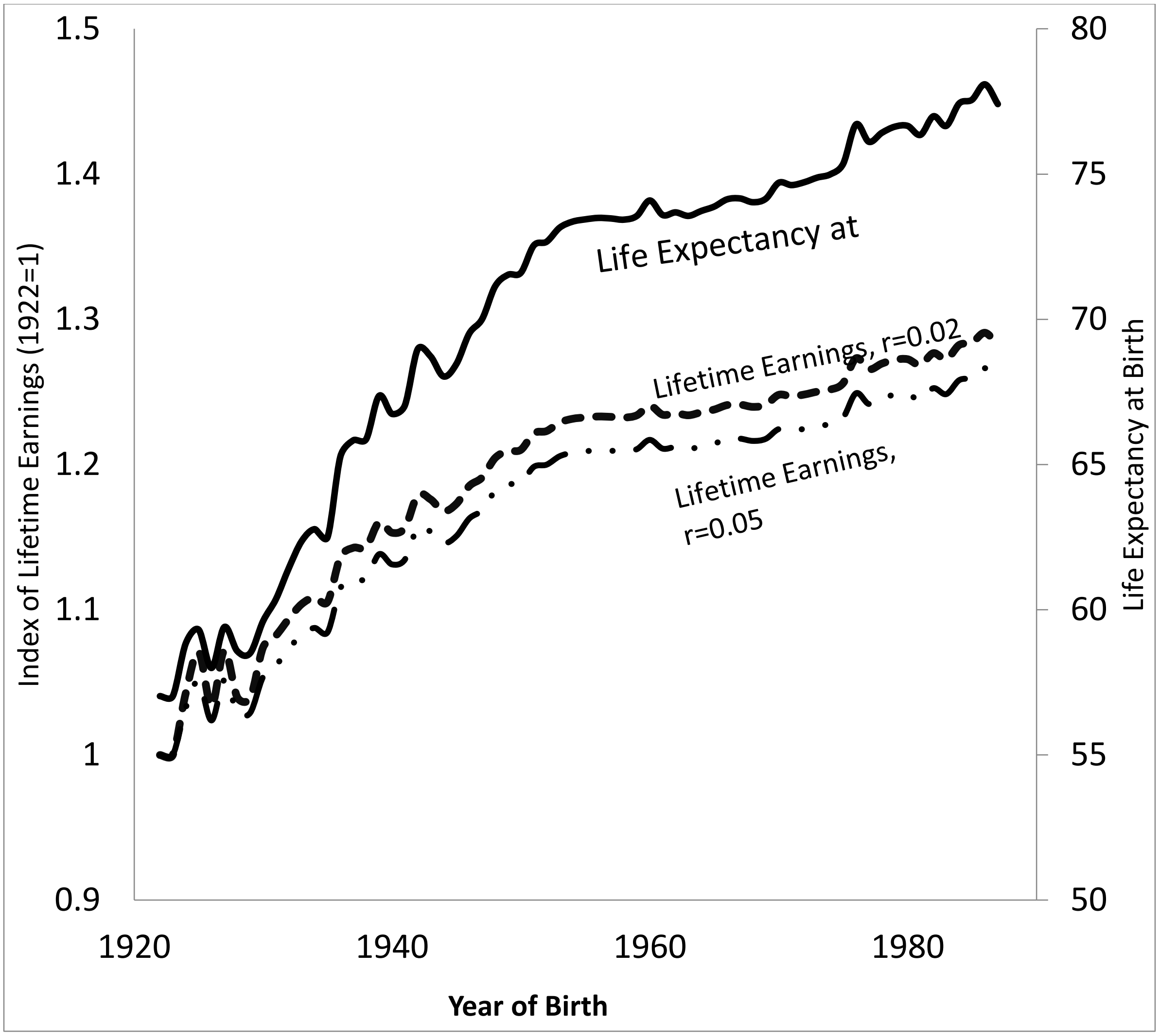
Jayachandran, S., and A. Lleras-Muney. 2009. “Life expectancy and human capital investments: evidence from maternal mortality declines.” *Quarterly Journal Economics* 124(1):349–398.

Lorentzen, Peter, J. McMillan, and R. Wacziarg. 2008. “Death and development.” *Journal of Economic Growth* 13(1): 81–124.

Pitt, M. M., M. R. Rosenzweig, and M.N. Hassan. 2012. "Human Capital investment and the gender division of labor in a brawn-based economy." *American Economic Review* 102(7): 3531-60.

Nunn, N., and D. Puga. 2012. “Ruggedness: the blessing of bad geography in Africa”, *Review of Economics and Statistics* 94(1):20–36.

University of East Anglia Climatic Research Unit (CRU); [Ian Harris]. Climatic Research Unit (CRU) Year-by-Year Variation of Selected Climate Variables by Country (CY), [Internet]. NCAS British Atmospheric Data Centre, 2013, *25th August, 2014*. Available at http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__DE_56531370-2613-11e3-9fca-00163e251233



Source: Chapter 3, Figure Four.
Note: We assume 1922 as the base year. We plot the implied present value of lifetime earnings. The lifetime earnings estimates are assumed for male residing in urban areas. The life expectancy at birth numbers are the maximum life expectancy enjoyed by a cohort across the countries, which is to capture what an average person would expect to enjoy staying on the frontier of health technology at the time of birth. While calculating the net present value of log of lifetime earnings, we try two different discount rates 2% and 5%. The period in the figure ranges from 1922 to 1987 as prior to 1922, the information on urban/rural residence is missing.

Figure 1: World Average Life Expectancy at Birth and Implied Lifetime Earnings Index

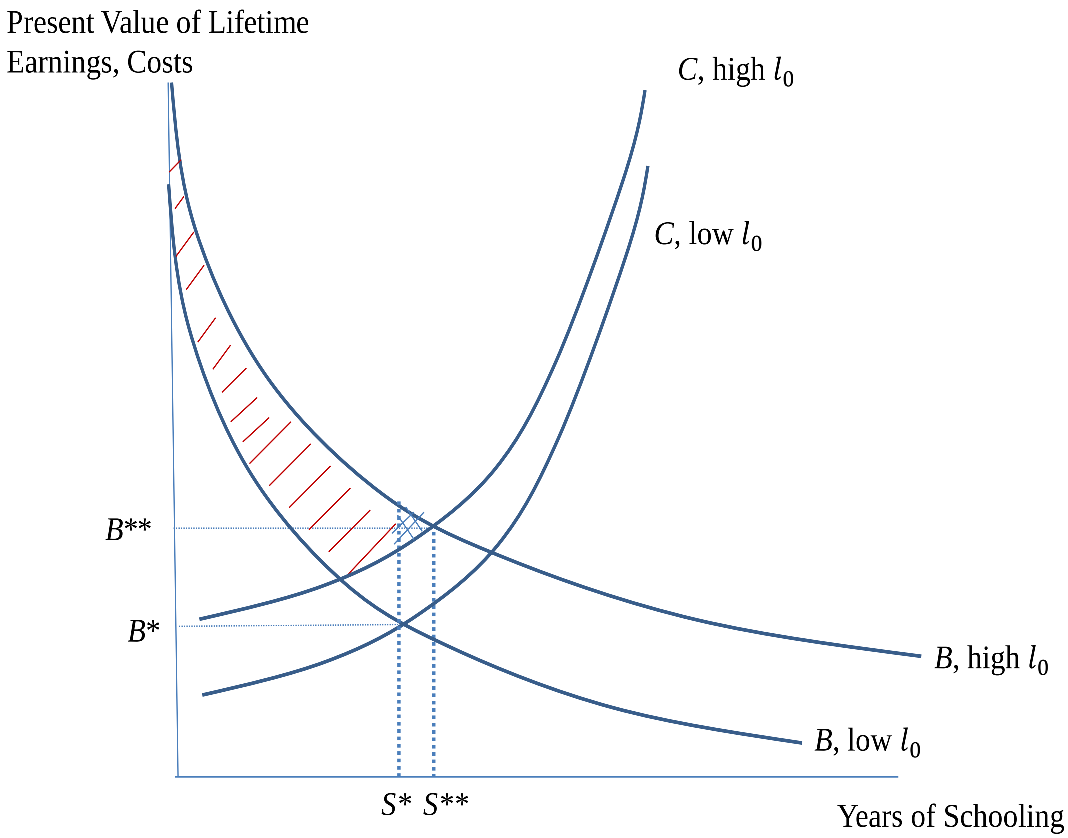


Figure 2: Benefits and Costs of Schooling in the Presence of Health Improvement

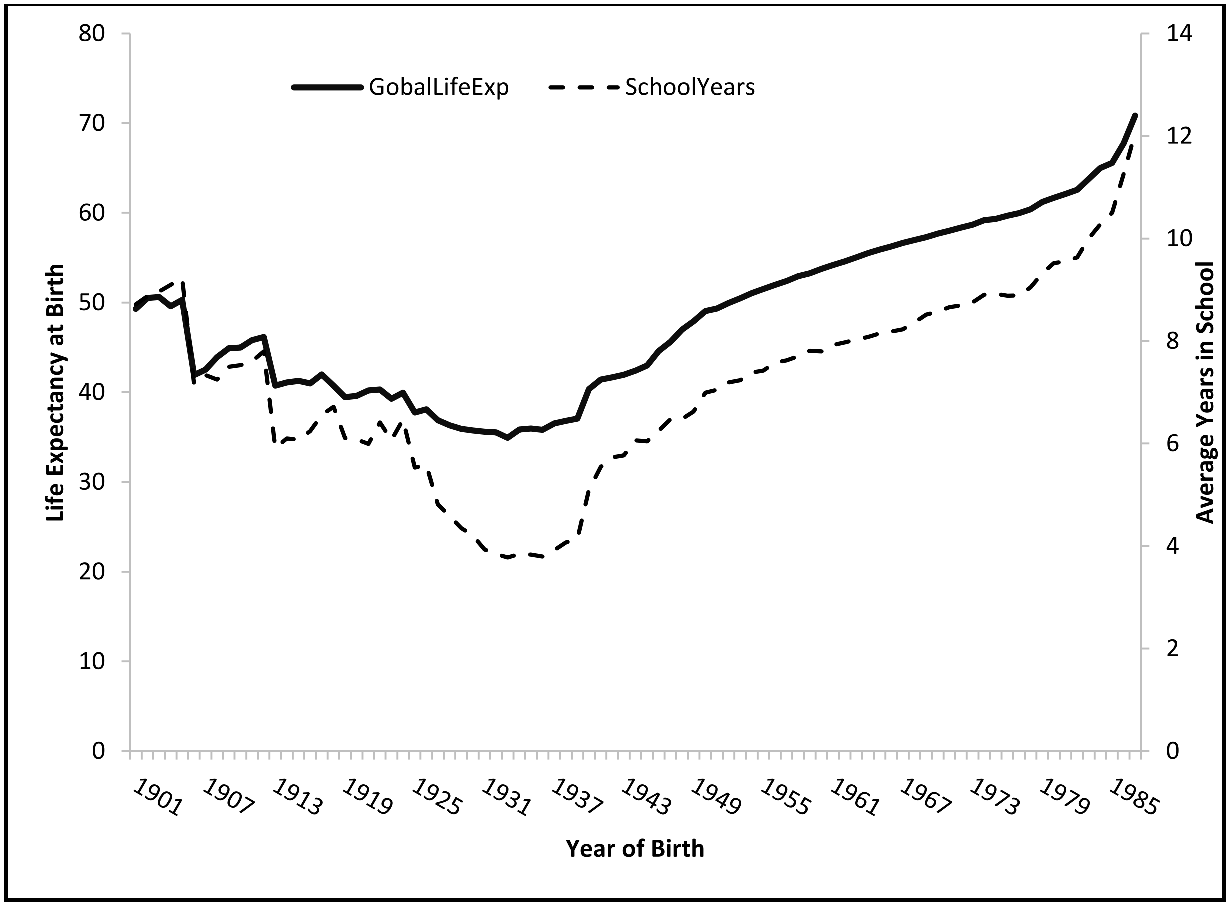


Figure 3: How Life Expectancy at Birth and Average Years in School Evolves Overtime

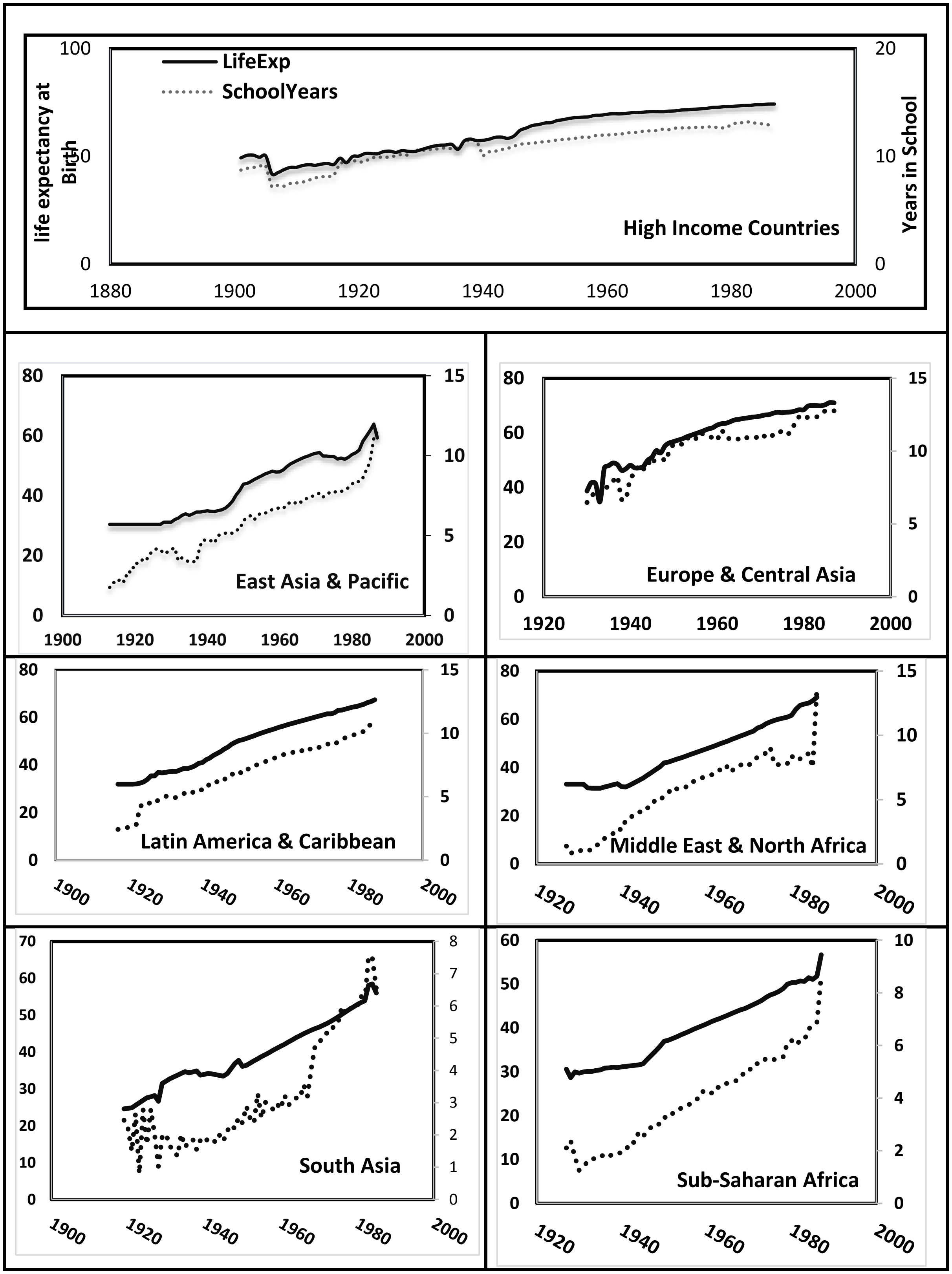
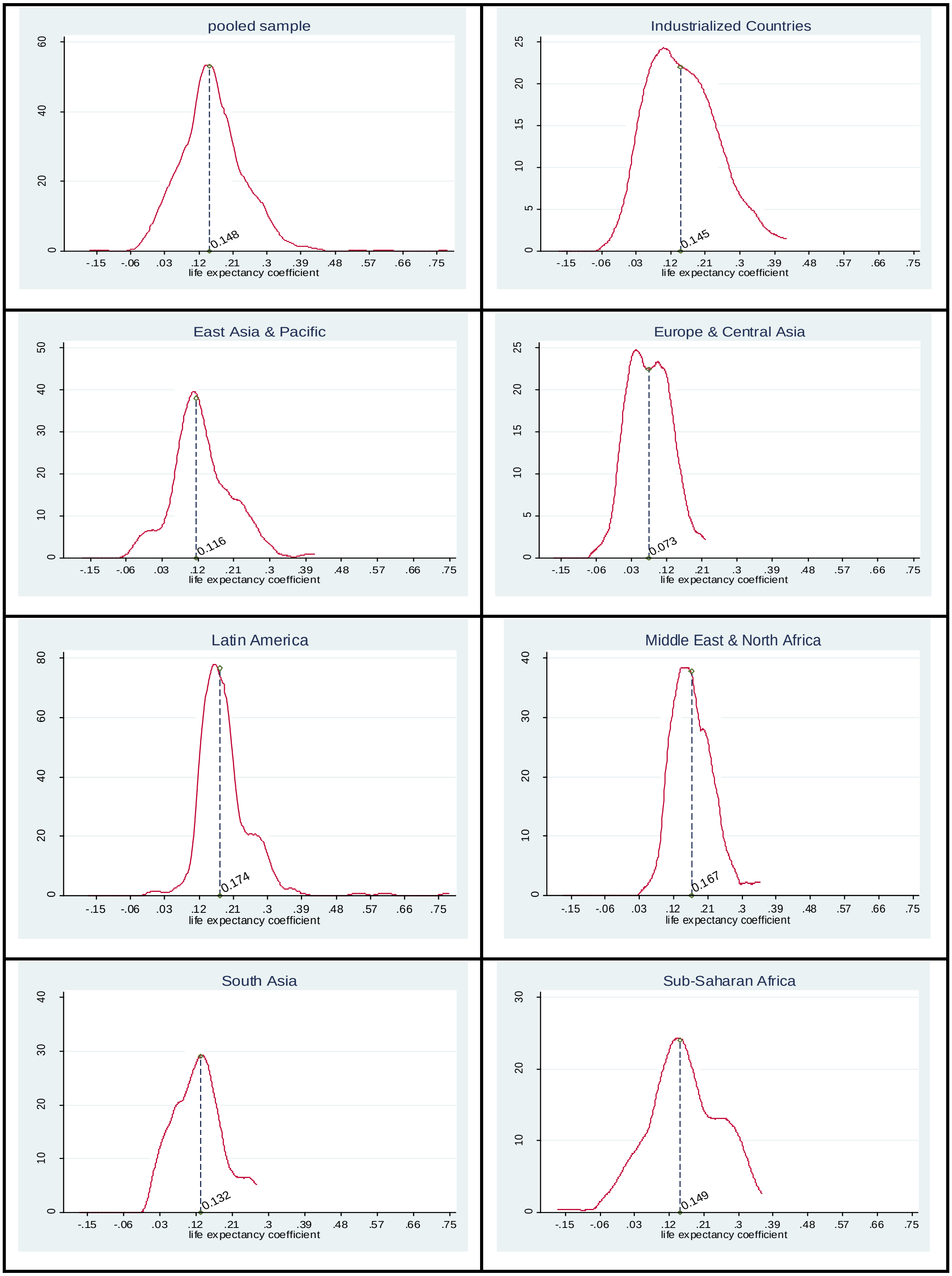


Figure 4: Life Expectancy at Birth and Average Years in School across Region and Time



Note: Median effect is indicated by the vertical line.

Figure 5: Kernel Density of Life Expectancy Effects on Years in School across Regions

Table 1: Survey Specific Estimates of Life Expectancy at Birth Effect on Schooling

Region	Number of Surveys	Positive		Negative		Life expectancy Effects on Schooling			
		significant	*insignificant*	*significant*	*insignificant*	*Mean*	*Std. Dev.*	*Min*	*Max*
High Income Countries	239	226	8	0	5	0.154	0.092	-0.017	0.420
Asia & Pacific	78	71	3	1	3	0.132	0.076	-0.022	0.413
East Asia & Central Europe	93	82	6	0	5	0.076	0.054	-0.041	0.218
Latin America	292	290	2	0	0	0.187	0.074	0.010	0.777
Middle East and North Africa	33	33	0	0	0	0.172	0.055	0.077	0.348
South Asia	49	49	0	0	0	0.128	0.062	0.034	0.27
Africa	135	123	5	2	5	0.16	0.093	-0.14	0.36
Total	919	874	24	3	18				
%		95.1%	2.6%	0.3%	2.0%				

Table 2: Life Expectancy at Birth and Education

	I	II	III	IV	V
% Urban		5.671***	5.655***	5.668***	5.671***
		[1.965]	[1.727]	[1.066]	[1.021]
% Male		1.671*	1.575**	1.269	1.323*
		[0.919]	[0.755]	[0.839]	[0.785]
Life Expectancy at Birth	0.153***	0.140***	0.138***	0.133***	0.134***
	[0.008]	[0.007]	[0.010]	[0.025]	[0.026]
Parents Life Expectancy			0.003	0.028	0.030*
			[0.025]	[0.017]	[0.018]
Cohort FE				YES	
Birth-Year FE					YES
Survey FE	YES	YES	YES	YES	YES
Constant	0.237	-3.211**	-3.162*	-3.635**	-3.821**
	[0.436]	[1.573]	[1.672]	[1.445]	[1.643]
N	6959	6143	5688	5688	5688
adj. R-square	0.985	0.984	0.984	0.985	0.985

Note: Significance level can be read as * p<0.1, ** p<0.05, *** p<0.01.

Table 3: Life Expectancy and Schooling across Male, Female, Urban, and Rural Group

	I		**II**		**III**		**IV**		**V**		**VI**	
	URBAN	RURAL	URBAN	RURAL	URBAN	RURAL	MALE	FEMALE	MALE	FEMAL E	MALE	FEMAL E
Life Expectancy at Birth	**0.149***** [0.010]	**0.142***** [0.009]	**0.119***** [0.023]	**0.127***** [0.026]	**0.118***** [0.024]	**0.128***** [0.028]	**0.119***** [0.009]	**0.159***** [0.011]	**0.11***** [0.024]	**0.154***** [0.026]	**0.11***** [0.026]	**0.154***** [0.028]
% Urban							5.771*** [1.643]	4.308*** [1.155]	5.61*** [1.137]	4.197*** [0.735]	5.59*** [1.022]	4.187*** [0.739]
% Male	2.730*** [0.639]	1.831** [0.705]	2.325*** [0.566]	1.683** [0.731]	2.331*** [0.532]	1.690** [0.704]						
Parent's Life Expectancy	-0.002 [0.022]	0.027 [0.023]	0.023 [0.018]	0.034* [0.021]	0.024 [0.019]	0.036* [0.021]	-0.013 [0.026]	0.022 [0.024]	0.018 [0.016]	0.037* [0.019]	0.021 [0.017]	0.039* [0.020]
Cohort FE			YES	YES					YES	YES		
Birth-Year FE					YES	YES					YES	YES
N	5687	5462	5687	5462	5687	5462	5681	5685	5681	5685	5681	5685

Note: Significance level can be read as * p<0.1, ** p<0.05, *** p<0.01. We estimate each specification for each group separately. An estimation on the appended male and female sample with an interaction of male-female indicator and life expectancy at birth shows that life expectancy coefficient statistically differs across male and female group. No such difference is found for the urban-rural sample.

Table 4: Region Specific Life Expectancy Effects on Schooling

	I							II						
	High Income Group	Asia & Pacific	East Asia & Central Europe	Latin America	Middle East and North Africa	South Asia	Sub-Saharan Africa	High Income Group	Asia & Pacific	East Asia & Central Europe	Latin America	Middle East and North Africa	South Asia	Sub-Saharan Africa
% Urban	-2.97 [1.872]	12.4*** [2.266]	7.5*** [1.846]	4.88*** [1.101]	11.7** [4.547]	19.0*** [2.99]	6.57*** [1.021]	-6.2*** [2.242]	10.5*** [1.509]	6.73*** [1.680]	4.71*** [1.102]	9.97*** [3.655]	17.14*** [3.392]	5.76*** [1.018]
% Male	-2.2*** [0.751]	3.63*** [1.268]	1.108 [0.919]	0.22 [0.611]	3.59** [1.778]	0.029 [0.626]	3.1*** [0.575]	-3.2*** [0.973]	3.05*** [0.769]	0.477 [1.152]	0.456 [0.545]	4.6*** [1.633]	-0.07 [1.008]	3.6*** [0.597]
Life Expectancy at Birth	**0.10***** [0.029]	**0.11***** [0.010]	**0.08***** [0.024]	**0.13***** [0.009]	**0.20***** [0.022]	**0.08***** [0.025]	**0.14***** [0.013]	**0.07***** [0.024]	**0.07***** [0.013]	**0.04**** [0.023]	**0.08***** [0.018]	**0.15***** [0.023]	0.028 [0.035]	**0.07***** [0.024]
Parent's Life Expectancy	-0.002 [0.009]	0.021 [0.024]	0.03*** [0.012]	0.04*** [0.013]	0.041** [0.018]	0.07*** [0.023]	0.034** [0.016]	-0.06** [0.024]	0.012 [0.023]	0.012 [0.014]	0.027 [0.019]	0.028 [0.021]	0.039 [0.027]	0.018 [0.020]
Birth-Year FE								YES	YES	YES	YES	YES	YES	YES
Survey FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	5688	5688	5688	5688	5688	5688	5688	5688	5688	5688	5688	5688	5688	5688

Note: Significance level can be read as * p<0.1, ** p<0.05, *** p<0.01.The region specific analysis adopted the World Bank classification based on income and region. We estimate each specification with an interaction of each of the control with region dummies to extract region specific estimates of life expectancy at birth. In specification II we control for birth-year specific fixed effect to control for differences in environment across birth cohorts.

Table 5: Life Expectancy at Higher Ages

	I	II	III	IV	V	VI
	LE at Birth	LE at 5	LE at Birth	LE at 10	LE at Birth	LE at 15
% Urban	5.865***	4.613***	5.875***	4.452***	5.805***	4.550***
	[1.214]	[1.536]	[1.131]	[1.494]	[1.095]	[1.495]
% Male	1.38	0.723	1.33	0.347	1.349	0.082
	[1.015]	[1.122]	[0.954]	[1.280]	[0.885]	[1.383]
Life Expectancy at Birth	**0.160*** **		**0.146*** **		**0.138*** **	
	[0.026]		[0.026]		[0.026]	
Life Expectancy at 5		**0.185** **				
		[0.075]				
Life Expectancy at10				**0.170*** **		
				[0.062]		
Life Expectancy at 15						**0.129** **
						[0.052]
Parents life Expectancy (25 years lag of Life Expectancy at Birth)	0.040***	0.052*	0.035**	0.044	0.029*	0.037
	[0.015]	[0.029]	[0.015]	[0.029]	[0.017]	[0.031]
Constant	-5.67***	-7.50	-4.714***	-5.56	-4.05**	-2.48
	[1.662]	[4.799]	[1.600]	[3.457]	[1.575]	[2.645]
Cohort FE	YES	YES	YES	YES	YES	YES
Survey FE	YES	YES	YES	YES	YES	YES
N	4453	4453	4982	4982	5302	5302
adj. R-square	0.987	0.981	0.985	0.979	0.985	0.978

Note: To facilitate comparison, we estimate life expectancy at birth effect in the sample for which data on life expectancy at higher ages are available. Standard errors are in brackets. Significance level can be read as * p<0.1, ** p<0.05, *** p<0.01.

Table 6: Effect of Life Expectancy at Birth in the Presence of Weather and Polity Scores

	(I) Weather	(II) Polity Data	(III) Incorporating Sample without Polity Data	(IV) Colonial Exposure & Polity
% Urban	5.679***	4.941***	5.496***	5.457***
	[1.058]	[1.257]	[1.096]	[1.080]
% Male	1.237	0.363	1.537**	1.608**
	[0.831]	[0.870]	[0.759]	[0.757]
Life Expectancy at Birth	**0.132*** **	**0.150*** **		
	[0.024]	[0.021]		
Have Polity Data			0.882**	0.756*
			[0.410]	[0.411]
(Do not have polity Data)* Life Expectancy at Birth			**0.145*** **	
			[0.026]	
(Have polity Data)* Life Expectancy at Birth			-0.013	
			[0.008]	
(Colony)*(Have Polity Data)* Life Expectancy at Birth				**0.135*** **
				[0.022]
(Never Colony)*(Do not have Polity Data)* Life Expectancy at Birth				-0.011
				[0.020]
(Colony)*(Do not have Polity Data)* Life Expectancy at Birth				0.011
				[0.008]
(Never Colony)*(Have Polity Data)* Life Expectancy at Birth				-0.022
				[0.017]
(Have polity Data)*Polity Score		-0.029	-0.019	0.008
		[0.036]	[0.031]	[0.030]
(Have polity Data)* Life Expectancy at Birth*Polity Score		0.001	0.000	
		[0.001]	[0.001]	
(Colony)*(Have Polity Data)* Life Expectancy at Birth*Polity Score				0.001
				[0.001]
(Never Colony) *(Have Polity Data) *Life Expectancy at Birth*Polity Score				0.001
				[0.001]
Parents Life Expectancy	0.028	0.036*	0.033**	0.034**
	[0.017]	[0.020]	[0.016]	[0.015]
Average Precipitation	-0.08	-0.076	-0.076	-0.079
	[0.052]	[0.047]	[0.050]	[0.049]
Average Temperature	0.049***	0.021	0.041**	0.044**
	[0.019]	[0.021]	[0.018]	[0.018]
Average Temperature*Average Precipitation	0.001	0.001	0.001	0.001
	[0.001]	[0.001]	[0.001]	[0.001]
Constant	-2.930*	-3.294*	-4.085**	-3.942**
	[1.648]	[1.810]	[1.680]	[1.659]
N	5602	3960	5602	5602
Adjusted R Square	0.985	0.987	0.985	0.986

Note: All specifications incorporate cohort and survey fixed effects. Standard errors are in brackets. Significance level can be read as * p<0.1, ** p<0.05, *** p<0.01. Excluding weather variables and utilizing only polity score increase our sample size by around 700 observations, however, this does not change the estimates that we observe in specification II.

Table 7: Life Expectancy at Birth Effects on Schooling, Individual Level analysis

	POOL	FEMALE	MALE	RURAL	URBAN
Urban	-2.141***	-2.247***	-2.007***		
	[0.099]	[0.098]	[0.100]		
Gender	0.486***			0.715***	0.337***
	[0.054]			[0.060]	[0.048]
Life Expectancy at Birth	**0.114*** **	**0.090*** **	**0.138*** **	**0.117*** **	**0.097*** **
	[0.009]	[0.009]	[0.010]	[0.009]	[0.009]
Parents Life Expectancy	0.029***	0.028***	0.028***	0.026***	0.033***
	[0.007]	[0.007]	[0.007]	[0.007]	[0.007]
Constant	2.259***	4.290***	0.698	-0.463	3.346***
	[0.606]	[0.598]	[0.663]	[0.568]	[0.640]
Birth Year Fixed Effect	YES	YES	YES	YES	YES
Survey Fixed Effect	YES	YES	YES	YES	YES
N	3953161	1901176	2051985	1512968	2440193
adjusted R-square	0.54	0.482	0.604	0.623	0.406
F	55.882	42.613	64.098	50.979	41.7

Note: Standard errors are in brackets. Significance level can be read as * p<0.1, ** p<0.05, *** p<0.01. Standard errors are clustered at the survey-cohort (survey specific birth-year) level.

APPENDIX D. APPENDIX TO CHAPTER 5

Table A1: List of Countries and Number of Surveys from Each Country

Country	Number of Surveys	Percent	Country	Number of Surveys	Percent
Afghanistan	2	0.22	Lebanon	2	0.22
Angola	2	0.22	Liberia	1	0.11
Albania	4	0.44	Sri Lanka	16	1.74
Argentina	20	2.18	Lesotho	1	0.11
Armenia	1	0.11	Lithuania	8	0.87
Australia	10	1.09	Luxembourg	9	0.98
Austria	9	0.98	Latvia	9	0.98
Azerbaijan	1	0.11	Morocco	2	0.22
Burundi	1	0.11	Moldavia	2	0.22
Belgium	8	0.87	Madagascar	5	0.54
Benin	1	0.11	Maldives	2	0.22
Burkina Faso	5	0.54	Mexico	13	1.41
Bangladesh	3	0.33	Macedonia	3	0.33
Bulgaria	9	0.98	Mali	2	0.22
The Bahamas	1	0.11	Malta	4	0.44
Bosnia-Herzegovina	2	0.22	Myanmar	2	0.22
Belarus	1	0.11	Mongolia	7	0.76
Belize	6	0.65	Mozambique	2	0.22
Bolivia	14	1.52	Mauritania	3	0.33
Brazil	28	3.05	Mauritius	12	1.31
Bhutan	2	0.22	Malawi	2	0.22
Botswana	1	0.11	Namibia	1	0.11
Canada	3	0.33	Niger	4	0.44
Switzerland	2	0.22	Nigeria	4	0.44
Chile	11	1.2	Nicaragua	5	0.54
China	1	0.11	Holland	8	0.87
Cote d'Ivoire	2	0.22	Norway	9	0.98
Cameroon	2	0.22	Nepal	5	0.54
Colombia	12	1.31	Pakistan	11	1.2
Comoros	1	0.11	Panama	19	2.07
Cape Verde	2	0.22	Peru	16	1.74
Costa Rica	21	2.29	Philippines	10	1.09
Cyprus	7	0.76	Papua New Guinea	3	0.33
Czech Republic	8	0.87	Poland	8	0.87
Germany	8	0.87	Puerto Rica	5	0.54
Djibouti	1	0.11	Portugal	9	0.98
Denmark	9	0.98	Paraguay	15	1.63
Dominican Republic	14	1.52	Romania	7	0.76
Ecuador	18	1.96	Russia	14	1.52
Spain	9	0.98	Rwanda	4	0.44
Estonia	9	0.98	Senegal	4	0.44
Ethiopia	9	0.98	Solomon Islands	2	0.22

Table A1 Continued

Country	Number Of Surveys	Percent	Country	Number Of Surveys	Percent
Finland	9	0.98	Sierra Leone	2	0.22
Fiji	1	0.11	El Salvador	15	1.63
France	9	0.98	Serbia	2	0.22
Micronesia, Fed. Sts.	1	0.11	Sao Tome and Principe	2	0.22
Gabon	1	0.11	Surinam	1	0.11
United Kingdom	8	0.87	Slovakia	9	0.98
Georgia	1	0.11	Slovenia	8	0.87
Ghana	4	0.44	Sweden	9	0.98
Guinea	2	0.22	Swaziland	2	0.22
Gambia, The	1	0.11	Syria	2	0.22
Greece	9	0.98	Chad	1	0.11
Guatemala	6	0.65	Togo	2	0.22
Guyana	1	0.11	Thailand	19	2.07
Honduras	20	2.18	Tajikistan	1	0.11
Croatia	3	0.33	Turkmenistan	1	0.11
Haiti	1	0.11	East Timor	2	0.22
Hungary	9	0.98	Tonga	1	0.11
Indonesia	13	1.41	Tunisia	3	0.33
India	8	0.87	Turkey	20	2.18
Ireland	6	0.65	Tanzania	10	1.09
Iran, Islamic Rep.	1	0.11	Uganda	4	0.44
Iraq	1	0.11	Ukraine	5	0.54
Iceland	9	0.98	Uruguay	19	2.07
Italy	9	0.98	USA	7	0.76
Jamaica	5	0.54	Venezuela	12	1.31
Jordan	8	0.87	Vietnam	7	0.76
Kazakhstan	1	0.11	West Bank and Gaza	12	1.31
Kenya	2	0.22	Yemen, Rep.	1	0.11
Kyrgyzstan	1	0.11	South Africa	26	2.83
Cambodia	5	0.54	Zaire	1	0.11
Kiribati	1	0.11	Zambia	4	0.44
Lao PDR	3	0.33			

Total Number of Countries 147; Total Number of Surveys 919

CHAPTER 6. ECONOMIC VALUATION OF ECOSYSTEM BENEFITS FROM CONSERVATION PRACTICES TARGETED IN IOWA NUTRIENT REDUCTION STRATEGY 2013: A NON MARKET VALUATION APPROACH

1. Introduction

Following the Environmental Protection Agency's recommendation, Iowa is the first state in the nation to develop a strategy paper to reduce nutrient loads through waterways to the Gulf of Mexico. The Iowa Nutrient Reduction Strategy 2013 sets a goal of reducing agricultural nonpoint source generated nitrogen (N) load by 41 percent and phosphorus (P) load by 29 percent in the waterways across 21 million acres of cropland in Iowa. The strategy paper evaluates the cost and performance of various agricultural conservation practices with different nitrate N and P load reductions. It develops several example cost scenarios incorporating various combinations of nutrient reduction practices, such as widespread adoption of conservation practices by farmers (reduced fertilizer application rate, adoption of cover crops, reduced tillage, and buffers etc.), land retirement, and wetland construction that can meet the specified target reduction. Out of these, three scenarios are predicted to achieve the targeted reduction of 41% N and 29% P. Table 1 presents these example scenarios, the agricultural conservation practices included under each scenario, and the estimated yearly implementation costs.

In addition to water quality improvement in Iowa and downstream waterbodies, the nutrient reduction practices will also offer a number of co-benefits through additional ecosystem services such as soil health improvement, greenhouse gas (GHG) emission reduction benefits, and wildlife habitat. However, the Nutrient Reduction Strategy (NRS) paper does not estimate the benefits and co-benefits to be derived with water quality improvement through implementation of the NRS. The valuation of these benefits and ecosystem services associated with proposed practices is the primary objective of this study.

A science-based systems approach to address the nutrient over-enrichment in waterways requires an understanding of both the costs and benefits of nutrient reduction technologies and implementation strategies. The estimation of the benefits and ecosystem services from the nutrient reduction practices is challenging since most, for example, soil health and water quality, are not routinely bought and sold.. This study uses economic tools of nonmarket valuation to estimate the benefits of water quality improvement that would result from the various nutrient reduction strategies. The co-benefits through additional ecosystem services are also included to the extent possible. We use primary data sets collected specifically for lake valuation purposes by Iowa State University, as well as by using benefit transfer methods, to assess these benefits.

2. Ecosystem Services from Nutrient Reduction Strategy

The benefits of reducing nutrient loads in Iowa's surface water include: improved water clarity, growth control of algae that negatively affect water-based recreation, minimization of dissolved oxygen that is problematic to aquatic biological diversity, and minimization of contamination occurrences in drinking water supplies. In addition to these water quality benefits, the nutrient reduction practices may offer a number of ecosystem services including increased opportunities for water-based outdoor recreation, aesthetic value, drinking water provision, wildlife recreation, improved soil quality, flood control, reduced global warming, biodiversity and endangered species protection, and pollination services.

We first evaluate the nutrient reduction practices to establish a link among agricultural conservation practices, changes in ecosystems, and resulting ecosystem services. We include six nutrient reduction practices in the evaluation because the example scenarios in the NRS particularly focus on these practices: maximum return to nitrogen rate (MRTN), wetland construction, cover crops, reduced tillage, buffers, and land retirement. A lit review on the agronomic and environmental effects of each practice is conducted. Other practices that are very effective in reducing nutrient transport to waterways (such as bioreactors, controlled drainage, application of nitrification inhibitor, and sidedressing) are not included in the review since their effects on the ecosystem, other than through improved water quality, is limited. An *appendix* with the review is available from the author upon request. The literature review suggests that each of the conservation practices results in multiple ecosystem services, mainly through reduced nutrients and reduced sediment transport into the lakes, streams, and rivers, reduced soil erosion, and sequestration of carbon from the atmosphere. We identify several final ecosystem services including water quality improvement, enhanced soil fertility, carbon sequestration and reduced greenhouse gas emission, pollination habitat, biodiversity, and enhanced wildlife habitat that would result from implementation of the suggested conservation practices.

Following Keeler and Polasky (2012), figure 1 presents a schematic depiction of how the ecosystem components respond to the agricultural conservation practices. Note that the link and pathways are established based on the literature review mentioned above. The relationship is complex. Some practices affect the ecosystem both through direct and indirect channels. For example, conservation tillage retains more crop residue which helps reduce soil erosion and improves soil quality by retaining topsoil. Reduced tillage further increases soil organic matter and reduces compaction which improves soil health. Similarly, wetlands contribute to biodiversity and endangered species protection by creating both wildlife habitat and pollinators' habitat. The last two columns in figure 1 (titled "ecosystem services" and "economic valuation") indicate the link between the ecosystem services and human welfare, and show how the effects of the ecosystems on human welfare can be measured.

3. Methodology for this Study

The direct local benefit of reduced N and P load under Iowa NRS is improved water quality in Iowa waterbodies, which will offer a number of ecosystem services. To derive the monetary value of benefits from water quality improvements, we will consider three use values- recreation opportunities, residential housing near the lakes, and drinking water purification cost. Additionally, we evaluate three co-benefits generated from nutrient reduction strategies: (a) offsite benefits from reduced soil erosion, (b) enhanced wildlife habitat, and (c) greenhouse gas emission reductions.

There are a number of additional ecosystem services that we are unable to quantify and monetize. For example, land retirement and wetland construction will have a positive impact on biodiversity and pollination but the magnitude of these effects is not known. Other examples of such excluded benefits are flood control, groundwater recharges, and non-use values of various ecosystem improvements. Finally, we do not attempt to quantify or monetize the on-farm benefits of conservation practices. For example, the agricultural conservation practices may affect soil productivity and result in higher crop yield. We exclude those benefits from this valuation exercise since those are private benefits, and the focus of this study is exclusively on external benefits.

3.1 Quantification of benefits from improved water quality

In the first step, we estimate the change in environmental effect, such as improved water quality, that is attributable to the adoption of new conservation practices. Next, we value these changes using nonmarket valuation methods. These valuations have been done using information from primary data sets collected specifically for valuation purposes (Iowa Lake Valuation Project, Iowa State University) as well as by using benefit transfer methods.

The Iowa Lakes Valuation Project is a large, multiyear project which collects a rich set of information on Iowans' lake visitation patterns and preferences, as well as demographics.[92] The survey has been administered five times in total, once in each of the four consecutive years 2002-2005, and again in 2009.[93] The second data set covers the usage and value of water quality improvements in Iowa's rivers and streams. Both of these data sets provide significant cross-sectional coverage of usage patterns. The values from these two data sets have been linked to estimates of water quality changes from a third dataset collected by the Limnology Lab at Iowa State University.[94] The surveys were designed to complement a lake database that includes water chemistry, biological analysis, and watershed geographic information systems data for 131 principal recreation lakes in Iowa. This combined effort resulted in detailed information on both the biological condition of Iowa lakes and the value and use of water quality improvements at those lakes. This data has been used to fit revealed preference models to estimate the benefits of reduced nutrients in this lake system.

A third source of information used to monetize the environmental benefits associated with water quality improvements is a meta-analysis on the value of clean water (Je, Kling, and Herriges, 2013). The value function from this study is utilized with the data set on biological conditions of Iowa lakes to assess the lake-specific hedonic value of water quality improvements

[92] Please see http://www.card.iastate.edu/lakes/ (last accessed on June 20th, 2015)

[93] The sixth round of the survey is conducted in 2014, and the data is currently under processing and review. We could not use those for the analysis in this report.

[94] Please see http://limnoweb.eeob.iastate.edu/minireport/. Last accessed on June 20th, 2015.

assuming water quality is a housing amenity. Finally, this willingness to pay measure for water quality improvements is combined with housing counts across all of Iowa's lakes to estimate the benefits of reduced nutrients translated into lake-adjacent housing prices.

Another welfare change from water quality improvement will come through reduced drinking water purification cost. To understand the implied benefits of nutrient reduction on municipal drinking water treatment cost, we exploit production and cost information from the state's largest water treatment facility, Des Moines Waterworks (DMWW), as a case study. We relate DMWW's source water quality with production cost data to calculate the potential savings from the reduced nutrient abatement requirement for the safe drinking water supply.

3.2 Quantification of benefits from reduced soil erosion

Many of the conservation practices will reduce soil erosion. We estimate the benefits from reduced soil erosion following a benefits transfer exercise. In most of the cases, the transfer of benefits is drawn from studies conducted in Iowa. However, in a few cases, we had to rely on studies in similar sites, e.g., Corn Belt or Midwestern states. First, we estimate how many acres of land will be treated by a specific conservation practice under each scenario. Next, based on the agronomic and environmental literature review, we draw a relevant low and high number on yearly soil erosion reduction rates from per acre adoption of that conservation practice. The formula to derive yearly total reduced soil erosion from a specific practice *j* is the following:

$$\text{Reduced Soil Erosion}_j = \text{Treated Acres}_j * \text{Rate of Soil Retention}_j, \quad (1)$$

where rate of soil retention is measured as tons per acre per year. The lower bound on soil retention rate for each of the agricultural conservation practices- the fall in soil erosion due to adoption of an acre of cover crops, reduced tillage, land retirement, and buffers- are adopted from RUSLE estimates based on Iowa.[95] The data is of high quality since it (i) uses actual land use data from 2006-2010 across all of the major land resource areas (MLRA) in Iowa, (ii) considers distance of cropland from waterbodies to adjust for the sediment delivery ratio. The data on soil erosion is on the low side since the model estimates soil erosion for the most conservative practice scenario compared to the baseline.

To assign a monetary value we adopt Hansen and Ribaudo's (2008) benefits measures of dollar-per ton soil for the Corn Belt region. The authors split the measure across 14 specific categories that benefit from reduction in soil erosion. Their dollar-per ton soil value considers welfare improvements due to (a) reservoir services(less sediment in reservoirs), (b) Navigation (shipping industry avoidance of damages from groundings), (c) water-based recreation (cleaner fresh water for recreation), (d) irrigation ditches (reduced cost of removing sediment and aquatic plants from irrigation channels), (e) road drainage(ditches less damage to and flooding of roads), (f) municipal water(lower sediment removal costs for water-treatment plants), (g) flood damages(reduced flooding and damage from flooding), (h) marine fisheries(improved catch rates for marine commercial fisheries), (i) freshwater fisheries (improved catch rates for freshwater commercial fisheries), (j)marine recreational(increased catch rates for marine recreational fishing), (k)municipal & industrial water use (reduced damages from salts and minerals dissolved from sediment), (l) steam power plants (reduced plant growth on heat exchangers), (m) soil productivity(reduced losses in soil productivity), and (n) dust cleaning (decrease in cleaning due to reduced wind-borne particulates).

The per ton soil erosion reduction was valued at $2.77 (in 2000 $) for the Corn Belt region. We adjust this value by excluding $1.01 due to soil productivity, $ 0.01 due to freshwater

[95] We thank Calvin Wolter for providing us with this data.

fisheries, \$0.18 due to municipal water treatment. The soil productivity is pure private benefit, while the last two categories will be partially captured while valuing the water quality improvement by their use value. This will save us from double counting problem in the valuation of ecosystem services (Keeler and Polasky, 2012). The adjusted per ton soil value will give us offsite benefits from erosion reduction. The per ton soil value is multiplied with the measure on total reduced soil erosion obtained in equation (1) to obtain the total value of offsite soil benefits. It can be expressed as

$$\text{Value of Reduced Soil Erosion}_j = \text{Reduced Soil Erosion in tons}_j * \$/\text{ton}. \quad (2)$$

3.3 Quantification of benefits from wildlife and carbon sequestration

Wildlife related benefits will mainly be derived from two practices, wetland creation and land retirement. We monetize the benefits by multiplying the treated by a measure of per-acre, or per household wildlife benefits obtained from a suitable study.

Our monetization of carbon benefits from nutrient reduction practices exploit the Social Cost of Carbon (SCC) from EPA (2013)[96]. SCC is a \$/ton measure that combines various damages including net agricultural productivity, human health, and property from a small increase in carbon dioxide. Note that we choose the SCC estimated at 3% discount rate.[97] We assume a homogenous rate of carbon sequestration across periods. To be consistent, when drawing the carbon sequestration rate by a practice from other studies, we focus on long term studies. To estimate the total carbon benefit from a practice *j*, we apply the following steps.

Step 1: Estimate how many acres will be treated under practice *j*.

Step 2: Use a low and high estimate on carbon sequestered per acre from suitable studies

[96] SCC estimates are available at http://www.epa.gov/climatechange/EPAactivities/economics/scc.html (last accessed on May 20th, 2015).

[97] The SCC estimates are provided up to 2050, based on this we extrapolate SCC values up to 2063 to match with our project life. Note that the cost estimates in NRS are obtained assuming a 50 year project life.

Step 3: Compute yearly total carbon sequestered (*CS*) by practice *j* as

$$CS_j = CS \text{ per unit of land under practice } j \ * \ \text{Total Land Areas Under Practice } j.$$

Step 4: Construct the total monetary benefits (*CB*) from carbon sequestered by practice *j*

$$CB_j = \sum_{t=2015}^{2063} SCC_t * CS_{jt}.$$

Since social damage functions are of convex shape, SCC are increasing overtime.

Step 5: Derive Equivalent Annual Carbon Benefits from Practice *j* assuming a 50-year project life.[98]

Note that reduction of carbon and other GHG emission from the atmosphere incorporates global spatial benefits-the reduced damages from global warming will improve welfare worldwide. To facilitate comparison and further analysis, we adjust the global monetized benefits from GHG reduction from each of the nutrient reduction practices in Iowa by Iowa's share of the global population (3 million/7 billion).

4. Benefits from Agricultural Conservation Practices

We estimate the reduced amount of erosion and sequestered carbon due to the adoption of a practice. Next, we assign monetary values to those following the method laid out in section three.

4.1 Benefits from construction of wetland

According to NRS, each 1000 acres of land will be treated with a wetland comprising 10 acres of pond and 35 acres of buffer surrounding the pond. As Table 2 shows, based on scenarios, this will treat 14%-61% crop acres out of the total 28.4 million cropland acres and, 0.17-0.4 million acres of cropland will go out of production. Note that eroded soils from all

[98] We assume a 4% discount rate while spreading the total carbon benefit across 50 years. NRS assumes a 4% discount rate while estimating EAC.

treated acres that previously ended up in the lakes, streams, and rives are now trapped by the wetland. We consider the baseline soil erosion rate in all these treated acres to calculate the potential reduction. This implies that soil erosion will decrease by up to 40 million tons per year.

The low and high values assume two different soil erosion rates, one from RUSLE estimates across all MLRAs in Iowa, which is 2.49 tons/acre, and another from Gleason *et al*. (2008), which is 4.45 tons/acre. For carbon sequestration, we choose the lower bound from the Gleason *et al*. (2008), which is 0.66 tons/acre/year and the upper bound of 0.95 tons/acre/year from Hansen *et al.* (2015). Yearly total carbon sequestered by the land converted into wetland lies in the range of 0.12-0.38 million tons.

Table 3 presents the monetary value of the benefits from the constructed wetland acres. The value of the offsite benefits of reduced soil erosion ranges from $20 million to $84 million per year based on the scenarios considered. In contrast to the soil erosion benefits, the recreation, aesthetic, and wildlife viewing benefits are based on value estimates obtained from Azevedo, Herriges, and Kling (2000).[99] They estimate the median Iowan household's yearly WTP to be $2.76-$6.76(in 2013 USD) for the wetland services. We adjust this value for land size, normalizing with respect to the wetland acres under NCS8 that converts the maximum acres into wetland. Total benefits from recreation, aesthetic, and wildlife view can reach up to $8.36 million under scenario NCS8. Finally, the GHG benefits from wetlands, considering only Iowans' share, is small. However, the GHG benefits can increase by $27 million if we consider worldwide damages from sequestered carbon, and reduction in other GHG gases from the constructed wetlands in Iowa. Total benefits lie in the range of $22-$120 million, and the largest benefits are generated under scenario NCS3 when global GHG benefits are included.

[99]The recreation, aesthetic, and wildlife viewing benefits are calculated based on constructed wetland acres rather than the treated acres.

4.2 Benefits from cover crops

Under two scenarios, NCS1 and NCS3, the strategy paper suggests that a total of 17-27 million acres of land will be treated by cover crops during the fallow period. For the reduced soil erosion rate, the lower bound is based on RUSLE estimates based on Iowa, which is extremely conservative since the model only considers treating Corn-Soybean (CS) acres under no-tillage with cover crops. The rate will be much higher if conventional tillage and Corn-Corn (CC) acres are treated with cover crops. Although Kasper *et al*. (2001) reports from a study in central Iowa that in CC acres, cover crops can reduce the erosion by 2.9 tons/acre, we choose the more conservative upper bound of 0.89 tons/acre from Schipanski *et al*. (2014). The low and high bounds of yearly carbon sequestration rates from cover crops, 0.99 tons/acre and 1.24 tons/acre, are chosen from Gonzalez-Ramirez *et al*. (2014) and Schipanski *et al*. (2014).

Table 4 reveals that cover crops can reduce soil erosion by up to 24 million tons per year and sequester carbon by up to 33 million tons per year under scenario NCS3. Table 5 presents the monetized value of reduced soil erosion and sequestered carbon. The annual offsite benefits from sequestered carbon and reduced erosion lie in the range of $23-$2440 million. If the lower bound on total benefits is considered, the majority of the benefit is coming through offsite benefits from reduced soil erosion. In contrast, 98% of the upper bound total benefit is derived from global GHG damage reduction.

4.3 Benefits from land retirement

The NRS considers land retirement under scenario NCS3. The plan is to retire 5% of the land currently under crop production, 1.14 million acres. The reduced soil erosion will come through two paths: (i) reduced erosion from the retired land, and (ii) soil trapped by this retired land from the surrounding land. We exclude the second path since we do not know the locations of the retired acres as well as size of surrounding crop acres they will treat. The RUSLE estimate, based on all MLRAs in Iowa, suggests an erosion reduction of 3 tons/acre assuming that lands under extended rotation will be converted into energy grasses (Miscanthus or Switchgrass). Note that this is conservative because the erosion rate in the row crop acres without extended rotation can be much higher, and we exclude sediment trapped from surrounding acres. The upper bound estimate is 4.45 tons/acre, drawn from Gleason *et al*. (2008). The carbon sequestration rates of 0.58tons/acre and 0.66 tons/acre are chosen from Gonzalez-Ramirez *et al*. (2014) and Gleason *et al*. (2008) respectively. Note that the upper bound we choose at 0.66 tons/acre is actually the lower bound reported in Gleason *et al*. (2008) for PPR region. As Table 6 shows, land retirement can reduce soil erosion by 3-5 million tons per year. The retired land sequesters carbon in the range of 0.66-0.75 million tons in every year.

Table 7 reports the monetized benefits from land retirement. The offsite benefit from reduced soil erosion amounts to $7-$11 million. The per acre recreation value of $72.71, chosen from Hansen *et al*. (2007), suggests that 1.14 million acres of retired land will offer a recreation and hunting benefit of $83 million. This estimate from Hansen *et al*. (2007) is conservative since it considers a limited set of recreation and hunting options in the CRP land. Finally, total local carbon benefits is 0.02 million, but the global benefits can be as high as $54 million. The total benefit is in the range of $90-$147 million.

4.4 Benefits from buffers

In one scenario, NCS8, the NRS includes a practice that 70% of all agricultural streams will have a vegetative buffer on each side of the streams that are not currently buffered. The buffers will be 35 feet wide and will cover 44,768 miles of agricultural streams (Table 14, NRS). Similar to land retirement, we only consider soil erosion prevented from the 0.4 million acres converted into buffers, and not the amount of soil erosion trapped from other acres. The rates on soil erosion and sequestered carbon are chosen from the same sources as we followed for the land retirement acres except the upper bound on carbon sequestration rate, which is chosen from the upper bound for wetland acres. Note that the buffers adjacent to the streams might sequester carbon at the same rate as the wetlands.

Table 8 shows that this practice will convert 380,000 acres of cropland into buffers. The soil erosion will be reduced by up to 1.7 million tons per year. The yearly sequestered carbon will range from 0.22 to 0.36 million tons. The monetized offsite benefits from reduced soil erosion, as reported in Table 9, will lie in the range of $2.26-$3.6 million. The local carbon benefit is marginal but the global benefit can increase up to $26 million. The total benefits lie in the range from $2.27-$29 million.

4.5 Benefits from reduced tillage

Conservation tillage is a widely adopted practice as it is very effective in reducing sediment transported phosphorus loss from the field. In example scenario NCS8, the NRS includes reduced tillage practice as follows:

(i) *Convert 90% of Conventional Tillage CC and CS Acres into Conservation tillage*: This will convert 7.66 million CC and CS acres of land (out of 8.5 million acres) currently under conventional tillage into conservation tillage.

(ii) *Convert 10% of Non-no-till CC and CS Acres into No-till*: Out of the total 16.2 million crop acres, which are currently under conventional or conservation tillage practice, 1.62 million acres will switch into no-till practice.

The RUSLE estimates, based on all MLRAs in Iowa, provide potential reduction in soil erosion for each type of change in tillage practice. Switching from conventional to no-till will reduce erosion by 2.48 tons from each treatment acre, while the erosion rate will fall by 1.45 tons per acre for switching from conventional to conservation tillage. The upper bound estimate is chosen from Zhaou, Al Kaisi, and Helmers (2009): for switching from conventional to conservation tillage (no-tillage), 2.7 tons/acre/year (3.05tons/acre/year) less soil will be eroded**.** For carbon sequestration we choose the estimates from USDA (2010) and Gonzalez-Ramirez (2014). The former reports a rate of 0.33 tons/acre/year for switching from conventional to conservation tillage and 0.64tons/acre/year for switching to no-till, while the latter reports an estimate for Iowa. The carbon is sequestered at the rate of 0.81 tons/acre for switching from baseline scenario to no-tillage system.

Table 10 shows the effects from reduced tillage practice. This can reduce soil erosion by 15-26 million tons in each year. Approximately 3.5-3.9 million tons of carbon will be sequestered in every year. Table 11 shows the total monetary benefit from reduced tillage lies in the range of $32.11-$329 million. The carbon benefit is relatively large because a large number of crop acres will be treated with either conservation tillage or no-till practices.

4.6 Benefits from nutrient application at MRTN rate

The NRS has considered several nutrient management options to ensure more efficient use of nitrogen and reduce nitrate loss through leaching and runoff. The practices are (a) to limit application of fertilizers to the Maximum Return to Nitrogen Rate (MRTN) - the rate of nitrogen application that maximizes the profit from crop production, (b) nitrification inhibitor that slows down the release of nitrogen in the field, and (c) sidedress-changes the timing of fertilizer application by application of Nitrogen N to the plant at the time when needed most. The MRTN has the potential to reduce nitrous oxide emission. In the following section, we will quantify the benefits from MRTN.

The strategy paper reports that the estimated average nitrogen application (commercial fertilizer and manure) to corn in a corn-soybean rotation is 151 lb/acre while the application rate is 201 lb/acre to corn in continuous corn in Iowa. Assuming a corn price of $5/bushel and a nitrogen price of $0.5/bushel, the MRTN for corn following soybean is 133 lb-N/acre and 190 lb/acre for corn following corn. The implication is that all MLRAs where the current rate is higher than this will have to adjust N application to follow MRTN. Application of the MRTN rate to all continuous corn and corn-soybean acres in Iowa would reduce nitrate-N loading by 28,000 tons/year, all of which can be attributed to reduced N fertilizer application. Adopting the biophysical relationship between nitrogen application and nitrous oxide emission from IPCC (2010) and Millar (2006) as the low and high bound, we calculate that MRTN adoption will reduce GHG emission by 1.14-5 ml Mt per year. As Table 12 reports, the implied benefit for Iowa from such a reduction in GHG emission can reach up to $0.15 million per year. However, the global GHG benefits can be as high as $353 million per year. Note that since a fixed MRTN rate is assumed, the benefits amounts do not vary across the scenarios.

4.7 Ecosystem services not quantified

The nutrient reduction practices included in the NRS will offer several other ecosystem services that we are not able to quantify due to lack of required information. Ecosystem services that are expected to stem from agricultural conservation practices but are missing in our analysis includes: pollination services, pest control, endangered species protection, wildlife habitat, and flood control among others.

Natural pollinators are critical to agriculture. Natural pollinators, such as insects and birds, add a value of US$ 190bl/year to agriculture and almost 15-30% of the average US diet depends on pollination services. The value of pollination services from honeybees is estimated in the range of $8-$16.4 billion while that from native bees is around $3.1 billion (Losey and Vaughan 2006). Availability of wild pollinators are in sharp decline and managed honeybees are exhibiting a high death rate since 2007. Pollinator habitat and health is jeopardized due to factors including monoculture cropping practices, natural pollinator habitat decline and fragmentation, pesticide and herbicide use, pests and migratory beekeeping practices (Ehmke et al. 2015). The conservation practices under Iowa NRS will improve the pollinators' habitat. In-field practices such as reduced tillage and cover crops will provide more floral resources, buffers and wetlands provided they include flowering strips, and retiring land from crop to CRP acres will improve habitat for pollinator services.

Valuation of pollinators' services is difficult to quantify due to three key missing links (i) the production function between pollinators and agricultural yield, (ii) evaluation of agricultural practices on pollinators' health, and (iii) societal values (people value pollinators for non-market reasons such as existence, floral and arboreal services, and biodiversity).

Following the same reasons, it is difficult to assign a monetary value to endangered species protection benefits. Based on literature search, Hansen et al. (2015) report that there are 76 wetland related endangered vertebrate species, but note the lack of a biophysical model to assess the value of imperiled species protection from wetland. Land retirement provides habitat for many threatened and endangered species which are difficult to quantify due to inadequate data (Sullivan 2004). Wildlife habitat and flood control are two additional services whose benefits are difficult to quantify, and those are left to future studies.

5. **Water Quality Benefits**

We consider three direct use values of improved water quality in this section. For the valuation purpose, we consider 131 major lakes in Iowa. Figure 2 shows the exact locations of these lakes in a map. When water quality improves due to the nutrient reduction strategy, recreationists from all over the state and residents living close to those lakes are expected to benefit directly. Similarly safe drinking water supply will be less costly due to reduced nutrients and sediments in the raw water. The three benefits we consider here were not included in the calculation of offsite benefits from per ton soil erosion reduction.[100]

5.1 Water quality benefits to local homeowners

We estimate the aesthetic value of water quality improvement to residents near major lakes in Iowa. To do so, we link the Iowa water quality database (Limnology Lab at Iowa State University)[101] with a meta-analysis (Ge, Kling, and Herriges 2013) on lake-adjacent households' willingness to pay for improved water quality, and adjacent property counts across the major lakes.

Step 1: Convert the raw water quality measurements of turbidity, dissolved oxygen, PH value, total nitrate and total phosphorus into quantile values (q) for all major lakes in Iowa.

Step 2: Calculate a water quality index from q values obtained from various water quality attributes using the National Sanitation Foundation water quality index formula, as follows: $WQI = \prod_{i=1}^{5} q_i^{w_i}$, where w_i is the weight. The adjusted weight assumed in the above formula is reported in Table 13.

[100] Note that in the previous section, while quantifying the offsite benefits from soil erosion reduction, the *per-ton-soil* benefits incorporate several benefits from water quality improvement, such as reduced dredging, reservoir services, and industrial waste treatment.

[101] Retrieved at http://limnology.eeob.iastate.edu/lakereport/default.aspx

Step 3: Using the above formula, derive two water quality indices for pre-NRS and post-NRS water quality attributes. We utilized 2004 water quality attributes to derive pre-NRS water quality index. Following targets set in the NRS, we assume there will be a uniform 40% reduction in nitrogen and 30% reduction in total phosphorus in all lakes in Iowa. We assume that other water quality attributes included in the formula (dissolved oxygen, pH value, and turbidity) will remain unchanged to obtain a lower bound estimate of the water quality index based on the NRS. Due to this nutrient reduction plan, the average water quality index in 131 lakes in Iowa will improve from 72.07 to 73.74. In short, average water quality index will at least increase by 1.68 once the nutrient reduction plan is adopted. Table 15 incorporates summary statistics on water quality changes due to NRS.

Step 4: Based on initial and improved water quality indices, we calculate the annual WTP per household per year for the proposed water quality improvement using the value function estimated for hedonic studies in the meta-analysis by Ge, Kling, and Herriges (2013). Only hedonic studies are included because we are considering that only households owning homes and living near the lakes will benefit from the locally improved water quality.

$$\begin{aligned}\mathrm{WTP} = {} & -2.67 * \mathrm{Initial\ Water\ Quality} + 4.48 * \Delta\mathrm{WaterQuality} + 27.94 * \mathrm{NortheastDummy} \\ & + 287.23 * (\mathrm{Lake\ Dummy}) + 4.69 * (\mathrm{Publicationdate}) + 284 \\ & * (\mathrm{InPersonDummy}) - 0.01 * (\mathrm{Income}) + 78.96 * (\mathrm{TotalValueDummy}) \\ & + 212.50 * \mathrm{Improvement\ Dummy} - 208.04 * \mathrm{ladderDummy} + 277.26 \\ & * \mathrm{CVDummy} + 217.88 * \mathrm{HedonicDummy} + 0.06 * \mathrm{SiteSize} - 0.004 \\ & * \mathrm{RegionSize}\end{aligned}$$

In the above equation, while deriving the hedonic values we assume $\mathrm{NortheastDummy} = 1$, $\mathrm{Lake\ Dummy} = 1$, $\mathrm{Publication\ Date} = 40(\mathrm{year\ 2010})$, $\mathrm{InPerson\ Dummy} =$

0, $\text{TotalValue Dummy} = 0$, $\text{Income} = 50{,}000$, $\text{Improvement Dummy} = 1$, $\text{Region Size} = 100 \text{ Square Miles}$, $\text{Ladder Dummy} = 0$, $\text{Site Size} = 10 \text{ Square Miles}$, $\text{CV Dummy} = 0$, $\text{Hedonic Dummy} = 1$, $\text{RegionSize} = 100 \text{ square miles around the lake}$.

Estimates from the above equation will translate the water quality changes into monetary values. The equation above implies that for a 1.67 point improvement in water quality, the average WTP per lakefront property per year is $655.27. Table 14 adds implied WTP for the improved water quality.

Step Five: We know the point location of each of the lake in Iowa. Extracting information from different sources that combine GIS and population censuses, we count the total number of housing units within half mile and one mile radius of each lake. Table 15 provides a breakdown of housing units by sources.

Step Six: We consider two radius distances, half mile and one mile, to obtain a lower bound and upper bound in count of housing units. Note that these counts are conservative since we are considering only residences near 131 lakes for which we have data from the Iowa Lakes Project, and we exclude a large number of housing units located near other lakes, and all rivers including the two major rivers, Mississippi and Missouri river. In our sample, 35 % of the lakes do not have any housing units within a half mile while 24% do not have any housing within a one mile radius from the lake. The value of the improved water quality for each lake is derived by multiplying the total number of housing units by mean willingness to pay obtained for that lake, as described in step five. The value of water quality improvement to local residents from the Iowa NRS is at least in the range of $14.6-$35.4 million in 2013 USD. The six most benefitted lakes from appreciation in aesthetic value are West Okoboji Lake, Saylorville Lake, Coralville

Lake, Storm Lake, Easter Lake and Clear Lake. The total aesthetic benefits accrued by residents surrounding these lakes are at least $1 million.

5.2 Recreation benefits

The cleaner water resulting from NRS will benefit outdoor recreationists who use the 131 lakes included in Iowa Lakes Project. To estimate the size of this benefit, we perform welfare analysis adopting the reveled preference method and utilizing the detail demographic and lake usage data from random household surveys conducted under the Iowa Lakes Project.[102]

Step One: We first convert the nitrate and phosphorus reduction from Iowa nutrient reduction strategy into a representative water quality measure. Following Egan *et al*. (2009), we consider *secchi* depth as a key measure of water quality. We utilize the baseline data from 130 lakes in Iowa to estimate the relationship among *secchi* depth, total nitrate, and total phosphorus with the following regression specification

$$Secchi = 1.69 + 0.0166 * TN - 0.004 * TP.$$

This translates the reduction of 42% nitrate and 30% phosphorus into corresponding change in *secchi* depth.

Step Two: Based on the baseline and predicted change in *secchi* depth, we employ a Random Utility Model to calculate the compensating variation under both linear and log specifications. The welfare estimates, as reported in Table 16, suggest that water quality improvement would generate recreation benefits in the range of $5-$22 million dollars.

[102] The modelling part for this section is done by Yongjie Ji, Postdoc Research Associate at CARD

5.3 Water quality and drinking water treatment costs

Drinking water in Iowa is obtained mainly from three sources: (i) groundwater from deep shallow wells, (ii) surface water from rivers, lakes, and reservoirs, and (iii) shallow groundwater that are under direct influence of surface water. In Iowa, approximately 90.3% of the total 3.05 million Iowans are served by public water supplies while the remaining 9.7% are served by private water systems. Although 92% of Iowa's water supply system uses groundwater as the primary source, approximately 45% Iowans are served by public water systems which collect source water from (ii) and (iii). Agricultural runoff carrying pollutants, such as nitrogen and phosphorus, often end up in lakes, streams, and rivers and contaminate the surface water. In addition, since nitrogen is highly mobile and soluble, it leaches easily through soil structure to reach drainage water systems, groundwater, and aquifers. The extent to which agricultural nitrogen contaminates groundwater depends on soil structure, surface and bedrock geology, especially soil crust and permeability.

5.3.1 *Implication of nutrient reduction strategy for safe drinking water in Iowa*

One direct benefit of improved water quality from Iowa Nutrient Reduction Strategy will be passed through to the drinking water treatment plants. Reduced nitrogen and phosphorus levels in the streams and rivers will reduce nitrate removal cost for the treatment plants drawing source water from streams, lakes and rivers.

Excessive nitrogen in water can cause blue baby syndrome. EPA regulates the maximum level of allowable nitrogen in drinking water, which is 10 milligrams per liter (mg/L). Even a nitrogen level of 2.5 mg/L in drinking water, much lower than the EPA recommended level, is associated with high risk of thyroid cancer (Ward *et al*. 2010). Phosphorus is responsible for algal blooms that make water smell bad and have dire consequences for aquatic life, human and animal health. Cyanobacteria, blue green algae, produces *cyanotoxin*s which have negative health consequences. Further, while treating water for algae with chemical disinfectants various disinfectants byproducts are generated that may cause cancer, birth defects, and damage the DNA (Naidenko, Craig, and Nils 2012; Villanueva *et al*. 2007).

An EPA assessment showed that between 1998-2005, 17% of Iowans were served with drinking water with nitrate levels above the recommended level. Total 250 public water supplies (PWS) were at risk of high nitrate contamination (Naidenko, Craig, and Nils 2012). In 2013, 125 PWS failed to comply with at least one health-based standard, affecting 10% of Iowans (Iowa DNR 2014). Most of these violations are related to coliform bacteria, and nitrate level. EWG's National Drinking Water Database based on 2004-2008 data showed that 50 water supply utilities in Iowa were violating the trihalomethanes or haloacetic acid standard, exposing around 62000 people (Naidenko, Craig, and Nils 2012). Examining 32 lakes in Iowa that were then used as source water by water treatment utilities, EWG found that 94% of the samples were detected with cyanobacteria at levels much higher than the recommended level by World Health Organization (WHO).

Drinking water utilities face significant cost from treating pollutants including nitrate and phosphorus that are directly caused by agricultural practices. Ribaudo *et al.* (2011) estimated the annual nitrate removal cost from drinking water sources across the US to be $4.8 billion, out of which agricultures' contribution is 1.7 billion dollar. The study implies that a 1% reduction of nitrate concentration in the source water would save $175 million per year. Naidenko, Craig, and Nils (2012) compile information on nitrate treatment and management costs. The water facilities cope with high nitrate levels in source water in several ways: by blending high nitrate wells with low nitrate wells , or shutting down wells if nitrate levels are high, building a nitrate removal facility, construction of holding ponds that will remove nitrates in natural way, and reverse osmosis of ground water. The reported per household construction cost of ion exchange and reverse osmosis systems for treating nitrate in drinking water can vary between $400-$1000 while the cost for drilling a new well can vary between $300-$400. The process of treating water for *cyanotoxins* is complex and highly expensive; the installation cost of a treatment facility to serve 100,000 people can vary from $4.4million to $56.6 million and annual operating costs can range from $0.5-$5.6 million (Naidenko, Craig, and Nils 2012).[103]

Private water wells are not subject to EPA regulatory limit. GEOSAM database records that there are currently 40,325 private water wells in Iowa.[104] Approximately 12% of Iowa's private water wells were detected with nitrate level above 10mg/L (University of Iowa Center for Health Effects of Environmental Contamination 2009). Keeler and Polasky (2014) provides an estimate of nitrate removal costs from private wells. Assuming that the least cost method of reverse osmosis will be adopted for nitrate treatment in drinking water, the cost to bring nitrates at least to the EPA level is $1790-$6160 per well, which is equivalent to $120-$414 per year.[105]. The cost of avoidance behavior, captured by purchase of bottled water for consumption, ranges in $241-$723 per individual per year.

The studies discussed above shed some light on the external cost of agriculture on drinking water treatment costs. If 17% of Iowans served by PWS are still experiencing high nitrate levels in their drinking water, the Naidenko, Craig, and Nils (2012) estimates on reverse osmosis systems for nitrate removal suggest an avoidance cost estimate of $84-$210 ml. If 12%

---

[103] In a high agricultural county, Fairmont, Minnesota, where 95% of land are under row crops, the cost for establishing a 5.4 million gallon per day treatment facility that will serve around 100000 people and treat source water for algal blooms, disinfection byproducts, and bad taste is $31.8 million.
[104] GEOSAM, Iowa Geological Survey. Available at http://geosam.iihr.uiowa.edu/search
[105] They assume a 20-year project life and 3% discount rate

of Iowa's 40,325 private wells still exhibit nitrate levels above 10mg/L, the Keeler and Polasky (2014) setting suggests that the total cost for Iowa will be $8.7-$30 ml.

5.3.2 ***Case Study on Des Moines Water Works***

Des Moines Water Works is the largest municipal water treatment plant in Iowa, and provides water supply to 0.5 million Central Iowans. Its three treatment plants are Fleur Drive Treatment Plant with a capacity of 100 million gallons per day (MGD), L.D. McMullen Treatment Plant at Maffitt Reservoir with a capacity of 25 MGD, and Saylorville Water Treatment Plant with a capacity of 10 MGD. The raw water at the Fleur Drive Plant comes from the Des Moines River, Raccoon River, and a Shallow Gallery system(a series of underground pipes located throughout Water Works Park next to the Raccoon River), where water must be pre-treated to remove sediment, organic matter and nitrate level. High nitrogen problem is common in all three sources at Fleur Drive Plant. Recently they have encountered similar nitrate problems in wells at McMullen plant. Based on the nitrate level on the source water, the system responds by switching from one river to the other, by maximizing use of the infiltration gallery system, using water stored in reservoirs, or adjusting production at the L. D. McMullen and Saylorville Water Treatment Plants. However, these ground water sources are reserved for high demand times and emergency conditions.

When the nitrate level is high, and the demand cannot be met by switching across plants and reservoirs, the facilities have to run their nitrate removal system. It has 8 vessels that are used to remove nitrate from source water, to bring it down to the EPA recommended level of 10mg/L. The treatment facilities did not have to run these vessels at all during the 2011-2012 time period. With the rise in nitrate level, in 2013 these vessels were operated for 74 days, 28 days in 2014, and all 26 days till January till 26th in 2015. The average cost per day for running this nitrate removal system is $7000. Facing a consistently high level of nitrate in Des Moines and Raccoon River and the resulting high treatment cost, DMWW has blamed agricultural discharge from the upstream. The debate on the impact of agriculture on water quality has been ignited once DMWW sued three counties for transporting N fertilizer into the river through tile drainage. An extract from DMWW website- *"A major conduit of nitrate pollution in the Raccoon River watershed is the artificial subsurface drainage system infrastructure, such as those created and managed by drainage districts. Des Moines Water Works recently filed a federal complaint against the Boards of Supervisors of Sac County, Buena Vista County, and Calhoun County, in their capacities as trustees of 10 drainage districts, for the discharge of nitrate pollutants into the Raccoon River"*[106].

The DMWW, being the largest municipal water treatment facility, has experienced an increase in the cost of nitrate treatment due to upstream water pollution. This has implications for other similar or smaller water treatment plants. Note that based on an extensive empirical analysis, Ribaudo *et al*. (2011) suggests that there is economies of scale in nitrate abatement cost. To understand the implied benefit of 40% reduction of Nitrogen in Iowa's waterways, we exploit information from Des Moines Waterworks (DMWW) to derive an estimate of abatement costs on treating drinking water for nitrate. The DMWW's savings from nitrate abatement cost will provide a relevant benefit estimate for Iowa from NRS. Following is a description of how we accomplish this.[107]

Step One: using daily data from January 1, 2012-January 26th, 2015 we predict the probability of treatment requirement based on the nitrate level in the Des Moines and Raccoon River, the total

---

[106] Retrieved at http://www.dmww.com/about-us/announcements/clean-water-act-litigation-faq.aspx (*last accessed on May 25th, 2015*).

[107] We thank Michael J. McCurnin, P.E., Director of Water Production, Des Moines Water Works for sharing some of their production information.

water processed(demand and treated water), producer price index, if the other two treatment plants were operating, and lag values of treatment. We estimate several specifications of the following binary model, and the results are reported in Table 17.

$$\begin{aligned}\text{Treated}_{d,m,t} = {} & \text{Nitrogen Level in Des Moines River}_{dmt} + \\ & \text{Nitrogen Level in Raccoon River}_{dmt} + \\ & \text{Total Water Processed}_{d,m,t} + \text{Producer Price Index}_{mt} + \\ & \text{Mcmullen Plant}_{d,m,t} + \text{Saylorville Plant}_{dmt} + \textstyle\sum_{l=1}^{3} l.\,\text{Treated}_{d,m,t} + \\ & \text{Unobserved}_{dmt},\end{aligned}$$

where *d*, *m*, and *t* denote day, month, and year.

Step Two: we consider specification IV as our preferred specification since it includes the dynamics and proceed based on the parameters obtained from this. The estimates suggest that the predicted probability of daily treatment is 14%. Based on the parameters obtained, we estimate the out of sample probability of treatment requirement if the mean nitrate level in the Des Moines River and Raccoon River falls by 40% due to the implementation of Iowa NRS. Panel (b) reports the results. This suggests that if the NRS is adopted, nitrate treatment requirement by DMWW will fall by 50%-100%.

Step three: Based on actual total treatment days in 2013, 2014, and average of 2013-2015, and average nitrate treatment cost per day, we develop three baseline cost scenarios as reported in Table 18. The baseline cost of treatment ranges from \$0.2-\$0.5 million. The 50%-100% reduced abatement requirement for nitrate due to the adoption of NRS will be translated into a yearly cost savings of \$0.1-\$0.5 million.

Step Four: The estimate above is obtained from the variable cost associated with operating the facility. The fixed cost of installing a nitrate treatment facility is \$7-\$183 million[108]. If we spread

[108]We gathered the replacement cost of nitrate removal facility from the following two sources. (i)http://www.desmoinesregister.com/story/money/agriculture/2015/05/14/water-works-nitrates-lawsuit/27331305/

this over 50 years, with a discount factor of 4% the equivalent annual cost will be $0.14-$3.7 million. Currently DMWW serves around 16% of the Iowa population. Assume that absent the NRS Iowa's surface and groundwater will eventually need treatment. We simulate the implied cost savings when 50% of Iowans need to have their drinking water treated. The estimated benefits, as reported in Table 19, range from $0.24 to $13.1 million per year.

The estimate derived from this case study is conservative since it considers only nitrate removal costs for one large municipality. Due to data limitations, we have not considered many other small treatment utilities which are experiencing similar problems, where the cost of abatement is much higher due to economies of scale. We do not include nitrate related health damages since nitrate levels below EPA recommended level still might have health implications. In addition, other nutrient removal costs, including that for *cyanotoxins*, chlorophylls, and bad smell, are not considered. Finally, the welfare cost of high nitrates in private wells are not included.

6. Comparison across Conservation Practices in terms of Benefits

Figures 3 to 5 shows the total benefits from reduced soil erosion, increased wildlife habitat, and carbon sequestration by conservation practices under each example scenario considered. Since SCC is calculated based on global damages, we translate the global carbon benefit into a local benefit by weighting the SCC by the share of 3 million Iowans' in 7 billion

(ii) http://www.dmww.com/upl/documents/water-quality/lab-reports/fact-sheets/nitrate-removal-facility.pdf (last accessed on May 19th, 2015).

The construction cost incurred by DMWW for the nitrate treatment facility in 1990-1991 was $4.1, which is $7 million at 2013 US$ value.

global population. That means when we are considering Iowan's share only, we are assigning a lower value to SCC. Under scenario NCS1, almost 65-70% of the benefits derive from wetlands, and another 30-34% from cover crops when local carbon benefits are considered. In this case, the MRTN's share is negligible. However, when we include the global carbon benefit and thus assign a higher value of carbon sequestration, cover crops' contribution to the lower bound of total benefit rises to 95%. This is because of the large amount of carbon sequestered by 17 million acres under cover crops in scenario NCS1. MRTN's contribution in total benefit reaches 5%.

Under scenario NCS3, when local carbon damage is considered, 5% of the retired land contributes to 48-60% of the total benefit. Similar contributions from cover crop acres and converted wetland acres are respectively 19%-24% and 12%-15%. However, when global damage from carbon is considered, cover crops' and MRTN's share in total benefit shows a sharp rise. Under scenario NCS8 with local carbon benefit, a larger share in total benefits comes from wetlands and tillage practices. However, when the global carbon benefit is considered, the lion's share of total benefit is taken up by reduced tillage practices, which is mainly because of a large coverage of acres under reduced tillage practice.

Tables 20 and 21 report total benefits by ecosystem services under each of the scenarios including benefits from improved opportunities for water-based recreation, local lake-adjacent housing values, and drinking water purification. Note that these three ecosystem services do not vary across scenarios since they are derived for a uniform reduction of 40% N and 29% P that is achieved under one of these three scenarios. The monetary benefit figures in Table 21 are 4-19 times larger compared to those in Table 20, which is mainly because of the higher carbon value assigned in the former case. Assuming a local carbon damage, the largest monetary benefits are derived from conservation practices included under NCS3 that lie in the range of $163-$246 million. The combined benefits under NCS1 in low and high scenarios are approximately 50% and 26% lower compared to those under NCS3. Finally, the combined benefits from low scenario under NCS8 are 34% lower compared to similar benefits under NCS3.

7. Conclusion

This study evaluates the potential external benefits from a number of agricultural conservation practices included under the Iowa Nutrient reduction strategy 2013. The estimation method follows a benefits transfer approach and excludes private benefits from the analysis. Besides the direct benefit of reduced nutrients in Iowa's waterbodies, these practices generate a large amount of benefits through other ecosystem services including reduced soil erosion, reduced carbon in the atmosphere, enhanced wildlife, and increased biodiversity. Iowans' welfare will improve from these ecosystem improvements through increased opportunities for outdoor recreation, aesthetic values of improved water quality, better quality for drinking water, and reduced greenhouse gas emission, among others.

The yearly aggregate monetary values for Iowans from these ecosystem services range from $88-$257 million. However, the benefit estimates are conservative for a number of reasons. First, we took a conservative approach when selecting values from the literature concerning rates of carbon sequestration, reduced soil erosion, and the monetization of ecosystem services using benefits transfer. Due to lack of supporting studies we could not include many benefits, such as pollinators, existence value associated with water quality and wildlife, biodiversity and endangered species protection etc. In addition, we focus exclusively on the benefits to be accrued to Iowans. We have not included the benefits to be accrued at the Gulf from improved aquatic life and reduction of the hypoxic zone. Inclusion of such benefits will make the actual benefit estimates larger. To put it in context, out estimates reveal that including the global benefits from additional carbon sequestered in Iowa can make the total benefit estimates 4 to 19 times larger.

The success of Iowa nutrient reduction strategy largely depends on farmers' adoption of targeted management practices at the field level. Incentives and payment vehicles may be necessary to incentivize them to participate in the program. However, since the source of funds is most likely to be public, to ensure the optimum use of public tax money in incentivizing nutrient management practices requires attaining the highest benefit per unit of money spent. In this regard, a detailed in-depth analysis of benefits derived from this strategy is imperative. The findings from this research will inform policymakers and stakeholders across the state and help them better understand the tradeoffs involved in policies that encourage conservation.

8. References

Azevedo, C., J.A. Herriges, and C.L. Kling. 2000. "Iowa Wetlands: Perceptions and Values." Report 00- SR91, Iowa State University, Center for Agricultural and Rural Development.

Egan, Kevin J., J.A. Herriges, C.L. Kling, and J.A. Downing. 2009. "Valuing Water Quality as a Function of Water Quality Measures." *American Journal of Agricultural Economics* 91:106-123.

Ehmke, M., C. Jones-Ritten, J. Shogren, and T. Panchalingam. 2015. "Integrating Ecological and Economic Considerations for Pollinator Habitat Policy." *Choices*, Agricultural and Applied Economic Association, 30(2).

Ge, Jiaqi, C.L. Kling, and J.A. Herriges. 2013. "How Much is Clean Water Worth? Valuing Water Quality Improvement Using A Meta-Analysis." Staff General Research Papers 36597, Iowa State University, Department of Economics.

Gleason, R., M. Laubhan, and N. Euliss Jr. (eds). 2008. Ecosystem Services Derived from Wetland Conservation Practices in the United States Prairie Pothole Region with an Emphasis on the U.S. Department of Agriculture Conservation Reserve and Wetland Reserve Programs, Professional Paper 1745, U.S. Department of the Interior, U.S. Geological Survey.

Gonzalez-Ramirez, J., A. Valcu, C.L. Kling, and P. Gassman. 2014. "Cropland Cash Rental Rates Estimates and Carbon Supply Curves for the Upper Mississippi River Basin." Center for Agricultural and Rural Development (CARD), Iowa State University.

Hansen, L., D. Hellerstein, M. Ribaudo, J. Williamson, D. Nulph, C. Loesch, and W. Crumpton. 2015. "Targeting Investments to Cost Effectively Restore and Protect Wetland Ecosystems: Some Economic Insights." ERR-183, U.S. Department of Agriculture, Economic Research Service.

Hansen, L., and M. Ribaudo. 2008. "Economic Measures of Soil Conservation Benefits: Regional Values for Policy Assessment." TB-1922, U.S. Department of Agriculture, Economic Research Service.

Hansen, L., 2007. "Conservation Reserve Program: Environmental Benefits Update." *Agricultural and Resource Economics Review* 36(2):267-280.

Iowa Department of Natural Resources. 2014. State of Iowa Public Drinking Water Program 2013 Annual Compliance Report. Environmental Services Division, Water Quality Bureau, Water Supply Engineering & Operations Sections.

Iowa Nutrient Reduction Strategy. 2013. "A Science and technology---based framework to assess and reduce nutrients to Iowa Waters and the Gulf of Mexico", Prepared by: Iowa Department of Agriculture and Land Stewardship, Iowa Department of Natural Resources, Iowa State University College of Agriculture and Life Sciences. Available at http://www.nutrientstrategy.iastate.edu/documents (*last accessed on June 28th, 2015*).

IPCC. 2006. 2006 IPCC Guidelines for National Greenhouse Gas Inventories. In H.S. Eggleston, L. Buendia, K. Miwa, T. Ngara, and K. Tanabe (Eds.), *National Greenhouse Gas Inventories Program* (Agriculture, Forestry and Other Land Use, Vol. 4). Kanagawa, Japan: IGES.

Kaspar, Tom, C., J.K. Radke, and J.M. Laflen. 2001. "Small grain cover crops and wheel traffic effects on infiltration, runoff, and erosion." *Journal of Soil and Water Conservation* 56:160-164.

Keeler, B.L., S. Polasky, K. Brauman, K. Johnson, J. Finlay, A. O'Neill, K. Kovacs, and B. Dalzell. 2012. "Linking water quality and well-being for improved assessment and valuation of ecosystem services." *Proceedings of the National Academy of Sciences* 109(45):18619–18624.

Keeler, B. L., and S. Polasky. 2014. "Land-Use Change and Costs to Rural Households: A Case Study in Groundwater Nitrate Contamination." *Environmental Research Letters* 9 (7).

Losey, J.E., and M. Vaughan. 2006. "The economic value of ecological services provided by insects." *BioScience* 56: 311–323.

Millar, N., G.P. Robertson, P.R. Grace, R. J. Gehl, and J.P. Hoben. 2010. "Nitrogen fertilizer management for nitrous oxide (N2O) mitigation in intensive corn (Maize) production: an emission reduction protocol for US Midwest agriculture." *Mitigation and Adaptation Strategies for Global Change* 15:185–204.

Naidenko, V. O., C. Craig, and B. Nils. 2012. "Troubled Waters Farm Pollution Threatens Drinking Water." Environmental Working Group. Available at http://www.ewg.org/research/troubled-waters (last accessed on May 30th, 2015).

Ribaudo, M., J. Delgado, L. Hansen, M. Livingston, R. Mosheim, and J. Williamson. 2011. "Nitrogen in Agricultural Systems: Implications for Conservation Policy." USDA Economic Research Report No. ERR-127, Economic Research Service, U.S. Department of Agriculture, Washington, D.C.

Schipanski, M. E., M. Barbercheck, M.R. Douglas, D.M. Finney, K. Haider, J.P. Kaye, A.R. Kemanian, D.A. Mortensen, M.R. Ryan, J. Tooker, and C. White. 2014. "A framework for evaluating ecosystem services provided by cover crops in agroecosystems." *Agricultural Systems* 125 (2014):12–22.

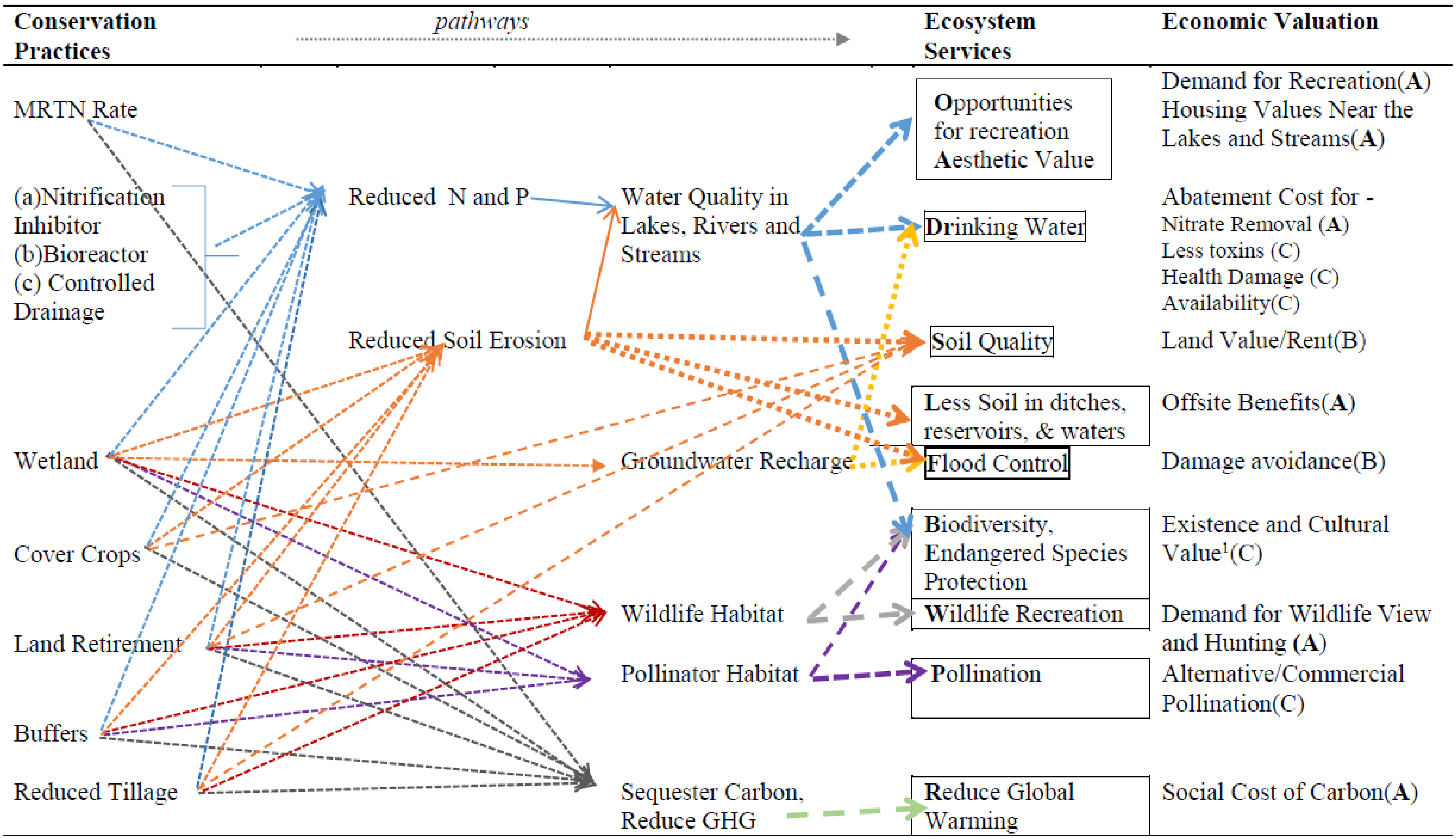
Sullivan, P., D. Hellerstein, L. Hansen, R. Johansson, S, Koenig, R. Lubowski, W. McBride, D. McGranahan, M. Roberts, S. Vogel, and S. Bucholtz. 2004. "The conservation reserve program: economic implications for rural America." Agricultural Economic Report (AER) No. 834, Economic Research Service, U.S. Department of Agriculture, Washington, D.C.

University of Iowa Center for Health Effects of Environmental Contamination. 2009. Iowa Statewide Rural Well Water Survey Phase 2 (SWRL2). Available: http://www.cheec.uiowa.edu/research/SWRL2.html (*last accessed on April 15th, 2015*).

Villanueva, C. M., K.P. Cantor, J.O. Grimalt, N. Malats, D. Silverman, and A. Tardon. 2007. "Bladder cancer and exposure to water disinfection by-products through ingestion, bathing, showering, and swimming in pools." *American Journal of Epidemiology* 165(2): 148-56.

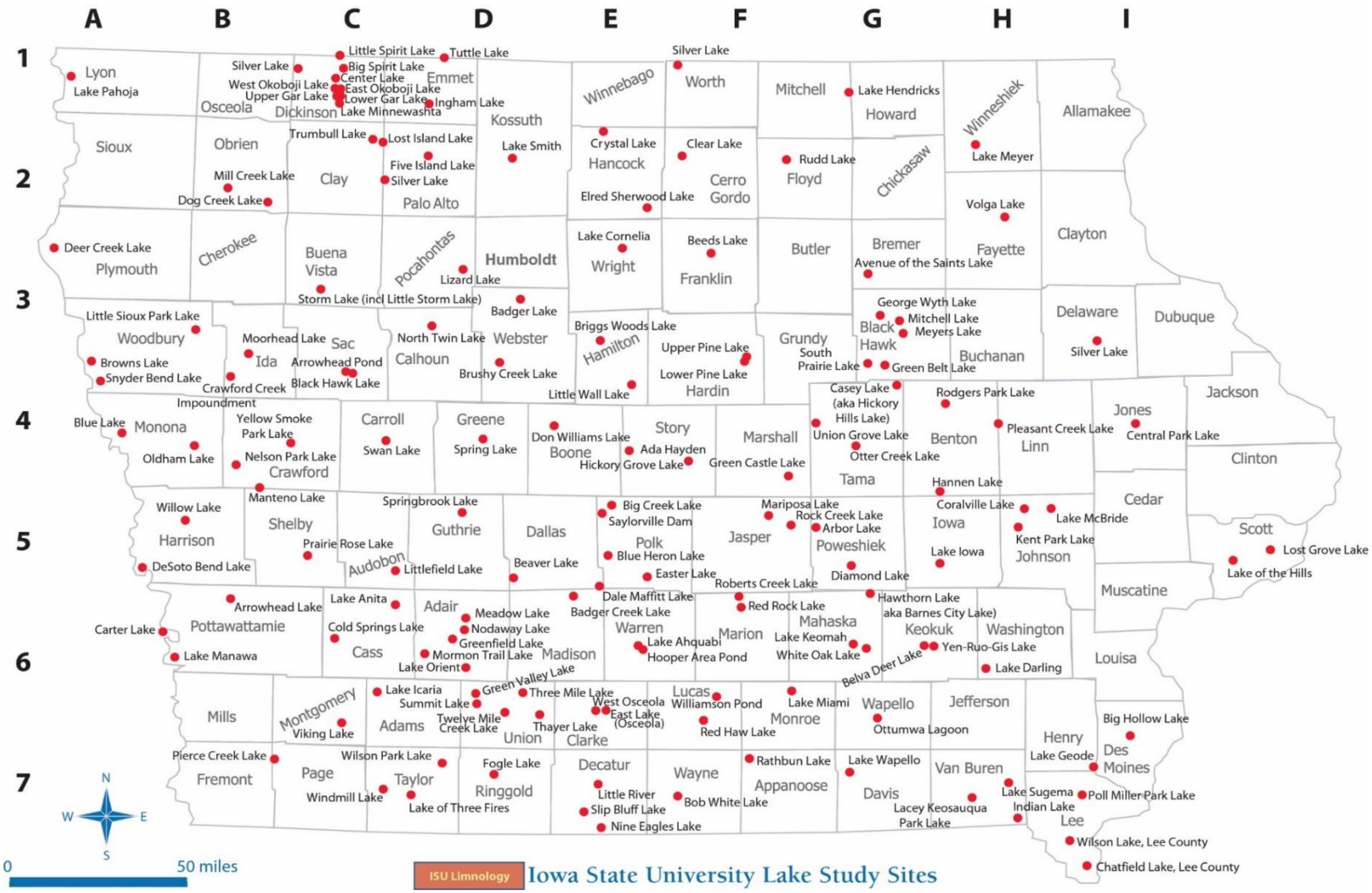
Ward, M. H., B. A. Kilfoy, P.J. Weyer, K.E. Anderson, A. Folsom, and J.R. Cerhan. 2010. "Nitrate intake and the risk of thyroid cancer and thyroid disease." *Epidemiology* 21(3): 389-95.

Zhou, X., M. Al-Kaisi, and M.J. Helmers. 2009. "Cost eff effectiveness of conservation practices in controlling water erosion in Iowa." *Soil & Tillage Research* 106:71–78.



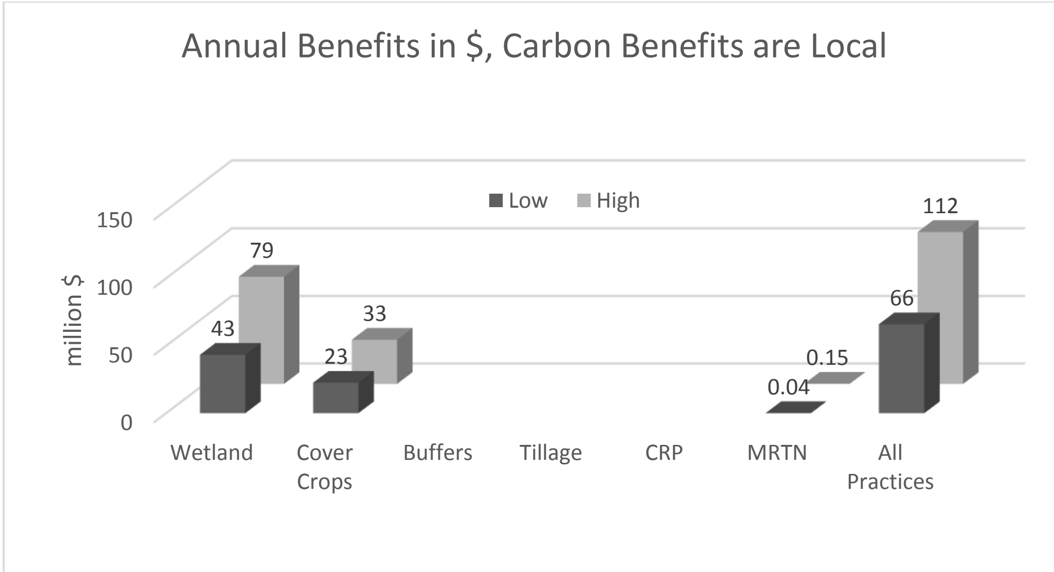
Note: A, B, and C stand for clarification of whether a valuation has been conducted here. A: Valuation at least partially undertaken here, B: Valuation is possible but not undertaken, C: Valuation is not undertaken due to lack of Data.

Figure 1: Link between Conservation Practices, Ecosystem, and Ecosystem Services



Source: CARD, Iowa State University

Figure 2: A Map with Major Lakes in Iowa



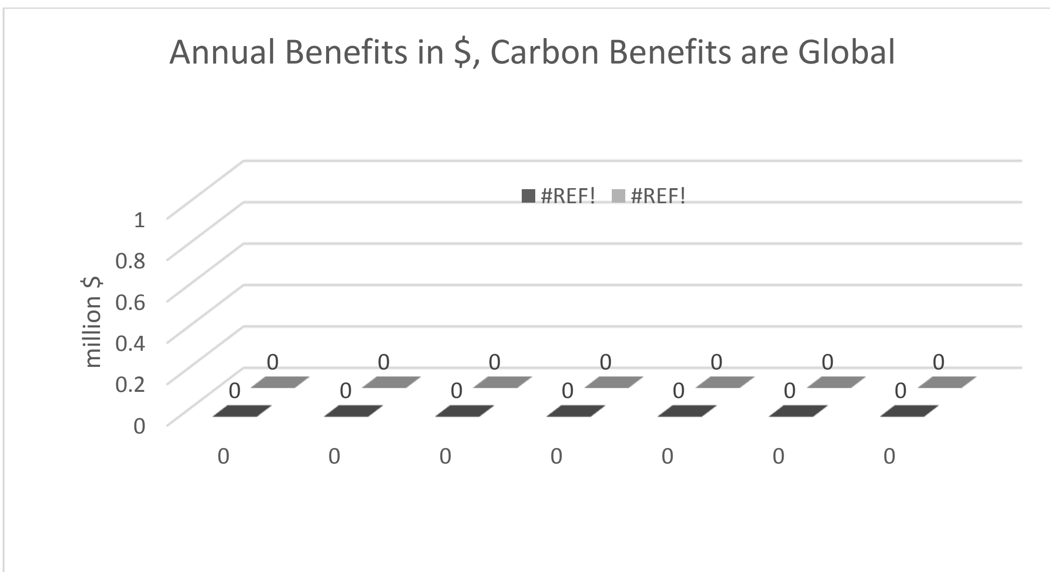


Figure 3: Benefits from Soil Erosion Reduction, Wildlife, and Carbon Sequestration by conservation Practices under Scenario NCS1

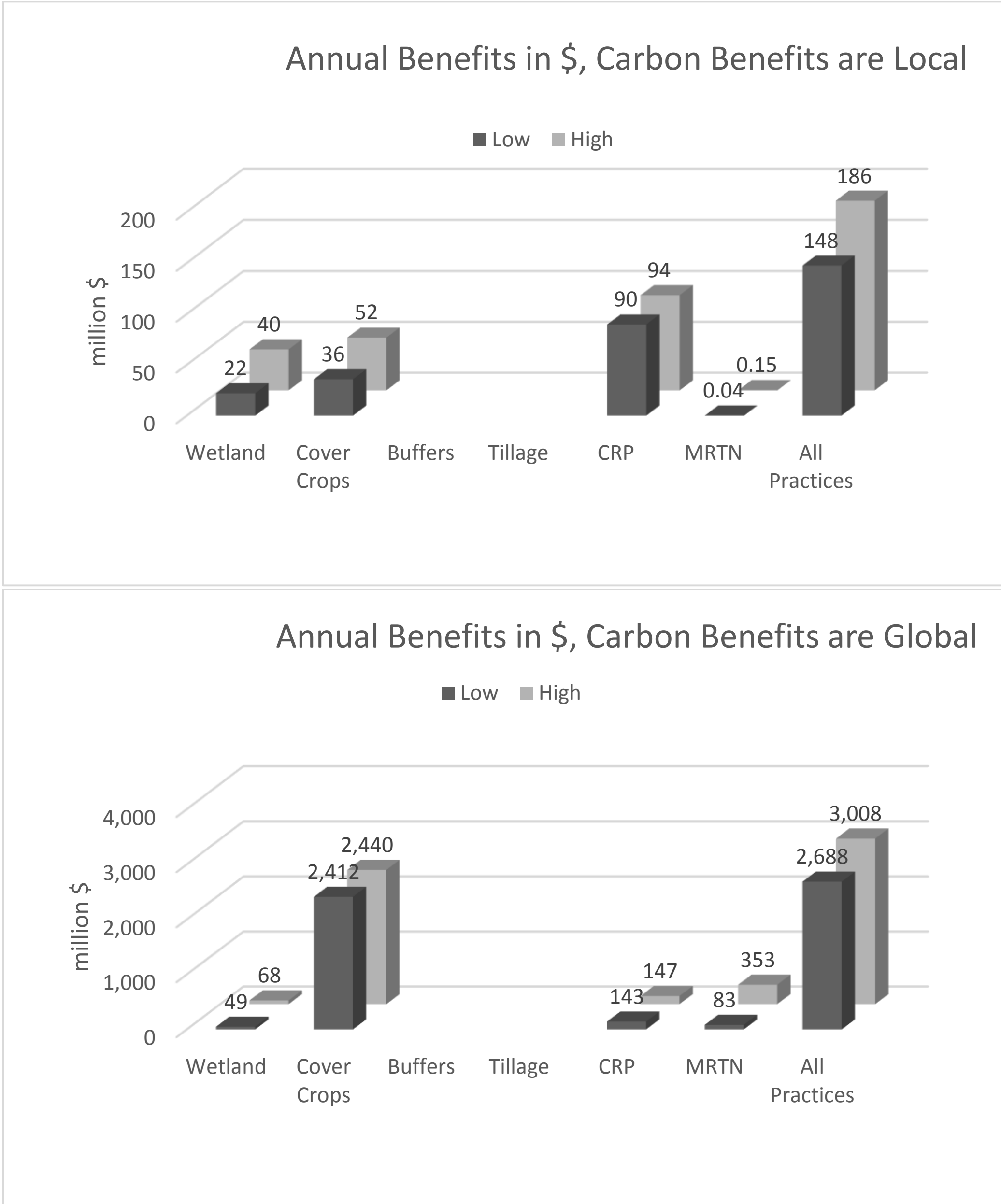
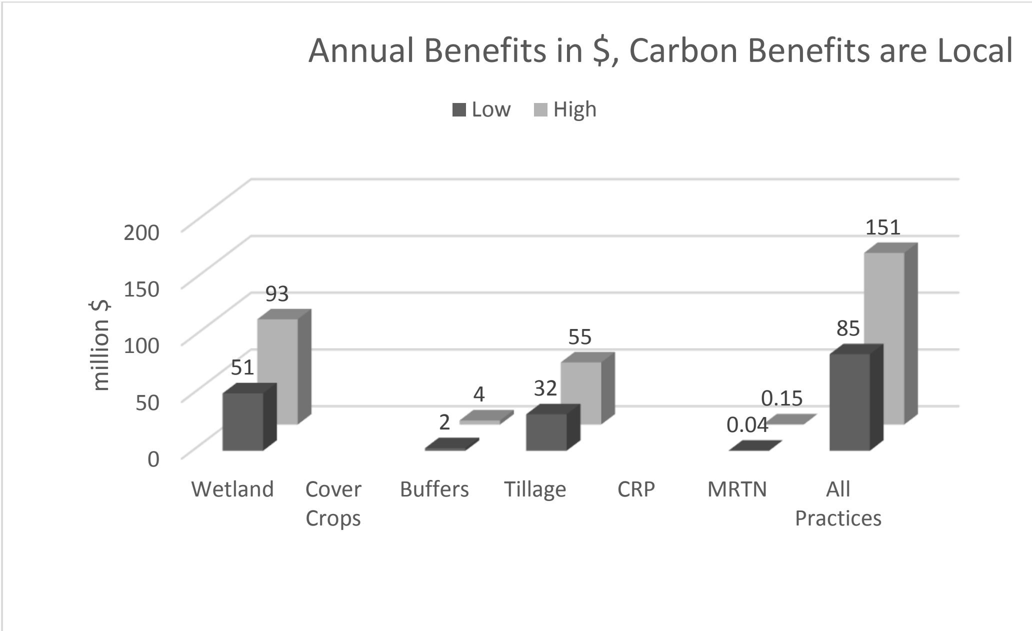


Figure 4: Benefits from Soil Erosion Reduction, Wildlife, and Carbon Sequestration by Conservation Practices under Scenario NCS3



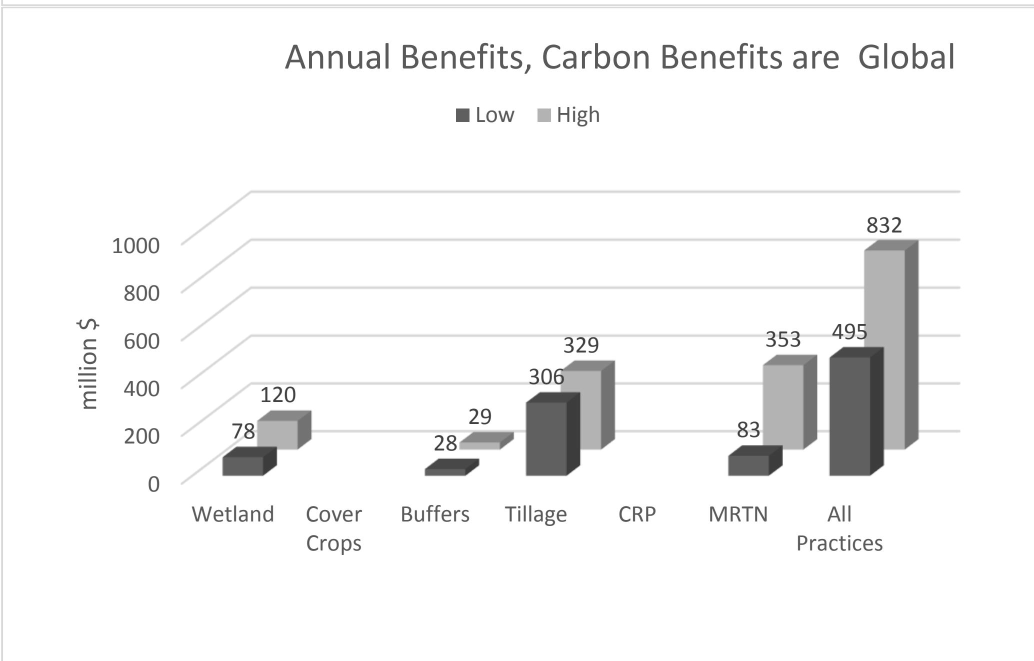


Figure 5: Benefits from Soil Erosion Reduction, Wildlife, and Carbon Sequestration by Conservation Practices under Scenario NCS8

Table 1: Three Scenarios Combining Practices to Achieve Nutrient Reduction Target

Scenarios	Practices Included	Nitrate N (%Reduction from baseline)	Phosphorus (%Reduction from baseline)	Initial Investment (million $)	Total Equal Annualized Cost EAC (million $/year)	Statewide Average EAC Costs ($/acre)
NCS1	(i) MRTN Rate, (ii) 60% Acreage with Cover Crop (iii) 27% of ag land treated with wetland (iv) 60% of drained land has bioreactor	42	30	3,218	756	36
NCS3	(i) MRTN Rate, (ii) 95% of acreage with Cover Crops (iii)34% of ag land in MLRA 103 and 104 treated with wetland, (iv)5% land retirement in all MLRAs	42	50	1,222	1,214	58
NCS8	(i) MRTN Rate, (ii) Inhibitor with all Fall Commercial N, (iii) Sidedress All Spring N, (iv)70% of all tile drained acres treated with bioreactor, (v)70% of all applicable land has controlled drainage, (vi) 31.5% of ag land treated with a wetland, (vii) 70% of all agricultural streams have a buffer *Phosphorus reduction practices*: (i)convert 90% of Conventional Tillage CS & CC acres to Conservation Till and (ii)Convert 10% of Non--No-till CS & CC ground to No--Till	42	29	4,041	77	4

Source: Table 5 in Iowa NRS. *EAC stands for Equal Annualized Cost (50 year life and 4% discount rate) and factors in the cost of any corn yield impact as well as the cost of physically implementing the practice. Average cost based on 21.009 million acres, of Corn-Corn and Corn-Soybean Rotation.

Table 2: Soil Erosion Reduction and Sequestered Carbon by Wetland Acres

Scenarios	Treated Acres of Land by wetland		Total converted wetland acres (ml acres)	Reduced Soil Erosion (ml tons/year)		Total sequestered carbon (ml tons/year)	
		Million acres		Low	High	Low	High
NCS1	27% of all Agricultural Land in all MLRAs	17	0.34	19.09	34.10	0.23	0.33
NCS3	34% of Agricultural Land in all MLRA 103 and 104	4	0.17	9.69	17.30	0.12	0.17
NCS8	31.5% of all Agricultural Land in all MLRAs	9	0.40	22.27	39.78	0.27	0.38

Note: Carbon Sequestration is calculated following the similar manner as that for soil erosion. Similarly, total carbon sequestration= Total treated acres*Per Acre carbon Sequestration.

Table 3: Monetary Value of Benefits from the Wetland Acres

	Offsite Benefits from Reduced Soil Erosion ($ ml/year)		Recreation, Wildlife View, Aesthetic ($ ml/year)		Value of Sequestered Carbon or Reduced GHG Emissions: Iowans' share ($ ml/year)		Global Benefits ($ ml/year)		Total Benefits ($ ml/year)	
	Low	High	Low	High	Low	High	Low	High	Low	High
NCS1	40.46	72.29	2.87	7.16	0.01	0.01	23.23	23.41	43.35	102.86
NCS3	20.53	36.68	1.46	3.63	0.01	0.01	27.10	27.31	22.00	67.63
NCS8	47.21	84.34	3.35	8.36	0.01	0.01	27.10	27.31	50.57	120.00

Note: The lower and upper bound of the total benefits is calculated respectively including lower bound on Iowan's share in GHG benefits and upper bound on global GHG benefits.

Table 4: Soil Erosion Reduction and Sequestered Carbon by Cover crop Acres

Scenarios	Treated Acres of Land by Cover Crops (million acres)		Reduced Soil Erosion (million tons/year)		Sequestered Carbon (million tons/year)	
			Low	High	Low	High
NCS1	60% of all Agricultural Land	17.03	10.35	15.15	16.86	21.11
NCS3	95% of all Agricultural Land	26.96	16.39	23.99	26.69	33.43
NCS8	N.A.	-	-	-	-	-

Table 5: Monetary Value of Benefits from the Cover Crop Acres

	Offsite Benefits from Reduced Soil Erosion ($ ml/year)		Value of Sequestered Carbon or Reduced GHG Emissions				Total Benefits ($ ml/year)	
			Iowans' share ($ ml/year)		Global Benefits ($ ml/year)			
	Low	High	Low	High	Low	High	Low	High
NCS1	21.95	32.13	0.64	0.65	1501.28	1508.85	22.59	1540.98
NCS3	34.75	50.87	1.02	1.02	2377.03	2389.01	35.77	2439.88
NCS8	-	-	-	-	-	-	-	-

Note: The lower and upper bound of the total benefits is calculated respectively including lower bound on Iowan's share in GHG benefits and including upper bound on global GHG benefits.

Table 6: Soil Erosion Reduction and Sequestered Carbon from Acres under Land Retirement

Scenarios	Acres of Land to be Retired (million acres)		Reduced Soil Erosion (million tons/year)		Sequestered Carbon (million tons/year)	
			Low	High	Low	High
NCS1	N.A.	-	-	-	-	-
NCS3	5% of Agricultural Land	1.14	3.20	5.07	0.66	0.75
NCS8	N.A.	-	-	-	-	-

Table 7: Monetary Value of Benefits from the Retired Acres

	Offsite Benefits from Reduced Soil Erosion ($ ml/year)		Recreation, Wildlife View, Aesthetic ($ ml/year)		Value of Sequestered Carbon or Reduced GHG Emissions				Total Benefits ($ ml/year)	
					Iowans' share ($ ml/year)		Global Benefits ($ ml/year)			
	Low	High	Low	High	Low	High	Low	High	Low	High
NCS1	-	-	-	-	-	-	-	-	-	-
NCS3	6.79	10.76	82.89	82.89	0.02	0.02	53.61	53.77	89.71	147.42
NCS8	-	-	-	-	-	-	-	-	-	-

Note: The lower and upper bound of the total benefits is calculated respectively including lower bound on Iowan's share in GHG benefits and upper bound on global GHG benefits.

Table 8: Soil Erosion Reduction and Sequestered Carbon from Acres under Buffers

Scenarios	Acres of Land to be Converted (million acres)		Reduced Soil Erosion (million tons/year)		Sequestered Carbon (million tons/year)	
			Low	High	Low	High
NCS1	N.A.		-	-	-	-
NCS3	N.A.		-	-	-	-
NCS8	70% of ag streams that are not currently buffered	0.38	1.07	1.69	0.22	0.36

Table 9: Soil Erosion Reduction and Sequestered Carbon from Acres under Buffers

	Offsite Benefits from Reduced Soil Erosion ($ ml/year)		Value of Sequestered Carbon or Reduced GHG Emissions				Total Benefits ($ ml/year)	
			Iowans' share ($ ml/year)		Global Benefits ($ ml/year)			
	Low	High	Low	High	Low	High	Low	High
NCS1	-	-	-	-	-	-	-	-
NCS3	-	-	-	-	-	-	-	-
NCS8	2.26	3.58	0.01	0.01	25.55	25.80	2.27	29.38

Note: The lower and upper bound of the total benefits is calculated respectively including lower bound on Iowan's share in GHG benefits and including upper bound on global GHG benefits.

Table 10: Soil Erosion Reduction and Sequestered Carbon from Conservation Tillage

Scenarios	Acres of Land to be Treated (million acres)		Reduced Soil Erosion (million tons/year)		Sequestered Carbon (million tons/year)	
			Low	High	Low	High
NCS1	N.A.		-	-	-	-
NCS3	N.A.		-	-	-	-
NCS8	Convert 90% of Conventional Tillage CC and CS Acres to Conservation tillage	7.66	11.08	20.83	2.53	2.53
	Convert 10% of Conventional Non-no-till CC and CS Acres to no-till	1.62	4.01	4.93	1.03	1.31
	Total	9.28	15.09	25.75	3.56	3.84

Table 11: Monetary Value of Benefits from the Conservation tillage

	Offsite Benefits from Reduced Soil Erosion ($ ml/year)		Value of Sequestered Carbon or Reduced GHG Emissions				Total Benefits ($ ml/year)	
			Iowans' share ($ ml/year)		Global Benefits ($ ml/year)			
	Low	High	Low	High	Low	High	Low	High
NCS1	-	-	-	-	-	-	-	-
NCS3	-	-	-	-	-	-	-	-
NCS8	31.99	54.60	0.12	0.12	273.86	274.35	32.11	328.94

Note: The lower and upper bound of the total benefits is calculated respectively including lower bound on Iowan's share in GHG benefits and including upper bound on global GHG benefits.

Table 12: Reduced GHG Emission from MRTN, and Implied Benefits in $

Average Nitrogen application			Reduced GHG emission by shifting to MRTN rate		Value of Sequestered Carbon or Reduced GHG Emissions			
			Low: Applying IPCC(2006) Formula	High: Applying Millar(2010) Formula	Iowa's Share ($ ml/year)		Global GHG Benefits ($ ml/year)	
	Current Rate (lb/acre)	Proposed MRTN (lb/acre)	Mt-CO_2 *e/year*		Low	High	Low	High
CC	201	190	262184	1678299				
CS	151	133	1167170	4942465				
Total			1.17 ml	4.94 ml	0.04	0.15	83.41	353.19

Method of Calculation

We show the calculation of GHG benefits for switching to MRTN rate for CC adopting the numbers from Millar (2010).

Carbon Emission under baseline: [(225.08*0.012*exp(0.00475*225.08))+1.47]*298*44/28= 4372.61 Kg CO_2 e/ha/year

Carbon Emission under MRTN: [(190*0.012*exp(0.00475*190))+1.47]*298*44/28= 3321.05 Kg CO_2 e/ha/year

The GHG emission is reduced by (4372.61-3321.05) = 1051.56 Kg CO_2 e/ha/year

Total Corn-Corn Land in Iowa= 1596013 hector

Total GHG emission reduction=(1596013*1051.56)/1000= 1678296.85 metric tons.

IPCC's linear formula is 1.47+(0.01*Fertilizer rate)*298*44/28.

Note that Both linear and non-linear formula assumed 1.47 kgCO_2 e/ha/year will be emitted under no nitrogen scenario.

Table 13: Weight assigned in construction of Water Quality Index

Parameter	Weight (adjusted)
Dissolved Oxygen	0.3
pH	0.2
Total Phosphorus	0.18
Total Nitrate	0.18
Turbidity	0.14

Table 14: Change in Water Quality and Iowan's Willingness to Pay

	Obs.	Mean	Std. Dev.	Min	Max
Water Quality Index Before the Plan	130	72.07	12.85	34.21	90.21
Water Quality Index After the Plan	130	73.74	12.13	35.04	90.31
Change in Water Quality	130	1.68	1.53	0.07	6.08
WTP(Hedonic)	130	655.27	38.27	599.81	752.46

Table 15: Information Sources Used to Extract Count of Total Housing Units near Lakes

	Sources	Total Number of Lakes	Housing Units within	
			Half mile	1 Mile
a	Parcel Data without property type information	3	2717	4936
	County Address Maps	68	16678	24644
	Parcel Data with property type info (Residential)	5	3516	8617
	Zillow and Google Maps	55	594	12396
	Total	131	23505	50593
b	Missouri Census Data Center[109]	131	23462	49938

Table 16: Compensating Variation in Iowa from Water Quality Improvement (in million $)

Specification	In million $ per year		
	NCS1 (42% less Nitrate and 30% less Phosphorus)	NCS3 (42% less Nitrate and 50% less Phosphorus)	NCS8 (42% less Nitrate and 29% less Phosphorus)
Linear form	5.34	9.62	5.34
Log form	13.89	22.44	12.7

[109] Source is Missouri Census Data Center, http://mcdc.missouri.edu/websas/caps10c.html (last accessed on July 17th, 2015). We know the latitude and longitude of each lake. Providing this input on Missouri Census Data Center extracts information from the census on total counts of housing units within half mile and 1 mile radius from the lake border.

Table 17: Nitrate Level in Source Water and Probability of Treatment

	I		II		III		IV	
	Both River	One River	Both River	One River	Both River	One River	Both River	One River
Nitrate Level in DSM River	0.684***	0.591***	0.268	0.165	0.352*	0.287	0.489**	0.500**
	(0.07)	(0.06)	(0.16)	(0.16)	(0.17)	(0.16)	[0.18]	[0.17]
Nitrate Level in Raccoon River	-0.086*		-0.083		-0.089		0.007	
	[-0.038]		[-0.078]		[-0.088]		[0.034]	
Lag of Nitrate Level in DSM River			0.418**	0.456**	0.484**	0.515**		
			[0.146]	[0.151]	[0.159]	[0.158]		
Lag of Nitrate Level in Raccoon River			0.017		0.062			
			[0.077]		[0.091]			
Treated Water (DD + Reserve adjustment)	0.045***	0.045***	0.044***	0.044***	0.092***	0.093***	0.037*	0.037*
	[0.009]	[0.009]	[0.009]	[0.009]	[0.015]	[0.015]	[0.018]	[0.018]
Producer Price Index	0.127***	0.119***	0.128***	0.123***	0.169***	0.166***	0.153	0.154
	[0.022]	[0.024]	[0.025]	[0.028]	[0.035]	[0.036]	[0.085]	[0.084]
McMullen Plant					-0.092*	-0.092*	-0.042	-0.042
					[-0.036]	[-0.036]	[-0.066]	[-0.065]
Saylorville Plant					-0.249*	-0.253*	-0.366*	-0.367*
					[-0.126]	[-0.123]	[-0.156]	[-0.156]
1st Lag of Treatment Indicator (Y_{t-2})							8.236***	8.284***
							[-1.063]	[-1.067]
2nd Lag of Treatment Indicator (Y_{t-2})							-5.724***	-5.777***
							[-1.098]	[-1.096]
3rd Lag of Treatment Indicator (Y_{t-3})							1.685**	1.687**
							[-0.554]	[-0.557]
Constant term	-26.95***	-25.84***	-27.25***	-26.57***	-35.34***	-34.93***	-29.42*	-29.58*
	[-3.569]	[-3.814]	[-4.115]	[-4.419]	[-5.787]	[-5.943]	[-13.41]	[-13.283]
N	879	897	763	788	763	788	878	896
Model Statistics								
pseudo R-square	0.70	0.70	0.69	0.69	0.72	0.73	0.95	0.95
AIC	225.44	227.84	220.65	219.13	202.80	199.82	53.18	51.18
Panel b: what will be the probability of treatment requirement in a certain day?								
Predicted Probability (In Sample)	0.14	0.14	0.16	0.16	0.16	0.15	0.14	0.14
Out Of Sample Prediction (40% Nitrate Reduction)	0.005	0.005	0.005	0.005	0.001	0.001	0.070	0.068
% Change	**-96.60**	**-96.74**	**-96.89**	**-96.99**	**-99.44**	**-99.46**	**-50.41**	**-50.90**
Updating Lag with initial Value at 0							0.000	0.000
% Change							**-99.99**	**-99.99**
Updating Lag with initial Value at mean=0.14							0.003	0.000
% Change							**-97.98**	**-99.98**

Note: Standard errors in brackets. Significance level can be read as * p<0.05, ** p<0.01, *** p<0.001.

Table 18: Estimated Cost Savings from Reduced Abatement Requirement of Nitrate

Baseline Cost Scenarios		Saving In treatment Costs		
		2013 (73 Days)	2014 (28 days)	Average 2012-2015 (26 Days)
		$518000	$196000	$240722
How the lag values were assessed while deriving change in probability of treatment requirement for nitrate	% Change in Nitrate Treatment Requirement	Saving In treatment Cost compared to Baseline		
		2013	2014	Average of 2013-2015
Dynamic Model: Both River				
Realized value of lag	-50.41	$257,573.36	$98,795.26	$91,738.46
Update the lag initial value at 0	-99.99	$510,932.07	$195,973.95	$181,975.81
Update the lag initial value at (Mean =0.14)	-97.98	$500,657.93	$192,033.18	$178,316.52
Dynamic Model : Des Moines River Only				
Realized Value of Lag	-50.90	$260,073.64	$99,754.27	$92,628.97
Update the lag Initial value at 0	-99.99	$510,928.84	$195,972.71	$181,974.66
Update the lag initial value (Mean=0.14)	-99.98	$510,921.47	$195,969.88	$181,972.03

Table 19: Drinking Water Treatment Benefits from Nutrient Reduction

	If Customers of DMWW (16% Iowans) are affected		If 50% of Iowans are affected	
	Min	Max	Min	Max
Operational Cost	0.09 ml	0.51 ml	0.29 ml	1.6 ml
Fixed Cost in $ (Equalized annual cost)	0.14 ml	3.7 ml	0.44 ml	11.5 ml
Annual Benefit in 2013 $	0.24 ml	4.2 ml	0.73 ml	13.1 ml

Table 20: Breakdown of Total Benefits by Conservation Practices and Ecosystem Services (Carbon Benefits are Local)

Categories	NCS1		NCS3		NCS8	
	Low	*High*	*Low*	*High*	*Low*	*High*
Wetland	40.46	72.29	20.53	36.68	47.21	84.34
Cover Crops	21.95	32.13	34.75	50.87		
Buffers					2.26	3.58
Tillage					31.99	54.60
Land Retirement						
MRTN	0.00	0.00	0.00	0.00	0.00	0.00
Reduced Soil Erosion	*62.41*	*104.42*	*55.29*	*87.55*	*81.46*	*142.52*
Wetland	2.87	7.16	1.46	3.63	3.35	8.36
Cover Crops	0.00	0.00	0.00	0.00		
Buffers						
Tillage						
Land Retirement			82.89	82.89		
MRTN	0.00	0.00	0.00	0.00	0.00	0.00
Recreation and Wildlife	*2.87*	*7.16*	*84.35*	*86.53*	*3.35*	*8.36*
Wetland	0.01	0.01	0.01	0.01	0.01	0.01
Cover Crops	0.64	0.65	1.02	1.02		
Buffers					0.01	0.01
Tillage					0.12	0.12
CRP			0.02	0.02	0.00	0.00
MRTN	0.04	0.15	0.04	0.15	0.04	0.15
Total Carbon Benefit	*0.69*	*0.81*	*1.09*	*1.21*	*0.18*	*0.29*
Benefits from 40% reduction in Nitrate N and Phosphorus in Iowa's Waterbodies						
Water Based Recreation	5.34	22.45	5.34	22.45	5.34	22.45
Residential Amenity	16.50	35.40	16.50	35.40	16.50	35.40
Drinking water Purification	0.24	13.10	0.24	13.10	0.24	13.10
Total	*88.06*	*183.34*	*162.81*	*246.24*	*107.08*	*222.12*

Table 21: Breakdown of Total Benefits by Conservation Practices and Ecosystem Services (Carbon Benefits are Global)

Categories	NCS1		NCS3		NCS8	
	Low	*High*	*Low*	*High*	*Low*	*High*
Wetland	40.46	72.29	20.53	36.68	47.21	84.34
Cover Crops	21.95	32.13	34.75	50.87		
Buffers					2.26	3.58
Tillage					31.99	54.60
Land Retirement						
MRTN	0.00	0.00	0.00	0.00	0.00	0.00
Reduced Soil Erosion	*62.41*	*104.42*	*55.29*	*87.55*	*81.46*	*142.52*
Wetland	2.87	7.16	1.46	3.63	3.35	8.36
Cover Crops	0.00	0.00	0.00	0.00		
Buffers					0.00	0.00
Tillage					0.00	0.00
Land Retirement			82.89	82.89		
MRTN	0.00	0.00	0.00	0.00	0.00	0.00
Recreation and Wildlife	*2.87*	*7.16*	*84.35*	*86.53*	*3.35*	*8.36*
Wetland	23.23	23.41	27.10	27.31	27.10	27.31
Cover Crops	1,501.28	1,508.85	2,377.03	2,389.01		
Buffers					25.55	25.80
Tillage					273.86	274.35
Land Retirement			53.61	53.77	0.00	0.00
MRTN	83.41	353.19	83.41	353.19	83.41	353.19
Total Carbon Benefit	*1,607.92*	*1,885.45*	*2,541.15*	*2,823.28*	*409.92*	*680.65*
Benefits from 40% reduction in Nitrate N and Phosphorus in Iowa's Waterbodies						
Water Based Recreation	5.34	22.45	5.34	22.45	5.34	22.45
Residential Amenity	16.50	35.40	16.50	35.40	16.50	35.40
Drinking water Purification	0.24	13.10	0.24	13.10	0.24	13.10
Total	*1,695.28*	*2,067.98*	*2,702.86*	*3,068.32*	*516.81*	*902.47*