

Estimating the Relationships among Education, Cognitive Ability, and Religion: Evidence from NLSY*

Han-Yen Kao[†]

Rutgers University

June 24, 2016

Abstract

This paper estimates the causal effect of education on religiosity in the United States using NLSY97. Fixed effects and instrumental variable method are used as identification strategies. Although cross-sectional ordinary least squares estimation shows a positive correlation between religious outcomes and educational attainment, both fixed effect models and IV estimation show statistically significant negative effects of education, even when cognitive test score is controlled. This suggests that conventional OLS omits factors that push both education and religiosity. (*JEL*: C23, C26, I29, Z12)

* I am most grateful to Anne Morrison Piehl for her guidance. I thank Costanza Biavaschi, Ira Gang, Zhutong Gu, Mark Killingsworth, Roger Klein, Carolyn Moehling, Ritu Sapra, Hilary Sigman, and the audience of the Empirical Micro Workshop at Rutgers University for helpful comments and suggestions. I am also in debt to the participants of the IRES 3rd Annual Graduate Student Workshop and the 14th Annual Conference of ASREC, especially my discussants Andrea Molle, Douglas Norton, and Sriya Iyer. All errors are my own.

[†] Department of Economics, Rutgers University – New Brunswick. 75 Hamilton St., New Brunswick, NJ 08901 (e-mail: hykao@econ.rutgers.edu).

1 Introduction

Education is the most central element of human capital in modern economy. Scholars have examined the private and social returns to education (e.g., Moretti, 2004). It has been shown to be related to better labor market outcomes (Card, 1999) and a number of nonproduction benefits, such as better health (Silles, 2009), less crime behavior (Lochner and Moretti, 2004), and additional positive outcomes (Lochner, 2011). Research establishing causal impacts of education has used multiple approaches to solve endogeneity concerns, including various instrument variable strategies, controlling family background or using family background as instrument, and estimates exploiting outcomes of twins (Card, 1999). Yet, there is little empirical research on the effect of education on religious outcomes.

Recently, economists have linked religion to several economic outcomes.² For example, Gruber (2005) finds evidence that higher level of religious participation is related to higher education and income, lower disability and welfare receipt, better health and marital outcomes. Studies have shown that religiosity has an effect on individual criminal behavior and crime rates,³ subjective well-being and happiness (e.g., Soydemir et al., 2004), suicide rate (e.g., Huang, 1996), altruistic behavior (e.g., Hrungr, 2004), several health outcomes (e.g., Deaton, 2011) and more.⁴

² See Chen and Hungerman (2014) and Iyer (2016) for the most recent introductions of this fast-growing field of research in the discipline of economics.

³ Baier and Wright (2001) provide evidence of deterrent effect of religion (religious beliefs and behaviors) on criminal behavior. On the contrary, Heaton (2006) uses past religiosity as instrument and finds no effect of religion on crime rates at the community level.

⁴ See Johnson, Tompkins, and Webb (2002) for a review of the relationships between religion and several health outcomes and well-being outcomes.

There are several possible mechanisms through which educational attainment could have causal effects on individual's religiosity, which is most commonly measured by religious attendance and self-reported religious beliefs. First, for attendance, education increases wages, and thus increases the opportunity cost of time spent on religious activities. Contrarily, Glaeser and Sacerdote (2008) suggest that the benefit from social connection, which includes church attendance, increases with education and thus provides educated people more incentive to participate in religious activities. Second, for religious beliefs, while the content of secular education may directly decrease religious belief, education may allow deeper understanding of the religious teachings and help people enhance their religious beliefs.

The economic study of the interaction of these two topics is relatively new, but long before the recent interest of economists, rather than education, psychologists have focused on the relationship between religiosity and intelligence.⁵ They find a consensus that religiosity is negatively related to intelligence (e.g., Lynn, Harvey, and Nyborg (2009) and Ganzach et al. (2013)). Recently, on the relationship between education and religiosity, psychologists Ganzach et al. (2013) show that when parent's religiosity (measured by parents' religious attendance) is high, children's educational attainment is associated with higher religiosity; when parent's religiosity is low, children's education is associated with lower religiosity. However, since they do not consider endogeneity of education in their models, the effects they suggest are not necessarily causal.

⁵ For example, Howells (1928) and Sinclair (1928) proposed negative relationship between religiosity and intelligence.

Evidence of causal effects of education on religion in the economic literature is growing but still limited.⁶ Cross-sectional data in the United States reveals a positive relationship between schooling years and measures of religiosity. In his ground-breaking introduction to the economics of religion, Iannaccone (1998) uses the General Social Survey (GSS) to investigate numerous theories on economics of religion. He finds a positive correlation between schooling years and church attendance. He also finds the more educated contribute more money in churches, perhaps unsurprising given the high correlation between education and income. Glaeser and Sacerdote (2008) show that people with higher education go to church more frequently overall, but the denominations with higher attendance rates have lower educated members. Arias-Vazquez (2012) shows that attending college attendance decreases one's religiosity (measured by self-reported religious attendance and importance) by estimating the individual fixed effects from the longitudinal Monitoring the Future (MTF) data. He also uses compulsory schooling laws and child labor laws to instrument for education in GSS data and shows similar results. Hungerman (2014) uses Canadian compulsory schooling laws as instrument for education, and finds that an additional year of schooling raises the probability of an individual reporting non-affiliated to any religious group by about 4.5 percentage points. In addition, Gervais and Norenzayan (2012) indicate that analytical thinking, which presumably grows with education, encourages religious disbelief.

An important concern is that religion may as well influence educational attainment. Gruber (2005) uses GSS and the 1990 census data to analyze the effect of religious market density, which is defined by “the share of the population in an area that

⁶ See Hungerman (2014) for a discussion of the economic literature on the educational effect on religion.

is of an individual's religion," on religious participation and economic outcomes. He uses area ancestral mix as an exogenous source of variation of religious preference. His findings suggest that higher density of religion market leads to higher religious attendance as well as higher education and other economic outcomes.

Empirically, the main challenge of this estimation is the endogeneity of education. It is possible that certain unobserved individual characteristics make people pursue more education and more (or less) religiosity at the same time. On the one hand, for example, Hungerman (2014) notes that tastes of civil participation enhance both attendance in school and religious activities. Moreover, people who prefer metaphysical thinking may also pursue more education and more religion. On the other hand, the findings of Gervais and Norenzayan (2012) imply people who tend to think analytically might pursue more education but less religion. In the former case, OLS is expected to overestimate the effect of education on religiosity. In the latter case, OLS underestimates its effect.

In this paper, I use two identification strategies, individual fixed effect models and instrumental variable method, to deal with the endogeneity of education as an explanatory variable of religiosity. First, I allow for individual fixed effects in the model to deal with endogeneity by removing the time-invariant heterogeneity. In addition to showing the effect of education on measures of religiosity, I show that controlling for parents' religiosity in the fixed effects model matters to the results. One worry is that fixed effects model may ignore the confounders that are not constant over time. For this reason, I develop alternative instrumental variable strategies to compensate this caveat. I adopt various measures of college accessibility as an exogenous source of variation. This

identification strategy relies on the assumption that the accessibility of college from an individual's residence is not related to his religiosity. More details are described in paper.

I use the confidential geocode data of 1997 National Longitudinal Survey of Youth (NLSY97 hereafter). The NLSY97 has the great benefits of questions on both religious beliefs and attendance, allowing some insight into possible mechanisms. Also, lots of control variables allowing psychological model of intelligence are paramount. I consider a measure of cognitive ability as another channel through which education may affect religiosity, measured by the score of Armed Force Qualifying Test (AFQT hereafter). Finally, there are sufficient observations to allow adopting IV strategy using geographic information. To do this, I apply for the confidential NLSY97 Geocode Data in order to obtain individuals' residence information. This allows me to calculate the college accessibility at the county level.

The structure of the rest of the paper is as follows. In section 2, I describe the NLSY97 data and the variables I use to measure the factors of interest. Section 3 shows the models of fixed effects. Section 4 provides empirical results. I show the pooled OLS and the panel regression results, along with robustness checks. Section 5 discusses the alternative identification strategy of instrument variables. Section 6 concludes.

2 Data

2.1 Overview

To conduct this estimation, three measures of the key variables are required: measures of religiosity, measures of education, and the measure of cognitive ability. The

NLSY97 includes rich set of questions regarding these aspects. Here in this paper, I include the most recent Round 15 of NLSY97, which has not been analyzed on this topic.⁷

It is necessary to define the term “religiosity” used throughout this paper before proceeding. While the sociological definition of religiosity consists of broad components,⁸ most papers in this education-religion literature measure religiosity by religious attendance and religious beliefs. At the same time, most surveys employed in this literature provide questions about attendance and self-reported beliefs. Although they do not capture every aspects of religiosity in its sociological definition, they are concrete, observable, and measurable. By this definition, they are not measurements of one’s spirituality and godliness, but measurements of how much an individual’s life style and ideology are related to religion.

The NLSY97 is a nationally representative survey started in 1997. It consists of 8,984 youths who were 12 to 17 years old when they were first interviewed in 1997. The longitudinal survey takes a yearly basis. Round 15 was conducted in 2011 and was released in August 2013. The survey contains various topics about the youth, including economic, social, psychological questions, and provides information about the individuals’ family background.

To measure educational attainment, I use the number of full-time schooling years completed as the main independent variable. College attendance is the other variable of interest. Arias-Vazquez (2012) uses MTF data to show that college attendance has a

⁷ The most recent data release, round 16, was released in 2016. It does not contain the questions regarding religious activities and beliefs. Therefore, round 15 is the newest data available on this topic to date.

⁸ For example, according to Cornwall et al. (1986), religiosity consists of cognitive component, affective component, and the behavioral component.

negative effect on religiosity. I check whether the same effect happens in NLSY97 sample by the models using highest degree as an alternative explanatory variable.

2.2 Measuring religiosity

Throughout this paper, the general level of religiosity includes two distinct elements: the frequency of religious attendance and an index of religious beliefs. Religious attendance captures how much time people spend on the most identifiable religious activity. The index of religious beliefs constructed from questions in the NLSY97 explicitly reveals the individuals' ideology about religion.

2.2.1 Religious attendance days

Religious participation is the most widely adopted form of religious practice. From 2000 to the most recent round, NLSY97 reports the respondents' religious attendance by asking "how often did you attend worship service in past 12 months."⁹ The answer is categorical and ranges from 1 (Never) to 8 (Everyday). According to this categorical response of religious attendance, I generate a continuous variable representing religious attendance days in the past year.^{10 11}

⁹ In the questionnaire, this variable describes the church attendance days in the last period, $t - 1$. In this paper, I run the regression of period t on this last period church attendance variable. The reason is that the educational attainment, highest grade "completed" or highest degree "received," are also describing the accumulative levels that have occurred in the past. In fact, even if I use the right-hand-side variables in period t to explain the church attendance in period t , the results do not change.

¹⁰ I use the following rule to construct the church attendance days. To avoid over-estimation, I choose the lower bound number when imputation is required. To be specific, those who report "never" are assigned zero days of church attendance. "Once or twice" is assigned one day per year. "Less than once a month (3-12 times)" is assigned three days. "About once a month" is assigned 12 days. "About twice a month" is assigned 24 days. "About once a week" is assigned 52 days. "Several times a week" is assigned 104 days. "Everyday" is assigned 365 days in a year.

Table 1 shows the percentage of religious attendance over the years using the sample of this paper.¹² As shown in Table 1, compared to 2000, nearly half of the respondents decreased religious attendance in 2011; about only one fifth increased attendance since 2000. The trends of each frequency of attendance are shown in Figure 1. From Table 1 and Figure 1, the number of people attending worship services for zero times increased the most (from 23% in 2000 to 35% in 2011); the number of people attending services for about once a week decreased the most (from 21% in 2000 to 12% in 2011).

2.2.2 Index of religious beliefs

I generate an index of religious beliefs from the responses to questions about religion. In 2002, 2005, 2008, and 2011, the NLSY97 records the responses to questions regarding attitudes to religion. There are five true-false questions regarding religious beliefs: 1. “I do not need religion to have good values (reverse coded);” 2. “Religious teachings are to be obeyed exactly as written;” 3. “I often ask God to help me make decisions;” 4. “God has nothing to do with what happens to me personally (reverse coded);” 5. “I pray more than once a day.”¹³ In each question, answer 1 means true and answer 0 means false. I conduct principle component factor analysis and use the correlation matrix to construct a weighted index of religious beliefs with mean zero and

¹¹ Since the dependent variable is categorical, an alternative is to conduct ordered probit estimation.

¹² The construction of the sample is described in section 2.4.

¹³ One may consider prayer as a religious practice instead of religious belief. I include the response to this question for two reasons. First, it is consistent with the variable of parents’ religiosity (described later). Second, factor analysis shows that “I pray more than once a day” accounts for a considerable variation of the index. In fact, when I run all the regressions excluding this question, the results do not change by this exclusion.

standard deviation one.¹⁴ Table 2 shows the percentage of religious answers over the years. In 2011, more than one-third of the sample reported decreasing religious belief since 2002, and only one-fourth of the sample reported increasing.

The religious practices and beliefs of youth could be highly affected by the family religious background. To control for family religious background, I use a variable that includes the responses to the same true-false questions from the parents of the youths reported in 1997.¹⁵ These questions were answered only by parents of the youths who were 14 years old or younger as of December 31, 1996. Therefore, in order to control for family religious background of the sample individuals, the sample is restricted to cover the cohorts of those who were 14 years old or younger before the first day of 1997. This restriction does not cause selection problem. The parents' religiosity variable is standardized to mean of zero and standard deviation of one.

2.3 Measuring cognitive ability

Cognitive ability is measured by the AFQT score,¹⁶ coming from the computer-adaptive form of Armed Services Vocational Aptitude Battery (CAT-ASVAB), which

¹⁴ Alternatively, I construct another unit weight index by simply adding up five answers (ranging from 0 to 5.) The results from both indices do not differ from each other.

¹⁵ This variable includes another binary-response question: "In the past 12 months, how often have you attended a worship service (like church or synagogue service or mass)?" "1" represents more than once or more and "0" otherwise. Therefore, this variable reveals parental background about both religious beliefs and practices. In fact, when family religious background is divided into religious attendance and religious beliefs and used to control corresponding religious variables, the results are similar and do not change the conclusion of this paper.

¹⁶ AFQT score is a percentile score coming from a weighted average of mathematical knowledge (MK), arithmetic reasoning (AR), word knowledge (WK), and paragraph comprehension (PC) in the ASVAB. In addition to these four areas, ASVAB also includes general science (GS), electronics information (EI), automotive and shop information (AS), mechanical comprehension (MC), and assembling objects (AO). These additional fields are closer to skills of the test takers and thus not included in the AFQT.

was conducted in the first round of the NLSY97. The AFQT scores are standardized to a variable with mean of zero and standard deviation of one.

Although the psychological debate is not the focus of this paper,¹⁷ there is room for discussion about the essential role of such IQ-like score (e.g., Heckman, 1995). While psychologists' initial focus is on the general inherited and less immutable intelligence, or the so-called "g", the measured cognitive ability is believed to grow with education. Therefore, the AFQT score itself should be better viewed as a measure of educational achievement (e.g., Cawley et al., 1997; Winship and Korenman, 1997; Hansen et al., 2004).

However, the AFQT score is still frequently employed as a proxy in social-science literatures to represent individual's cognitive ability, aptitude, or intelligence.¹⁸ Winship and Korenman (1997) states that "both education and early IQ have important effects on later intelligence as measured by the AFQT."¹⁹ In addition, according to Cawley et al. (1997), the AFQT and "g" are highly correlated. In this paper, the correlation coefficient between 2011 schooling years and the AFQT score is 0.59.²⁰ It preliminarily reveals that the AFQT score is not an inappropriate measure of cognitive ability.

¹⁷ There was a debate on the correlation between intelligence, race, and achievement after the book "*The Bell Curve*" by Herrnstein and Murray was published in 1994.

¹⁸ For example, treating the AFQT score as ability, Carneiro and Heckman (2002) control for ability to estimate the relationship between family income and college attendance. Altonji et al. (2012) compare cognitive skills, represented by the AFQT scores, across the cohorts in NLSY79 (1979 National Longitudinal Survey of Youth) and NLSY97.

¹⁹ Page 231 in Winship and Korenman (1997).

²⁰ For the full sample in the NLSY97, the correlation coefficient is almost the same with the sample I construct in this paper.

In summary, as long as we do not overplay the interpretation of the AFQT, it provides a different angle from schooling years and highest degrees to take intellectual ability into account. We do see interesting difference from the results.

2.4 Sample description

This section describes how I construct the sample of this paper. What determines the sample size and the number of observation is the availability of the key variables. In order to account for the effect of the educational variables on the various religious variables, and to control for family religious background and other important demographic characteristics, missing values inevitably have effects on the sample size.

As described in section 2.1, the biggest drop of the sample size from the full NLSY97 sample is due to missing values of parents' religiosity, which results from the fact that only the parents of youth of 14 years old or younger have the responses to the questions about religion. It does not raise sample selection issue. The second largest drop is due to the missing values of the AFQT score. The relevant religious variables do not differ by inclusion or exclusion of AFQT scores.²¹

Table 3 summarizes the effects of the missing data on the sample size. It shows that I control for the religion-related variables, which include religious preference in 1997,

²¹ Considering the representativeness of the NLSY, Altonji et al. (2012) discuss the attrition and missing data in NLSY79 and NLSY97. They find small difference between the full sample and the stayers in certain observable characteristics according to attrition. They show that some of the observable variables for those who with AFQT scores differ from those who without AFQT scores. However, since the observable results of AFQT takers and the full sample have similar features, the differences are acceptable (see Altonji et al. (2012), Table B2 and Table B3 in Appendix B). Moreover, they provide evidence that the joint distributions of the AFQT score and some observable variables do not depend on the age at which the individuals took the test (see Altonji et al. (2012), Table B4 in Appendix B). In this paper, since the cohorts included are smaller than the full NLSY97 sample, the problem of test-taking age would be even smaller.

parents' religiosity, as well as demographic characteristics including race, marital status, living in urban area or not, and having children or not. The final sample size in this paper is 4,480 and the sample accounts for 49.9 percent of the full NLSY97 sample. While the final sample represents just fewer than 50 percent of the overall sample of the NLSY97, it is 83 percent of those who were eligible, i.e., 14 years old and younger as of 12/31/1996, for the relevant bank of questions.²²

As for the number of observations, when running religious attendance regression, there are observations of 12 years, which include 45,439 observations. The numbers of observations per individual is 10.1. For the index of religious beliefs, since the questions about religion were asked in 2002, 2005, 2008, and 2011, the observation years are restricted to these four years. The total number of observations is 13,371. The number of observations per individual is 3.0.

2.5 Descriptive Statistics

Table 4a shows the 2011 cross-sectional descriptive statistics of the key educational and religious variables. I report the means and standard deviations of religious attendance days, the index of religious beliefs, the parents' religiosity index (reported in 1997), highest grades completed, and the AFQT scores (reported in 1997). The average schooling years in 2011 is 13.8 years. About 36 percent of the sample have bachelor's degree or above. Religious attendance is on average 15.5 days per year or 1.3 days per month in the past year.

²² In fact, not all people in my sample were 14 or less in the first round. First, it is natural that some people who were 14 years old in 1996 became 15 in 1997. Second, in the constructed sample, there are a small part of people aged 16 or 17 in round 1 of the interview, which should have been taken in 1997. This was because some of the round 1 interview was taken in 1998 due to administrative reasons.

Table 4b shows the descriptive statistics of the control variables of the constructed sample. Age, gender, marital status, childbearing, living area (urban area or not), and religious preference in 1997 are reported. The mean age of the sample in 2011 is 28.3 years old. Forty-three percent of the sample are married, and 45 percent of the sample have children. Seventy-eight percent of the sample live in urban areas. Sixty-one percent of the sample are whites. As for their early religious preference, Roman Catholic and Protestants of all denominations account for 85 percent of the sample.

3 Econometric Model

3.1 Identification strategy: individual fixed effects

To estimate the effect of education on religiosity, the key challenge is endogeneity of education. In particular, there might be factors affecting both education and religiosity. Some of these factors are observable. For example, as Arias-Vazquez (2012) points out, children from religious families may be more likely to stay in high schools and thus are more likely to go to college, since religious families might be better at preventing high school children from dropping out. A solution to this problem is to control for parents' religiosity.

There might also be some unobserved individual characteristics affecting both education and religiosity. For instance, people with a preference for metaphysical thinking might pursue more education and more religion. When a model ignores such a confounding factor, the effect of education on religion will be upward biased. On the contrary, Gervais and Norenzayan (2012) and Mocan and Pogorelova (2014) argue that

people who tend to think analytically might pursue more education and less religion. If their argument is correct, the effect of education on religion will be downward biased. Therefore, the expected sign of the bias is ambiguous.

Assuming the unobserved variables are constant over time, the longitudinal property of the panel data allows me to estimate a fixed effects model. In particular, I assume the unobserved individual characteristics affecting education and religiosity are included in time-invariant individual heterogeneity term ϕ_i in the following model:

$$R_{i,t} = \alpha Edu_{i,t} + \beta X_{i,t} + \phi_i + \tau_t + \epsilon_{i,t}. \quad (1)$$

For individual i in survey year t , the dependent variable $R_{i,t}$ is religious attendance days or the index of religious beliefs, depending on specification. Both religious attendance days and the index of religious beliefs are treated as continuous variables. $Edu_{i,t}$ is educational attainment, namely, highest grade completed or highest degree received. $X_{i,t}$ is a set of individual-level control variables including age, gender, race, marital status, childbearing, religious preference, rural or urban area dummy, and parents' religiosity. τ_t is year effect. ϕ_i is the time invariant heterogeneity. Because of the identification assumption that the unobserved characteristics ϕ_i is time-invariant for each person i , it has no subscript t . That is, the confounder of the effect of education on religiosity is constant over the years. $\epsilon_{i,t}$ is the error term.^{23,24}

²³ In fact, a cohort effect C_c could be included in the model. However, since the sample used in this paper only consists of three birth-year cohort, the cohort effect is expected to be minor. Therefore, I only include age in $X_{i,t}$ as a control variable.

²⁴ Ideally, adding an individual-specific trend, $g_i * t$, to account for the source of heterogeneity would help to gain the results that are closer to causal effects. It is possible that there is a systematic difference of religious trend between people with different education level. However, since the highest grade completed increases one year per year of age for most people, the individual specific trend would only identify non-traditional students whose age and schooling year do not increase at the same time, which is a particular group and generates concerns about

For those control variables that are time variant, they are included in fixed effects models. For those time invariant ones, they are only included in pooled OLS model, which is shown in section 4.1 as a benchmark. In the fixed effects models, selected time-invariant control variables are used to divide the sample into different set of subsamples. Specifically, I divide the sample by parent's religiosity, cognitive ability, and gender, and run the fixed effects models separately. The results of subsample fixed effects model are shown in section 4.3.

3.2 Long change fixed effects estimation

In most cases, the highest grade completed and the individuals' age increase one-for-one simultaneously, at least up to some point when they finish education. One might worry that this mechanical manner in fixed effects model would reduce or limit the variation that can be identified. Therefore, I deal with this one-year-by-one-year relationship between highest grade completed and age in two ways.

First, I include highest degree received as an alternative explanatory variable. Unlike highest grades, the time interval between one's degrees and age is less prone to have one-to-one relationship.

Second, to check whether this relationship cause any problems on the estimation, I run the same fixed effects model with only the observations in the first and the last years. We expect the results to be sharper than what is in the model with full observations, as the changes of the religious variables and the educational years are greater. The discussion of the results is shown in section 4.2.3.

sample selection and external validity. Therefore, I choose not to include this $g_i * t$ term in this paper.

4 Empirical Results

4.1 Benchmark results: pooled OLS

First of all, I examine the general correlations between the measures of religiosity and education without dealing with endogeneity. By running the conventional pooled OLS, Table 5 shows the relationship between religious attendance days and educational variables, including highest grade completed, highest degree received, and the AFQT score. Table 6 runs the same model with the dependent variable being the index of religious beliefs.

Table 5 confirms the positive relationship between religious participation and schooling at the cross-sectional level, as appears in Iannaccone (1998) and Gruber (2005). Religious attendance days per year is positively and significantly correlated with highest grade completed and highest degree received. The AFQT score also has positive and significant relationship with religious attendance. Not surprisingly, parents' religiosity has a strong and significant positive impact on children's church attendance. Controlling for parents' religiosity (model (iv) to (vi)), the relationship between educational attainment and cognitive ability becomes a little smaller but still positive and significant. Moreover, it can also be seen that married people and female goes to church more than non-married and male, respectively. Living in urban area is generally negatively related to church attendance. Also, church attendance days decreases with age. The ones who have children appear to have less church attendance days.

In Table 6, the dependent variable is the index of religious beliefs. What is interesting is that, although the relationship between religious beliefs and highest degrees are mixed and somewhat less significant, higher AFQT score is significantly associated with lower index of religious beliefs, even when parent's religiosity is controlled. This relationship confirms the findings of most psychological research. Again, married people and female has higher degree of religious beliefs than non-married and male. People live in urban areas have lower level of religious beliefs than those who live in rural areas. The index lowers with age. The ones who have children appear to have slightly higher religious beliefs index.

In models (vii) and (viii) of Table 5 and Table 6, the AFQT score is included as well as educational attainment as an explanatory variable. In both cases, highest grade completed is still significant. The AFQT becomes less significant in Table 5 while educational attainment is still significant. Highest degree becomes less significant in Table 6 while AFQT and highest grade completed are still significant. It gives an idea that although without causal links, cognitive ability and educational attainment have separate relationships with religiosity.

4.2 Panel models

In model (1), running pooled OLS means that we ignore individual effect ϕ_i , or assume that the individual effect ϕ_i is uncorrelated with the explanatory variables $Edu_{i,t}$ and $X_{i,t}$. However, as discussed before, we worry that there would be some omitted factors affecting individual's religiosity. Fixed effects analysis allows us to consider individual time-invariant heterogeneity ϕ_i , and gives us a consistent estimation.

4.2.1 Fixed effects estimation

Table 7 shows the results of the fixed effects estimation. In column (i), the highest grade completed has a negative and significant effect on church attendance days. This is consistent with recent research which suggests negative causal effect of educational attainment on religious attendance days. Each schooling year makes an individual decrease church attendance per year by 1.15 days. With the average highest grade completed being 13.76, on average, it is equivalent to 15.8 days less of religious attendance per year after one completes education.

Column (ii) uses the highest degree received to explain the variation in religiosity. Similar to the effect of highest grade completed, the effect of highest degree received on church attendance is opposite to pooled OLS results, and shows that the higher degree received, the less often the youths attend worship services. Compares to the reference group, high school graduates, college attendance makes people decrease about two days less attending religious service. The difference between none degree and high school graduate is even larger. It reports more than seven days less per year after one gets high school degree.

The effects of educational attainment on the index of religious beliefs are shown in column (iii) and (iv). For highest grade completed, each schooling year causes an individual to decrease religious-belief index by 0.019. It is equivalent to an average decrease of religiosity index by a quarter of standard deviation after one completes education. There is also a negative effect on the individuals' index of religious beliefs for

each higher degree received. College attendance decreases people's religious belief index by 0.06 standard deviation.

4.2.2 Discussion

The results of fixed effects estimation tell us that the OLS overestimates the effect of educational attainment on religious attendance and religious beliefs. In particular, assuming endogeneity of education is a result of omitted variable bias which is fully captured by the individual fixed effects, this implies that the OLS biases the estimate upward. The results of pooled OLS on the index of religious beliefs are already negative. Compare to OLS, the fixed effects estimates on the index are even smaller. Therefore, the OLS estimates on the index of religious beliefs are again biased upward.

The direction of the bias of the OLS estimation suggests that the effects of unobserved individual characteristics on education and on church attendance are of the same direction. That is, if the omitted variable increases education, it increases measures of religiosity as well. As for religious attendance, a possible explanation is the benefit from social connection increases with educational attainment and in the form of religious attendance (Glaeser and Sacerdote (2008)), or the tastes of civic participation (Hungerman (2014)). As for religious beliefs, it could be the individual's taste of metaphysical thinking or certain subjects such as philosophy and theology.

4.2.3 Long period change model

Table 8 reports the results of the same models on the observations of only the beginning year and the last year available. The source of variation here is from the long-

period changes in religious attendance and beliefs, and the long-period changes in educational attainment. In other words, the immediate responses in these behaviors are not required in this model to identify the effects. If the responses on religious variables for some reason take some time to reflect on the effect of education, then this model would capture the effects more precisely.

In Table 8, almost all the coefficients of the educational attainment variables are greater than the ones in the original fixed effects model shown in Table 7, which includes the observations in all available years. This alleviates the worry that the identification could only come from a peculiar group. Therefore, it further reassures the idea that in general the educational attainment reduces religious attendance and beliefs.

4.3 Time-invariant control variables

In the fixed effects models, some important time-invariant control variables are not applicable. Therefore, the sample is divided into subsamples in various ways to show how these constant control variables have impacts on the estimation. I run the fixed effects models on the sample with stronger family religious background versus weaker background, using the median of parents' religiosity as the threshold. Similarly, higher cognitive ability and lower cognitive ability are defined by above and below the median of the AFQT score. I also include regressions on only male and female. Table 9a shows the effects by subsamples on church attendance. Table 9b shows the effects by subsamples on the index of religious beliefs. In these cases, the other demographic

control variables do not account for the differences we see in the effects of educational attainment.²⁵

4.3.1 Subsample by family religious background

When dividing the sample by the family religious background, it shows that the negative effects of educational attainment on religious participation and religious beliefs are much stronger for those who come from families that are more religious.

For religious attendance, according to column (ii) and column (iii) in Table 9a, the difference between the effects on people from stronger and weaker family religious background is about 1 day for each additional schooling year, or 13.8 days after they complete education. Similarly, the highest degree received also sees difference by religious background. The coefficients of highest degree received on people from weaker religious families are smaller and insignificant. For the index of religious beliefs, as shown in column (ii) and column (iii) of Table 9b, the difference of the effects on the index has a similar pattern.

In summary, educational attainment influences people's attitudes toward religion. The effect is stronger for people from more religious families. This result does not support the findings in Ganzach et al. (2013), which argues that when the religious background is strong, education increases individuals' religiosity, and vice versa.

²⁵ In fact, as shown in Table 7, most control variables are not significant in fixed effect models. The reason might be that the periods are not long enough to account for enough variation to show their influences. The control variables tend to have higher significance when I run the subsamples of which the educational attainment has lower explanatory power. For instance, in model (iii) of Table 9a, the control variables have higher *p*-values, while in Table 7 they are insignificant as in the full sample. However, when they are significant, the signs are consistent with the full sample model. Thus we do not consider the differences come from the control variables.

4.3.2 Subsample by AFQT scores

In column (iv) and column (v) of Table 9a and 9b, the sample is divided into subsamples by their AFQT scores. In Table 9a, the effects of the highest grade completed on church attendance for higher and lower AFQT scorers are both significant, and the effect for people with lower AFQT scores is somewhat larger, but the difference is not as large as it is when divided by religious background. As for the highest degree, the effect of college attendance on church attendance is higher for people with higher AFQT scores, while the effect for people with lower AFQT scores is less precise. However, none of them are larger than the effect of college attendance on church attendance for the full sample.

In Table 9b, on the contrary, the highest grade has slightly larger and more significant negative effect on religious beliefs for people with higher AFQT scores, while the coefficients of college attendance along with other degrees are larger (in absolute values) for the low AFQT group.

In short, the subsample analysis by the AFQT scores is less consistent and less informative than by family religious background. It suggests that if cognitive ability has a causal effect on religiosity, what is shown in the fixed effect model is not enough to correct the omitted variable bias from OLS. Therefore, an obvious causal link between cognitive ability and religiosity is not found. Again, it does not support the findings in Ganzach et al. (2013), which suggests a negative direct effect of intelligence on religiosity.

4.3.3 Subsample by gender

Subsample by gender shows that there is a clear distinction between the effects on male and female. The negative effects of educational attainment on religiosity are stronger for male. Each higher grade completed causes male participate about 0.3 days per year less than female, or 4.1 days per year less than female after graduation. Effects of education on religious beliefs are also greater for men. This is consistent with the fact shown in OLS estimation, in which the coefficients of male dummy are always negative.

5 Alternative Identification: Instrument Variables

If the omitted variable bias cannot be completely erased by the individual fixed effects due to unobserved variables that are not constant over time, the fixed effects method may not be able to fully capture the causal effect of education. Therefore, I propose using various measures of accessibility of college as instrument variables to estimate the effect of education on religious outcomes. To begin with, I will discuss some efforts done in recent years.

5.1 Literature review and candidate instruments

First, Hungerman (2014), Arias-Vazquez (2012), and Mocan and Pogorelova (2014) use compulsory schooling laws to instrument education. All of them find negative effect of education on religion and find OLS biased upward. However, LATE (local average treatment effect) interpretation implies that the estimation can only apply to people with education under high school level. The limit of using compulsory schooling

laws as instrument is that the effect of higher education on religiosity, which is the interest of many, is not captured.

Second, Kortt et al. (2012) use parents' education as IV of education. They find positive relationship between education and religious attendance under OLS estimation, but find no causal relationship once they treat education as endogenous and use parents' education as instrument. However, since parents' own education has an impact on their own religion, which further affects youths' religiosity, parents' religiosity might not be exogenous enough.

Another candidate instrument for education is college proximity. Addressing the estimation of returns to schooling, Card (1995) argues that the differences in the accessibility of college could be a source of variation that is exogenous to the labor market outcomes. The rationale behind this method is that the cost of transportation affects the opportunity cost of getting education, especially for people from low income families. The identification assumption is thus college proximity influences labor market earnings only through its impact on the individuals' education decision. He also uses the interaction term of college proximity and family educational background as an alternative instrument. Both instruments he employs are shown to be effective to correct the bias of OLS when estimating the effect of education on labor market outcomes.

5.2 Instrument variable: college accessibility

In order to capture the effect of higher education on religiosity, I adopt a similar instrument variable to Card's college proximity – college accessibility. Below I describe how I construct the two measures of college accessibility.

With the county-level NLSY97 Geocode data, I construct two measures of college accessibility. First, compare to counties with fewer total seats available (relative to total young population), counties with more seats available in colleges provide more accessibility of education to the potential students. Using the 2000 census data, I calculate the ratio of number of seats (enrollment) in colleges relative to the population from 18 to 29 years old in the individual's residential county.

Second, in order to consider the variation of geographical education premium, I consider the percentage of people with bachelor's degree or higher in a county. Because the percentage of people go to college could still be endogenous to educational attainment, I use this variable dated back to 1990 to alleviate the worry. The rationale is that a certain part of difference in educational attainment comes from geographical difference. Recognizing this cannot remove endogeneity entirely, I view this instrument as a last attempt.

Table 10 lists the potential instrument variables with their descriptive statistics. The last column displays the first-stage coefficients of the instruments in models for schooling years, controlling for other individual characteristics. Both candidate IV's seem to have direct impact on the educational attainment of the individuals in the sample.

The exclusion restriction requirement of college accessibility depends on whether the accessibility of college in the residential area is uncorrelated to its religious outcomes. There are possibilities that this exclusion restriction may not hold. One is the possibility that people of certain religious level have selected to live in areas that have more colleges.²⁷ If families that care more about education select to live in counties that

²⁷ I thank Douglas Norton for bringing up this issue. Another one is secular competition for religious attendance.

provide more seats in college, and if their emphasis on education cause them to have less religiosity, then children from these families might have higher education and lower religiosity at the same time. In this case, OLS biases the effect of education downward, and vice versa. Taking this caveat into account, Rouse (1995) argues that if the location decision is made by the parents, not the children, then the caveat is alleviated and the validity of this instrument cannot be ignored. The bottom line is the exogeneity needs to be recognized conditional on families' background and geographic effects.

A second caveat of this instrument is that religious institutes in counties with higher college accessibility may face more secular competition. Activities associated with college life, no matter on campus or off campus, may have an impact on the attendance rates of religious institutes. If this is true, when a county has more college students, the number of religious adherents might be lower.

If at least one of the caveats is true, then the number of college students should be correlated with the number of religious adherents. To check this, I use the 2000 county-level data of Religious Congregations and Membership Study from theARDA.com to match with the Census data. I check whether there is a significant correlation between number of college students and the number of adherents by county,²⁸ controlling for the population. There is no correlation between the two.²⁹ Therefore, it shows the accessibility of college has no direct effect on religious outcomes.

5.3 Model and results

²⁸ I adapt the "adjusted total number of adherents" computed by Finke and Scheitle (2005).

²⁹ The results of the regression are not shown in the paper and are available from the author upon request.

Given the college accessibility has no direct effect on religious outcomes, it can therefore be used to instrument for education to explain religiosity. The IV estimation model would be as follows:

$$R_i = \delta Edu_i + \gamma X_i + \mu_i, \quad (2)$$

$$Edu_i = \theta X_i + \pi C_i + \varepsilon_i. \quad (3)$$

Again, R_i is individual i 's religious outcome. Edu_i is i 's educational attainment. X_i is a vector of characteristics, not including i 's college accessibility, C_i . The principle of relevance holds when the coefficient π is not equal to zero, as shown in the last column of Table 10. Exclusion restriction holds when $E(C_i \varepsilon_i) = 0$. The satisfaction of exclusion restriction and relevance of college proximity guarantees that IV estimator generates unbiased estimate of δ .

I conduct the IV estimation using the most recent round (2011 data) of the NLSY97, since it has the most complete educational attainment. The results are shown in Table 11. Column (1) shows the coefficients of conventional OLS estimation. As in Table 5 and Table 6, conventional OLS shows positive relationship between schooling years and religious attendance, and the relationship between schooling years and the index of religious beliefs. The IV estimates, however, gives opposite estimation of the effects of educational attainment. In panel A, the effect of an additional schooling year appears to decrease religious attendance days by 3.82 to 4.86 days per year, but the coefficients are imprecise. In panel B, the IV estimation reveals that an additional year causes a decrease in the index of religious beliefs by 0.5 to 0.72 standard deviation. The IV estimation reverses the sign of the coefficient which is positive in OLS. Both estimates are significant.

The IV estimations are consistent with the findings of the fixed effects model. Like the fixed effects estimation, and IV estimations show that OLS upward biases the effect of educational attainment on religious outcomes. Both strategies provide evidences that the hidden factors which cause endogeneity drive educational attainment and religious outcomes to the same direction, and that educational attainment causes religiosity to decrease.

6 Conclusion

This paper provides additional evidence to the literature of the relationship between education and religion with cognitive ability considered, using a nationally representative dataset that is less exploited on this topic. Consistent with the findings of the literature using cross-sectional estimation on GSS data, the pooled OLS results show positive relationships between educational attainment and religiosity. In order to deal with endogeneity, I adopt fixed effects models and use two measures of college accessibility to instrument for education. The results suggest that the educational attainment has negative impacts on both religious attendance and beliefs. Cognitive ability, which is positively related to educational attainment, only has an imprecise negative link with religious beliefs, but not with religious attendance.

The fixed effects models suggest that each schooling year decreases 1.2 to 2.2 religious attendance days per year, or 15.8 to 29.7 days less after one completes education, and decreases the index of religious beliefs by 0.019 to 0.029, or 0.25 to 0.4 standard deviations after graduation. These results are especially stronger for those who

come from families with stronger religious background. IV estimation gives larger estimates. Each schooling year is associated with 3.8 to 4.8 less religious attendance days per year, and a decrease in the index of religious beliefs by 0.5 to 0.7 standard deviations. The fixed effects results are robust to the length of periods between observations.

At the cross-sectional level, cognitive ability represented by the AFQT scores has no significant relationship with religious attendance days, but interestingly has a strong negative relationship with religious beliefs. However, the fixed effects estimates of the effects of educational attainment on religiosity show little difference between high AFQT scorers and low AFQT scorers. Cognitive ability does not show meaningful impacts on religiosity in IV estimations either.

There are some thoughts for future research agenda. First, a branch of literature of economics of religion is based on structural modeling of time allocation. Azzi and Ehrenberg (1975) set a classic model of religious participation that is in the fashion of labor supply model and is followed by many theoretical and empirical works. When it comes to time allocation, taking educational background, income, occupation, and industry into consideration, the opportunity cost of church attendance can be modeled. Using the individual time allocation in micro-data to conduct further research is an interesting way to investigate this question.

Second, about the choice of model, if one believes that previous religious status determines later religiosity, then a lagged responses model would be a proper choice. However, the choice between fixed effects model and lagged response model requires strong assumption about how religiosity is determined. Even though one alternatively assumes it is past religiosity that determines current religious activity instead of

unobserved characteristics, we still need alternative identification strategy such as instrument variable to give an interval of the causal effect estimation.³⁰

Third, most papers in this literature look at American data. However, different countries have different educational systems and cultural backgrounds that convey various views of beliefs. It requires an overall international investigation to call the findings of this literature conclusive.³¹

Here I quote a sentence in Glaeser and Sacerdote (2008) to end this paper. It perhaps represents the deep incentive to study this topic. “The strong and interesting relationships between education and religion serve to remind us that human capital is not just about income and that investment in schooling has a wide range of interesting side effects beyond making people richer.”³²

³⁰ Arellano and Bond (1991) provide a GMM estimator for models include both fixed effect and lagged dependent variable.

³¹ A psychological paper Meisenberg et al. (2012) addresses this topic at the country level. They find negative relationships between education and religion, and conclude that the effects differ by country and by religion. The results they show are correlations, not causal effects.

³² Glaeser and Sacerdote (2008), pp. 189-190.

Reference

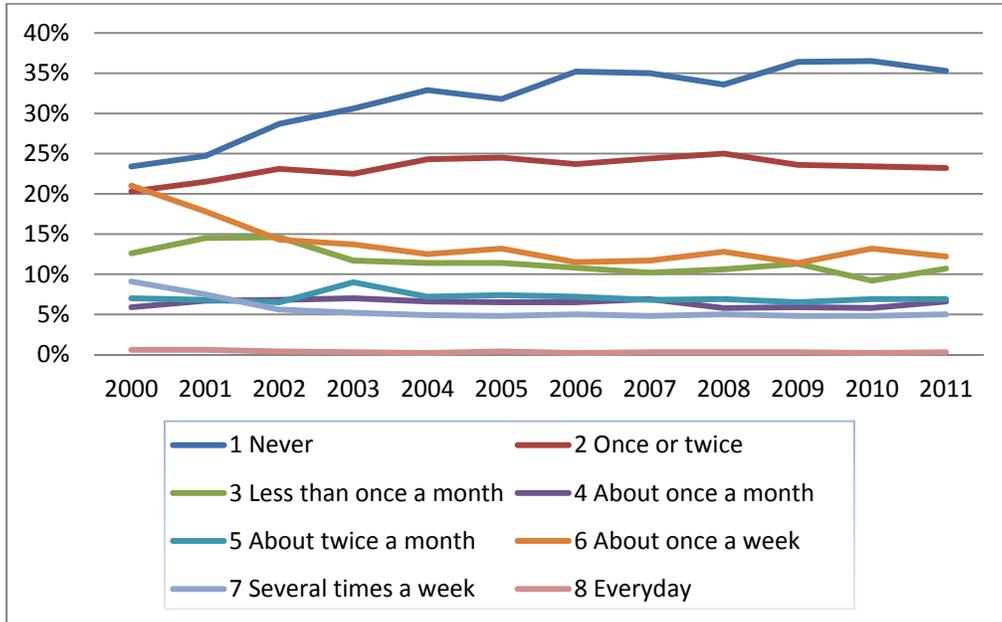
- Altonji, J.G., P. Bharadwaj, and F. Lange (2012). Changes in Characteristics of American Youth: Implications for Adult Outcomes. *Journal of Labor Economics*, 30(4), 783-828.
- Arellano, M., and S. Bond (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and An Application to Employment Equations. *Review of Economic Studies*, 58, 277-297.
- Arias-Vazquez, F.J. (2012). A Note on the Effect of Education on Religiosity. *Economics Letters*, 117(3), 895-897.
- Azzi, C., and R. Ehrenberg (1975). Household Allocation of Time and Church Attendance. *The Journal of Political Economy*, 83(1), 27-56.
- Baier, C.J. and B. Wright (2001). “If You Love Me, Keep My Commandments”: A Meta-analysis of the Effect of Religion on Crime. *Journal of Research in Crime and Delinquency*. 38(1), 3-21.
- Card, D. (1999). The Causal Effect of Education on Earnings, in *Handbook of Labor Economics*, 3A, 1801-1863, ed. by O. Ashenfelter and D. Card. Amsterdam and New York: North Holland.
- Carneiro, P. and J.J. Heckman (2002). The Evidence on Credit Constraints in Post-Secondary Schooling. *The Economic Journal*, 112(Oct.), 705-734.
- Cawley, J., K. Conneely, J.J. Heckman, and E. Vytlačil (1997). Cognitive Ability, Wages, and Meritocracy. B. Devlin, S. Fienberg, D. Resnick, and K. Roeder (Eds.), *Intelligence, Genes, and Success: Scientists Respond to The Bell Curve*, New York: Copernicus Press, 179–192.

- Chen, D.L. and D.M. Hungerman (2014). Economics, Religion, and Culture: A Brief Introduction. *Journal of Economic Behavior & Organization*, 104, 1-3.
- Cornwall, M., S.L. Albrecht, P.H. Cunningham, and B.L. Pitcher (1986). The Dimensions of Religiosity: A Conceptual Model with an Empirical Test. *Review of Religious Research*, 27(3), 226-244.
- Deaton, A.S. (2011). Aging, Religion, and Health. *Explorations in the Economics of Aging*. Chicago: University of Chicago Press, 237-262.
- Finke R. and C.P. Scheitle (2005). Accounting for the Uncounted: Computing Correctives for the 2000 RCMS Data. *Review of Religious Research*, 5-22.
- Ganzach, Y., S. Ellis, and C. Gotlibovski (2013). On Intelligence Education and Religious Beliefs. *Intelligence*, 41(2), 121-128.
- Gervais, W.M. and A. Norenzayan (2012). Analytic Thinking Promotes Religious Disbelief. *Science*, 336(6080), 493-496.
- Glaeser, E.L. and B.I. Sacerdote (2008). Education and Religion. *Journal of Human Capital*, 2(2), 188-215.
- Gruber, J.H. (2005). Religious Market Structure, Religious Participation, and Outcomes: Is Religion Good for You? *Advances in Economic Analysis & Policy*, 5(1), Art. 5.
- Hansen, K.T., J.J. Heckman, and K.J. Mullen (2004). The Effect of Schooling and Ability on Achievement Test Scores. *Journal of Econometrics*, 121(1), 39-98.
- Heckman, J.J. (1995). Lessons from the Bell Curve. *Journal of Political Economy*, 103(5), 1091-1120.
- Herrnstein, R.J. and C. Murray (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.

- Hrung, W.B. (2004). After-life Consumption and Charitable Giving. *American Journal of Economics and Sociology*, 63, 731-45.
- Huang, W. (1996). Religion, Culture, Economic and Sociological Correlates of Suicide Rates: A Cross-national Analysis. *Applied Economics Letters*, 3, 779-82.
- Hungerman, D.M. (2014). The Effect of Education on Religion: Evidence from Compulsory Schooling Laws. *Journal of Economic Behavior & Organization*, 104, 52-63.
- Iannaccone, L.R. (1998). Introduction to the Economics of Religion. *Journal of Economic Literature*, 36(3), 1465-1495.
- Iyer, S. (2016). The New Economics of Religion. *Journal of Economic Literature*, 54(2), in press.
- Kortt, M.A., B. Dollery, and S. Pervan (2012). Religion and Education: Recent Evidence from the United States. *Applied Economics Letters*, 19(12), 1175-1178.
- Lochner, L. (2011). Nonproduction Benefits of Education: Crime, Health, and Good Citizenship. *Handbook of the Economics of Education*, 4, 183.
- Lochner, L. and E. Moretti (2004). The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports. *American Economic Review*, 94(1), 155-189.
- Lynn, R., J. Harvey, and H. Nyborg. (2009). Average Intelligence Predicts Atheism Rates across 137 Nations. *Intelligence*, 37(1), 11-15.
- Meisenberg, G., H. Rindermann, H. Patel, and M.A. Woodley (2012). Is It Smart to Believe in God? The Relationship of Religiosity with Education and Intelligence. *Temas em Psicologia*, 20(1), 101-120.

- Mocan, N. and L. Pogorelova (2014). Compulsory Schooling Laws and Formation of Beliefs: Education, Religion and Superstition. NBER Working Paper 20557. <http://www.nber.org/papers/w20557>
- Moretti, E. (2004). Estimating the Social Return to Education: Evidence from Longitudinal and Repeated Cross-sectional Data. *Journal of Econometrics*, 121, 175-212.
- Rouse, C.E. (1995). Democratization or Diversion? The Effect of Community Colleges on Educational Attainment. *Journal of Business & Economic Statistics*, 13(2), 217-224.
- Silles, M. (2009). The Causal Effect of Education on Health: Evidence from the United Kingdom. *Economics of Education Review*, 28, 122-128.
- Soydemir, G.A., E. Bastida, and G. Gonzalez (2004). The Impact of Religiosity on Self-assessments of Health and Happiness: Evidence from the US Southwest. *Applied Economics*, 36, 665-72.
- Winship, C. and S. Korenman (1997). Does Staying in School Make You Smarter? The Effect of Education on IQ in The Bell Curve. B. Devlin, S. Fienberg, D.P. Resnick, and K. Roeder (Eds.), *Intelligence, Genes, and Success: Scientists Respond to The Bell Curve*, New York: Copernicus Press, 215–234.

Figure 1. Percentage of religious attendance over the years.



Note: This figure includes the constructed sample of this paper. The sample construction is in section 2.4.

Table 1. Percentage of religious attendance over the years.

Religious Attendance in past 12 months	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
1. Never	23.4%	24.7%	28.7%	30.6%	32.9%	31.8%	35.2%	35.0%	33.6%	36.4%	36.5%	35.3%
2. Once or twice	20.3%	21.5%	23.1%	22.5%	24.3%	24.5%	23.7%	24.4%	25.0%	23.6%	23.4%	23.2%
3. Less than once a month / 3-12 times	12.6%	14.5%	14.6%	11.7%	11.4%	11.4%	10.8%	10.2%	10.6%	11.3%	9.2%	10.7%
4. About once a month / 12 times	5.9%	6.7%	6.8%	7.0%	6.6%	6.5%	6.5%	6.9%	5.8%	5.9%	5.8%	6.6%
5. About twice a month / 24 times	7.0%	6.8%	6.5%	9.0%	7.2%	7.4%	7.2%	6.8%	6.9%	6.5%	6.9%	6.9%
6. About once a week	21.0%	17.8%	14.3%	13.7%	12.5%	13.2%	11.5%	11.7%	12.8%	11.4%	13.2%	12.2%
7. Several times a week	9.1%	7.5%	5.6%	5.2%	4.9%	4.8%	5.0%	4.8%	5.0%	4.8%	4.8%	5%
8. Everyday	0.6%	0.6%	0.4%	0.3%	0.2%	0.4%	0.2%	0.3%	0.3%	0.3%	0.2%	0.3%
% remain unchanged since last year	-	45.3%	46.5%	47.9%	49.3%	52.2%	53.1%	53.5%	53.7%	55.2%	55.8%	57.2%
% increased since last year	-	22.0%	18.9%	23.2%	20.6%	22.7%	19.6%	22.3%	23.6%	20.1%	21.8%	20.9%
% decreased since last year	-	32.6%	34.5%	28.9%	30.2%	25.1%	27.3%	24.2%	22.6%	24.7%	22.4%	21.9%
In 2011, % remain unchanged since 2000						29.5%						
In 2011, % increased since 2000						21.3%						
In 2011, % decreased since 2000						47.3%						

Note: The percentage is based on valid observations from the constructed sample.

Table 2. Percentage of answers about religious beliefs and index over the years.

Religious beliefs questions (True=1, False=0)	2002	2005	2008	2011
1. "I don't need religion to have good values." (reverse coded)	50.6%	46.9%	43.7%	44.0%
2. "Religious teachings should be obeyed exactly as written in every situation."	37.9%	35.1%	34.9%	35.2%
3. "I often ask God to help me make decisions."	71.4%	71.5%	69.5%	68.6%
4. "God has nothing to do with what happens to me personally." (reverse coded)	77.9%	79.6%	77.4%	75.0%
5. "I pray more than once a day."	35.3%	35.7%	37.6%	40.8%
<hr/>				
Trend of index of religious beliefs over the years	2002	2005	2008	2011
% remain unchanged since last time	-	46.1%	48.5%	53.7%
% increased since last time	-	28.7%	30.2%	24.5%
% decreased since last time	-	31.3%	32.3%	25.8%
<hr/>				
In 2011, % remain unchanged since 2002		40.1%		
In 2011, % increased since 2002		24.2%		
In 2011, % decreased since 2002		35.7%		
<hr/>				
Four-question index of religious beliefs (sum of the four answers) (Standard deviation)	2.38 (1.26)	2.36 (1.29)	2.28 (1.31)	2.27 (1.36)

Note: 1. Percentages are based on valid observations from the constructed sample.

2. Question 5, "I pray more than once a day," is not included in the estimation of this paper.

Table 3. Effects of Missing Data on Sample Sizes

	Religious attendance analysis			Index of religious beliefs analysis	
	Sample sizes	Percentage		Sample sizes	Percentage
NLSY97 full sample	8,984	100.00%	NLSY97 full sample	8,984	100.00%
Exclude if race missing	8,904	99.11%	Exclude if race missing	8,904	99.11%
Exclude if 1997 religious preference missing	8,786	97.80%	Exclude if 1997 religious preference missing	8,786	97.80%
Exclude if (A) religious attendance missing	8,565	95.34%	Exclude if (B) religiosity index missing	8,120	90.38%
Exclude if (C) highest grade completed missing	8,559	95.27%	Exclude if (C) highest grade completed missing	8,120	90.38%
Exclude if (D) highest degree missing	8,559	95.27%	Exclude if (D) highest degree missing	8,098	90.14%
Exclude if (C) and (D)	8,551	95.18%	Exclude if (C) and (D)	8,067	89.79%
Exclude if parent's religiosity missing	5,719	63.66%	Exclude if parent's religiosity missing	5,458	60.75%
Exclude if AFQT missing	4,737	52.73%	Exclude if AFQT missing	4,546	50.60%
Exclude if urban, marital status, or number of children missing	4,716	52.49%	Exclude if urban, marital status, or number of children missing	4,483	49.90%
Exclude if (B) Religiosity index missing	4,480	49.87%	Exclude if (A) church attendance missing	4,480	49.87%
Subsample of 14 years old and young as of 12/31/1996	5,419	100%	Subsample of 14 years old and young as of 12/31/1996	5,419	100%
Percentage of the 14 and younger subsample	4,480	82.67%	Percentage of the 14 and younger subsample	4,480	82.67%

Notes: 1. The large amount of missing data on parents' religiosity is that only the parents of youths younger than 14 years old responded to the questions about attitudes, which questions about religion belong to.

Table 4a. Sample descriptive statistics of the key educational/religious variables in Round 15 (2011)

	Mean	Standard deviation	Number of valid data
Religious attendance days in past 12 months	15.53	32.43	3,897
Index of religious beliefs	-0.02	1.01	3,184
Parents' religiosity index	0	1.01	4,480
Highest grade completed as of the survey date	13.76	2.87	3,957
AFQT score	46.33	29.41	4,480
Highest degree received	Frequency	Percentage	
None	341	8.54%	
High school	2,211	55.36%	
Bachelor	1,192	29.84%	
Master	203	5.08%	
PhD	47	1.18%	
Total	3,994	100%	

Note: 1. Some numbers of valid observation in 2011 are less than the sample size 4,480 due to missing values.

2. "High school" includes GED and High school diploma.

3. "Bachelor" includes Associate/Junior college (AA) and Bachelor's degree (BA, BS).

4. "Master" includes MA and MS.

5. "PhD" includes PhD and Professional degree (DDS, JD, MD).

Table 4b. Sample descriptive statistics of the control variables in Round 15 (2011)

	Mean	Standard deviation	Number of valid data
Age	28.34	1.31	4,480
Gender (male equals one, zero otherwise)	0.51	0.50	4,480
Marital status (married equals one, zero otherwise)	0.43	0.50	3,988
Having any children (equals one, zero otherwise)	0.45	0.50	3,995
Living in urban area (equals one, zero otherwise)	0.78	0.41	3,946
<hr/>			
Race	Frequency	Percentage	
White	2,762	61.65%	
Black or African American	1,163	25.96%	
American Indian, Eskimo, or Aleut	33	0.74%	
Asian or Pacific Islander	56	1.25%	
Else	466	10.40%	
Total	4,480	100%	
<hr/>			
Current religious preference in 1997	Frequency	Percentage	
Roman Catholic	1,275	28.46%	
Protestants	2,554	56.99%	
Jewish	37	0.83%	
Other Christian	77	1.72%	
Muslim	16	0.36%	
Hindu/Buddhist	4	0.09%	
Others	377	8.41%	
None	140	3.13%	
Total	4,480	100%	

- Note: 1. Some numbers of valid observation in 2011 are less than the sample size 4,480 due to missing values.
 2. "Protestants" include Baptist, Methodist, Lutheran, Presbyterian, Episcopal/Anglican, United Church of Christ, Disciples of Christ, Reformed Church in America, Holiness, Pentecostal, Nondenominational Christian, and other Protestant.
 3. "Other Christian" includes Mormon, Eastern Orthodox, and Unitarian.
 4. "Others" include Native American Tribal Religion, Bah'ai, Greek, Roman, Norse, .etc. Mythology, Satanic, Wicca/Witchcraft/Magic/Pagan, Scientology, other Eastern, and others.
 5. "None" include Agnostic, Atheist, and no religion/personal philosophy.

Table 5. Pooled OLS results on religious attendance

Dependent variable	Religious attendance days per year							
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Highest Grade	.952*** (.151)	-	-	.873*** (.147)	-	-	.806*** (.142)	-
Highest Degree								
1. None	-	-.434 (.660)	-	-	-.174 (.652)	-	-	.308 (.653)
2. Bachelor	-	4.636*** (.808)	-	-	4.277*** (.790)	-	-	3.762*** (.775)
3. Master	-	7.939*** (2.458)	-	-	7.410*** (2.429)	-	-	6.611*** (2.356)
4. PhD	-	-2.187 (2.190)	-	-	-2.623 (2.244)	-	-	-3.610 (2.285)
AFQT	-	-	1.196*** (.417)	-	-	1.167*** (.407)	.354 (.411)	.793* (.412)
Male	-1.924** (.760)	-2.186*** (.751)	-2.477*** (.743)	-2.135*** (.741)	-2.387*** (.732)	-2.633*** (.725)	-2.133*** (.742)	-2.347*** (.735)
Age	-1.027*** (.076)	-.988*** (.078)	-.769*** (.066)	-1.009*** (.075)	-.966*** (.077)	-.775*** (.065)	-.996*** (.075)	-.932*** (.076)
Married	7.622*** (.778)	7.536*** (.783)	7.627*** (.784)	7.152*** (.768)	7.074*** (.772)	7.147*** (.773)	7.135*** (.768)	7.051*** (.773)
Parent's religiosity	-	-	-	5.526*** (.361)	5.544*** (.362)	5.584*** (.364)	5.529*** (.362)	5.550*** (.363)
Any children	-1.890*** (.711)	-2.372*** (.698)	-2.786*** (.705)	-1.845*** (.699)	-2.309*** (.688)	-2.630*** (.693)	-1.781** (.711)	-2.106*** (.706)
Urban area	-3.250*** (.786)	-3.190*** (.790)	-3.141*** (.790)	-2.816*** (.766)	-2.759*** (.770)	-2.720*** (.770)	-2.835*** (.765)	-2.816*** (.767)
Sample size	4,480	4,480	4,480	4,480	4,480	4,480	4,480	4,480
Observations	45,439	45,439	45,439	45,439	45,439	45,439	45,439	45,439
R-squared	0.0919	0.0911	0.0896	0.1136	0.1130	0.1118	0.1136	0.1134

Notes: 1. Individual clustered robust standard errors are in the parentheses. All models control for race and religious preferences.

2. The reference group of the highest degree received is high school graduate.

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

Table 6. Pooled OLS results on index of religious beliefs

Dependent variable	Index of religious beliefs							
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Highest Grade	-.010*	-	-	-.014***	-	-	.020***	-
	(.005)			(.005)			(.005)	
Highest Degree								
1. None	-	.014	-	-	.030	-	-	-.092***
		(.029)			(.028)			(.029)
2. Bachelor	-	-.035	-	-	-.054**	-	-	.053*
		(.029)			(.027)			(.027)
3. Master	-	-.010	-	-	-.116*	-	-	.055
		(.064)			(.062)			(.063)
4. PhD	-	-.335***	-	-	-.352***	-	-	-.134
		(.124)			(.130)			(.128)
AFQT	-	-	-.156***	-	-	-.158***	-.179***	-.172***
			(.014)			(.013)	(.014)	(.014)
Male	-.185***	-.184***	-.203***	-.194***	-.191***	-.208***	-.196***	-.200***
	(.025)	(.025)	(.025)	(.024)	(.024)	(.023)	(.024)	(.024)
Age	-.012***	-.011***	-.012***	-.012***	-.011***	-.012***	-.017***	-.017***
	(.003)	(.003)	(.002)	(.003)	(.003)	(.002)	(.003)	(.003)
Married	.199***	.200***	.208***	.180***	.182***	.188***	.185***	.184***
	(.027)	(.027)	(.027)	(.027)	(.027)	(.026)	(.026)	(.026)
Parent's religiosity	-	-	-	.255***	.255***	.255***	.253***	.253***
				(.013)	(.013)	(.013)	(.013)	(.013)
Any children	.107***	.109***	.052**	.108***	.112***	.059**	.079***	.072***
	(.026)	(.026)	(.025)	(.025)	(.025)	(.024)	(.024)	(.024)
Urban area	-.106***	-.106***	-.094***	-.088***	-.089***	-.077***	-.081***	-.079***
	(.025)	(.025)	(.024)	(.023)	(.023)	(.023)	(.023)	(.023)
Sample size	4,480	4,480	4,480	4,480	4,480	4,480	4,480	4,480
Observations	13,371	13,371	13,371	13,371	13,371	13,371	13,371	13,371
R-squared	0.2334	0.2336	0.2515	0.2854	0.2855	0.3034	0.3048	0.3047

Notes: 1. Individual clustered robust standard errors are in the parentheses. All models control for race and religious preferences.

2. The reference group of the highest degree received is high school graduate.

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

Table 7. Fixed effects results on religious attendance and index of religious beliefs

Dependent variable	Religious attendance days per year		Index of religious beliefs	
	(i)	(ii)	(iii)	(iv)
Highest Grade	-1.150*** (.119)	-	-.019*** (.005)	-
Highest Degree				
1. None	-	7.304*** (.698)	-	-.051 (.032)
2. Bachelor	-	-1.991*** (.560)	-	-.062*** (.021)
3. Master	-	-1.696 (1.417)	-	-.132*** (.046)
4. PhD	-	-5.857*** (1.960)	-	-.222** (.107)
Age	.255 (.511)	.135 (.511)	-.012 (.023)	-.009 (.023)
Married	.699 (.652)	.436 (.656)	.031 (.022)	0.036 (.022)
Any children	-.407 (.570)	-.112 (.570)	.029 (.021)	.028 (.021)
Urban area	-1.100** (.503)	-1.046** (.504)	-.034* (.020)	-.035* (.020)
Year	-.602 (.504)	-.404 (.507)	-.005 (.022)	-.000 (.023)
Observations	45,439	45,439	13,371	13,371
Sample size	4,480	4,480	4,480	4,480
R-squared: within	0.0116	0.0140	0.0070	0.0073
between	0.0084	0.0042	0.0116	0.0082
overall	0.0001	0.0010	0.0131	0.0096

Notes: 1. Individual clustered robust standard errors are in the parentheses.

2. The reference group of the highest degree received is high school graduate.

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

Table 8. Fixed effects results on religious attendance per year and religiosity index: robustness check

Dependent variable	Religious attendance days per year (Only 2000 and 2011)		Index of religious beliefs (Only 2002 and 2011)	
	(i)	(ii)	(iii)	(iv)
Highest Grade	-2.157*** (.309)	-	-.029*** (.007)	-
Highest Degree				
1. None	-	8.151*** (1.643)	-	-.023 (.046)
2. Bachelor	-	-4.886*** (1.778)	-	-.115*** (.034)
3. Master	-	-7.845** (3.066)	-	-.156** (.065)
4. PhD	-	-11.621** (5.630)	-	-.357*** (.138)
Age	.829 (1.893)	.287 (1.913)	-.003 (.038)	.008 (.038)
Married	1.237 (1.766)	.588 (1.765)	.031 (.035)	.038 (.036)
Any children	-.106 (1.783)	.712 (1.756)	.039 (.035)	.040 (.034)
Urban area	-1.742 (1.647)	-1.848 (1.647)	-.072** (.033)	-.073** (.033)
Year	-.968 (1.862)	-.393 (1.890)	-.003 (.037)	-.015 (.038)
Observations	6,821	6,821	7,235	7,235
Sample size	4,233	4,233	4,357	4,357
R-squared: within	0.0462	0.0429	0.0177	0.0188
between	0.0004	0.0004	0.0083	0.0079
overall	0.0021	0.0060	0.0116	0.0117

- Notes: 1. Individual clustered robust standard errors are in the parentheses.
2. The reference group of the highest degree received is high school graduate.
3. The sample was 14 years old and had parents' religiosity report in 1997.
4. The observation of church attendance days only includes 2000 and 2011. The observation of index of religious beliefs only includes 2002 and 2011.
*: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$.

Table 9a. Subsamples for the fixed effects estimation: Religious attendance days per year

Subsample	Full sample (i)	Stronger religious background (ii)	Weaker religious background (iii)	Higher AFQT scores (iv)	Lower AFQT scores (v)	Male (vi)	Female (vii)
Highest Grade	-1.150*** (.119)	-1.581*** (.169)	-.545*** (.162)	-1.049*** (.134)	-1.122*** (.255)	-1.309*** (.166)	-1.032*** (.169)
Observations	45,439	26,637	18,802	23,825	21,614	22,774	22,665
Sample size	4,480	2,613	1,867	2,312	2,168	2,288	2,192
R ² : within	0.0116	0.0189	0.0037	0.0246	0.0050	0.0133	0.0102
between	0.0084	0.0202	0.0013	0.0065	0.0179	0.0098	0.0053
overall	0.0001	0.0000	0.0022	0.0015	0.0009	0.0000	0.0004
Highest Degree							
1. None	7.304*** (.698)	9.869*** (.934)	3.863*** (1.040)	8.673*** (.755)	5.982*** (1.152)	6.412*** (.954)	8.219*** (1.022)
2. Bachelor	-1.991*** (.560)	-3.357*** (.820)	.024 (.666)	-1.328** (.625)	-1.045 (1.258)	-2.617*** (.826)	-1.520** (.769)
3. Master	-1.696 (1.417)	-3.653* (2.114)	1.267 (1.551)	-1.040 (1.483)	6.205 (5.487)	-5.551** (2.752)	.368 (1.579)
4. PhD	-5.857*** (1.960)	-8.242*** (2.820)	-1.886 (2.007)	-4.042** (2.037)	-12.053* (7.290)	-7.820** (3.252)	-4.653* (2.435)
Observations	45,439	26,637	18,802	23,825	21,614	22,774	22,665
Sample size	4,480	2,613	1,867	2,312	2,168	2,288	2,192
R ² : within	0.0140	0.0226	0.0050	0.0305	0.0065	0.0141	0.0147
between	0.0042	0.0108	0.0021	0.0027	0.0085	0.0032	0.0052
overall	0.0010	0.0005	0.0030	0.0049	0.0000	0.0012	0.0008

Notes: 1. Strong family religious background subsample includes people of whose parents' religiosity indices are higher than median, and vice versa.

2. Higher cognitive ability subsample includes people of whose AFQT scores are higher than median score, and vice versa.

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

Table 9b. Subsamples for the fixed effects estimation: Index of religious beliefs

Subsample	Full sample	Stronger religious background	Weaker religious background	Higher AFQT scores	Lower AFQT scores	Male	Female
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
Highest Grade	-0.019*** (.005)	-.025*** (.006)	-.011 (.007)	-.020*** (.006)	-.005 (.008)	-.023*** (.007)	-.017*** (.006)
Observations	13,371	8,025	5,346	7,133	6,238	6,694	6,677
Sample size	4,480	2,613	1,867	2,312	2,168	2,288	2,192
R ² :							
within	0.0070	0.0109	0.0038	0.0180	0.0006	0.0102	0.0069
between	0.0116	0.0016	0.0415	0.0029	0.0003	0.0025	0.0177
overall	0.0131	0.0051	0.0312	0.0072	0.0001	0.0037	0.0160
Highest Degree							
1. None	-.051 (.032)	-.078* (.042)	-.013 (.048)	-.007 (.050)	-.064 (.041)	-.056 (.046)	-.049 (.043)
2. Bachelor	-.062*** (.021)	-.075*** (.027)	-.040 (.033)	-.013 (.024)	-.121*** (.047)	-.074** (.031)	-.058** (.028)
3. Master	-.132*** (.046)	-.194*** (.053)	-.024 (.084)	-.063 (.051)	-.253 (.190)	-.144** (.071)	-.136** (.061)
4. PhD	-.222** (.107)	-.289* (.160)	-.125 (.118)	-.137 (.113)	-.752*** (.050)	-.170 (.175)	-.267** (.136)
Observations	13,371	8,025	5,346	7,133	6,238	6,694	6,677
Sample size	4,480	2,613	1,867	2,312	2,168	2,288	2,192
R ² :							
within	0.0073	0.0119	0.0036	0.0159	0.0041	0.0099	0.0080
between	0.0082	0.0026	0.0219	0.0164	0.0007	0.0000	0.0127
overall	0.0096	0.0068	0.0172	0.0157	0.0012	0.0005	0.0119

Notes: 1. Strong religious background subsample includes people of whose parents' religiosity indices are higher than median, and vice versa.

2. Higher cognitive ability subsample includes people of whose AFQT scores are higher than median score, and vice versa.

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

Table 10. Instrument variables for education using college accessibility

IV's	Description	Mean	Standard Deviation	Number of Observations	First-stage coefficient on schooling years
(1) Seats-young population ratio in 2000	Number of seats (enrollment) in colleges in the individual's residential county ^{1,2} , divided by the number of population from 18 to 29 years old.	35.31%	11.02%	2,970	0.872** (0.354)
(2) BA percent in 1990	Percentage with bachelor's degree or higher, in the population of 25-years old and older in 1990 ³ .	19.31%	7.97%	2,970	0.013*** (0.005)

Notes: 1. Generated from 2000 Census data.
 2. A robustness check about the choice of year is conducted.
 3. From NLSY97 Geocode data.
 4. Standard errors are in the parentheses in the last column.

Table 11. Instrumental variable estimates: college accessibility

	(1)	(2)	(3)
Estimator	OLS	2SLS IV: studentPCT young2000	2SLS IV: BApercent1990
<u>Panel A: Church Attendance Days as dependent variable</u>			
Coefficients	0.969*** (0.249)	-3.820 (5.443)	-4.857 (5.577)
Number of observations	3,647	3,647	3,647
<u>Panel B: Index of Religiosity as dependent variable</u>			
Coefficients	0.026*** (0.007)	-0.723** (0.320)	-0.509** (0.210)
Number of observations	2,970	2,970	2,970

Note: 1. Robust standard errors are in the parentheses.
 2. This is conducted cross-sectionally using the 2011 data.