The Journal of Business Inquiry 2014, 13, Issue 1 (Special Issue), 48-60
http:www.uvu.edu/woodbury/jbi/articles/
ISSN 2155-4072

The Impact of Religion on Youth Outcomes

By BRYSON POPE, JOSEPH PRICE AND DEAN R. LILLARD*

We use data from several nationally representative datasets to estimate the relationship between church attendance and risky behaviors and whether these associations vary when one accounts for selective participation. We use various empirical methods including propensity score matching, sibling and family fixed-effects models, and instrumental variables models that exploit cross-state variation in blue laws. Our results across the different approaches converge into a general pattern that youth with higher church attendance are less likely to commit property or violent crimes, smoke, drink, use drugs, or receive a traffic ticket.

Keywords: Church Attendance, Risky Behaviors, Crime, Blue Laws

JEL Classification: Z12, I31, K42

I. Introduction

Past research has consistently shown a correlation between participation in religious activities and positive social behavior among youth (Mocan and Rees, 2005; Johnson *et al.,* 2000). Religiously involved youth are less likely to use alcohol, cigarettes, and drugs, have pre-marital sex, and commit crimes (Sabia, 2006). The literature recognizes that these simple associations do not identify whether religious participation has a causal effect on these positive outcomes. This is due to the inability of observing many potentially influencing factors pertaining to youth and their families.

There are a number of mechanisms through which religious groups could plausibly have a causal effect on youth outcomes. First, religious groups transmit values and help youth internalize moral messages and norms. Second, members and leaders of religious groups monitor youth behavior and serve as examples for youth to emulate. Third, religious groups provide an institutional setting that provides activities for youth that pull them away from harmful influences and behaviors. For one reason or another, it is clear that parents and children choose to participate in religious activity. Therefore, the usual issues of self-selected samples make it challenging to statistically identify whether religious participation actually has a causal effect on youth outcomes.

To confront that challenge, we follow a few well-established statistical methods. First, we compare outcomes of observationally similar individuals who participated in religious groups with different levels of intensity. We also use multivariate regression and propensity score matching methods, models that include either individual and family fixed effects, and an instrumental variable approach developed by Gruber and Hungerman (2006) that exploits temporal and cross-

---

*Bryson Pope: Department of Economics, Brigham Young University. E-mail: brysonrpope@gmail.com; Joseph Price: Department of Economics, Brigham Young University, 162 FOB, Brigham Young University, Provo, UT 84602. E-mail: joseph_price@byu.edu. Phone: (801) 422-5296, Fax: (801) 422-0194; Dean R. Lillard: Department of Human Sciences, Ohio State University. E-mail: lillard.13@osu.edu.

The authors gratefully acknowledge helpful comments from participants at the National Poverty Conference and the Religious Faith and Social and Applied Sciences Conference.

sectional variation in whether stores are open for business on Sundays (blue laws). We use data from the National Longitudinal Surveys of Youth 1979 (NLSY79), the Children of the National Longitudinal Surveys of Youth 1979 (CNLSY79), and Monitoring the Future (MTF). The richness of these nationally representative data allows us to implement the above methods and control for many observed characteristics. Although non-identical results are to be expected, the use of various different statistical techniques provides different approaches to finding the direction and magnitude of the influence religion has on youth outcomes.

II. Methods and Existing Evidence

Many previous studies have presented estimates of the simple or conditional correlation between religious participation and positive outcomes. Few, however, use methods or have data needed to credibly estimate whether this relationship is of a causal nature. The correlational studies typically fail to control for the potential of selection bias. Below we give a brief review of the literature and methods used by the few studies that rigorously address selection bias with regard to religion and crime.

Correlational Evidence

Correlational studies typically compare outcomes of people who participate differently in religious activities but who are observationally equivalent across a set of individual or family characteristics. For example, studies find that, controlling for background characteristics, youth who are more religiously involved are less likely to commit crimes (Mocan and Rees, 2005; Johnson *et al,.* 2000).

Many correlational studies focus on particular groups of interest. In their meta-analysis of 60 studies, Baier and Wright (2001) found that correlations between religiosity and outcomes were consistently stronger for blacks. Dehejia *et al.* (2007) estimate how negative outcomes vary with economic and social disadvantage (low income, low education), and then examine whether those associations differed for youth who participated in religious activities more and less frequently. They find that the negative association between disadvantage and negative outcomes is weaker among youth who participate more in religious activities. These associations are consistent with the hypothesis that religious involvement may confer benefits but do not prove a causal relationship.

Causal Evidence and Methods Used

A. Matching Estimators

Matching estimators are based on the idea that one should compare differences in behavior of observationally similar individuals to reduce the influence of unobservables. Although matching estimators are similar in nature to traditional regression estimates, the actual comparison between churchgoers and non-churchgoers is made more explicit. We use one of the most common matching methods, propensity score matching. [1]

To compute a propensity score we regress whether or not a person goes to church (in our case at least monthly) on a set of observed matching characteristics. In general this set should not

[1] Rosenbaum and Rubin (1983) provide the theorem that shows one can use this method to reduce the dimensionality of the set of data one uses to compare individuals.

include any characteristic that might be affected by the treatment (here church attendance).[2] We use the coefficients from the full sample to predict the probability that each respondent attends church. We then divide the sample into those who report they attend church monthly and those who report they do not attend church, or attend church less than monthly. Finally, we use the propensity score of each person who attends religious services to match him/her to a person who does not. Each attending person is matched to the non-attender whose propensity score is closest in value. Under the conditions described in Rosenbaum and Rubin (1985), Heckman, Ichimura, and Todd (1998) and Heckman *et al.* (1998), the causal effect of attending church at least monthly is given by the difference between the average outcome of attenders and matched non-attenders.

The economic literature is replete with studies that have used similar propensity score matching techniques as the study at hand. Lundquist (2004), for example, uses propensity matching scores for an individual's propensity to marry, and find divergent trends in nuptial patterns between black people and white people disappear in the military. Jalan and Ravallion (2003) also use Rosenbaum and Rubin's (1983) propensity score matching approach to estimate the distribution of net income gains from an Argentinean workfare program.

It is important to note that there are a number of different ways in which calculated propensity scores can be matched. We have chosen to use the nearest neighbor matching method. Other options may include caliper matching, stratification matching, and difference-in-difference matching involving kernels and local linear weights. Neighbor matching is attractive in its simplicity and its well accepted intuitive approach.

B. Fixed-Effects Models

While matching estimators permit one to use a rich set of covariates when comparing individuals, it does not resolve the fundamental problem of the possible existence of unobserved factors that determine the behavior of interest and other outcomes being studied. A second method solves this problem for a subset of unobserved factors that are shared in common by people who also share factors that do not vary over time. This method is referred to as the "fixed-effects" method.

To implement the fixed-effects method, one estimates models that compare differences in religious participation and outcomes across members of the same group. In such models, any unobserved characteristic that determines both participation and the outcome of interest will be "swept out" as long as that characteristic is common to all members of the group. One might, for example, compare the religious participation and outcomes of siblings. The key critical assumption in these models is that the shared environment of the siblings fully captures the influence of unobserved factors that determine both religious participation and outcomes of interest.

When data contain information on the same individual over time, it is also possible to estimate a model that includes individual fixed effects. These models test whether an individual changes his behavior over time as he participates more (or less) in religious activities. Social scientists commonly estimate models with individual fixed effects to try to estimate the causal effect of choices to marry (Korenman and Neumark, 1991) or engage in sex during adolescence (Sabia, 2006). Such models only account for unobserved determinants that do not vary over time.

[2] Even among this set one can only use variables in a range with common support in the treated and untreated states. Heckman, Ichimura, and Todd (1998) and Heckman *et al.* (1998) describe conditions that define the set of variables one can use to match.

While fixed-effects models have attractive properties, they also have shortcomings.[3] Fixed-effects estimators only account for time-invariant omitted variables,[4] and require sufficiently many of the same individual observations over time or that include a sufficient number of groups with two or more members. Furthermore, an implicit assumption of the fixed-effects estimator is that the variation in the variable of interest (here church attendance) measures a randomly assigned change. However, this assumption is unlikely to be met. It is unlikely, for example, that individuals who change their religious attendance are representative of a randomly selected individual from the general population. Similarly, one suspects that, if siblings attend religious services with different frequency, they probably differ in unobservable ways (even when they share the same family environment). Finally, the fixed-effects estimator may be too conservative because it ignores any differences in behavior across individuals that are correlated religious attendance. It is possible that some of the cross-individual variation in religious participation results in causal changes in outcomes. The fixed-effects model, however, assumes that variation has no causal effect.

C. Instrumental Variables

Researchers also try to estimate causal effects by the method of instrumental variables (IV). The method requires finding a variable (instrument) that is correlated with the explanatory variable of interest and that is uncorrelated with the outcome being studied. That is, the IV must affect the outcome of interest only through the explanatory variable. This is a difficult task. Also, individuals must experience different levels of the instrument, and the variation in exposure to the instrument must cause individuals to participate in religious services with different frequency. The model is implemented in two steps. First, one runs a model of participation in religious services as a function of the instrument (and other covariates). Second, one uses the predicted value of religious participation in a model of the behavior of interest. Under the above assumptions, the IV estimator identifies the causal effect of religious involvement.

The few IV studies that investigate the impact of religion have used three different instruments: the repeal of blue laws which allowed stores to be open on Sundays, the fraction of people in a given area who are of the same ethnicity, and the fraction of residents who adhere to a particular religion. For example, Gruber and Hungerman (2006) use differences in the date that state governments repealed laws that prohibit stores to operate on Sunday. This instrument rests on the assumption that individuals spend time on activities that give them the most utility. By banning stores from operating on Sundays, state governments raise the cost of shopping (and thereby lower the relative cost of attending church). The reverse effect occurs when governments repeal blue laws. It becomes cheaper to shop and therefore church is a less attractive way to spend time for people on the margin. They find that church attendance drops following the repeal of blue laws. They then use NLSY79 data to show that people who were religious before blue laws were repealed, drank more and were more likely to use drugs after the repeal of blue laws allowed stores to remain open on Sundays.

In other work, Gruber (2005) instruments religious participation with the fraction of people in an individual's area who are of the same ethnicity. Using GSS data he shows that people are more likely to go to church if they live near more people of the same ethnicity. Using ethnic

[3] For an overview, see Currie (2003).
[4] Ruhm (2005) notes that in fact using fixed effects in the face of a time-varying omitted variable may even aggravate the bias.

concentration as a proxy for religious affiliation, he finds that people living in areas with more people of the same ethnicity have more income, higher education, are more likely to marry, are less likely to divorce, and experience lower levels of welfare receipt and disability.

In his study of county-level crime rates, Heaton (2006) uses county-level religiosity from 1916 to instrument for current religious participation in the same county. He finds that while county-level current religious adherence predicts measures of county-level crimes committed in models estimated by Ordinary Least Squares, coefficients from models estimated by IV show no impact of religion on crime.

Each of these instruments for religious participation has its drawbacks. In some cases the first stage relationship between the instruments and the religious participation is statistically weak. In the case of blue laws, there have been no state-level changes in blue laws since 1992 so using this instrument for more recent data would require using law changes at the county or local level.

III. Data

We use data from the National Longitudinal Survey of Youth of 1979 (NLSY79), the Panel Study of Income Dynamics (PSID), and the Monitoring the Future (MTF) surveys. We rely on the NLSY79 and PSID data to generate our main results. The NLSY79 and PSID are nationally representative surveys that track individuals over years including years in which many state governments repealed blue laws. Both surveys identify each individual’s state of residence along with many other characteristics. In addition, outcome data of siblings (NLSY79, PSID) and other relatives (PSID) are tracked. Although the MTF lacks state identifiers and has no information on outcomes of siblings, results from the MTF supplement our main results in informative ways.

The National Longitudinal Study of Youth 1979 (NLSY79)

The NLSY79 is a nationally representative sample of 12,686 youth who were between the ages of 14 and 22 in 1979. In 1979, the respondents provided detailed information about their family background, including the religious affiliation of their parents. The NLSY79 contains religious participation measures in 1979 and then detailed information about illegal activities in 1980. Both violent crime and property crime are measures based on self-reported behavior and are described in more detail in Levitt and Lochner (2001).

Church attendance is reported in the NLSY79 on a six-point scale: never, infrequently, monthly, two to three times a month, weekly, and more than weekly. For the analysis, we convert the six-point scale into a measure of number of days per year, which we standardize by dividing by the standard deviation across the sample.

The Children of National Longitudinal Study of Youth 1979 (CNLSY79)

Starting in 1986, children of the female NLSY79 respondents were surveyed. This sample, known as the Children of the NLSY79 (CNLSY79) allows us to observe the religious participation and outcomes of a second generation. Together with data on religious participation of the parents of the NLSY79, we can examine religious participation across three generations. More importantly, the data allow us to carry out our analysis exploiting differences within an individual over time and differences between siblings.

One downside is that since the original respondents to the NLSY79 were 14-21 in 1979, the only older youth that we observe in the sample are those who were born to relatively younger mothers. One way that we can address this issue is by looking at outcomes among youth at even younger ages. While a ten-year old does not really fit the focus of our studies on youth, it is possible that the patterns that start to develop at these younger ages will influence behavior of individuals in the focus age range in the rest of our paper.

The CNLSY79 provides data on several important indicators of youth outcomes. The first is the behavior problem index (BPI) which is based on the parent's report of whether the child exhibits certain behaviors such as being aggressive, fighting, etc. The responses to these questions are aggregated into the BPI with higher scores representing worse behavior. The second is the PIAT-reading raw score which is a widely used measure of child cognitive achievement. This is a particularly appropriate measure when looking at the impact of religion since many congregations put a great deal of emphasis on personal scripture study. Academic achievement is likely to displace other bad behaviors because of the increased prospects of post-secondary schooling, thus increasing the opportunity cost of deviant behavior. The third is a self-reported measure of whether the individual has ever had a physical fight or intentionally harmed another person.

The Panel Study of Income Dynamics (PSID)

The PSID is a household-based survey that began in 1968 with a survey of about 5,000 households. Each member of the 1968 household and all offspring of any member of those households has been followed and interviewed. Annual interviews were conducted from 1968 through 1997. Since 1997 surveys are conducted biennially. We use data from 1968 through 2003. As of 2003, data have been collected from more than 60,000 individuals.

Information on religiosity and religious participation varies widely across the years the PSID has operated. The PSID collected data on the frequency that the head of household attended church from 1968-1972. From 1970-1976 and from 1981-2003 the PSID collected data on the religious preference of the head of each household in the sample. In addition, the religious preference of wives of household heads was asked from 1985-2003. Starting in 1994 the PSID collected more detailed measures of affiliation, including specific Protestant denomination affiliation. These data are available for all individuals who participated at least once over the full duration of the PSID surveys.

Additional information on religiosity is available for a subset of PSID respondents who participated in the Child Development Supplement (CDS) in 1997 and 2002. Because the PSID asks questions in intermittent years, we created several measures of religious affiliation. We first identified "own affiliation" for all PSID sample members who were either a "head" or a "wife" in the survey year religious affiliation was asked. We do so iteratively, working back from the more detailed denominational information in the 1994 and later surveys to the broader response categories used in earlier years. In coding the data, we allowed respondents to identify multiple religions across different survey years (to allow for the possibility that a person changed religions). We next created an "imputed" religious affiliation measure that was coded for each son or daughter of a PSID head or wife for whom affiliation was collected. This coding assigns the religion of the mother and/or father to the child living with them. Here again, we coded all reported affiliations. We thus measure multiple affiliations when the mother and father were of different religions and when one or both of them changes religious affiliation over the period we observe them.

In our multivariate analysis we categorize individuals who identify with some religion according to whether or not the religion treats Sunday as a holy day. We define a person to observe a holy day other than Sunday if he or she reports any affiliation with Judaism, Jehovah's Witnesses, Seventh Day Adventists, other non-Christian religions or any other religion.

In order to analyze the behavior of the largest number of CDS respondents as possible, we use the frequency of church attendance of the father and/or mother (whichever is greater) rather than the child's reported church attendance. Results are similar, however, when we use the child's reported frequency of attendance.

We also categorize PSID respondents according to whether their state of residence (at each age) required retail stores to close on Sundays (i.e. had a "blue law" in force). To categorize the presence of a blue law we use information reported in Laband and Heinbuch (1987), Goos (2005), and Gruber and Hungerman (2006).

Monitoring the Future (MTF)

Monitoring the Future (MTF) is a cross-sectional school-based survey that the University of Michigan administers each year to 15,000-19,000 students. From 1975 to 1990 the MTF only surveyed students in 12^{th} grade. Since 1991, the MTF also surveys students in grades 8 and 10. The survey asks students to report how frequently they attend church. Responses are limited to four categories: never, a few times a year, monthly, and weekly. We convert these responses into days per year. The survey also asks students to report whether they have consumed cigarettes, alcohol, and drugs over the last 30 days, and at what quantity. The MTF also asks each student if he has ever received a (traffic) ticket. We use these data as a proxy measure of risky behavior.

MTF data can only be used as supplementary analyses because the study does not release data to identify siblings and state of residence or to track individuals over time. Given the size of the samples collected for the MTF, providing access to such data could greatly benefit researchers and increase use of the MTF data. Summary statistics of these surveys with regards to religiosity are presented in Table 1.

Table 1: Summary Statistics

	Frequency of Religious Attendance			
CNLSY79	Never	Few times	Monthly	Weekly
BPI	91.849	84.369	80.222	78.414
PIAT-reading	53.596	55.078	55.976	57.467
Hurt someone	0.326	0.382	0.319	0.286
N(BPI)	841	1,083	608	2,422
NLSY79				
Violent crime	0.192	0.156	0.131	0.126
Property crime	0.122	0.087	0.057	0.050
N(violent crime)	1,140	2,168	860	2,549

Table 1: Summary Statistics: Continues

	Frequency of Religious Attendance			
	Never	Few times	Monthly	Weekly
MTF				
have had a ticket	0.436	0.405	0.394	0.331
have smoked	0.720	0.742	0.717	0.616
smoked last 30 days	0.375	0.363	0.327	0.219
used alcohol	0.933	0.953	0.942	0.872
alcohol last 30 days	0.706	0.724	0.701	0.566
used drugs	0.670	0.668	0.612	0.470
drugs last 30 days	0.388	0.358	0.302	0.193
N(ticket)	9,129	29,010	13,203	23,390

Notes: Each column includes the individuals from the dataset that reported that frequency of religious attendance. The sample size varies a little bit across outcome measures and we report the sample size of one outcome measure (indicate in parenthesis) from each dataset.

IV. Results

This section provides some empirical estimates of the impact of religious participation on a wide range of youth outcomes using the datasets described. We recognize that our models likely fail to capture important controls for youth outcomes and that our measure of religion (frequency of church attendance) fails to capture the full picture of religious involvement (Regnerus, 2003). However, our goal here is to describe the implementation of each of these five approaches in sufficient detail so that they may be more easily used by other researchers, and additionally to give an indication of the apparent general trend that emerges regarding the relationship between youth outcomes and church attendance.

Each of these methods address in different ways the fact that people choose to go to church and that the factors that influence this decision may bias our estimates. Since we cannot randomly assign people to attend church, we are never sure whether the correlations we observe between church attendance and good behavior is due to selection into church attendance or a causal impact of religion on behavior. In the regression framework, this simply means that the error term (the unexplained factors of our outcome) is positively correlated with the likelihood that someone attends church. This creates an upward bias in our estimate of church attendance on youth outcomes and causes us to attribute a larger impact to church attendance than we should. Each of the methods described below provide different solutions to this problem, though each method comes with its own set of advantages, drawbacks, and unique data requirements. This multi-faceted approach is also based on the recommendation of Currie (2003) that when dealing with the sample selection problem, it is preferable to use a number of techniques.

A. Multivariate Regression

The most common approach to the selection problem is to control directly for the factors that lead people to attend church more often. Higher levels of church attendance is associated with being female, from the south, in a larger family, and having a mother who is married, is more educated, and does not work.

Our results (Table 2) suggest that youth experience better outcomes if they attend church more often. This association is statistically significant at the 5 percent level for 10 of the 11 undesirable outcomes. To put these results into context, consider how behaviors would change if one could exogenously increase church attendance by one standard deviation. If the estimated associations were causal, a one standard deviation increase in church attendance would cause the probability of smoking, using drugs, or drinking alcohol in the past 30 days to fall by about 6-7 percent.

B. Matching Estimators

An alternative is to simply match each church-attending youth with a non-church attending youth who otherwise have similar observable characteristics. We can then compare outcomes between this matched set of individuals. In this way, we attempt to establish the counterfactual with the key assumption that all unobservables that influence the decision to go to church are identical between the two individuals.

Table 2 provides our estimates of the differences in youth outcomes using the propensity score matching strategy. [5] Our results indicate that this method also suggests that youth who go to church often have better outcomes. Furthermore, the magnitude of the benefit attributed to church is even larger than the OLS estimate. For example, the PSM estimate shows a decrease of 0.012 percentage points (8.1 percent) to commit a violent crime and a 0.064 percentage point decrease (9.3 percent) in the likelihood of a frequently religious youth to drink alcohol sometime in their life. These are both larger than our OLS estimate. Furthermore, the magnitude of these coefficients seems reasonable when compared with the results reported by Mocan and Rees (2005). They report a 0.013 (6.8 percent) marginal effect for young males who report having no religion for juvenile assault (nearly identical to our results for violent crime), 0.032 on male juvenile theft, 0.011 on male juvenile robberies, and 0.027 for juveniles selling of drugs. Similarly, results from Sabia (2006) indicate that youth who attend church services weekly decrease their chance of committing suicide with marginal effects ranging from 0.007 to 0.043. The magnitude of our results for the effect of religion on juvenile violent and property crimes, then, seems plausible when compared with the existing literature.

The difference in the coefficients across the various methods used is not surprising given the different specifications used in each model. However, a general trend emerges. Although the various models produce different estimates, with few exceptions the sign value on the coefficients indicates frequent religious attendance decreases negative outcomes in youth. The main disadvantage of the matching estimators is that we can only match on observable characteristics in our data. There are likely many other factors that influence why people go to church (and also affect their outcomes) that we cannot observe. One way to capture some of these unobserved factors is by comparing personal characteristics to an individual who has the same unobserved characteristics such as a sibling or the individual's past or future self.

[5] We use all of the default options for the psmatch2 command in STATA. There are a number of ways in which the matching strategies can be implemented giving a possible wide range or estimates (and thus subject to some misuse or abuse). We chose the propensity score matching estimator since it appears to be the most widely used in social sciences.

Table 2: Estimates for Frequent Religious Involvement in Youth on Negative Outcomes

Dataset	Outcome	Mean	OLS	PSM	Family fixed effects	Individual fixed effects
CNLSY79	BPI	82.22	-1.143*	18.330**	0.018	0.470
			[0.522]	[3.284]	[0.578]	[0.550]
	PIAT-reading	56.10	0.316**	2.716**	0.246*	0.341**
			[0.110]	[0.762]	[0.104]	[0.111]
	hurt someone	0.323	-0.004	-0.032	-0.013*	-0.005
			[0.003]	[0.021]	[0.006]	[0.005]
NLSY79	violent	0.148	-0.009*	-0.012	0.021	--
			[0.004]	[0.012]	[0.012]	--
	property	0.075	-0.013**	-0.031**	0.003	--
			[0.003]	[0.009]	[0.009]	--
MTF	tickets	0.384	-0.038**	-0.024	--	--
			[0.002]	[0.020]	--	--
	drink (life)	0.690	-0.036**	-0.064**	--	--
			[0.001]	[0.010]	--	--
	drink (30 days)	0.306	-0.071**	-0.138**	--	--
			[0.001]	[0.018]	--	--
	smoke (life)	0.920	-0.056**	-0.107**	--	--
			[0.001]	[0.020]	--	--
	smoke (30 days)	0.661	-0.067**	-0.129**	--	--
			[0.001]	[0.017]	--	--
	Drug (life)	0.587	-0.091**	-0.176**	--	--
			[0.001]	[0.018]	--	--
	Drug (30 days)	0.292	-0.077**	-0.145**	--	--
			[0.001]	[0.016]	--	--

Notes: Each column include controls for age, parent's marital status, education, and work status. Standard errors are clustered at the individual level. Statistical significance of 1% and 5% are denoted by **, and * respectively.

C. Individual and Family Fixed Effects

We specify and estimate models with fixed effects (i.e. separate intercepts) either for each individual or for siblings. In the individual fixed-effect model the effect of religious attendance is identified for individuals who change how often they attend church over successive survey waves (or who report different attendance frequencies). Individuals who never attend and individuals who always report the same frequency of attendance are dropped from the analysis. In this specification, the coefficient on religious attendance measures how much behavior changes as frequency of attendance changes, holding constant all time-invariant person-specific characteristics.

In the family fixed-effect model, the effect of religious attendance is identified not only for individuals who change their attendance over time but also for siblings who (consistently) attend with different frequency. In these models all individuals remain in the sample as long as the sample includes data for at least one other sibling. The family fixed-effects model controls for time-invariant characteristics that siblings experience in common. The set of such characteristics might include shared genetic endowments, shared attitudes or cultural beliefs parents transmitted to them, and shared environmental influences.

The fixed-effects specification demands much of the data but produces more conservative estimates of the effect of religious attendance. Thus comparing siblings is likely better than comparing two unrelated individuals even when there is a host of control variables. It will also capture the time-invariant unobserved characteristics that multivariate regression or matching strategies cannot address. We are able to observe siblings in the NLSY79 and CNLSY79 datasets. In the CNLSY79 data we also observe repeated measures of the same individual in terms of both their frequency of church attendance and various outcomes.

The OLS regression in Table 2 is a simple Ordinary Least Squares multivariate regression. The fourth column limits the sample to only one observation for each individual (the one with the highest age) and includes a set of family fixed effects. This approach controls for all of the time-invariant characteristics of a particular family that influence all of the children in the home in the same way.

The final column contains coefficients obtained when individual level fixed effects are used which control for all of the time-invariant characteristics that influence the outcomes of an individual. This is equivalent to a cross-sectional regression in which we were able to include controls for every single characteristic of the individual that does not change over time (i.e. parent's education, race, age at birth; prenatal care; environmental factors during their childhood, etc.). More importantly it also controls for all of the time-invariant unobservable factors that might influence outcomes for which we would never be able to adequately measure with current methods.

The OLS results indicate a positive association between church attendance and good youth outcomes. A unit increase in church attendance leads to a 1.143 drop in the BPI (2 percent) and a 0.316 point rise in the reading test score (1 percent). In addition, a standard deviation increase in church attendance leads to a 0.9 percentage point increase in the probability of committing a violent crime (6 percent) and a 1.3 percentage point increase for property crime (17 percent). When we include the individual fixed effects we find a positive impact for the reading scores and the likelihood of harming others, while reading scores remain positive and significant.

D. Instrument for Church Attendance

Table 3 reports results of a model in which we exploit variation in the presence of blue laws across different states and over various years. Since blue laws in the United States are always designed to restrict what types of stores can be open on Sunday, we interact our measure of the presence of a blue law with whether or not the individual belongs to a religious denomination for which Sunday is their day of worship or holy day. In Table 3 we present results estimated with (Column 2) and without (Column 1) state fixed effects. The notes on Table 3 include a list of all of the additional control variables that are included in both regressions.

We find that while youth are not more likely to smoke in states either with or without blue laws, those whose religion treats Sunday as a holy day are less likely to start smoking in states that do not allow businesses to operate on Sundays. The results remain (marginally) significant in column 2 when we estimate these models with state fixed effects. The model with state fixed effects identifies the effect for states that either changed their laws or for individuals who moved between states with and without blue laws.

Table 3: Effect of Religious Affiliation and Blue Laws on the Probability a Youth Starts to Smoke

	(1)	(2)
Blue law * Sunday is holy day	-0.0103**	-0.0090*
	(0.0050)	(0.0051)
State has blue law	0.0087	0.0116**
	(0.0055)	(0.0058)
Sunday is holy day	-0.0072	-0.0058
	(0.0049)	(0.0049)
Mean outcome	0.030	0.030
State FE?	No	Yes
N	98,935	98,770

Notes: The coefficients reported are the marginal effects from a probit regression. Sample restricted to PSID respondents age 5 to 30 who have not yet begun to smoke. Statistical significance of 5% and 10% are denoted by **, and * respectively. All models control for age, highest grade completed to date, religious affiliation, number of reported religions, whether a female respondent gave birth or conceived in the calendar year, the fraction of the year she spent pregnant, marital status (married, single, divorced), cigarette price, sex, race (black, Hispanic, other), the average of household income over all observed years, and a year trend.

V. Conclusion

The goal of this paper has been to illustrate some of the tools used by economists to address the problem of selection. The need for these tools arise because individuals can choose whether or not to go to church and the unobservable factors that make a person more likely to go to church also likely influence their behavior directly. This selection problem will lead to biased estimates and cause us to attribute a larger effect to church attendance than we should. Each of these methods that we have described in this paper provides distinct advantages over the traditional OLS approach, but specific limitations may prohibit their use in particular datasets or settings.

It is also important to mention that the absence of a significant effect when using these tools is not proof that religion has no effect on the outcomes of youth. These methods often require the analysis to focus on a subset of the sample, making precise inference difficult. However, when we do detect significant effects using these methods we can be more confident of ascribing them to a causal interpretation. For each method used, our results suggest that more frequent church attendance has a real impact on youth behavior, specifically on current substance use (smoking, alcohol, and drug). Future exploration using these methods on a variety of datasets will allow for more precise estimates of the magnitude of these effects.

References

Baier, Colin J., and Bradley R. E. 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.

Currie, Janet. 2003. "When Do We Know What We Think We Know? Determining Causality." In *Work, Family, Health and Well-Being*, ed. Suzanne M. Bianchi, Lynne M. Casper and Roselind Betkowitz King, 275-92.

Dehejia, Rajeev, Thomas DeLeire, Erzo F.P. Luttmer, and Joshua Mitchell. 2007. "The Role of Religious and Social Organizations in the Lives of Disadvantaged Youth." NBER Working Paper No. 13369.

Goos, Maarten. 2005. "The Impact of Shop Closing Hours on Labor and Product Markets." http://cep.lse.ac.uk/stokerochford/papers/new/Goos.pdf.

Gruber, Jonathan. 2005. "Religious Market Structure, Religious Participation, and Outcomes: Is Religion Good for You?" NBER Working Paper No. 11377.

Gruber, Jonathan, and Daniel M. Hungerman. 2006. "The Church vs the Mall: What Happens When Religion Faces Increased Secular Competition?" NBER Working Paper No. 12410.

Heaton, Paul. 2006. "Does Religion Really Reduce Crime?" *The Journal of Law and Economics*, 49(1): 147-72.

Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd. 1998. "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 66(5): 1017-98.

Heckman, James J., Hidehiko Ichimura, and Petra Todd. 1998. "Matching As An Econometric Evaluation Estimator," *Review of Economic Studies*, 65(2): 261-94.

Jalan, Jyotsna, and Martin Ravallion. 2003. "Estimating the benefit incidence of an antipoverty program by propensity-score matching." *Journal of Business & Economic Statistics*, 21(1): 19-30.

Johnson, Byron R., David B. Larson, Spencer De Li, and Sung Joon Jang. 2000. "Escaping from the Crime of Inner Cities: Church Attendance and Religious Salience Among Disadvantaged Youth." *Justice Quarterly*, 17(2): 377-91.

Korenman, Sanders, and David Neumark. 1991. "Does Marriage Really Make Men More Productive?" *The Journal of Human Resources*, 26(2): 282-307.

Laband, David N., and Deborah Hendry Heinbuch. 1987. *Blue Laws: The History, Economics, and Politics of Sunday-Closing Laws*. Lexington, MA: Lexington Books.

Levitt, Steven D., and Lance Lochner. 2001. "The Determinants of Juvenile Crime." In *Risky Behavior Among Youths*, ed. Jonathan Gruber. Chicago. IL: University of Chicago Press, 327-73.

Lundquist, Jennifer Hickes. 2004. "When Race Makes No Difference: Marriage and the Military." *Social Forces*, 83(2): 731-57.

Mocan H. Naci, and Daniel I. Rees. 2005. "Economic Conditions, Deterrence and Juvenile Crime: Evidence from Micro Data." *American Law and Economics Review*, 7(2): 319-49.

Regnerus, Mark D. 2003. "Religion and Positive Adolescent Outcomes: A Review of Research and Theory." *Review of Religious Research*, 44(4): 394-413.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika*, 70(1): 41-55.

Rosenbaum, Paul R., and Donald B. Rubin. 1985. "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score." *American Statistician*, 39(1): 33-38.

Ruhm, Christopher J. 2005. "Maternal Employment and Adolescent Development," IZA Discussion Papers 1673, Institute for the Study of Labor (IZA).

Sabia, Joseph. 2006. "Does Early Adolescent Sex Cause Depressive Symptoms?" *Journal of Policy Analysis and Management*, 25(4): 803-25.