Aggregation Reversals

and the

Social Formation of Beliefs

by

Edward L. Glaeser

Harvard University and NBER

and

Bruce Sacerdote[*]

Dartmouth and NBER

October 19, 2007

Abstract

In the past two elections, richer people were more likely to vote Republican while richer states were more likely to vote Democratic. This switch is an aggregation reversal, where an individual relationship, like income and Republicanism, is reversed at some level of aggregation. Aggregation reversals can occur when an independent variable impacts an outcome both directly and indirectly through a correlation with beliefs. For example, income increases the desire for low taxes but decreases belief in Republican social causes. If beliefs are learned socially, then aggregation can magnify the connection between the independent variable and beliefs, which can cause an aggregation reversal. We estimate the model's parameters for three examples of aggregation reversals, and show with these parameters that the model predicts the observed reversals.

I. Introduction

Richer people are more likely to vote Republican: Figure 1 shows this connection using 2000 data from the National Annenberg Election Survey. Figure 2 shows that richer states are more likely to vote Democratic in the same election. More educated people attend church more regularly, as shown in Figure 3, but as Figure 4 shows, more educated denominations have far lower attendance rates (Glaeser and Sacerdote, 2002). Figure 5 shows the positive individual level relationship between income and service in the military; Figure 6 shows the negative state level relationship between the same two variables.[1]

These three examples are aggregation reversals where a statistical relationship at the individual level is reversed at some level of aggregation, like the state or denomination. On a statistical level, aggregation reversals illustrate the ecological fallacy that aggregate relationships easily inform us about individual level parameters (e.g. King, 1997). In this literature, aggregation reversals can result from an unusual distribution of omitted variables across units of aggregation. But this insight does not help us to understand the economic causes of aggregation reversals could help us to generate predictions about when these reversals might occur.[2]

In Section II, we present two closely related models of aggregation reversals. While we start with a particularly simple model that creates aggregation reversals through two related decisions, we focus on an economic model of aggregation reversals that relies on the social formation of beliefs. That model assumes one exogenous variable, which could be education or income, and one endogenous outcome, like voting Republican or Church attendance. In either model, two ingredients are needed to generate an aggregation reversal. First, the exogenous variable must have two different effects on the outcome that go in opposite directions. For example, higher incomes might be

---

[1] Data sources are described in the data appendix.
[2] One way of understanding the difference between our approach and the ecological inference literature is that this literature sees aggregate information as a means of inferring individual level parameters, while we see aggregate relationships as intrinsically interesting.

associated both with a dislike for taxes and more liberal social beliefs. The dislike of taxes pushes richer people towards Republicanism; the liberal social views push them away from Republicanism. Second, there must be a social multiplier that operates on the indirect channel. In our belief model, this social multiplier exists because beliefs reflect social learning which means that the aggregate relationship between beliefs and the exogenous variable is much stronger than the individual relationship between those two variables.

None of these assumptions will always hold. Exogenous variables are often unrelated to beliefs. Often, variables that increase the direct returns to some activity will also be correlated with beliefs that make the activity even more attractive. In that case, the social formation of beliefs will create a more standard social multiplier (as in Glaeser, Scheinkman and Sacerdote, 2003). The unusual nature of our assumptions is compatible with the fact that aggregation reversals are themselves unusual.

If these two assumptions do hold, then the model predicts when aggregation reversal will occur. For an aggregation reversal, the ratio of the direct effect of the exogenous variable on the outcome to the opposite indirect effect must lie between an upper and lower bound. The lower bound is somewhat greater than one. The upper bound is the social multiplier related to the indirect effect, which in our primary model lies in the formation of beliefs. When this social multiplier is larger because beliefs strongly reflect social influence, then aggregation reversals become more likely. Aggregation reversals are also more likely when the beliefs correlate strongly with the exogenous variable. Greater sorting across groups on the basis of the exogenous variable makes a reversal less likely.

In Section III, we present an empirical methodology to calibrate our model. We use individual level relationships and the aggregate relationship between beliefs and the exogenous variable to predict the aggregate relationship between the exogenous variable and the outcome. The goal of this calibration is to show the plausibility of our model, not to reject alternative explanations for the aggregation reversal.

In Section IV, we look at the relationship between income and voting for George Bush in 2000. The rich were more likely to vote for Bush but less likely to agree with Republican social policies on prayer in schools or limits on abortion. The negative relationship between social policy views and income is much stronger at the state level than at the individual level, which suggests a sizable social multiplier. Finally, views on social policy are highly correlated with voting for Bush, which implies that these beliefs impact the outcome. Together these facts predict an aggregation reversal that is larger than the one we empirically observe.

In Section V, we return to the relationship between education and religion discussed by Glaeser and Sacerdote (2002). Education is positively associated with church attendance but negatively associated with belief in the devil or the literal truth of the Bible. The denominational correlation between education and these beliefs is much stronger than the individual correlation, which again suggests a large social multiplier. These beliefs are also highly correlated with attendance and seem to be important determinants of religious observance. Our parameter estimates again predict an aggregation reversal that is bigger than the one that we see in the data.

In Section VI, we examine the relationship between income and military service. Income is positively correlated with military service but negatively associated with pro-military beliefs. The income-belief relationship gets stronger at the state level, but the implied social multiplier is smaller than in our two other examples. The relationship between pro-military beliefs and military service is also weaker than the relationship between religious beliefs and religious attendance or social policy beliefs and voting Republican. Our results are mixed. Using one of our belief measures, we almost exactly predict the actual aggregation reversal. The other belief measure fails to predict any aggregation reversal. Section VII concludes.

II. The Social Formation of Beliefs and Aggregation Reversals

To illustrate the basic mechanism that can explain aggregation reversals, we first present an extremely stylized model where individuals choose two variables Y_i and A_i. We will later treat Y_i as an outcome and A_i as a belief, but we now treat them as two symmetric choice variables.   Individuals maximize:

$$(1)\quad -\frac{1}{2}\left(Y_i - \lambda_Y \hat{Y}_j - \beta X_i - \epsilon_i\right)^2 - \frac{1}{2}\left(A_i - \lambda_A \hat{A}_j - \delta X_i - \varepsilon_i\right)^2 + \gamma A_i Y_i,$$

where X_i is an exogenous variable, $\hat{Y}_j$ is the community average of Y_i in community j, $\hat{A}_j$ is the community average of A_i in community j, ϵ_i and ε_i are noise terms, and β, δ, γ, λ_Y, and λ_A are parameters. We have made two important assumptions. First, people have social preferences and they want their choices of both Y and X to be close to the community average. Second, the choice of A and the choice of Y are linked through an assumed complementarity between the two variables.   We also make a technical assumption that $(1-\lambda_A)(1-\lambda_Y) > \gamma^2, 1 > \lambda_Y$, and $1 > \lambda_A$ to ensure that second order conditions will hold and that the system will be stable.   With these assumptions, optimization yields that

$$(2)\quad Y_i = \frac{\beta + \gamma\delta}{1-\gamma^2} X_i + \frac{\gamma\delta(\lambda_A + \lambda_Y - \lambda_A\lambda_Y) + \beta(\gamma^2\lambda_A + \lambda_Y - \lambda_A\lambda_Y)}{((1-\lambda_A)(1-\lambda_Y) - \gamma^2)(1-\gamma^2)} \hat{X}_j + noise \text{ and}$$

$$(3)\quad \hat{Y}_j = \frac{\beta(1-\lambda_A) + \gamma\delta}{(1-\lambda_A)(1-\lambda_Y) - \gamma^2} \hat{X}_j + noise$$

where $\hat{X}_j$ is the community average of X_i, in community j. An aggregation reversal occurs when the individual level relationship between X and Y has a sign different from the aggregate relationship between X and Y. Such a reversal requires β to have the opposite sign from $\gamma\delta$ (if they were of the same sign, then we would be looking at a

social multiplier effect), so we assume that while β and γ are positive, δ is negative. The "X" variable must increase the return to Y and decrease the return to A. With this assumption, a reversal occurs if:

$$(4) \quad \frac{1}{1-\lambda_A} > \frac{\beta}{\gamma|\delta|} > 1 + \frac{(1-\gamma^2)\lambda_A \frac{Cov(X_i,\hat{X}_j)}{Var(X_i)}}{(1-\lambda_A)(1-\lambda_Y)-\gamma^2+\frac{Cov(X_i,\hat{X}_j)}{Var(X_i)}(\gamma^2\lambda_A+\lambda_Y-\lambda_A\lambda_Y)}$$

The term $\frac{\beta}{\gamma|\delta|}$ is the ratio of the direct impact that the exogenous variable, X, has on the endogenous variable, Y, with the indirect impact that X has on Y working through the second choice variable A. Equation (4) implies that aggregation reversals require $\frac{\beta}{\gamma|\delta|}$ must be smaller than $\frac{1}{1-\lambda_A}$, which would be the social multiplier for A if Y was fixed. If Y were fixed, then $\frac{1}{1-\lambda_A}$ would be the ratio of the within-group individual level relationship between X and A and the across-group relationship between X and A.   The first part of the inequality is therefore more likely to hold when social preferences are particularly important in the formation of A.

The ratio $\frac{\beta}{\gamma|\delta|}$ must also be greater than one plus a term that will be small if the degree of sorting on the basis of X across groups is small.   As the degree of sorting becomes perfect, i.e. as $\frac{Cov(X_i,\hat{X}_i)}{Var(X_i)}$ becomes close to one, the left hand side converges to $\frac{1}{1-\lambda_A}$ and an aggregation reversal becomes impossible.

This stylized model suggests that aggregation reversal require that an exogenous variable impacts the outcome through two separate channels and that these channels work in opposite directions.   Aggregation reversals also need a significant social multiplier operating in the second channel and the degree of sorting on the basis of the exogenous variable to be small.

We now turn to a related model that emphasizes the social formation of beliefs. There are few areas where the power of social influence has been so strongly observed than in the formation of beliefs, so this is a particularly natural place to look for the conditions needed to create aggregation reversals.

The Model with Beliefs

We now assume that individuals continue to choose a value of Y_i and that their objective function is to minimize the expectation of $-\frac{1}{2}(Y_i - \beta X_i)^2 + \gamma A_i Y_i$. We have eliminated the noise term, the direct social influence on the choice of Y, and the choice of A. We now think of A as an unknown variable that individual's must estimate. Utility maximization implies $Y_i = \gamma E(A) + \beta X_i$. The variable X_i continues to represent any variable, perhaps income or education, which might plausibly increase (or decrease) the benefits of the activity; β and γ are parameters are assumed to be positive so that X_i has a direct impact on the choice of Y_i. To produce an aggregation reversal, X_i must also have an indirect impact through the expectation of A.

The expectation of A is based on a common prior, a private signal and the communication of those signals within a community group. The common prior for A has mean zero and variance σ_A^2. Each individual receives a private signal equal to $A + cX_i + \kappa + \eta_i$, where η_i is an individual-specific normally distributed error term with mean zero and variance σ_η^2, and κ is a normally distributed common error term with mean zero and variance σ_κ^2. The cX_i term reflects the possibility that signals may be correlated with individual background attributes. Through communication, people learn the signals of a set of Z neighbors. The neighbors' signals are garbled in transmission so that individual i learns $A + cX_z + \kappa + \eta_z + \mu_z^i$ from individual z, where μ_z^i is the mean-zero garbling with variance σ_μ^2.

The communication of signals is one way of capturing the social learning that is critical to our model. This social learning creates the possibility for a correlation between an exogenous variable and a belief to grow with the level of aggregation. Asch (1955) pioneered a psychological literature showing that base their stated beliefs on the statements of others, even about something so seemingly obvious as the length of a line. While some authors have suggested that Asch's results show just a tendency to conform in statements, not in beliefs, Berns et al. (2005) use brain scans to show that statements by peers about shapes cause activity in the parts of the brain associated with spatial analysis not in the parts of the brain associated with social relations. Bandura (1977) is a classic text on the importance of social learning. Within economics, Merlo and Schotter (2001) is just one of the many papers on the power of social learning in experimental settings.[3] There is abundant evidence on the importance of learning from people around us.

We consider two cases, both of which are structured to eliminate learning about c. First, we assume that $\sigma_\kappa^2 > 0$, that c is known to equal zero, and that σ_η^2 is a function of X, $v(X_i)$. Second, we assume that $\kappa = 0$ and c>0, but that individuals have an incorrect assessment of c, denoted $\overline{c}$, that they fail to update. In this case, we assume that σ_η^2 is a constant. The first case assumes that individual attributes do not bias signals, but create more or less precision. In this case, actors are fully rational.[4] The second case assumes that the individual's attributes bias the signal and this bias is not fully corrected. This second case assumes limited rationality. In both cases, the values of X are known.

In the first case, standard signal extraction implies that the estimate of A equals:

[3] There is also an extensive theoretical literature on social learning in economics following Kandori, Mailaith and Rob (1993) and Ellison and Fudenberg (1993, 1995).
[4] A variant of this assumption might be that an individual attribute causes over-optimism about the precision of the signal, as in Scheinkman and Xiong (2003).

$$E(A) = \sigma_A^2 \frac{\left(P_i(X_i) + \sum_z P_z(X_z)\right)(A + \kappa) + P_i(X_i)\eta_i + \sum_z P_z(X_z)\left(\eta_z + \mu_z^i\right)}{1 + \left(\sigma_A^2 + \sigma_\kappa^2\right)\left(P_i(X_i) + \sum_z P_z(X_z)\right)},$$

where $P_i(X_i) = \dfrac{1}{v(X_i)}$ and $P_z(X_z) = \dfrac{1}{v(X_z) + \sigma_\mu^2}$. We will use a linear approximation around $X_i = X_z = 0$ to relate this formula to the "X" terms which determine these precisions:

(5') $E(A) \approx K_1 + \delta_1 X_i + \delta_1 \dfrac{ZP_z'(0)}{P_i'(0)} \dfrac{\sum_z X_z}{Z} + noise$,

where $K_1 = \dfrac{(A + \kappa)(P_i(0) + ZP_z(0))}{\Delta}$, $\Delta = \dfrac{1}{\sigma_A^2}\left(1 + \left(\sigma_A^2 + \sigma_\kappa^2\right)(P_i(0) + ZP_z(0))\right)$, and

$\delta_1 = \dfrac{P_i'(0)(A + \kappa)}{\sigma_A^2 \Delta^2}$. If we assume that the average characteristics of the communicating neighbors equal the average characteristics of people in community j, denoted $\hat{X}_j$, then the equation can be rewritten:

(5") $E(A) \approx K_1 + \delta_1 X_i + \lambda_1 \delta_1 \hat{X}_j + noise$,

where $\lambda_1 = \dfrac{ZP_z'(0)}{P_i'(0)}$. Values of "X" that reduce the variance of signals will change posterior beliefs and, on average, increase the accuracy of those beliefs. Of course, in settings when that error term is sufficiently large, changes in "X" that increase the precision of signals can actually make realized ex post assessments more biased.

In the second case where $\kappa = 0$ and $c > 0$, we assume that individuals believe that c equals $\bar{c}$ with certainty. We continue to use the same precision notation as above,

$P_i = \dfrac{1}{\sigma_\eta^2}$ and $P_z = \dfrac{1}{\sigma_\eta^2 + \sigma_\mu^2}$, but in this case, these are constants, not functions of X. In this case, the signal extraction formula delivers:

$$(6)\quad E(A) = \frac{\dfrac{A + (c - \bar{c})x_i + \eta_i}{\sigma_\eta^2} + \sum_z \dfrac{(A + (c - \bar{c})x_z + \eta_z + \mu_z^i)}{\sigma_\eta^2 + \sigma_\mu^2}}{\dfrac{1}{\sigma_A^2} + \dfrac{1}{\sigma_\eta^2} + \dfrac{Z}{\sigma_\eta^2 + \sigma_\mu^2}}.$$

Using $\dfrac{\sum_z X_z}{Z} = \hat{X}_j$, we have

$$(6')\quad E(A) = K_2 + \delta_2 X_i + \lambda_2 \delta_2 \hat{X}_j + noise$$

where $K_2 = \dfrac{A(P_i + ZP_z)}{\dfrac{1}{\sigma_A^2} + P_i + ZP_z}$, $\delta_2 = \dfrac{(c - \bar{c})P_i}{\dfrac{1}{\sigma_A^2} + P_i + ZP_z}$ and $\lambda_2 = Z\dfrac{P_z}{P_i}$.

Both models of learning deliver a linear formula where the expected value of A equals a constant, plus a slope parameter times the individual's own "X" characteristic plus a multiplier times that slope parameter times the average value of "X" in the individual's group. In the first learning process, the linear formulation is an approximation to a fully rational non-linear signal extraction formula. In the second formulation, the linear formulation is an exact representation of an imperfectly rational learning process.

We will use the linear formulation $E(A) = K + \delta X_i + \lambda \delta \hat{X}_j + \varepsilon_i$, which might reflect either learning process. This formula implies that average belief in community j equals $K + (1 + \lambda)\delta \hat{X}_j$ plus any noise that is not averaged away. The social learning means that the impact of "X" on group level beliefs will be stronger than the impact of "X" on individual beliefs. The term $1 + \lambda$ is the ratio of the impact of "X" on beliefs at the

group level divided by the impact of "X" on beliefs within groups, which we refer to as the social multiplier.

While our signal extraction description describes only one form of social learning, the reduced form relationship $E(A) = K + \delta X_i + \lambda \delta \hat{X}_j + \varepsilon_i$ does seem to capture the robust empirical relationships between many beliefs and exogenous variables, such as the connections between education and belief in heaven or income and the belief that abortion is wrong. Gentzkow and Shapiro (2004) show the relation between different forms of education in the Islamic world and beliefs about who is responsible for the September 11, 2001, attack on the World Trade Center. DiTella, Galiani, and Schargrodsky (2007) find that beliefs about capitalism are altered with an allocation of property rights. Our belief formation equation is simply meant to admit the possibility that beliefs are correlated with individual and aggregate characteristics.

With this formulation of beliefs, individual outcomes satisfy:

(8) $\quad Y_i = \gamma K + (\gamma \delta + \beta) X_i + \gamma \lambda \delta \hat{X}_j + \varepsilon_i,$

and group outcomes equal: $\hat{Y}_j = \gamma K + (\gamma \delta (1 + \lambda) + \beta) \hat{X}_j + \hat{\varepsilon}_j$. Table 1 lists the model's predictions about the coefficients when outcomes are regressed on the exogenous variable.

Since γ and β are always positive, when δ is also positive, then the impact of social learning will be to create a social multiplier where aggregate coefficients are bigger than individual coefficients. In that probably more common case, people with higher values of "X" choose higher levels of "Y" both because of the direct effect and the indirect effect through beliefs. The belief effect becomes larger at the aggregate level because of social learning and this causes the aggregate coefficient to be larger than the individual coefficient.

The more interesting case that can explain aggregation reversals occurs when δ is negative, and the indirect effect of "X" on "Y" that works through beliefs goes in the opposite direction of the direct effect of "X" on "Y". An aggregation reversal requires the individual-level regression coefficient of "Y" on "X" to be positive, which means that

$\beta > \gamma|\delta|\left(1+\lambda\dfrac{Cov(X_i,\hat{X}_j)}{Var(X_i)}\right)$. An aggregation reversal also requires the aggregate

regression coefficient of $\hat{Y}$ on $\hat{X}$ to be negative, which requires $\gamma|\delta|(1+\lambda)>\beta$. Putting these conditions together implies that an aggregation reversal requires:

$$(9) \qquad\qquad 1+\lambda > \frac{\beta}{\gamma|\delta|} > 1+\lambda\frac{Cov(X_i,\hat{X}_j)}{Var(X_i)}$$

This inequality is closely analogous to equation (4) in the previous version of the model. Just as before, aggregation reversals require that the ratio of direct to indirect impact of X, $\frac{\beta}{\gamma|\delta|}$, must fall within a range. The upper bound on the range is again a social multiplier, but in this case the social multiplier refers to the connection between the exogenous variable X and beliefs about A. The lower bound is again one plus a term that is small when $\dfrac{Cov(X_i,\hat{X}_j)}{Var(X_i)}$ is small. When $\dfrac{Cov(X_i,\hat{X}_j)}{Var(X_i)}$ is close to zero, then the inequality means that the ratio of direct to indirect effects must be greater than one and less than the social multiplier.

The model predicts when we should expect aggregation reversals. First, as we assumed above, there must be a correlation between the exogenous variable and beliefs that works in the opposite direction as the direct relationship between the exogenous variable and the outcome. Second, there must be a sizable social multiplier in the formation of beliefs, so that the relationship between beliefs and the exogenous variable is larger at the aggregate level than at the individual level. In the next section, we discuss our empirical strategy

for estimating these parameters to assess whether the model can explain the aggregation reversals shown in Figures 1 through 6.

III. Discrete Outcomes and Empirical Implementation

In this section, we discuss an empirical approach to our model of aggregation reversals. Our approach is to estimate the key parameters using a set of empirical moments that excludes the aggregate relationship between exogenous variables and outcomes and then to see if those parameters predict the aggregation reversal. This approach is certainly not an attempt to refute any alternative models of the observed aggregation reversals. Our goal is just to show that our model is at least a plausible explanation of these phenomena.

In keeping with the model, we will continue to focus on one independent variable in each case, either income or education. The model is aimed at explaining an aggregation reversal of a univariate relationship, but we neither mean to suggest that other variables do not matter, nor to suggest that our coefficients are causal. In all cases, we are comfortable with the interpretation that the estimated univariate relationship reflects both the impact of the independent variable and other omitted variables that are correlated with that variable. The relationship between income and Republicanism that we examine surely reflects many things that are correlated with income. This is not a problem for the model and we will not try to isolate any effects that reflect income alone.

We presented the model in a continuous formulation to make it more intuitive, but our examples will involve discrete outcomes and discrete beliefs. To move from the theory to the data, we assume that the observed outcome variable takes on a value a zero or one, which captures voting for Bush or attending church or owning a gun. We assume that individuals choose an outcome of one when $Y_i = \gamma K + (\gamma\delta + \beta)X_i + \gamma\lambda\delta\hat{X}_j + \varepsilon_i,$ is positive. We assume that all of the relevant noise terms are normally distributed, and let $F_z(.)$ denote the cumulative distribution function and $f_z(.)$ denote the density for a normal random variable z.

The expected share of people living in place j who choose the positive outcome will equal $F_{z_1}\left(\gamma K + (\gamma\delta(1+\lambda) + \beta)\hat{X}_j\right)$, where $z_1 = (\gamma\delta + \beta)(X_i - \hat{X}_j) + \varepsilon_i$. We assume that z_1 has the same distribution in each community. The average marginal effect of $\hat{X}_j$ on the share of the population that chooses one will equal $(\gamma\delta(1+\lambda) + \beta)\int_j f_{z_1}\left(\left(\gamma K + (\gamma\delta(1+\lambda) + \beta)\hat{X}_j\right)\right)\frac{1}{J}dj$, if there are measure J communities, so the average marginal effect of $\hat{X}_j$ on $\hat{Y}_j$ continues to be $\gamma\delta(1+\lambda) + \beta$ as shown in Table 1.

Within a community, for a given value of X, the share of people who choose one will equal $F_{\varepsilon_i}\left(\gamma K + (\gamma\delta + \beta)X_i + \gamma\lambda\delta\hat{X}_j\right)$. If the distribution of X within the community is characterized by a density function g(X), the estimated marginal effect of "X" is $(\gamma\delta + \beta)\int_{x_i} f_{\varepsilon_i}\left(\gamma K + (\gamma\delta + \beta)x_i + \gamma\lambda\delta\hat{X}_j\right)g(x_i)dx_i$, so the average effect of x within communities is again $\gamma\delta + \beta$. We will use the estimated within group coefficient to provide us with an estimate of $\gamma\delta + \beta$.

The estimated marginal effect of X on Y across the entire population equals $\gamma\delta + \beta + \gamma\delta\lambda\dfrac{\partial E(\hat{X}_j)}{\partial X_i}$, and $\dfrac{\partial E(\hat{X}_j)}{\partial X_i}$ will again equal $\dfrac{Cov(X_i, \hat{X}_j)}{Var(X_i)}$. As such, the existence of discrete outcome variables does require Probit estimation techniques, but it does not change the connection between the parameters of the model and relationship between exogenous variable and outcomes.

We assume that the "X" variable is mean zero and variance one and we will normalize our independent variables appropriately in the empirical work. The parameter $\dfrac{Cov(X_i, \hat{X}_j)}{Var(X_i)}$ can be measured directly from the data, as it reflects the share of variation in "X" that is across group rather than within group. This parameter can be estimated with the share of the variation in "X" that is explained by group dummies. We will scale

14

our "X" variables so that they have a mean of zero and a variance of one, so the missing parameter is just the $Cov(X_i, \hat{X}_j)$.

We also assume that we do not measure beliefs directly but rather see several discrete measures of beliefs that take on values of one when $\psi_k E(A) + \xi_i^k$ is positive. For example, in the case of religious belief, we have questions on statements about belief in the devil and beliefs in the literal truth of the Bible. The parameter ψ_k reflects the relationship between the relevant beliefs and the particular discrete measure. The term ξ_i^k is an error term specific to the person and the measure of beliefs. By scaling δ and the error terms appropriately, we can always ensure that the latent belief variable has a variance of one. We also normalize $1 - \psi_k^2 = Var(\xi_i^k)$ so that the variance of the belief proxies also equals one. These assumptions are innocuous scaling assumptions about unobserved latent variables that drive the zero-one decision.

Within a given group, the share of the population that answers yes to belief question k is $F_{z_2}\left(\psi_k (K + (1 + \lambda)\delta \hat{X}_j)\right)$ where $z_2 = \psi_k\left(\delta(X_i - \hat{X}_j) + \varepsilon_i\right) + \xi_i^k$. The estimated marginal effect of $\hat{X}_j$ will be $\psi_k \delta(1 + \lambda) \int_j f_{z_2}\left(\psi_k (K + (1 + \lambda)\delta \hat{X}_j)\right) \frac{1}{J} dj$ and the average effect will be $\psi_k \delta(1 + \lambda)$. Within groups, the estimated marginal effect of x on beliefs will be $\psi_k \delta$ so the ratio will again give us the social multiplier. As such, we use the within group belief regressions to estimate $\psi_k \delta$ and the ratio of the coefficient on "X" from within group and across group regressions to estimate the social multiplier.

We estimate the parameter ψ_k by assuming that we have two potential proxies for the underlying beliefs. If ψ_k is the same for both of these proxies, then covariance of the two underlying normal variables is equal to ψ_k^2. This covariance is empirically implied by the means and covariance of the two discrete measures of underlying beliefs. If the two

values of ψ_k are not the same, then the potential range for ψ_k is between the underlying covariance and one and we can consider this entire range for this variable.

Our final moment is the relationship between beliefs and outcomes. In both cases, we observe a discrete proxy for the relevant underlying variables Y_i and $\psi_k E(A) + \xi_i^k$. Again, we can use the means of the two discrete variables and their covariance to estimate the underlying covariance of the two normal variables. The model predicts that this underlying covariance will equal $\psi_k \left(\gamma + \beta \delta (1 + \lambda Cov(X_i, \hat{X}_j)) \right)$.

Table 1 lists these set of predictions that we use to estimate the parameters. We need estimates of six parameters: β, γ, δ, λ, ψ_k and $Cov(X_i, \hat{X}_j)$. We use six moments to estimate these parameters: (1) the within group relationship between exogenous variable and outcome, (2) the amount of variation in the exogenous variable that is within group, (3) the within-group effect of the exogenous variable on the proxy for beliefs, (4) the across-group effect of the exogenous variable on the proxy for beliefs, (5) the correlation of different proxies for beliefs and (6) the individual level covariance between the proxy for beliefs and the outcome variable. We will then see whether these parameters predict an aggregation reversal and whether they come close to predicting the observed aggregate relationship between the outcome and the exogenous variable.

The square root of the covariance of the latent belief variables provides our estimate of ψ_k. The share of variation in X that is explained by group dummies estimates $Cov(X_i, \hat{X}_j)$. The ratio of the aggregation relationship between the exogenous variable and the belief proxy and the within-group relationship between the exogenous variable and the belief proxy delivers $1 + \lambda$. The ratio of the within-group relationship between the exogenous variable and the belief proxy and the square root of the covariance of the latent belief variables delivers δ.

The value of β is found by subtracting δ/ψ_k times the estimated covariance of the latent belief proxy and the latent outcome from the estimated within group marginal effect of the exogenous variable on the outcome and then dividing by $1-\delta^2-\delta^2\lambda Cov(X_i,\hat{X}_j)$. The value of γ is found by subtracting this estimate of β from the estimated within-group effect of the exogenous variable on the outcome and then dividing by the estimate of δ.

IV.   Income and Republicanism

The aggregation reversal that occurs in the relationship between income and Republicanism is quite striking. At the individual level, there is a modest positive relationship between earnings and voting for President Bush in 2000. According to the National Annenberg Election Study of 2000, 55 percent of the top quintile of the income distribution voted for the Republican in 2000; 36 percent of the bottom quintile of the income distributed voted for him in the same year. This positive income-Republicanism relationship certainly corresponds with popular notions of Republicanism, but those notions are seemingly contradicted by the profoundly negative relationship between income and Republicanism at the state level shown in Figure 2.   The correlation coefficient is $-.57$.

The model seems to have a reasonable chance of explaining this aggregation reversal because of the multiple aspects of Republicanism.   In post-Reagan America, Republicanism has been associated with both lower taxes, which presumably appeal to the wealthy and conservative social stances (Glaeser, Ponzetto and Shapiro, 2005). Republicans have regularly championed school prayer and limits on abortion. The appeal of these social stances depends on conservative social beliefs that are not positively correlated with income.   In the Annenberg data set, the correlation between income and the propensity to say that the government should not put limits on abortion is 10 percent. The combination of a direct relationship between income and the financial returns from low tax Republican policies and an indirect relationship where income decreases

conservative social beliefs which decreases the support for Republicanism suggests that our model may indeed explain the observed aggregation reversal.

To test this hypothesis, we use the Annenberg data to estimate the parameters of the model. We will then see whether these estimated parameters predict an aggregation reversal and the coefficient seen in the aggregate data. We estimate an aggregate marginal effect of log income on the propensity to vote for Bush of -.274 using Federal Election Commission voting records and 2000 Census data on income, with a standard error of .058. This estimation is done using nonlinear least squares to fit the data to a cumulative normal function. This reflects our assumption that the underlying heterogeneity in political preferences is normally distributed. The ordinary least squares estimate is -.272 with a standard error of .057.

We have two different belief variables that we use in our estimation: survey responses to questions as to whether "the federal government should not put limits on abortion," and "the government should not support prayer in schools." In both cases, we have reduced the answers to these questions to taking on two values. The first column of Table 2 shows the results using the first question; the second column shows the results using the second question.

Two of our empirical moments are independent of the choice of belief variable. The marginal impact of income on voting for Bush, our estimate of $\gamma\delta + \beta$, is found using a Probit equation with state fixed effects on the Annenberg data. As shown in the first row of Table 2, we estimate .079 for this coefficient. The estimate of $\dfrac{Cov(X_i, \hat{X}_j)}{Var(X_i)}$ is based on the amount of variation in income explained by state fixed effects. The second row of Table 2 shows our estimate of .023 for this parameter. We have also estimated this sorting parameter using the Census Individual Public Use Micro Samples (IPUMS) and found a similar parameter estimate of .022.

As described above, we need the correlation between the two questions to form our estimate of ψ_k which is based on the correlation of the two variables. The third row of Table 2 shows the estimated covariance of the latent belief variables of .143 which is our estimate of ψ_k^2.

The remaining rows in the table give estimates that differ between the two belief questions. In the fifth row, we show the estimated impact of income on beliefs within states, which is our estimate of $\psi_k \delta$. This estimate is -0.041 in the case of support for limits on abortion and -0.079 in the case of opposition to prayer in school. Income is more strongly negatively correlated with beliefs about prayer in school than with beliefs about limits on abortion. The sixth row shows the associated value of δ which is found by dividing this estimate by our estimate of ψ_k. This parameter estimate is -.109 in the case of the abortion question and -.210 for the prayer question.

The seventh row gives the state level marginal impact of the logarithm of income on the share of the population that answers yes to the two questions. In both cases, the estimated coefficient explodes. In the case of limits on abortion, the group-level coefficient rises to -.399 and in the case of school prayer the group level coefficient rises to -.477. The much stronger aggregate relationship supports the hypothesis of a social multiplier, although sorting on omitted variables could explain some part of the observed group level relationship.

In the eighth row, we report the value of λ implied by the ratio of group level effects to individual effects. In the case of abortion, this parameter is 8.73. In the case of school prayer it is 5.02. We also find evidence for social multipliers when we examine relationships at lower levels of aggregation. Perhaps it is surprising to estimate social multipliers that are this strong, but these findings are certainly in line with the social psychology literature that argues that beliefs like these are very much the product of social interactions.

In the ninth row, we show the estimated covariance of the latent outcome variable (support for Bush) and the latent belief variable. There is a higher covariance in the case of the abortion question, .219, then in the case of the school prayer question, .148. This higher correlation reflects the stronger connection between Republicanism and views on abortion than between Republicanism and views on school prayer.

The tenth and eleventh rows give our estimates of γ and β, the impact of beliefs on voting for Bush and the direct effect of income on voting for Bush. In both cases, beliefs seem to have a much stronger impact on voting Republican than income. We estimate γ coefficients of .598 and .432 for the abortion and prayer questions respectively. When we use the abortion question, we estimate a direct effect of income (β) of .144. When we use the prayer question, we estimate a higher direct effect of income of .170.

The twelfth row gives the predicted group level relationship between income and voting for Bush implied by these parameters. In both cases, the model predicts a healthy aggregation reversal. Using the abortion question, we predict a negative relationship of -.488 and with the prayer relationship we predict a negative relationship of -.377. This predicted relationship should be compared to the actual aggregate relationship of -.274 shown in the thirteenth row.

The predicted value found by using the school prayer question is reasonably close to the actual aggregate relationship; the prediction based on the abortion question is far too negative. Although given the imprecision of our estimates, it is not clear that we should make too much of these findings. We suspect that we are predicting too much of an aggregation reversal, because we are overestimating the size of the social multiplier in the case of the abortion question. For example, if we used the social multiplier estimated using the school prayer question but continued to use all of the other moments of the abortion question, then our predicted aggregate relationship would be much closer to the actual aggregate relationship.

V. Religion and Education

We now revisit the connection between education and religion discussed in Glaeser and Sacerdote (2002). At the individual level, more educated people attend church more often. In the General Social Survey, 53 percent of college graduates attend church once per month or more while 45 percent of high school dropouts attend that frequently. However, more educated denominations are far less religious than less educated denominations. Figure 4 shows the correlation between average years of education and church attendance across denominations.

As in the case of voting Republican, there are at least two different reasons to go to church. First, going to church is a conventional social activity that connects people within a community and provides certain type of services for children. Second, going to church is thought by some to yield otherworldly returns, such as going to heaven. Religious beliefs are surely based on social influences. Where else would most people come up with their views about the afterlife? These two different functions of church attendance—one of which is highly dependent on socially formed beliefs and one of which is not—suggest that our model can possibly explain the aggregation reversal.

Glaeser and Sacerdote (2002) provide evidence that education predicts participation in almost every formal social activity at the individual level. Education predicts membership in political clubs, fraternal clubs, hobbyist associations and even sports clubs. In this light, it seems unsurprising that education also predicts membership in religious groups and attendance at church. There are several possible explanations of this fact. Group membership may be seen as a form of investment in social capital and people who live investing in human capital may also see returns to social capital. Glaeser, Ponzetto and Shleifer (2006) argue that education includes heavy doses of socialization that increases the ability to interact effectively with others. If education directly increases the returns to social activity and if church-going is a social activity,

then this creates a direct effect where education should increase the amount of church attendance.

Religion is, of course, not just another social club. Most religious groups also promise some forms of otherworldly returns to religious adherence. Yet education generally predicts less belief in the supernatural. For example, belief in heaven, the devil and the literal truth of the bible all decline strong with years of education. There are several possible interpretations of this phenomenon. One view is that these phenomena are at odds with modern science and more education naturally includes more of the science that disproves religious belief. An alternative view argues that secular education is often anti-clerical and the negative relation between education and religion reflects the impact of secularist indoctrination. We take no view in this debate, but simply note the robust negative relationship between years of schooling and religious beliefs.

Whatever negative effects exist between education and religious beliefs at the individual level, they do appear to be magnified at the denomination level. Figure 9 shows the negative relationship between average years of schooling in a denomination and belief in the devil. This magnification may well reflect the social formation of beliefs. Educated denominations contain people who are less likely to be strong believers and who speak those views regularly. The more educated denominations are also more likely to have educated religious leaders who are less likely to be strong believers and who themselves determine the basic tenets of the denomination. More educated denominations are more likely to have sermons discussing nuanced ethical issues rather than strong statements affirming the literal truth of the bible or the damnation of non-believers.

We will use belief in the devil and the literal truth of the bible for our calibration. While these variables take on 4 and 3 values respectively in the General Social Survey, we reduce them to binary variables based on the mean response. Our calibrations appear in Table 3. In the first row, we describe the within denomination relationship between education and attendance of .031. Education has been normalized to have a standard deviation of one, so this coefficient means that a one standard deviation increase in

education is associated with a 3.1 percent increase in the probability that an individual will attend church regularly.

In the second row, we give the share of the variation in education that is explained by denomination fixed effects: 5.6 percent. Education is more closely linked to denomination than income is to states. The third row gives the correlation of latent belief variables implied by the correlation of our two proxies for beliefs and the fourth row shows that this implies a value for ψ_j of .412.

The fifth row gives the marginal impact of education on belief within denominations. In both cases, more educated people are less likely to have strong religious beliefs. A one standard deviation increase in education is associated with a 3.2 percent decrease in the propensity to say that the bible is the literal truth and an 11.8 percent decrease in the propensity to believe in the devil. When we divide these values by our estimate of ψ_k in the sixth row, we estimate of δ to be -.078 and -.287 for the bible and devil question respectively.

In the seventh and eighth rows, we show the results of the cross denomination regressions of beliefs on education and the implied social multiplier. In the case of the bible question, we estimate an aggregate coefficient of -.6 and a value for λ of 17.7 In the case of the devil question, somewhat appropriately, we estimate an aggregate coefficient of -.666 and estimate λ to be 4.6. The denomination level coefficients are almost identical and the difference in the social multiplier is driven primarily by the lower individual level relationship between education and belief in the literal truth of the bible.

The ninth row delivers the implied covariance between the latent belief variable and the latent attendance variable which is .148 in the case of the bible question and .13 in the case of the devil question. The ninth row shows that these covariances imply almost identical values of γ for the two belief questions: .369 and .365. In both cases, higher beliefs are strongly associated with greater propensity to attend church. The similar

estimates of γ imply quite different estimated values for β: .06 for the bible question and .139 for the devil question. The difference is driven by the much stronger correlation between education and belief in the devil than between education and belief in the literal truth of the bible.

The twelfth and thirteenth rows give the predicted and actual aggregate relationships between education and religious attendance. In both cases, the parameters predict a robust aggregation reversal. With the bible question, our parameters predict an aggregation reversal of -.479 and, with the devil question, we predict an aggregation reversal of -.454. The actual aggregate relationship is -.314. We again predict too negative a relationship, although our standard errors are sufficiently big that these differences are only modestly statistically different.

There are many possible explanations for our overly negative predictions. In the case of the bible question, we suspect that our estimate of the social multiplier is too high. In the case of the devil question, we may have overestimated the magnitude of δ, the connection between education and beliefs. As before, we think the lesson of this calibration is that the model is more effective at predicting the sign of the aggregate effect than at accurately predicting the actual size of the effect.

VI.   Income and Military

We now turn to our final aggregation reversal: the relationship between income and military service. Figure 5 shows the positive relationship between veteran status, which is defined as having had some military service, and family income at the individual level. This relationship we show uses data from the General Social Survey, but the Census shows similar results. This relationship surely reflects some amount of both treatment and selection, but when we look at parental occupation status in the General Social

Survey, we also find an -.02 percent correlation between higher income occupations and entering into the army. People from the lowest socioeconomic backgrounds tend not to join the army, either because they are not allowed in or because they choose not to join.

Figure 6 shows the relationship between income and military service at the state level using the 2000 census. In this case, we look only at people born after 1950 which means that we are looking almost mainly at people who have volunteered. Alternative cutoff dates make little difference. People from lower income states are much more likely to join the military than people from higher income states. The Northeast trio of Massachusetts, Connecticut and New Jersey are particularly unlikely to have members in the military.

One explanation for this phenomenon that corresponds well with standard prejudices about red and blue states is that people in high income areas are less likely to believe in army and military service. At the individual level, perhaps, relatively anti-war teachers perhaps tell students that fighting in the military is bad. Perhaps, high levels of income correlate with a dislike of the self-sacrifice that military service entails.

We use two belief variables to look at enthusiasm for the army. The first belief variable is a question about whether respondents think that the army is a good experience for men. The second belief variable is whether respondents have confidence in the army as an institution. In both cases, income and education are negatively correlated with the pro-military belief. In both cases, the correlations between income and beliefs at the state level are significantly stronger than the correlations at income and beliefs at the individual level.

Table 4 yields the parameter estimates for this example. The first row gives the within state relationship between income and military service. A one standard deviation in log of income is associated with a 5.9 percent increase in the probability of having a veteran in the household. The second column provides the variation in income explained by state dummies which is the same parameter that appeared in our first example.

The third and fourth columns display the estimated covariance of the latent belief variables and the implied value of ψ_k: .335. These belief variables are less correlated than the belief variables in the previous examples. The fifth row shows the estimated impact of income on beliefs within states and the sixth row shows the implied value of δ. We estimate that a one standard deviation increase in income is associated with a 10 percent decrease in pro-military beliefs, using the first question, and an eight percent decrease in pro-military beliefs when we use the second question. In this case, the estimates of δ using the two different questions seem reasonably close.

The seventh and eighth rows give the estimated impact of income on pro-military beliefs across states and the implied values for λ. The implied values of λ are 2.6 and 5.9. The confidence in the army question appears to have a much stronger social multiplier.

The ninth row gives the implied covariance of the latent belief variable with the outcome. In the case of the question about the military being a good experience, the correlation is robust. The tenth row shows that this correlation implies that γ equals .58 for the question about the benefits of military service. The latent correlation between confidence in the army and military service is almost zero and this implies a value of .037 for γ. This small connection between beliefs and outcomes makes an aggregation reversal unlikely.

In the eleventh row, we show the implied values of β which is .12 when we use the question about the military being a good experience and .06 when we use confidence in the army. These higher values in the case of the good experience come from a higher estimate of γ. When the impact of beliefs is estimated to be higher, then the direct effect of income must also be higher.

Finally in the twelfth and thirteenth rows, we look at the predicted and actual aggregate relationship between income and military service. In the case of the good experience

question, we predict an aggregate relationship of -.098 when the actual relationship is -.092. In the case of the confidence in the army question, our low estimate of γ means we fail to estimate an aggregate reversal and predict an aggregate coefficient of .042.

VII.   Conclusion

This paper presents a model of aggregation reversals where individual relationships have the opposite sign from group relationships. In the model, aggregation reversals occur when the exogenous variable impacts the outcomes through two channels. In one of the channels, there is a social multiplier, so the aggregate relationship between the exogenous variable and the outcome increases at higher levels of aggregation. We focus on the case where the exogenous variable is correlated with beliefs, because we believe social influence is a critical determinant of most beliefs. The model predicts that aggregation reversals will occur when the ratio of the direct effect of the exogenous variable on the outcome to the belief effect negative and is somewhat greater in absolute value than one and less than the social multiplier.

We applied the model to three examples of aggregation reversals. We did not try to rule out other alternative theories or identify causal parameters. Instead, we asked whether parameter values that were estimated from individual level relationships and the aggregate relationship between beliefs and the exogenous variable would predict the aggregation reversal that we see in the data. Our work does not rule out alternative explanations, but rather tries to establish some degree of plausibility for our theory.

We examined the relationship between income and voting for George Bush in 2000, the relationship between education and religious attendance and the relationship between income and military service. We used two different belief variables for each aggregation reversal yielding six different predictions about aggregate relationships. Our calibration exercise had successes and failures. In five out of six cases, the parameters did predict an aggregation reversal. Except for the one case where we don't predict an aggregation reversal, we estimate a large relationship between beliefs and outcomes and a large social

27

multiplier in the formation of beliefs. Together these tend to predict an aggregation reversal. In many cases, we predict too much of a reversal because our estimated social multiplier is quite large.

The paper suggests that the social formation of beliefs and the resulting social multipliers may be significant determinants of important aggregate phenomena. These three aggregation reversals are hardly the only such cases that exist. Figures 7 and 8 show the aggregation reversal in the relationship between gun ownership and income. In all of the cases we examined, we found substantial social multipliers in beliefs, but omitted variables that could be driving these findings. We hope that future empirical work will put more effort on more cleanly identified estimates of the magnitude of social learning.

References

Asch, Solomon E. (1955). "Opinions and Social Pressure", *Scientific American*, 193(5): 31-35.

Bandura, Albert (1977). *Social Learning Theory*. New York: General Learning Press.

Barberis, Nicholas, Andrei Shleifer and Robert Vishny (1998). "A Model of Investor Sentiment", *Journal of Financial Economics*, Vol. 49: 307-343.

Berns, G., J. Chappelow, C. Zink, G. Pagnoni, M. Martin-Skurski, J. Richards (2005). "Neurobiological Correlates of Social Conformity and Independence During Mental Rotation", *Biological Psychiatry*, Vol. 58, no. 3 (August): 245-253.

Di Tella, Rafael, Sebastian F. Galiani, and Ernesto S. Schargrodsky (2007). "The Formation of Beliefs: Evidence from the Allocation of Land Titles to Squatters", *Quarterly Journal of Economics* (forthcoming).

Ellison, Glenn and Drew Fudenberg (1995). "Word-of-Mouth Communication and Social Learning", *The Quarterly Journal of Economics*, Vol. 110, no. 1 (February): 93-125.

Ellison, Glenn and Drew Fudenberg (1993). "Rules of Thumb for Social Learning", *The Journal of Political Economy*, Vol. 101, no. 4 (August): 612-643.

Gentzkow, Matthew A. and Jesse M. Shapiro (2004). "Media, Education and Anti-Americanism in the Muslim World", *The Journal of Economic Perspectives*, Vol. 18, no. 3 (Summer): 117-133.

Glaeser, Edward L., Giacomo Ponzetto, and Jesse M. Shapiro (2005). "Strategic Extremism: Why Republicans and Democrats Divide on Religious Values", *Quarterly Journal of Economics*, Vol. 120, no. 4 (November): 1283-1330.

Glaeser, Edward L., Giacomo Ponzetto, and Andrei Shleifer (2006). "Why Does Democracy Need Education?", NBER Working Paper 12128, March 2006.

Glaeser, Edward L. and Bruce I. Sacerdote (2002). "Education and Religions", NBER Working Paper 8080, February 2002.

Glaeser, Edward L., Bruce Sacerdote, and Jose Scheinkman (2003). "The Social Multiplier", *The Journal of the European Economic Association*, Vol. 1, no. 2: 345-353.

Kandori, Michihiro, George J. Mailath, and Rafael Rob (1993). "Learning, Mutation and Long Run Equilibria in Games", *Econometrica*, Vol. 61, no. 1 (January): 29-56.

King, Gary (1997). *A Solution to the Ecological Inference Problem*. Princeton: Princeton University Press.

Merlo, Antonio M. and Andrew Schotter (2001). "Learning By Not Doing: An Experimental Investigation of Observational Learning", PIER Working Paper No. 01-040, October 2001.

Murphy, Kevin M. and Andrei Shleifer (2004) "Persuasion in Politics" *American Economic Review* 94(2): 435-439.

Scheinkman, Jose A. and Wei Xiong (2003) "Overconfidence and Speculative Bubbles," *Journal of Political Economy,* 1183-1219.

Table 1: Predictions of the Model		
Marginal Impacts of the Exogenous Variable on the Outcome Variable	Individual-Level	$\gamma\delta + \beta + \gamma\delta\lambda Cov(X_i, \hat{X}_j)$
	Individual-Level within Group	$\gamma\delta + \beta$
	Group Level	$\gamma\delta(1+\lambda) + \beta$
Variation in X explained by group level dummies		$Cov(X_i, \hat{X}_j)$
Marginal Impact of the Exogenous Variable on Proxies for Belief	Individual-Level within Group	$\psi_k\delta$
	Group Level	$(1+\lambda)\psi_k\delta$
Covariance of latent proxies for belief		ψ_k^2
Covariance of latent proxies for belief with latent outcome variable		$\psi_k\left(\gamma + \beta\delta + \beta\delta\lambda Cov(X_i, \hat{X}_j)\right)$

Table 2: Parameters for the Relationship between Income and Voting for Bush		
	Abortion Question	School Prayer Question
Within-state impact of income on voting for Bush ($\gamma\delta + \beta$)	.079 (.008)	.079 (.008)
Variation in Income explained by states $\dfrac{Cov(X_i, \hat{X}_j)}{Var(X_i)}$	.023 (.004)	.023 (.004)
Implied covariance of latent beliefs ψ_k^2	.143 (.025)	.143 (.025)
Implied value of ψ_k	.378 (.034)	.378 (.034)
Marginal Impact of Income on Belief within States ($\psi_k\delta$)	-.041 (.007)	-.079 (.013)
Implied value of δ	-.109 (.023)	-.210 (.039)
Marginal Impact of Income on Belief across States ($(1+\lambda)\psi_j\delta$)	-.399 (.007)	-.477 (.010)
Implied value of λ	8.73 (1.714)	5.018 (.939)
Implied covariance of latent belief with latent outcome $\psi_k\left(\gamma + \beta\delta(1+\lambda Cov(X_i, \hat{X}_j))\right)$	.219 (.015)	.148 (.027)
Implied value of γ	.598 (.070)	.432 (.082)
Implied value of β	.144 (.020)	.170 (.029)
Implied value of aggregate relationship $\gamma\delta(1+\lambda)+\beta$	-.488 (.116)	-.377 (.112)
Estimated value of aggregate relationship $\gamma\delta(1+\lambda)+\beta$	-.274 (.058)	-.274 (.058)

Table 3: Parameters for the Relationship between Education and Attendance		
Belief Question:	Bible is the Literal Truth	Belief in the Devil
Within-state impact of on education on attendance ($\gamma\delta + \beta$)	.031 (.003)	.031 (.003)
Variation in Education explained by denominations $\dfrac{Cov(X_i,\hat{X}_j)}{Var(X_i)}$	.056 (.002)	.056 (.002)
Implied covariance of latent beliefs ψ_k^2	.170 (.027)	.170 (.027)
Implied value of ψ_k	.412 (.033)	.412 (.033)
Marginal Impact of Income on Belief within denominations ($\psi_k\delta$)	-.032 (.015)	-.118 (.010)
Implied value of δ	-.078 (.036)	-.287 (.036)
Marginal Impact of Income on Belief across denomination ($(1+\lambda)\psi_j\delta$)	-.602 (.004)	-.666 (.002)
Implied value of λ	17.731 (113.18)	4.628 (.517)
Implied covariance of latent belief with latent outcome $\psi_k\left(\gamma + \beta\delta(1+\lambda Cov(X_i,\hat{X}_j))\right)$	.148 (.026)	.130 (.021)
Implied value of γ	.369 (.070)	.365 (.068)
Implied value of β	.060 (.015)	.136 (.027)
Implied value of aggregate relationship $\gamma\delta(1+\lambda)+\beta$	-.479 (.128)	-.454 (.123)
Estimated value of aggregate relationship $\gamma\delta(1+\lambda)+\beta$	-.314 (.070)	-.314 (.070)

	Military is a good experience	Confidence in the Army
Within-state correlation between income and military service $(\gamma\delta + \beta)$	.059 (.003)	.059 (.003)
Variation in Income explained by states $\dfrac{Cov(X_i, \hat{X}_j)}{Var(X_i)}$	.026 (.002)	.026 (.002)
Implied covariance of latent beliefs ψ_k^2	.112 (.031)	.112 (.031)
Implied value of ψ_k	.335 (.047)	.335 (.047)
Marginal Impact of Income on Belief within States ($\psi_k\delta$)	-.035 (.015)	-.027 (.003)
Implied value of δ	-.104 (.047)	-.082 (.016)
Marginal Impact of Income on Belief across States ($(1+\lambda)\psi_j\delta$)	-.125 (.003)	-.190 (.002)
Implied value of λ	2.577 (23.04)	5.923 (.758)
Implied covariance of latent belief with latent outcome $\psi_k\left(\gamma + \beta\delta(1+\lambda Cov(X_i, \hat{X}_j))\right)$	.191 (.023)	.011 (.009)
Implied value of γ	.583 (.109)	.037 (.028)
Implied value of β	.119 (.036)	.062 (.004)
Implied value of aggregate relationship $\gamma\delta(1+\lambda)+\beta$	-.098 (.065)	.041 (.015)
Estimated value of aggregate relationship $\gamma\delta(1+\lambda)+\beta$	-.092 (.028)	-.092 (.028)

Table 4: Parameters for the Relationship between Military Service and Income

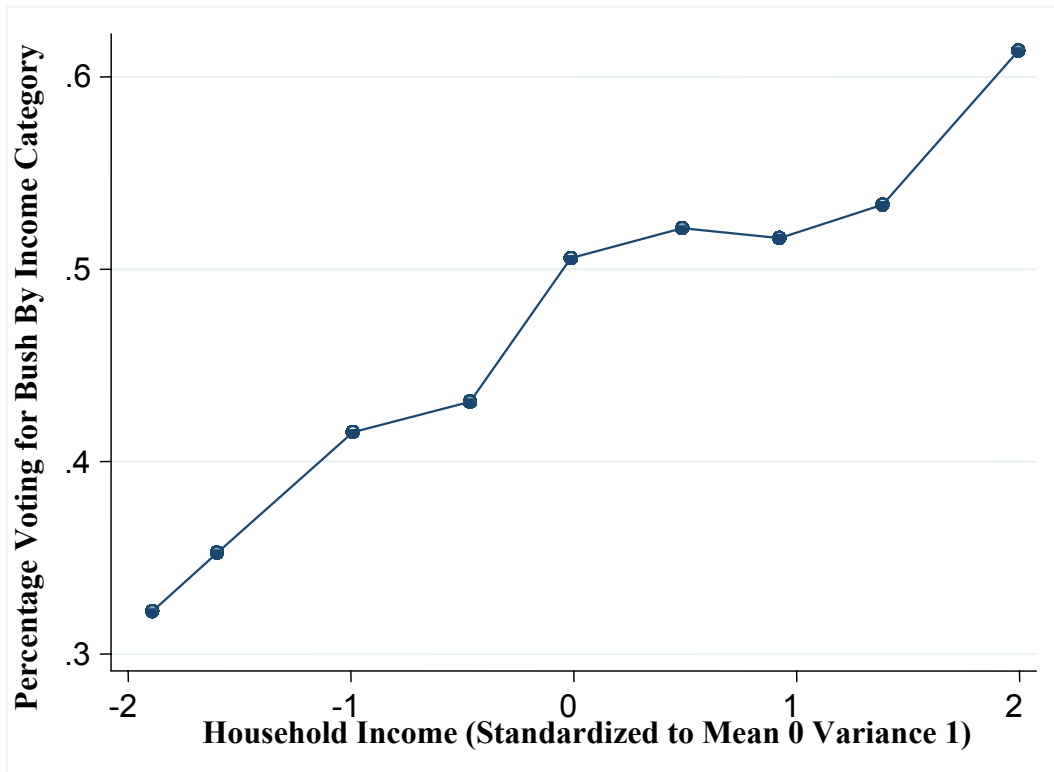## Figure 1: Voting for Bush and Income: Micro Data

Figure 2: Voting for Bush and Income: State Data
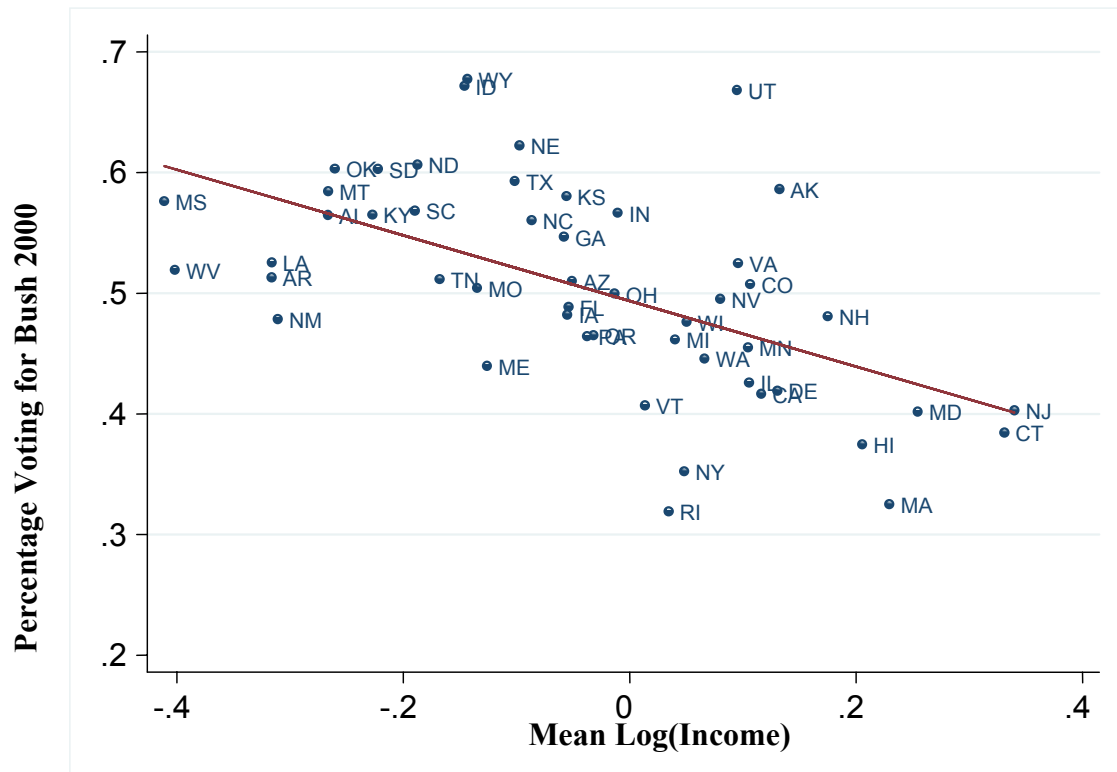
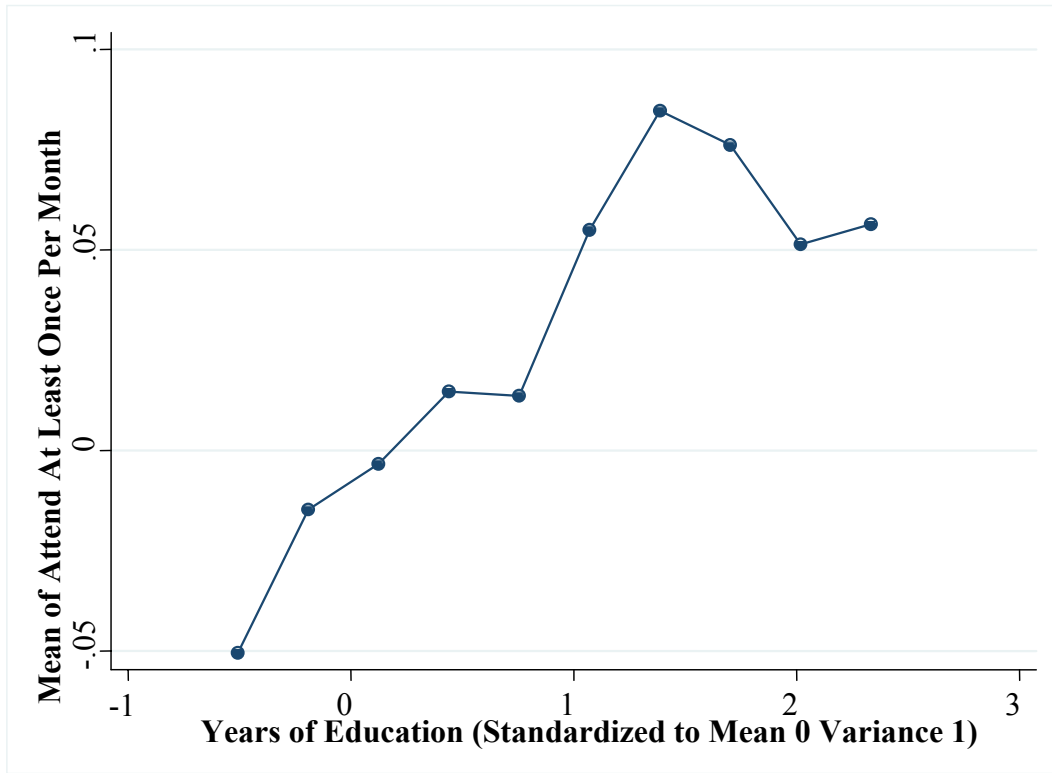**Figure 3: Religious Attendance and Education: Micro Data**

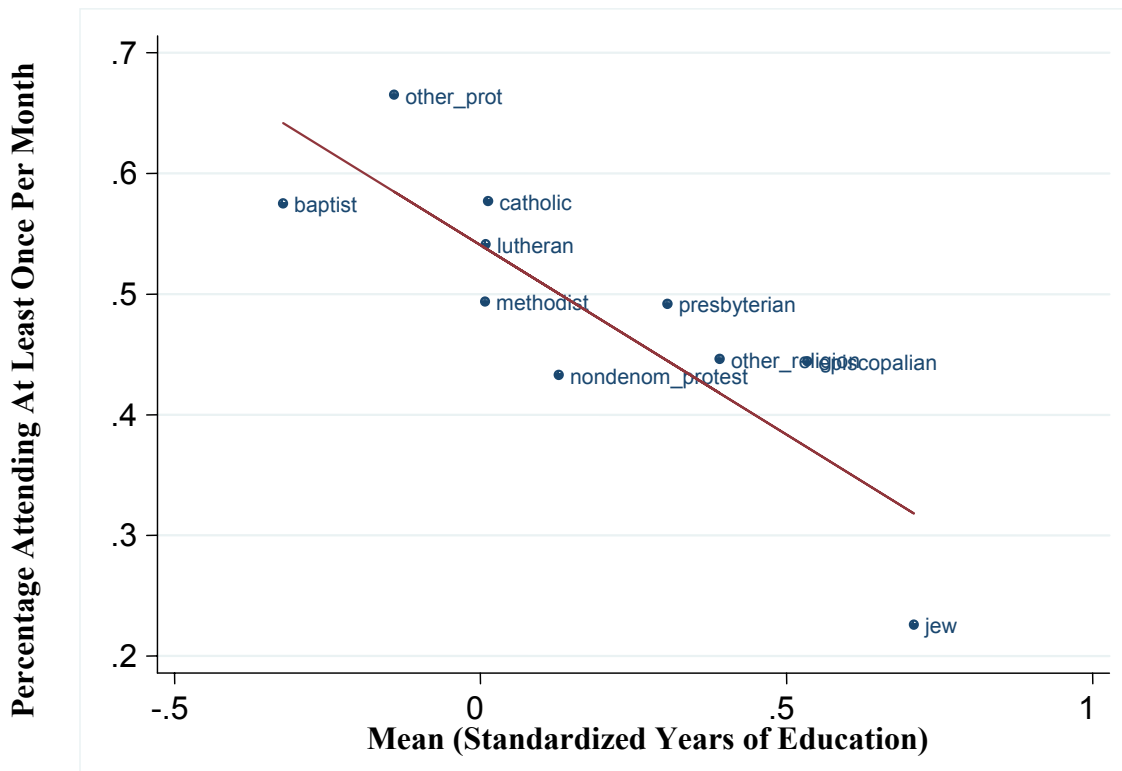**Figure 4: Religious Attendance and Education: Denomination Data**

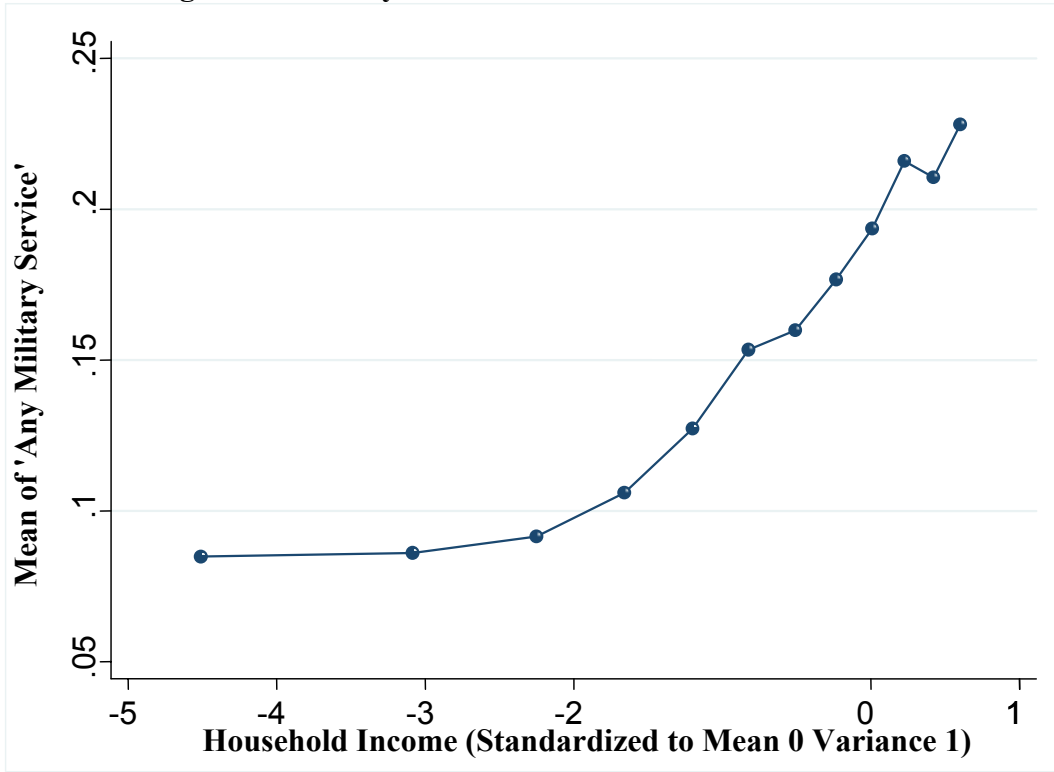**Figure 5: Military Service and Income at the Individual Level**

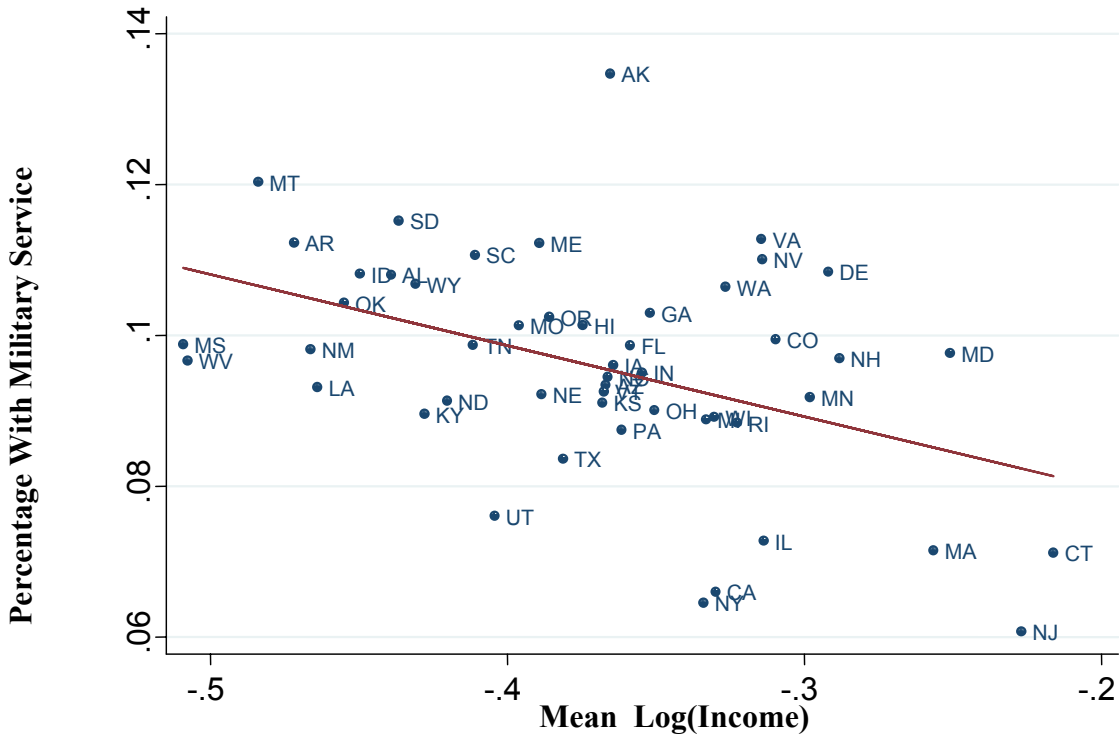**Figure 6: Military Service and Income at State Level**
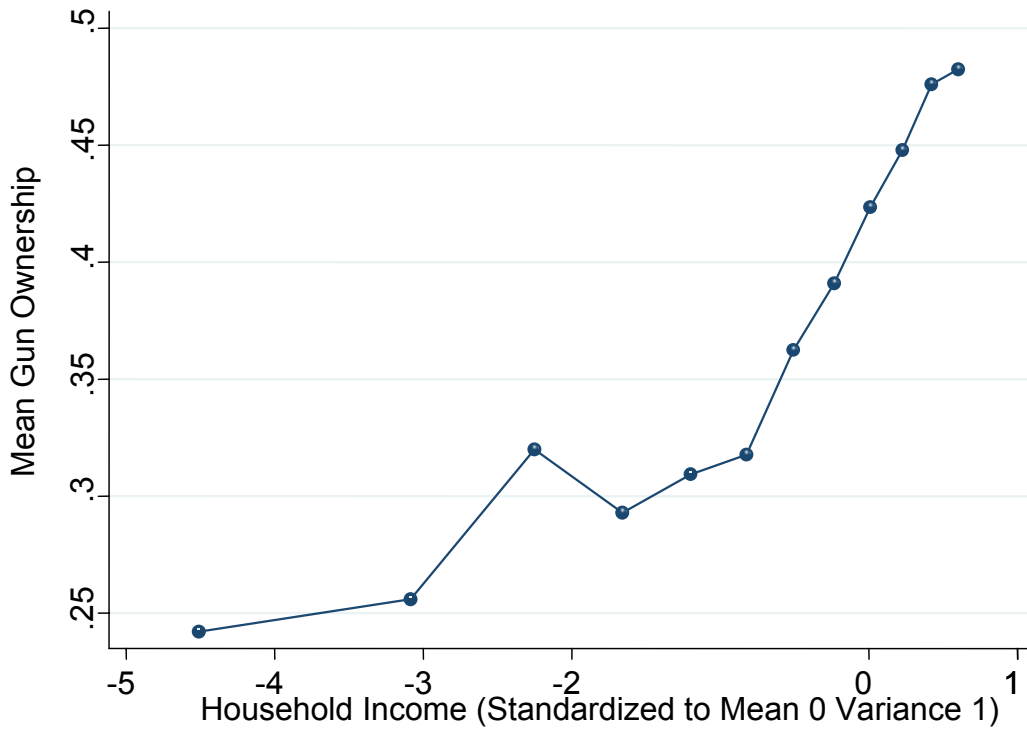
Figure 7: Gun Ownership and Income: Micro Data
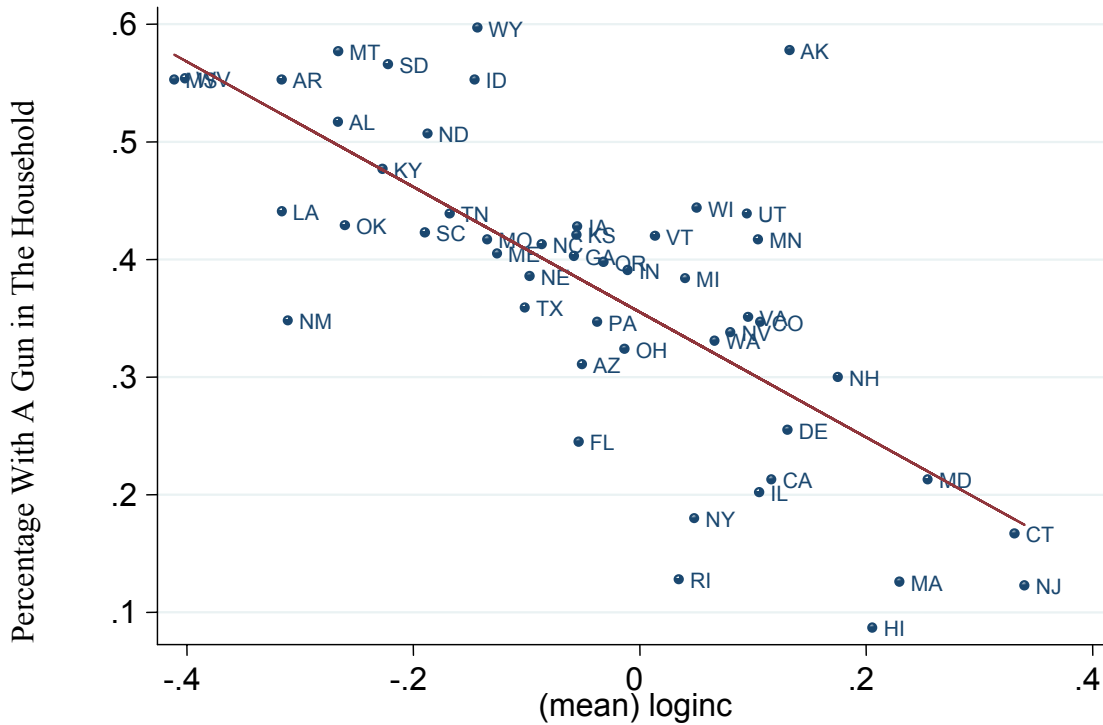

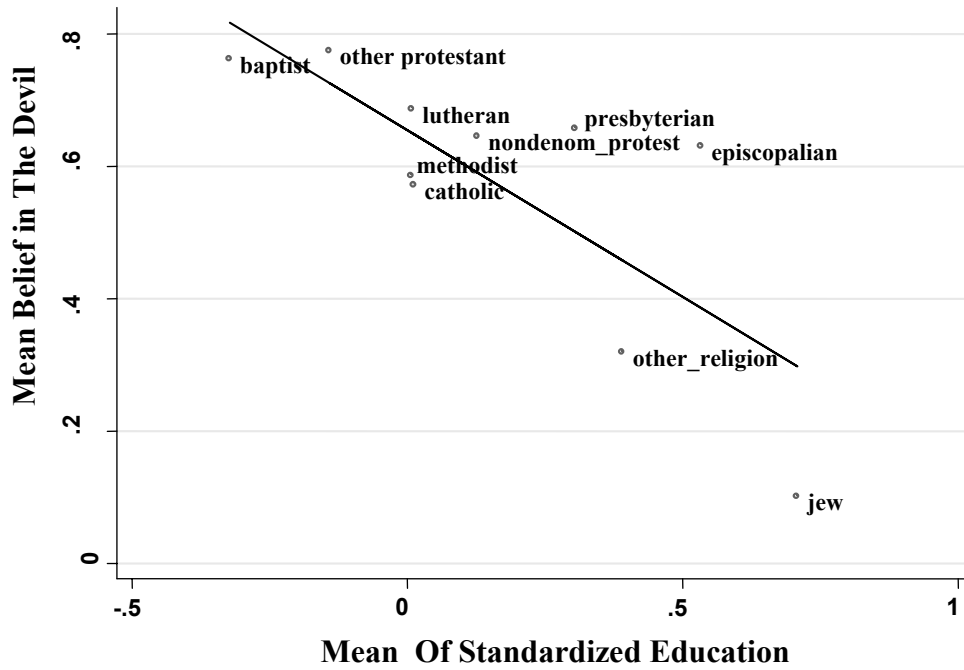**Figure 8: Gun Ownership and Income: State Level**

Figure 9: Belief In the Devil By Standardized Education Level

40

Data Appendix

Voting and Income: State level voting results in 2000 are from the Federal Election Commission. State level income for 2000 is calculated from the Individual Public Use Micro data available at ipums.org and uses the household income measure.

Micro data are from the National Annenberg Election Survey 2000. The data CD Rom accompanies the volume Romer, Daniel, Kate Kenski, Paul Waldman, Christopher Adasiewicz and Kathleen Hall Jamieson, *Capturing Campaign Dynamic,* (New York: Oxford University Press), 2004. We use only the data collection waves that occurred after the election and code as voting for Bush all those who report voting for Bush. We code as 0 those voting for Gore or some other candidate. The abortion variable is coded so that a 1 is given to those who respond that the federal government should not restrict abortion. Those who respond that the federal govt should restrict are coded as 0. Those who say the federal government should NOT allow school prayer are coded as a 1 and those who say the federal government should allow school prayer are coded as 0.

Attendance and Education: We use the General Social Survey 2004 data. For attendance, we use the variable ATTEND and code as a 1 those who attend religious services once per month or more. For beliefs about the devil we use the variable DEVIL and code as a 1 those who say "yes definitely" or "yes probably" when asked if they believe in the Devil. Our recoded variable has a mean of 62 percent. For beliefs about the bible, we code as 1 those who respond that "The Bible is the actual word of God and it is to be taken literally, word for word." Our recode variable has a mean of 30 percent.

For denomination, we divide respondents into 10 denominations using the variables RELIG4 and DENOM. Our 10 denominations and frequencies are

Religion	Freq.	Percent	Cum.
Baptist	9,547	20.64	20.64
Catholic	11,417	24.69	45.33
Episcopalian	1,103	2.38	47.71
Jew	973	2.10	49.82
Lutheran	3,127	6.76	56.58
Methodist	4,662	10.08	66.66
No Religion	4,292	9.28	75.94
Nondenom_Protest	1,965	4.25	80.19
Other_Prot	5,930	12.82	93.01
Other_Religion	1,340	2.90	95.91
Presbyterian	1,893	4.09	100.00
Total	46,249	100.00	

Military Status and Income: We use the General Social Survey 2004 data. We use the inflation adjusted income number from the variable REALINC. We use the variable VETYEARS to infer whether the respondent has any military experience. The mean of our variable for military service is 18.7 percent.   For beliefs we use MILOK and code as a 1 anyone who believes that the military offers a good experience for men. The mean of our recoded variable is 62.4 percent. We also use CONARMY and code as a 1 anyone who has "a great deal of confidence in the military." This variable has a mean of 62 percent.