

What's the Matter with the Middle Class?
State Differences in the Effects of Education and Income on Party Choice

David Weakliem
University of Connecticut

Robert Biggert
Assumption College

August 10, 2007

Abstract: Regional differences in party support have attracted a good deal of attention since the 2000 election. A striking feature of the current pattern is that Democratic support is higher in more affluent states. At the individual level, income is associated with Republican support, but in a recent paper, Gelman et al. (2006) find that this relationship is weaker in more affluent states. In affluent states, people with high and low incomes both tend to vote Democratic; in poorer states, people with low incomes vote Democratic while people with high incomes vote Republican. This paper extends Gelman et al.'s analysis by considering both education and income. We find that the effects of income and college education both vary among states, in a largely independent manner. Variation in the effects of college education is related to the educational composition of the state: where college education is more common, it is more strongly associated with support for the Democrats. Overall, regional differences are largest in the middle classes, contrary to the claims of some popular and theoretical accounts. There is some evidence that a pattern of weaker class divisions is associated with more support for the Democrats.

What's the Matter with the Middle Class?
State Differences in the Effects of Education and Income on Party Choice

Regional differences in support for the Democratic and Republican parties have received a great deal of attention in recent years. References to "red" (Republican) and "blue" (Democratic) states have become common in both scholarly and popular discussion. Although regional differences in party support have existed throughout the history of the United States, their nature has changed over the last half-century: the South has gone from overwhelmingly Democratic to strongly Republican and the Northeast, particularly New England, from mostly Republican to mostly Democratic. A striking feature of the contemporary pattern is that support for the Democrats is higher in more affluent states. Since the Democrats have traditionally been regarded the party of the working class and the poor, this point has attracted a good deal of attention. Some popular accounts have concluded that it represents a change in the relationship between economic position and voting: that Republicans are now the party of the lower and middle classes and Democrats are the party of the affluent (see Frank 2006 for examples). However, as Robinson (1950) noted long ago, individual and group level relationships are not necessarily the same. At the individual level, higher income continues to be associated with Republican voting, as has been true since at least the 1930s.

The regional differences in voting patterns, however, are of interest in their own right, not merely as an illustration of the ecological fallacy. Although regional differences have not received as much scholarly attention as individual differences, there have been several relevant studies. Lipset (1963, p. 273-8) argues that there is a general tendency for poorer regions to be more conservative, which he explains by the

persistence of traditional social relations in those regions: "the social structure remains in some part the way it was before the age of capitalism... The positions of rich and poor are defined as the natural order of things and are supported by personal, family, and local loyalties...." Although pre-capitalist social relations no longer prevail in even the poorest parts of the United States, it is possible that the historical legacies of the regions created enduring differences in values and opinions. Although regional differences in partisanship have changed, Peltzman's (1985) analysis of Congressional representatives suggests that the same ideological pattern has endured for more than a century: the Northeast has been the most liberal region and the South has been the most conservative, as Lipset's hypothesis suggests. This account suggests that changes in voting patterns result from increasing ideological differentiation between the major parties at the national level, particularly on issues of race.

A second question is the relationship between individual-level and regional differences. One possibility is that there is some general regional effect that applies to all groups equally. For example, all kinds of people in Mississippi--black and white, rich and poor, men and women--may be more inclined to vote Republican than their counterparts in Massachusetts. However, both popular and theoretical accounts often suggest that the pattern of regional differences is not uniform in all groups--that regional differences are stronger among some kinds of people, or that certain combinations of region and social position are distinctive. Journalistic analysis of specific elections often claim that groups defined by a combination of geographical and social characteristics, such as "Reagan Democrats," "soccer moms," or "NASCAR dads," showed a particularly strong swing towards one party. There is certainly a good deal of

exaggeration in such accounts, but the perceptions of close observers should not be dismissed out of hand.

Some theoretical accounts also imply that regional differences will not be uniform across groups. Similarly, Hertzler (1939, p. 26) suggested that "the upper classes have . . . a greater degree of universality in their attitudes, values, and interests, and are less bound by habitation." Similarly, Berelson, Lazarsfeld, and McPhee (1954) held that the lower classes were more sensitive to local context than the middle classes, so that their voting patterns would show more regional variation. More recently, Przeworski and Sprague (1986) argued that voters, particularly working-class voters, respond to the class composition of support for the parties. For example, if a traditionally working-class party gains votes from the middle classes, working-class voters may feel less identification with the party and become less likely to vote for it. Huckfeldt and Kohfeld (1989) applied this idea to race, arguing that some white voters would turn against a party that received high levels of black support.

Empirical analysis of the relationship between individual and regional differences is difficult because most national surveys do not contain sufficiently large samples to provide accurate estimates of voting patterns below the national level, but there has been some recent work on this question, notably Gelman, Shor, Bafumi, and Park (2006). They find a pattern that is the opposite of what Hertzler (1939) predicted: regional differences are larger among more affluent voters. That is, upper-income voters in poor states are overwhelmingly Republican, while upper-income voters in rich states are evenly divided. Another way to express their finding is that the effect of income is stronger in poor states and than in affluent states. They are primarily interested in identifying and describing these patterns, rather than explaining them, but they

suggest that economic issues may have less salience in more affluent states (Gelman et al. p.26).

Although Gelman et al. (2006) offer a thorough and sophisticated analysis, they focus exclusively on income--they start by noting the state-level correlation between wealth and Democratic support and ask whether it holds at the individual level. Previous research, however, suggests that the effects of education may also vary. Weakliem and Biggert (1999) find significant regional differences in the effect of education on a variety of attitudes. Weakliem (2002) finds some national differences in the effect of education; specifically, the tendency for education to be associated with liberal views on non-economic issues seems to be stronger in more affluent nations. Moreover, Hertzler's (1939) hypothesis is particularly relevant to educational groups, since more educated people are more likely to obtain information from the national media and to have a wider geographic circle of acquaintances. On the other hand, a broader perspective may lead to differentiation rather than uniformity by making people more conscious of their distinctive regional characteristics. For example, Southerners who read the New York Times might be motivated to articulate and defend cultural differences of which they would otherwise only be dimly aware.

This study offers an analysis of the relationship between individual and regional differences in voting choices, using data from the 2004 presidential election. It focuses on the effects of income and college education, but also gives some attention to race and gender differences.

Models

The hypotheses discussed above involve the presence or absence of interactions between region and other characteristics. Terms like the "gender gap," "class voting,"

or "income polarization" refer to the association between a variable and voting choices. In order to have a general term, we will refer to all such associations as "polarization." The presence of interactions means that the strength and possibly even the direction of polarization will be different in different regions. The discussion in this section will assume that polarization is always in the same direction, so that the meaning of "more" or "less" polarization is unambiguous, but the models place no restriction on the direction of polarization.

An issue that is implicit in many accounts of voting patterns, but rarely addressed directly is the relationship between polarization and overall support for the parties. For example, do the Democrats do better where class voting is high, or low, or is there no tendency in either direction? Popular accounts often assume that there is some relationship--that polarization benefits one party over the other. For example, a recent New York Times article states that "there may be serious risks for the Democrats" in adopting a style that would appeal to men, since they "have historically relied on a gender gap advantage--with women" (Lizza 2007, p. 3). The implication is that the Democrats benefit from the gender difference in party supports, so that a reduction of this difference could hurt their overall performance. Lipset (1981, pp. 512-5), in reviewing Presidential elections since the 1940s, argues that the Democratic party has been more successful when the level of class polarization has been higher.

Logically, there is no necessary connection between polarization and overall performance. A decline of the gender gap could result from increased support among men, or reduced support among women, or some combination of the two. However, there may be an empirical connection between them--for example, the analysis of Gelman et al. implies that in the 2004 election, support for the Democrats was higher

where income polarization was lower. The existence of such a relationship could have important implications for party strategy. For example, a negative relationship between polarization and Democratic vote might be interpreted to mean that Democrats gain by appealing to higher-income people and Republicans by appealing to lower-income people--that is, parties do better by reaching out rather than focusing on their "base." As discussed below, care must be taken in drawing conclusions of this kind, but the issue is clearly important.

To simplify the discussion, we will consider the interaction of region with a variable comprised of only two groups, which we will call the "upper" and "lower" classes. There are a variety of ways that region could be defined, but since the analysis will involve states, we will refer to the units as states. The standard way of expressing interaction is:

$$y_{ik}=y_i+u_k+v_{ik} \quad (1)$$

If v_{ik} is 0 for all i and k , there is no interaction: the mean of the dependent variable for any combination of class and region is simply an additive function of effects of region and class.

It is difficult to see the possible connection between polarization and overall support in this expression. However, interactions can be approached in an alternative manner by giving separate equations for regional effects in different classes. This is mathematically equivalent to the conventional treatment, but makes it easier to see the conditions under which polarization will be correlated with overall support. One equation represents state differences among the upper class, the other represents the state differences among the lower classes. The idea of a general state effect applying to all groups--that is, no interaction between class and state--can be written:

$$y_{1k}=y_1+u_k \quad (2a)$$

$$y_{2k}=y_2+u_k \quad (2b)$$

The first equation gives the proportion of the upper class in a state voting Republican (y_{1k}) as a function of some baseline for the class and a state effect, and the second gives the proportion of the lower classes voting Republican as a function of a baseline and the same state effect. Because the state effects are the same for both classes, the difference between the classes ("class voting") will be identical in every state: y_1-y_2 .

Suppose that state effects are not identical:

$$y_{1k}=y_1+u_{1k} \quad (3a)$$

$$y_{2k}=y_2+u_{2k} \quad (3b)$$

In this case, the level of class voting in a state will be $u_{1k}-u_{2k}$. If the population of each state is 50% upper class and 50% lower class, the overall level of support for the Republicans will be $((y_1+y_2)+(u_{1k}+u_{2k}))/2$. Since the first part of this expression is a constant, the covariance between class voting and overall support will be expectation of $(u_{1k}+u_{2k})(u_{1k}-u_{2k})/2$, or:

$$(\text{Var}(u_{1k})-\text{Var}(u_{2k}))/2 \quad (4)$$

Thus, the level of class voting and overall party support will be correlated if the variances of u_1 and u_2 differ. The Berelson, Lazarsfeld, and McPhee (1954) hypothesis of smaller regional variance in the upper classes implies a negative covariance between Republican support and class voting--that is, Republicans will do better where class voting is lower. The point can be illustrated by taking an extreme case in which there the Republican share of the upper-class vote is the same in every state--that is, $\text{Var}(u_{1k})=0$. In that case, state differences in overall performance would depend entirely

on differences in performance among the lower class. The Republicans would do better when they did better in the lower class--that is, where class voting was lower.

The arithmetic becomes more complicated if the groups are not of equal size, since the covariance term does not cancel out, and variation in the larger class will have more influence on vote totals. Hence, it is unlikely that there will be any general rule about the association between class voting and overall support--the model simply shows that a relationship is possible under a variety of conditions.

However, the variables in the above equations have meaningful interpretations in the context of theories of voting behavior. Hypotheses about the relative size of the variances have been discussed above, and the degree of correlation between u_{1k} and u_{2k} is also of interest. If there are general state-level factors that affect both classes in the same way, a positive correlation will result. Conversely, a correlation near zero would suggest that the state-level factors influencing the votes of the different classes were very different. A negative correlation could indicate that the groups react in opposite ways to the same factors, or that each reacts against the other, as Pzeworski and Sprague (1986) and Huckfeldt and Kohfeld (1989). In this sense, the correlation between variation in different groups can be regarded as an index of social consensus, as Lipset (1963, p. 15) suggests.

Data

In order to get reasonably precise estimates of polarization at the state level, it is necessary to have large samples in each state. In standard national surveys, the samples in each state are roughly proportional to the population of the state, so that the estimates for large states are quite precise, but those for small states are not. Exit polls, because they are intended to predict the vote in the Electoral Collage, contain large

samples for even the small states. In most elections, including 2004, there is a separate exit poll for each state in addition to the national exit polls. There is some variation in the questions asked in the state exit polls, and in recent elections not all of them have had questions on education, which is the key variable in this analysis. In 2004, twenty-one of the state exit polls in twenty-one included a question on educational attainment. The states are from all parts of the country and are not distinctive in any obvious way. Each state sample includes about 2,000 cases.

The dependent variable for the analysis is reported vote for Kerry or Bush--the small number who chose other candidates or reported that they did not vote for President are excluded. Independent variables are dummy variables for black and other non-white race, a dummy for men, and variables for income and college education. Income is measured by eight categories: under \$15,000, \$15-30,000, \$30-50,000, \$50-75,000, \$75-100,000, \$100-150,000, \$150-200,000, and over \$200,000. Education was originally measured by five categories, but preliminary examination indicated that there was little difference in voting choice among the three lowest (not a high school graduate, high school graduate, and some college). Hence, we collapse it into a variable with three values: 0 for not a college graduate, 1 for college graduates, and 2 for graduate education.

Results

Results from logistic regression models applied to the national sample are shown in Table 1. The models include the main effects of state and the demographic variables, but no interactions. Because of the large sample size, all of the estimates are highly significant, so we omit the standard errors. A positive sign means that an increase in the independent variable increases the probability of voting Republican. In the first model,

which includes race, gender, and income, the estimated effect of income is .130. To judge the size of this effect, we can take a hypothetical person with an income of 45,000 (category 3) and characteristics giving him a 50% chance of voting Republican. A person with the same characteristics but an income of under \$15,000 would have a 43.5% chance of voting Republican, while a person with those characteristics and an income of over \$200,000 would have a 66% chance of voting Republican. The second model adds a control for education, which has a negative sign, meaning that people with more education are less likely to vote Republican after controlling for the other variables. The estimated effect of education becomes about 50% larger after controlling for college education. The third model uses dummy variables for each category of the education and income variables. The results of this model show that higher income is uniformly associated with higher chances of voting Republican, and higher levels of education with lower chances. Although pattern of the estimates is not precisely linear, the departures are not very striking, so we will continue to use the original form of these variables in subsequent analyses.

The next phase of the analysis involves fitting the second model in each state separately. This is equivalent to allowing the effects of all of the independent variables to differ by state. The means, standard deviations, and minimum and maximum values of the state estimates are shown in Table 2. Of course, the state estimates are all based on samples, so some of the differences in estimates reflect sampling variation rather than real differences among the states. Because the state samples are all about the same size, the standard errors of the state-level estimates of gender, income, and education are also about the same in every state: typically about .12, .04, and .08 respectively. The observed standard deviations across states thus suggest that the between-state

differences are not all sampling error--there is some real variation in the effects of all of these variables, particularly college education. Because the number of people who are black or "other" race differs substantially among the states, comparison is more difficult for race. For example, the estimate for black race is large and positive (pro-Republican) in South Dakota, but not significantly different from zero, as there were only nine black respondents in a sample of about 1,500.

Figure 1 gives a scatterplot of the estimated effects of income and education in each state. There is virtually no correlation between them--that is, knowing whether the effect of college education is strong or weak in a particular state does not help one to predict whether the effect of income is strong or weak. The estimated effect of income is unusually small in two states, New Jersey and Washington, where it appears to have almost no effect on voting choice. In the other nineteen states, income substantially increases the chance of voting Republican. Education is associated with more support for the Democrats in every state. Although there is a wide range of estimates, there are no notable outliers--the largest absolute z-score is only 1.5. The smallest (most pro-Democratic) effects of education are found in California, New York, Oregon, and North Carolina, while the largest are found in Michigan, Missouri, Pennsylvania, and Ohio. There is a fairly clear geographical distinction between the states with larger and smaller effects--the pro-Democratic effect tends to be stronger on the Atlantic and Pacific coasts, and weaker in the middle of the country. Although this distinction is not useful in an explanatory sense, it suggests that the state differences are systematic rather than completely idiosyncratic.

Gelman et al. (2006) found a significant relationship between per-capita income and the effect of income on voting. With the present estimates, the correlation is $-.36$ --

that is, the effect of income is smaller in more affluent states. Because Gelman et al. (2006) did not consider education, a closer replication involves fitting model 1 in each state separately, in which case the correlation of estimated income effects and per-capita income is $-.45$. The correlation between per capita income and the estimated effect of college education is $-.25$. Because the effects of education are typically negative (pro-Democratic), this means that the effect of education is stronger in more affluent states, or one might say that the effects of both education and income shift in a pro-Democratic direction in richer states.

However, affluence is only one of the state-level characteristics that might be related to the effects of income and education. We considered a number of variables, but found that per-capita income and the proportion of college graduates were the only ones that approached statistical significance. Table 3 shows the results of regressions of the estimated effects of education and income on per-capita income and the proportion of college graduates. Because of the small number of states, there is a good deal of uncertainty, but the proportion of college graduates has a significant relationship to the effect of education: the more college graduates, the stronger the pro-Democratic effect. Per-capital income has a negative association with the estimated effect of income, but it does not approach statistical significance. Overall, the state-level variation in the effect of income remains unexplained, and examination of Figure 1 does not suggest any obvious candidates.

Turning the relationship between income and educational polarization and overall support for the parties, examination of Figure 1 shows most of the states in the upper right of the figure supported Bush, while most of those in the lower left supported Kerry. In a regression of the Republican share of the vote on educational and income

polarization, both had positive estimates with t-ratios of slightly over 1.0. That is, the Republicans did best in states in which the pro-Republican effect of income was large and the pro-Democratic effect of education was small. One might describe this as the "traditional" pattern, where the middle class is Republican relative to the working class. However, the significance is not strong enough to make this anything more than a conjecture.

Finally, we can consider the similarity of the state effects for different groups. Table 4 shows summary statistics for the predicted values of different types of voters in the different states. Five of the examples are white women with various combinations of education and income--low education (not a college graduate) and low income (under \$15,000), low education and high income (over \$200,000), high education (graduate school) and low income, and college graduates earning \$50-75,000. The final example is white male college graduates earning \$50-75,000. Examples for non-whites are not considered because the values in some states are strongly affected by sampling error. The first column shows the mean predicted probability of supporting the Republican and the second shows the standard deviation of that predicted probability. For example, white women with low income and low education were predicted to have a .48 chance of voting Republican in the average state. Because floor or ceiling effects might have a discernable influence in some cases, the third column shows the standard deviation of the logit of the predicted probability. Using either measure, it is those with high levels of income and education who show the most variation. Contrary to the predictions of Hertzler (1939), regional differences are higher among the upper classes.

Table 5 gives the correlations among the state-level predictions for the groups listed in Table 4. The correlations are all positive--that is, in states where one group is

Republican relative to the national average, all others will tend to be relatively Republican as well. However, the correlations between people with low and high levels of education are considerably weaker than the others. For example, the correlation between the predicted proportion voting Republican for high income people without a college degree and high income people with graduate education is only .185. To a large extent, people with low and high levels of education vary independently. In contrast, the correlation between the voting patterns of men and women is .869--that is, in states in which men tend to vote Republican relative to men in other states, women tend to vote Republican relative to women in other states.

Discussion

This analysis supports Gelman et al.'s (2006) finding that there are different patterns of cleavages in the states. However, we find that the differences involve education as well as income. Moreover, the differences in the effects of education are related to the educational composition of the state: in states with more college graduates, the pro-Democratic effect of education is stronger. As a result of generational replacement, the proportion of adults with college degrees is almost certain to rise over the next several decades. Hence, if the relationship found here holds generally, the pro-Democratic effect of education is likely to grow stronger.

Although much of the media discussion of regional differences has focused on the working classes, we find that there is more regional variation among people with high income and education. Whether we should say that something is the matter with the middle classes of Louisiana and Ohio, where people with graduate education and high incomes are overwhelmingly Republican, or with those in Oregon and Washington, where they are overwhelmingly Democratic, is a matter of opinion. But the regional

differences are primarily differences of the middle class.

References

- Berelson, Bernard, Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting*. Chicago: University of Chicago Press.
- Gelman, Andrew, Boris Shor, Joseph Bafumi, and David Park. 2006. "Rich State, Poor State, Red State, Blue State: What's the Matter With Connecticut?". Unpublished ms.
- Hertzler, J. O. 1939. "Some Sociological Aspects of American Regionalism." *Social Forces* 18: 17-29.
- Huckfeldt, Robert and Carol Weitzel Kohfeld. 1989. *Race and the Decline of Class in American Politics*. Urbana: University of Illinois Press.
- Lipset, Seymour Martin. 1963. *Political Man*. Garden City, NY: Doubleday.
- Lizza, Ryan. 2007. "Invasion of the Alpha Male Democrat." *New York Times*, January 7.
- Peltzman, Sam. 1985. "An Economic Interpretation of the History of Congressional Voting in the Twentieth Century." *American Economic Review* 75: 656-75.
- Adam Przeworski and John Sprague. 1986. *Paper Stones: A History of Electoral Socialism*. Chicago: University of Chicago Press.
- Robinson, William S. 1950. "Ecological Correlations and the Behavior of Individuals." *American Sociological Review* 15: 351-7.
- Weakliem, David L. 2002. "Education and Political Opinions: An International Comparison." *International Journal of Public Opinion Research* 14: 141-57.
- Weakliem, David L. and Robert Biggert. 1999. "Region and Political Opinion in the Contemporary United States." *Social Forces* 77: 863-86.

Table 1
Effects of race, gender, income and education on presidential vote, national sample

	(1)	(2)	(3)
Black	-2.66	-2.72	-2.71
Other Race	-0.54	-0.55	-0.54
Male	0.27	0.28	0.27
Income	0.130	0.193	
Education		-0.341	
Income=1			-1.42
Income=2			-1.16
Income=3			-0.84
Income=4			-0.54
Income=5			-0.47
Income=6			-0.40
Income=7			-0.31
Income=8			_____
No College			0.70
College Grad.			0.43
Grad School			_____

Notes: see text for definitions of variables. Dummy variables for state are also included, but coefficient estimates are omitted to save space.

Table 2
Summary statistics for state-level estimates of effects

	Mean	SD	Min	Max
Black	-1.16	1.03	-3.23	1.44
Other Race	-0.68	0.31	-1.15	-0.08
Male	0.20	0.17	-0.04	0.72
Income	0.14	0.06	0.01	0.25
College	-0.31	0.19	-0.66	-0.03

Table 3
Estimated effects of state characteristics on income and educational polarization

	Income	College Education
Constant	.3144 (.1103)	-.0082 (.2872)
Per-capita income (thousands)	-.0048 (.0057)	.0244 (.0148)
% Graduates	-.0006 (.0049)	-.0385*** (.0128)
R2	.127	.373
Sig	.294	.015

Table 4
State-level means and standard deviations of predicted Republican support for selected groups

Group	Mean	SD	SD (logit)
Low education, low income	.48	.076	.314
Low education, high income	.70	.093	.464
High education, high income	.57	.131	.558
High education, low income	.33	.102	.464
Middle educ. & income	.50	.084	.350
Middle educ. & income, men	.55	.067	.290

Note: examples are whites; all except the last are women

Table 5
Correlations among state-level predicted values of predicted Republican support in selected groups

Correlations

		11	12	13	14	15	16
11	low education, low income	1	.396	.287	.618(**)	.706(**)	.714(**)
12	low education, high income	.396	1	.749(**)	.185	.713(**)	.553(**)
13	high education, high income	.287	.749(**)	1	.640(**)	.879(**)	.694(**)
14	high education, low income	.618(**)	.185	.640(**)	1	.811(**)	.749(**)
15	Middle educ & income	.706(**)	.713(**)	.879(**)	.811(**)	1	.869(**)
16	Middle educ & income, men	.714(**)	.553(**)	.694(**)	.749(**)	.869(**)	1

** Correlation is significant at the 0.01 level (2-tailed)

Figure 1

