The determination of United States voter turnout: A panel*

Nicolas de Roos†
Sydney University

Abstract

A panel of off-year general elections is employed to examine the institutional determinants of voter turnout in the United States. We employ instrumental variables and correlated random effects to contend with unobservable fixed determinants of voter turnout. We find only mixed support for the role of institutional factors in influencing turnout. Some support is found for the notion that constraining the powers of policy makers and the voting public tends to reduce turnout. Finally, the variation in cross-section regression results suggests that some care is required in examining individual cross-sections.

*I would like to thank Michael Boozer, Don Andrews, Yianis Sarifidis, Max Schanzenbach, and Hajime Katayama for valuable comments and suggestions. Any errors remain my own. This article is a revised version of a paper submitted as partial completion of my Yale Ph.D.

†Nicolas de Roos: School of Economics and Political Science, Merewether Building H04, University of Sydney, Sydney, NSW 2006, Australia, e-mail: n.deroos@econ.usyd.edu.au, phone: +61 (2) 9351 7079, fax: +61 (2) 9351 4341.

This paper employs panel data techniques to examine the determinants of voter turnout across states in the United States. Previous studies have examined this question using cross-section data, time-series data, survey data, simulations, and experimental methods. The application of panel data delivers two primary benefits. First, by accounting for unobservable characteristics of individual states, it permits consistent estimation of parameters of interest. Second, the additional scope of the data set improves the robustness of the results. The paper serves two broad purposes. First, in a cross-section study of voter participation across states in the 1982 general election, Merrifield (1993) finds a prominent role for many institutional factors in the determination of voter turnout across states. The present paper examines the robustness of this finding to the inclusion of data on multiple elections and the inclusion of state-specific unobserved effects. Second, the paper illustrates the potential benefits of panel data methods for this type of problem.

The paper finds mixed support for the role of institutional factors in influencing turnout. Some support is found for the influence of constraints on the powers of policy makers and voters. Registration requirements also tend to discourage turnout. Little evidence is found for a role of the other institutional variables in determining turnout. The results suggest that focussing on a single cross-section could prove misleading.

The rest of the paper is structured as follows. Below, a selection of the literature on voter participation is briefly discussed. Section 2 introduces the data used in the study. Section 3 outlines the econometric methodology used. The paper's results are discussed in section 4, and section 5 concludes.

1 Related Literature

According to the rational voter theory, an individual will choose to vote if the benefits from voting exceed the costs. Defining P as the probability of casting a decisive vote, B as the expected benefit of altering the election outcome (investment benefits), D as the consumption (intrinsic) benefits of voting, and C as the costs of voting, an individual will vote if[1]

$$PB + D > C. \tag{1}$$

Consumption benefits to voting comprise factors creating a perceived obligation to vote including civic pride, peer pressure, political trust and party loyalty. Investment benefits reflect perceptions about the likely policies of the candidates, the impact of such policies on voters, and the ability of policy makers to implement such policies. However, the probability of an individual casting a decisive vote is small.[2] It could therefore be argued that the investment benefits to voting might play only a minor role. We proceed on the presumption that investment benefits may indeed play a role. People may also vote if they think voting may influence the policies of the victor.

The literature is divided on the relative importance of institutional and cultural determinants of voter participation. Merrifield (1993) argues that the investment benefits to voting are influenced by institutional and political factors. He finds that institutional factors that determine the powers of legislators and voters are important determinants of voter turnout across states in the United States. Jackman (1987) finds that institutional factors are important determinants of

[1]Downs (1957) and Riker and Ordeshook (1968) are the classic references.

[2]See, for example, Mulligan and Hunter (2003) and Gelman et al. (1998).

turnout across industrial democracies. An alternative explanation is that some cultures are more participatory. Interest in politics, pride in political institutions, and feelings of civic duty may all influence an individual's likelihood of voting. Abramson and Aldrich (1982) argue that an erosion of partisanship and a belief that the responsiveness of government has waned account for most of the decline in voter turnout since the 1960s in the United States.

Matsusaka (1995) extends the rational voter theory to account for imperfect information. He argues that people are more likely to abstain from voting if they are unable to evaluate the candidates due to a lack of information on, for example, the likely policies of the candidates and the consequences of those policies. According to this theory, people with higher levels of education and income, older people, married people, and long time residents in an area may be more likely to vote.

A variety of methods have been used to investigate the determinants of voter turnout. For example, Merrifield (1993) and Jackman (1987) examine institutional determinants of voter turnout using a cross section of states and countries, respectively. Knack (1994) also employs cross section regression analysis, applied to individual survey data. Matsusaka (1993) uses a time series of California ballot propositions to investigate the relationship between election closeness and voter turnout. Fair and Dworkin (1993) use simulations to test a version of Downs' (1957) rational voter model. Abramson and Aldrich (1982) employ a discrete choice model to analyse the effect of partisanship and government responsiveness on turnout using survey data.

The present study focusses on the institutional and political determinants of voter turnout. The omission of any important determinants of turnout could lead to inconsistent parameter estimates if any of these omitted factors are related to the explanatory variables of interest. Cross-section regression methods cannot address this problem. We use a panel and employ instrumental variables methods and a correlated random effects model to deal with this issue. Fornos et al. (2004) have also recently employed panel data to explain voter turnout in Latin America. However, they do not explicitly control for unobserved determinants of turnout.

2 Data

A list of the explanatory variables, their sources, and descriptive statistics are included in the Data Appendices. Our interest is primarily in the institutional and political variables. The political party of the governor is included to test the idea that Republicans are more likely to vote than Democrats. The requirement of pre-election day registration is included to examine the cost of registration. Merrifield (1993) suggests that political parties can influence the perceived obligation to vote. To examine this issue, control of both houses and the governorship is included as a proxy for the affect of partisanship on voter mobilisation. The remaining institutional variables reflect constraints on policy makers or voters which, as argued by Jackman (1987), tend to reduce turnout. The number of legislators and elected members of the executive branch will tend to reduce the power of individual legislators and elected members. Line-item veto authority tends to make governors more powerful, but constrains the power of legislators. Similarly, the inclusion of a constitutional provision for legislation by petition and ballot initiative tends to make voters more powerful, but constrains legislators.

Control variables include weather variables, demographic variables, and socioeconomic variables. Rainfall and the deviation from the normal daily maximum temperature on election day are included to control for any costs to voting perceived by individuals due to election day weather conditions. Demographic variables include the rate of population growth in the preceding 4 years, the percentage of state population above the age of 65, and the percentage of state population below the age of 18. These variables control for the possibility that an individual is more likely to vote the longer she has been in a particular area, and the older she is. Finally, the state unemployment rate and the percentage of people below the poverty line are controlled for.

The data set covers the 50 states of the United States for the 7 off-year (non-Presidential) general election years between 1974 and 1998. No turnout data is available for Louisiana in 1994. Hence, an unbalanced panel is constructed and used for estimation.[3] Poverty data was unavailable in the years 1974, 1978, and 1982. Instead, data from the years 1975, 1979, and 1984 was used.

3 Model specification

It is likely that a number of institutional and cultural determinants of turnout are not captured by the econometric model. In particular, no cultural factors are included in the data. Some of these determinants are likely to differ systematically across states. Moreover, both institutional and cultural determinants are likely to display inertia. We therefore allow for both time-varying and fixed unobserved determinants of voter turnout. Moreover, aggregating to the state level, the individual unobserved fixed determinants of turnout are allowed to differ by state. The paper employs panel data techniques using data at the state level to account for these unobserved fixed effects.

The following panel data model is considered:

$$\begin{aligned} Y_{st} &= X_{st}\beta + Z_s\gamma + \varepsilon_{st} \qquad (2) \\ \varepsilon_{st} &= \alpha_s + u_{st}, \end{aligned}$$

where Y_{st} denotes turnout in off-year congressional elections in state s at time t, with $s \in \{1,2,\ldots,S\}$ and $t \in \{1,2,\ldots,T\}$.[4] The columns of X and Z capture time-varying and fixed demographic, environmental, and institutional determinants of voter turnout, respectively. β and γ are k- and g-vectors of coefficients, respectively. The composite error term, ε, reflects fixed and time-varying unobservable state-specific determinants of turnout, α and u, respectively. These unobserved variables reflect cultural determinants and any omitted institutional determinants of voter turnout.

Our estimation strategy for equation (2) depends on whether the explanatory variables are independent of the unobserved fixed effect, α. If the explanatory variables are independent of the fixed effect, consistent and efficient estimates of β and γ can be obtained by generalised least squares (GLS) under the standard random effects model. If they are correlated with the fixed effect, then GLS estimates are inconsistent and we can obtain consistent and efficient estimates of β, but not γ, by ordinary least squares (OLS) under the fixed effects model. We might expect a number of the explanatory variables to be correlated with unobserved cultural and institutional

[3]The balanced panel omitting Louisiana was also examined, yielding similar results.

[4]Note that the basic econometric model employed is linear, while equation (1) suggests a non-linear relationship. That is, the perceived probability of influencing the election result is not included in the data set. If this probability varies across states, then investment benefits will not enter the model in linear fashion. In addition, aggregation to the state level is unlikely to preserve linearity. However, it is unclear what the exact functional form would be at the state level. For convenience and consistency with the literature, we therefore adopt a liner model.

determinants of turnout. We return to this issue in sections 3.2 and 3.3, below, when we discuss instrumental variables and correlated random effects estimation, respectively.

In the present context, pooled OLS regression suffers from two main shortcomings. First, in the presence of unobservable fixed effects, OLS does not exploit the error structure of equation (2) to provide efficient estimation. Second, if the unobservable fixed effects are correlated with the explanatory variables, then, just like the GLS estimates, OLS estimates will be biased and inconsistent. To test for the presence of an unobserved fixed effect, we calculate the following statistic:

$$\frac{\sum_{s=1}^{S}\sum_{t=1}^{T-1}\sum_{r=t+1}^{T}\hat{\varepsilon}_{st}\hat{\varepsilon}_{sr}}{\left(\sum_{s=1}^{S}\left(\sum_{t=1}^{T-1}\sum_{r=t+1}^{T}\hat{\varepsilon}_{st}\hat{\varepsilon}_{sr}\right)^{2}\right)^{1/2}}, \tag{3}$$

where $\hat{\varepsilon}$ are the residuals from the pooled OLS regression.[5] Under the null hypothesis of no serial correlation of the residuals, this statistic has an asymptotic standard normal distribution.

3.1 Random effects and fixed effects

More formally, if equation (4), below, holds, the GLS estimator of the random effects model provides consistent and efficient estimates of β and γ.

$$E\left(\alpha_s|X_{st},Z_s\right)=0 \tag{4}$$

To construct this estimator, we first define the following orthogonal projection operators:

$$\begin{aligned} P_V &\equiv \frac{1}{T}I_S\bigotimes l_T l_T' \\ Q_V &\equiv I_{ST}-P_V, \end{aligned}$$

where l_T is a T-vector of ones and I_S is an identity matrix of dimension S. P_V transforms data into group means, while Q_V transforms data into deviations from group means. We assume that the idiosyncratic errors, u_{st}, are serially uncorrelated and have constant unconditional variance across time, σ_u^2, and that the unobserved effect, α_s, is homoskedastic, with variance σ_α^2. Then, if equation (4) holds, the covariance matrix of residuals from equation (2) is given by

$$\mathrm{cov}(\varepsilon_{st}|X_{st},Z_s)=\sigma_u^2 I_{ST}+T\sigma_\alpha^2 P_V\equiv\Omega.$$

Given Ω, equation (2) can then be estimated by GLS. The covariance matrix, Ω, can be estimated through use of within-group and between-group variation in the data.

The within or fixed effects estimator, $\hat{\beta}^W$, is obtained by transforming equation (2) by Q_V to obtain

$$Y_{st}-\overline{Y}_s=\left(X_{st}-\overline{X}_s\right)\beta+u_{st}-\overline{u}_s, \tag{5}$$

where $\overline{Y}_s=1/T\sum_t Y_{st}$. The sum of squared residuals from equation (5) can be used to obtain unbiased and consistent estimates of σ_u^2.

The between estimator is obtained by transforming equation (2) by P_V to obtain

$$\overline{Y}_s=\overline{X}_s\beta+Z_s\gamma+\alpha_s+\overline{u}_s. \tag{6}$$

[5] See, for example, Wooldridge (2002).

Provided equation (4) holds, the sum of squared residuals from equation (6) yields unbiased and consistent estimates for the variance of $\bar{\varepsilon}_s = \alpha_s + \bar{u}_s$, $\sigma^2_{\bar{\varepsilon}} = \sigma^2_\alpha + (1/T)\sigma^2_u$. Letting $\theta \equiv (\sigma^2_u/(\sigma^2_u + T\sigma^2_\alpha))^{1/2}$, we can define the operator

$$P \equiv I_{ST} - (1-\theta)P_V.$$

We can estimate θ by $\hat{\theta} = \hat{\sigma}^2_u / T\hat{\sigma}^2_{\bar{\varepsilon}}$. Equation (2) can then be estimated by feasible GLS by applying the transformation, P, and estimating equation (7), below, by OLS, to obtain $\hat{\beta}^{GLS}$ and $\hat{\gamma}^{GLS}$.[6]

$$Y_{st} - \left(1-\hat{\theta}\right)\bar{Y_s} = \left(X_{st} - \left(1-\hat{\theta}\right)\bar{X_s}\right) + \hat{\theta}Z_s\gamma + \hat{\theta}\alpha_s + \left(u_{st} - \left(1-\hat{\theta}\right)\bar{u_s}\right) \tag{7}$$

If equation (4) fails, then $\hat{\beta}^{GLS}$ and $\hat{\gamma}^{GLS}$ will be biased and inconsistent. Consistent estimates of β could then be obtained by using the within estimator, $\hat{\beta}^W$. In order to test equation (4), we employ a specification test suggested by Hausman (1978). The null hypothesis is that equation (4) holds. Defining $\hat{q} \equiv \hat{\beta}^W - \hat{\beta}^{GLS}$, form the test statistic

$$\hat{q}'\left(\operatorname{var}\left(\hat{\beta}^W\right) - \operatorname{var}\left(\hat{\beta}^{GLS}\right)\right)^{-1}\hat{q}. \tag{8}$$

Under the null hypothesis, this has the χ^2 distribution with k degrees of freedom.

Rejection of the specification test implies the GLS estimator provides biased estimates of β. However, the alternative, the within estimator, has two main shortcomings. First, the within transformation eliminates all time-invariant variables, so γ cannot be estimated. Second, the within estimator ignores variation across states. Consequently, two alternative estimators are considered below. The first method involves providing instruments for those columns of X and Z that are potentially correlated with α_s. The second alternative is to model the correlation between the fixed effect and the explanatory variables directly using the correlated random effects model.

3.2 Instrumental variables

A common problem in instrumental variable (IV) estimation is to find appropriate instruments, excluded from equation (2), that are uncorrelated with α_s. We follow Hausman and Taylor (1981) and partition X and Z into columns that are exogenous to α and columns that are endogenous to α. The exogenous columns of X and Z then serve two functions; they produce unbiased estimates of elements of β and γ, and they act as instruments for the endogenous columns.

We might expect the following explanatory variables to be correlated with the fixed effect: the political party of the governor; whether or not one political party controls both houses and the governorship; the unemployment rate; and the fraction of the state population living below the poverty line. The political party of the governor in a particular state will reflect characteristics and views of the voters of that state which could potentially also determine attitudes toward electoral participation. A similar argument could be made for the dummy variable capturing whether one political party controls both houses and the governorship. The rate of unemployment and the fraction of state population below the poverty line could both reflect economic or social characteristics of the state population which are likely to influence attitudes towards participation. The remaining variables relate to weather, demographics, registration requirements,

[6] Hausman and Taylor (1981) establish that the resultant estimator satisfies the Gauss-Markov assumptions.

and elements of state constitutions. We proceed under the assumption that these are independent of the unobserved fixed effects.

Under these assumptions, all the columns of Z are exogenous to α. Partition X into X^x and X^e, exogenous and endogenous components of X, respectively. Then, form the instrument set $A \equiv [Q_V \vdots X^x \vdots Z]$. Defining the projection matrix $P_A \equiv A(A'A)^{-1}A'$, we can then estimate

$$P_A P Y_{st} = P_A P X_{st} \beta + P_A P Z_s \gamma + P_A P \varepsilon_{st} \tag{9}$$

to obtain unbiased and consistent estimates, $\hat{\beta}^{IV}$ and $\hat{\gamma}^{IV}$.

3.3 Correlated random effects

Next we consider the correlated random effects (CRE) model (Chamberlain; 1982, 1984), in which we explicitly model the correlation between the fixed effect, α, and the endogenous explanatory variables, X^e. Consider a linear relationship between α and the columns of X^e:

$$\begin{aligned} \alpha_s &= X^e_{s1}\lambda_1 + \ldots + X^e_{sT}\lambda_T + \upsilon_s \\ &= X^e_s \lambda + \upsilon_s \end{aligned}$$

where λ is a vector of correlation coefficients, X^e_s is a vector of endogenous time-varying explanatory variables for state s, and the error term, υ, is orthogonal to X^e by construction. Substituting into (2) yields

$$Y_{st} = X_{st}\beta + Z_s\gamma + X^e_s\lambda + \upsilon_s + u_{st}. \tag{10}$$

We first estimate the unrestricted reduced form,

$$Y_{st} = X^x_{st}\beta_x + Z_s\gamma + X^e_s\pi_s + \upsilon_s + u_{st}. \tag{11}$$

where $\pi = (\pi_1, \ldots, \pi_S)$ contains the unrestricted coefficients on the endogenous explanatory variables. Equation (11) is estimated by GLS, yielding $\hat{\pi}^{GLS}$. We can then test the restrictions implied by (10) and obtain the restricted coefficients by estimating

$$\hat{\pi}^{GLS} = R\phi, \tag{12}$$

where R is a matrix of linear restrictions implied by equation (10), and $\hat{\phi}$ contains the restricted correlated random effects estimates. A χ^2 test is then employed to test the restrictions. Jakubson (1991) notes that this test is an omnibus test with the potential to detect misspecification in many directions and may have low power in detecting misspecification in any particular direction. Hence, a rejection should be treated with some caution.

4 Results

4.1 Cross-section results

Table 1 presents results of the cross-section OLS regressions covering the seven non-presidential elections in the sample. These results serve as a point of comparison on three fronts. First, we can compare the results of Merryfield (1993) with the 1982 cross-section results. Second, we can examine the robustness of the results across the cross-sections. Finally, we can discuss the panel data model in light of these results.

Table 1: Cross-section regressions

Election year	1974	1978	1982	1986	1990	1994	1998
Constant	-38.70	-4.282	7.267	23.25	41.44#	39.83	88.20#
	(35.52)	(40.19)	(15.32)	(19.06)	(24.01)	(39.54)	(45.45)
Area	-0.045*	0.045#	0.049*	0.116**	0.047#	0.032	0.046
	(0.021)	(0.026)	(0.019)	(0.032)	(0.023)	(0.026)	(0.034)
Line Item Veto	-3.880	-2.148	-6.319**	-5.730*	-9.139**	-2.351	-4.657
	(2.428)	(2.010)	(2.308)	(2.328)	(3.031)	(2.564)	(3.366)
Initiative	4.204	2.810	4.200#	5.639**	7.543**	4.354*	2.317
	(3.069)	(2.192)	(2.397)	(1.874)	(2.386)	(1.635)	(2.718)
Control	-1.021	-1.166	0.347	-4.938*	1.359	-1.264	4.457#
	(2.305)	(3.115)	(1.816)	(2.026)	(2.262)	(2.076)	(2.546)
Executive officers	-0.520	-0.680	-0.445	-0.364	0.280	-0.662#	0.708
	(0.381)	(0.446)	(0.451)	(0.451)	(0.600)	(0.372)	(0.534)
Governor's Party	-0.708	3.475	-2.521	2.908	0.660	2.670	1.559
	(1.995)	(3.226)	(1.664)	(1.721)	(2.295)	(1.954)	(3.455)
Legislators	-0.018	-0.005	-0.023#	-0.022	-0.021	-0.010	-0.032*
	(0.019)	(0.013)	(0.013)	(0.018)	(0.014)	(0.015)	(0.015)
Registration	-0.652	-4.459	-4.583#	-6.453*	-8.030	-6.660*	-9.628*
	(6.139)	(5.208)	(2.281)	(2.460)	(6.906)	(3.193)	(3.956)
Population growth	-0.402	-0.983**	-0.781**	-1.276**	-0.708**	-0.940**	-0.988**
	(0.250)	(0.345)	(0.109)	(0.213)	(0.243)	(0.313)	(0.346)
Poverty	-1.677**	-1.484*	-0.785#	0.614	-1.078**	-0.305	-0.776
	(0.263)	(0.567)	(0.413)	(0.388)	(0.393)	(0.381)	(0.591)
Rain	-8.402	-11.32	-5.523	-2.189	3.173	-6.569	-5.005
	(6.057)	(12.59)	(3.959)	(1.771)	(4.438)	(11.79)	(6.396)
Temperature	-0.213	0.205	-0.159	0.111	0.111	-0.361*	-0.287
	(0.200)	(0.175)	(0.095)	(0.147)	(0.152)	(0.151)	(0.205)
Unemployment	0.281	-0.585	0.164	-3.451**	-0.988	-1.660	-0.497
	(0.825)	(0.924)	(0.459)	(0.779)	(0.871)	(1.010)	(1.455)
Senior	1.282	0.934	1.305*	0.719	0.908	0.728	-1.130
	(0.852)	(1.036)	(0.583)	(0.572)	(0.771)	(1.014)	(1.462)
Youth	2.977**	2.129*	1.740**	1.275*	0.605	0.774	-0.554
	(0.926)	(1.026)	(0.401)	(0.551)	(0.710)	(1.142)	(1.209)
Observations	50	50	50	50	50	49	50
R^2	0.695	0.634	0.757	0.764	0.649	0.650	0.537
$\bar{R}^2$	0.573	0.473	0.650	0.660	0.495	0.492	0.332
Dep var sum of sq	82116	72245	87357	74147	74111	83224	73501
Sum of sq errors	1531	1503	958.5	1064	1875	1076	2137

Notes: Figures in parentheses are heteroskedasticity-robust standard errors (with degrees of freedom correction). #,*,** indicate significance at the 10%, 5%, 1% level for two-tailed tests.

We might expect the OLS results for the 1982 election to resemble those of Merrifield (1993). Differences arise from differences in the set of explanatory variables considered, and some differences in the measurement of explanatory variables. Some of Merrifield's political/institutional variables were omitted from the present study because of difficulties in constructing a consistent or reliable series covering all the election years. These include an indicator of whether a balanced budget amendment is in the state constitution; an indicator for the presence of a limit on state general obligation debt in the state's constitution; and the income elasticity of income and sales tax funds.Merrifield also included an indicator variable for the existence of a unicameral state legislature and nonpartisan elections. Nebraska is the only state with a unicameral legislature and nonpartisan elections, so this variable is in effect a dummy variable for Nebraska. The variable was therefore omitted. There were also minor differences in measurement between Merrifield (1993) and the present study in the following variables: the difference between the actual and normal temperature on election day in degrees F; the rate of population growth; and the percentage of families below the poverty line.

Differences in explanatory variables between the Merrifield study and the present paper result in only minor differences in results for the majority of variables. However, coefficient estimates for several variables differ noticeably. First, control of both legislative houses and the governorship by one party significantly (both in a statistical and a practical sense) increases the voting propensity in Merrifield's study, but appears to have little impact in the present study. Second, the effect of a pre-election day registration requirement is substantial in the present study, but not in Merrifield's. Finally, the statistical significance of the relationships is generally less pronounced in the present study.

Comparison of the cross-section results for all elections does not reveal a consistent set of determinants of voter turnout. We do see a relatively robust relationship between voter turnout and the following explanatory variables: the holding of a line item veto authority by the Governor; the existence of a constitutional provision for direct or indirect legislation by petition and ballot initiatives; a pre-election day registration requirement; the population growth rate; the percentage of families below the poverty line; and the state land area. However, for the remaining explanatory variables, a robust significant relationship is not apparent. We next turn to the panel data results to shed further light on the determinants of voter turnout.

4.2 Panel results

Table 2 presents estimates of β and γ under each of the panel data models considered. The columns contain estimates of, respectively, the pooled OLS regression; the within or fixed effects estimator based on equation (5); the GLS estimator based on equation (7); the instrumental variables estimator based on equation (9); and the correlated random effects estimator based on equations (11) and (12). The first 4 listed variables are fixed through time, while the remainder are time-varying.

Pooled OLS estimates are contained in column 1. Based on the test statistic of equation (3), the null of no serial correlation in the residuals is rejected, implying the presence of fixed unobservable determinants of voter turnout. Incorporating only within variation, the fixed effects model (column 2) explains very little of the variation in the data. Time variation in voter turnout is significantly affected by time variation in measures of poverty, election day temperature, and unemployment. Other explanatory variables have no statistically significant impact. In the random effects model (column 3), we incorporate both within and between variation. However, the Hausman specification test (equation (8)) suggests that some of the explanatory variables

Table 2: Panel regressions

Model	Pooled OLS[a]	Fixed effects	Random effects (GLS)	Instrumental variables	Correlated random effects
Constant	27.12		27.86**	26.80**	9.558
	(18.13)		(9.176)	(9.188)	(19.68)
Area	0.032#		0.023	0.022	0.056**
	(0.016)		(0.015)	(0.015)	(0.014)
Line Item Veto	-4.211*		-3.817#	-3.959#	-7.304**
	(1.703)		(2.113)	(2.114)	(2.263)
Initiative	4.885**		6.041**	6.119**	3.145*
	(1.621)		(1.569)	(1.571)	(1.440)
Control	-0.620	0.211	-0.011	-0.071	0.310
	(0.790)	(0.584)	(0.572)	(0.581)	(0.295)
Executive officers	-0.309	-0.132	-0.426	-0.462	-0.494#
	(0.238)	(0.482)	(0.274)	(0.275)	(0.283)
Governor's party	-0.132	-0.295	-0.269	-0.351	-1.484**
	(0.954)	(0.556)	(0.549)	(0.554)	(0.283)
Legislators	-0.025**	0.028	-0.015	-0.015	0.021#
	(0.008)	(0.035)	(0.012)	(0.012)	(0.011)
Registration	-6.923*	1.313	-1.214	-1.251	-6.647**
	(2.756)	(1.433)	(1.308)	(1.309)	(1.880)
Population growth	-0.525**	-0.006	-0.143#	-0.136#	-0.517**
	(0.160)	(0.080)	(0.074)	(0.074)	(0.147)
Poverty	-0.997**	-0.286#	-0.556**	-0.486**	-0.358**
	(0.208)	(0.143)	(0.125)	(0.134)	(0.073)
Rain	-0.647	0.890	0.630	0.484	-1.102
	(1.597)	(0.942)	(0.930)	(0.932)	(1.799)
Temperature	0.025	0.067*	0.065*	0.061#	-0.093
	(0.053)	(0.032)	(0.032)	(0.032)	(0.070)
Unemployment	0.599*	0.623**	0.717**	0.753**	1.490**
	(0.229)	(0.158)	(0.151)	(0.153)	(0.071)
Senior	0.909	0.312	0.710#	0.714#	1.266**
	(0.813)	(0.506)	(0.376)	(0.377)	(0.417)
Youth	0.882**	0.144	0.398*	0.409*	1.254*
	(0.279)	(0.225)	(0.179)	(0.180)	(0.546)
R^2	0.441	0.104			
$\overline{R}^2$	0.416	-0.086			
Specification tests	3.273[b]		71.52[c]		998.8[d]
	{0.001}		{0.000}		{0.000}

Notes: Figures in parentheses are standard errors; figures in brackets {} are p-values.
#,*,** indicate significance at the 10%, 5%, 1% level for two-tailed tests.
[a] Robust standard errors are reported for the pooled OLS regression.
[b] Test for the presence of an unobserved effect. Refer to equation (3).
[c] Hausman (1978) test for correlation between α_s and X_{st}. Refer to equation (8).
[d] Omnibus test of the correlated random effects restrictions of equation (12).

are correlated with the fixed effect. Therefore, both the random effects model and pooled OLS are misspecified. To overcome this problem while making use of the between-groups variation and providing estimates of γ, the instrumental variables and correlated random effects methods outlined above are considered.

Column 4 of Table 2 contains instrumental variables estimates based on equation (9) under the assumption that the variables Control, Governor's party, Poverty, and Unemployment are correlated with the fixed effect. The instrumental variables estimates are very similar to the random effects estimates, suggesting that the instruments are in fact correlated with the endogenous variables. Many of the variables appear to influence voter turnout. However, of the political/institutional variables, a statistically significant relationship is found only for the existence of a Line Item Veto Authority, and the Initiative variable.

The final column contains estimates based on the correlated random effects model using the same set of endogenous variables. Coefficient estimates for the exogenous variables are obtained from the unrestricted reduced form given by equation (11). Estimates for endogenous variables are obtained by imposing the restrictions of equation (12). According to this model, with the exception of the Control variable, all of the political/institutional variables provide a statistically significant impact on voter turnout.

A direct comparison of the results of the IV and CRE models suggests they are generally similar, with a few notable exceptions. Qualitatively different results are found for the variables Governor's party, Legislators, Registration, and Temperature. Differences in results reflect the different assumptions of the two procedures. IV estimates depend on the properties of the instruments: the extent to which they are correlated with the endogenous variables, and whether they are truly orthogonal to the fixed effect. CRE estimates depend on the precision of the unrestricted estimates, and the validity of the restrictions implied by equation (12). The omnibus test implies an overwhelming rejection of the restrictions of equation (12). In addition, we might expect the unrestricted coefficient vector to be somewhat imprecisely estimated; a total of 208 coefficients are estimated from 349 observations. This suggests the CRE estimates may be treated with some caution.

Of the political variables, Table 2 suggests that the number of executive officials and the dummy variable for the governor's line-item veto authority both have a negative influence on turnout, while the dummy variable for ballot initiatives has a positive effect on voter turnout. These results are consistent with Merrifield (1993). A negative coefficient on the number of executive officials is consistent with the notion that raising the number of elected officials reduces the power of each, reducing the investment benefits from voting. Merrifield interprets a negative coefficient on the line-item veto variable as evidence that constraining state legislators' powers has a greater impact on turnout than constraining the powers of the governor. A positive coefficient on the ballot initiative variable provides some support for the suggestion that empowering voters tends to increase turnout.

No evidence is found for a role for the Control variable. Mixed evidence is found for the influence of the following variables on turnout: the political party of the governor; the number of legislators; and a pre-election day registration requirement. The CRE estimates suggest that with a Republican governor more voters are mobilised, but this is not strongly supported by IV estimates. The CRE model suggests a positive relationship between turnout and the number of legislators, while a negative relationship is suggested by the IV estimates. Finally, the CRE estimates suggest that a pre-election day registration requirement substantially discourages voter participation, while the IV estimates only suggest a mild negative influence.

Of the non-political variables, a positive relationship with turnout is suggested for the unemployment rate and both the fraction of people above 65 and below 18, while a negative relationship is suggested for the rate of population growth and the percentage of people below the poverty line. The strong positive impact of unemployment on turnout suggests that this issue may be important in motivating voters. By contrast, poverty tends to reduce turnout. The demographic results are consistent with the idea that older people and married people are more likely to vote, and that people are more likely to vote the longer they have lived in a particular area. Contrary to the findings of Merrifield (1993), but in line with Knack (1994), no clear relationship is found between turnout and the weather on election day.[7]

A comparison of the cross-section and panel results reveals that the results of individual cross-section regressions can be misleading. For example, a significant negative influence of the Control variable is found for the 1986 election, but a significant positive influence is found for 1998. By contrast, the panel data results are inconclusive for this variable. More generally, there is a great deal of variation in results across the different election years. In addition, the extra observations available in a panel provide additional identification power. For example, the poverty variable and the number of executive officers are found to have a significant negative influence on turnout using the full panel, while the results of the cross-section regressions are mixed. For the unemployment variable, the within variation appears to provide much of the identification.

5 Conclusions

This paper finds mixed support for Merrifield's (1993) result that institutional and political variables have an influential role in the American voter's electoral participation decision. Some evidence is found in support of the conclusion that constraining the powers of policy makers and the voting public tends to reduce turnout. A robust role was not found for the political party of the governor and little support was found for the notion that a political party controlling both houses and the governorship tends to mobilise voters to participate. Finally, mixed support is found for a negative impact on turnout stemming from pre-election day registration requirements.

A great deal of variation was found in the cross-section regression results indicating that a focus on a single cross section can potentially be misleading. For robustness, both an instrumental variables and a correlated random effects approach were considered for dealing with the correlation of some of the explanatory variables with unobserved persistent determinants of turnout. In some instances, these approaches yielded qualitatively different results. Possible future work could examine these cases further, and expand the scope of the study to incorporate cultural determinants of voter turnout.

References

[1] Abramson, P. R. and Aldrich, J. R. (1982). The decline of electoral participation in America. *American Political Science Review* 76: 502-521.

[7] One might expect that the effect of election day temperature would differ for states with different weather conditions. To examine this issue, an alternative specification was considered in which the temperature variable was split by interacting it with a dummy variable for those states with a normal temperature above 70, and a dummy variable for those states below 70. Even with this change, no robust role was found for the temperature variable.

[2] Briffault, R. (1996). *Balancing acts: The reality behind state balanced budget requirements*. New York: The Twentieth Century Fund Press.

[3] Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of Econometrics* 18: 5-46.

[4] Chamberlain, G. (1984). Panel data. In Z. Griliches and M. Intriligator (Eds.), *Handbook of Econometrics*, Volume 2, 1247-1318, Amsterdam: North Holland.

[5] Downs, A. (1957). *An economic theory of democracy*. New York: Harper and Row.

[6] Fair, J. and Dworkin, J. B. (1993). Determinants of voter participation: Some simulation results. *Public Choice* 77: 823-834.

[7] Fornos, C. A.; Power, T. J; and Garand, J. C. (2004). Explaining voter turnout in Latin America, 1980 to 2000. *Comparative Political Studies* 37(8): 909:940.

[8] Gelman, A.; King, G.; and Boscardin, W. (1998). Estimating the probability of events that have never occurred: When is your vote decisive?. *Journal of the American Statistical Association* 93(441): 1-9.

[9] Hausman, J. (1978). Specification tests in econometrics. *Econometrica* 46: 1251-1272.

[10] Hausman, J. and Taylor, W. (1981). Panel data and unobservable individual effects. *Econometrica* 49: 1377-1398.

[11] Jackman, R. W. (1987). Political institutions and voter turnout in the industrial democracies. *American Political Science Review* 81: 405-423.

[12] Jakubson, G. (1991). Estimation and testing of the union wage effect using panel data. *Review of Economic Studies* 58: 971-991.

[13] Knack, S. (1994). Does rain help the Republicans? Theory and evidence on turnout and the vote. *Public Choice* 79: 187-209.

[14] Matsusaka, J. G. (1993). Election closeness and voter turnout: Evidence from California ballot propositions. *Public Choice* 76: 313-334.

[15] Matsusaka, J. G. (1995). Explaining voter turnout patterns: An information theory. *Public Choice* 84: 91-117.

[16] Merrifield, J. (1993). The institutional and political factors that influence voter turnout. *Public Choice* 77: 657-667.

[17] Mulligan, C. and Hunter, C. (2003). The empirical frequency of a pivotal vote. *Public Choice* 116: 31-54.

[18] Riker, W. H. and Ordeshook, P. C. (1968). A theory of the calculus of voting. *American Political Science Review* 62: 25-42.

[19] Stigler, G. J. (1972). Economic competition and political competition. *Public Choice* 13: 91-106.

[20] Wooldridge, J. (2002). *Econometric analysis of cross section and panel data*. Cambridge: The MIT Press.

Data Appendix 1: Data sources

Variable	Description	Source
Turnout	Percentage of voting age population voting in the general election	(SA)
Area	Land area (millions of acres)	(SA)
Line Item Veto	Governor has line item veto authority (1=yes)	(TBS)
Initiative	Constitutional provision for direct/indirect legislation by petition and ballot initiatives (1=yes)	(DBM)
Control	One political party controls both houses of the legislature and the governorship (1=yes)	(TBS)
Executive officers	Number of separately elected executive branch officials	(CSG)
Governor's party	Political party of the governor (1=Democrat, 0.5=Independent, 0=Republican)	(TBS)
Legislators	Total number of state senators and representatives	(TBS)
Registration	Pre-election day registration requirement (1=yes)	(TBS)
Population growth	Population growth rate over the previous 4 years	(SA)
Poverty	Percentage of people below the poverty line	(SA)
Rain	Election day rainfall in the state's largest city (inches of water)	(NCDC)
Temperature	The difference between the actual and normal maximum temperature on election day (degrees F)	(NCDC)
Unemployment	The unemployment rate	(SA)
Senior	Percentage of the population aged 65 years and over	(SA)
Youth	Percentage of the population aged below 18	(SA)
Source key		
CSG	Directory of elected state officials, Council of State Governments	
DBM	Magleby, D. B. (1984), *Direct legislation: Voting on ballot propositions in the US*. Baltimore: Johns Hopkins Press.	
NCDC	National Climate Data Center	
SA	Statistical Abstracts of the United States (various years)	
TBS	The Book of the States (various years)	

Data Appendix 2: Descriptive statistics

Variable	Observations	Mean	Standard deviation	Minimum	Maximum
Turnout	349	38.421	9.514	3.500	59.400
Area	50	45.244	54.892	0.669	364.885
Line Item Veto	50	0.860	0.351	0.000	1.000
Initiative	50	0.460	0.503	0.000	1.000
Control	350	0.463	0.499	0.000	1.000
Executive officers	350	6.011	2.344	1.000	12.000
Governor’s party	350	0.591	0.489	0.000	1.000
Legislators	350	149.649	60.516	49.000	424.000
Registration	350	0.923	0.267	0.000	1.000
Population	350	4.726	5.119	-8.468	31.832
Poverty	350	12.911	4.012	3.700	26.600
Rain	350	0.113	0.275	0.000	2.170
Temperature	350	-0.130	8.680	-25.000	23.000
Unemployment	350	6.109	2.189	2.200	15.500
Senior	350	11.712	2.308	2.374	18.400
Youth	350	27.979	3.143	22.308	38.576