DEPARTMENT OF INTERNATIONAL AND EUROPEAN ECONOMIC STUDIES

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

State-Dependent Effect on Voter Turnout: The Case of US House Elections

Panagiotis Th. Konstantinou

Theodore Panagiotidis

Costas Roumanias

Working Paper Series

19-02

January 2019

State-Dependent Effect on Voter Turnout: The Case of US House Elections*

Panagiotis Th. Konstantinou†
Athens University of Economics and Business

Theodore Panagiotidis‡
University of Macedonia

Costas Roumanias§
Athens University of Economics and Business,

Abstract

In models of voter participation, the effects of election margin and campaign expenditure can be shown to be state-dependent – varying with low/high turnout. We empirically assess these implications for observed turnout, employing data from US House elections from 2000 to 2008 by means of quantile regression analysis. We document that the effects of expected election margin and campaign spending on turnout are state-dependent: the later is positive and decreasing, whereas the former is negative and U-shaped. Other determinants' influence on turnout (e.g. education, population density) is also shown to vary across the conditional distribution of turnout rate. Our findings are robust to a number of extensions.

JEL Classification: C21, D72
Keywords: Voter Turnout, Election Margin, Campaign Expenditure, Quantile Regression.
Current Version: January 25, 2019.

---

*We have benefited from discussions and comments on earlier drafts by Charalambos Christou, Evangelos Dioikitopoulos, Kevin Dowd, Kostas Drakos, Kevin Featherstone, Justin Fisher, Christos Genakos, Charles Grant, Nikitas Konstantinides, Sophocles Mavroides, Nassos Roussias, Anja Shortland, Francesco Trebbi, as well as seminar participants at Bath, Brunel and UOM, and conference participants at the Eastern Economic Association meeting. We wish to thank two anonymous referees for helpful and insightful comments. We particularly wish to thank Gary Jacobson for sharing his data on house elections, without which part of the estimations would have been impossible.

†Department of International and European Economic Studies, Athens University of Economics and Business, Athens, Greece.

‡Department of Economics, University of Macedonia, Thessaloniki, Greece.

§**Corresponding Author**. Department of International and European Economic Studies, Athens University of Economics and Business, 76 Patission Street, GR – 104 34 Athens, Greece. Email: roumanias@aueb.gr. Tel.+30-210-8203197, Fax: +30-210-8214945.

"This is going to be a turnout election"
President Barack Obama[1]

1 Introduction

Since Downs' (1957) seminal contribution, voter turnout has spurred a vivid debate on both theoretical and applied levels. From an empirical point of view, turnout and the factors that determine it have received considerable attention in policy circles, particularly around election times. One strand of the literature focuses on the importance of personal communication and social pressure in mobilizing voters to participate by means of experimental studies (e.g. Gerber and Green, 1999, 2000; Gerber, Green and Larimer, 2008). Another strand of the literature looks at the importance of political interest and sense of civic duty (Blais and St-Vincent, 2010) as well as the importance of belonging to large groups (Oberholzer-Gee and Waldfogel, 2005). A third strand of the literature attempts to identify factors that determine turnout within a country in the tradition of Cox and Munger (1989) at a more aggregate level (see e.g. Gentzkow, 2006; Shachar and Nalebuff, 1999; Washington, 2006).[2]

We follow the latter line of research, but approach voter participation from a different perspective. To motivate our empirical work, we note that a class of theoretical models predict that campaign spending increases, and expected election margin reduces turnout. However these effects depend on whether turnout is low or high i.e. they can be shown to be *state-dependent*. Relaxing the assumption of "symmetry", we employ quantile regression (QR) techniques, which allow modeling the entire conditional distribution of turnout rates. More importantly, QR allows the effects of key covariates of interest (e.g. expenditure and electoral margin) to vary at different (conditional) quantiles of the observed rate of participation. In a nutshell, we are addressing the following questions, (*i*) 'does money matter?', (*ii*) 'does expected closeness of a race matter?' for mobilizing voters, and (*iii*) 'when do these matter the most?', when turnout is high or low?

In particular, a class of the existing theoretical models of turnout imply that the partial effects of spending should be decreasing along the distribution of turnout, whereas the effects of margin should follow a U-shaped curve. Contrasting these implications with data from House elections over the period 2000–2008 is instructive.[3] We choose congressional elections for two reasons. First, the existence of 435 congressional districts in each election cycle in the US allows us to work with a relatively large number of observations over the time span we study, whereas examining Senate or Gubernatorial elections would leave us with with a much smaller sample. Second, in Senate or Gubernatorial elections the results would be "aggregated" and some of the heterogeneity we uncover might have been lost – we do control for concurrent elections as a robustness check of our results.

The fact that we assess the above theoretical results by casting our analysis within a quantile regression framework, produces a novel set of empirical results and bridges a gap in the existing literature, which focuses usually on average effects, masking potentially important

[1]Quote from the National Post on November 5, 2012.

[2]Another group of papers following Powell (1986) has focused on identifying the differences in turnout among countries (see the relevant discussion in Blais (2006)).

[3]The specific time span is dictated by data availability.

variation in the effect of campaign spending and expected margin across the turnout distribution. More importantly, our evidence is compatible with predictions stemming from existing theoretical models (see e.g. Herrera, Morelli and Palfrey, 2014; Palfrey and Rosenthal, 1985; Xefteris, 2018, *inter alia*), regarding the effect of spending and expected closeness on turnout. In some detail, we find that campaign spending has a positive and significant effect only for elections in which turnout is low — the effect becoming effectively zero (insignificant) for higher turnout rates. The effect of the margin of victory is negative and significant, and displays a U-shaped pattern for increasing quantiles of the distribution of turnout rate, in resonance with existing theoretical models. Finally, we provide new evidence on the effects other key controls have on the conditional distribution of turnout rate, and show that these effects are not uniform for different quantiles of its distribution, thereby extending the existing empirical results in the literature.

Apart from confirming existing theories on predictions that have so far, by and large, gone unnoticed in the literature, our results point towards important policy implications regarding voter turnout. In particular, insofar as turnout is affected in asymmetric ways across its conditional distribution by different variables, our results suggest that when considering policies aimed towards increasing voter participation, choosing the policies that best fit specific races might be imperative. For example, instruments that might help increase participation in races where turnout is expected to be high might perform poorly in races where turnout is expected to be low and vice versa.

Our empirical analysis relates to studies that model aggregate turnout rates.[4] For instance, Cox and Munger (1989) use campaign finance data from the Federal Election Commission (FEC) to examine the impact of expenditures and closeness on turnout in the 1982 US House elections. A more recent empirical literature examines how turnout depends on socio-geographical factors, such as constituency race, age, sex, education level, population density, etc. For example Gentzkow (2006), using county-level data, examines the importance of media in voter turnout and, in particular, how the gradual introduction of television on American soil affected turnout,[5] whereas Washington (2006) focuses on the positive effect Black Democratic candidates have on voter turnout. The approach we take departs from the existing literature in that it explicitly models the entire conditional distribution of voter participation rates, providing a richer set of empirical findings than those obtained by simply focusing on the conditional mean of the distribution of the rates of voter turnout.

The rest of the paper is structured as follows: Section 2 briefly reviews the theoretical predictions of existing theoretical models on turnout, our empirical methodology and an overview of the data we employ. Our empirical results are discussed in Section 3 which also contains some extensions and robustness analysis. The last section concludes.

[4] There is another strand of the literature that focuses on individual-level turnout. See for instance Matsusaka and Palda (1999) and Oberholzer-Gee and Waldfogel (2005, 2006) and the meta-analysis of Smets and van Ham (2013).

[5] Oberholzer-Gee and Waldfogel (2006) employing data on individuals, examine how introducing local news in Spanish motivates Hispanic voters.

2 Theoretical Predictions, Data and Empirical Methodology

2.1 Turnout, Spending and Electoral Margin: Theory and Prior Empirical Findings

Various theoretical studies focus on different factors which motivate electoral participation. Matakos, Troumpounis and Xefteris (2015, 2016) discuss the effect of polarization on turnout. Group-based models of voter turnout, such as Shachar and Nalebuff (1999) and Uhlaner (1989), emphasize the role of political leaders in mobilizing party followers to vote. Leaders coordinate voters with close ideological positions to vote in an attempt to maximize the party's probability of winning or its vote share. Dhillon and Peralta (2002) and Feddersen (2004) provide excellent surveys of the theoretical literature on voter turnout.

In a comprehensive survey, Geys (2006) looks into the aggregate-level empirical literature on voter turnout and its most significant determinants. Electorate size and electoral closeness are found significant frequently: turnout is higher when the election is close and when the size of the electorate is smaller. For instance, Arnold (2018) using data from mayoral elections from the German state of Bavaria concludes that electoral closeness matters as an increase in closeness by a standard deviation increases turnout by 1.27 percentage points. The effect of electoral margin (or closeness) in congressional elections has also been assessed in early contributions like Dawson and Zinser (1976) as well as in Crain, Leavens and Abbot (1987) and Cox and Munger (1989), where a negative relation between turnout rates and margins is documented. Moreover, turnout rates appear to be positively affected by a more stable population structure; to increase as campaign expenditures increase; and to be strongly affected by the legal framework (institutions such as compulsory voting, registration procedures etc.).[6] On the other hand, a positive effect of campaign spending on turnout in congressional elections is documented in Cox and Munger (1989).[7]

An extensive part of the theoretical literature on voter turnout focuses on the first or second order effects of key explanatory variables on turnout. The results are often readily interpretable in terms of how marginal effects might vary across the dependent variable's (turnout) distribution. For example, a large volume of theoretical literature (Ledyard, 1984; Palfrey and Rosenthal, 1985; Herrera et al., 2014; Arzumanyan and Polborn, 2017; Xefteris, 2018) predicts that the marginal effects of spending are positive and decreasing in magnitude for larger levels of spending. Following from the monotonic relationship of turnout and spending,ceteris paribus, the marginal effects of spending are expected to be more pronounced for lower levels of turnout, implying that we should observe empirically decreasing marginal effects across the turnout distribution.

Moreover, Herrera et al. (2014) examine turnout levels for different levels of probability

[6] Geys (2006) also assesses previous findings by means of a meta-analysis and he discusses how the determinants of turnout can be grouped in three categories: (*i*) socioeconomic, (*ii*) political and (*iii*) institutional variables, the first two of which are included in our work. The last one is employed for cross-country regressions and is not helpful in our case. See also Blais (2006) for another review.

[7] The positive effect of campaign spending has also been documented in early work by Settle and Abrams (1976) for presidential elections, in Caldeira and Patterson (1982) for state legislative elections and for gubernatorial elections in Patterson and Caldeira (1983).

that a voter has a preference for a party. Given that our data come from House elections (i.e. from a clearly majoritarian system, electoral margins are a function of the probability that a voter has a preference for one of the parties. Herrera et al.'s (2014) figure 1 implies that as turnout increases, the partial effects of margin (probability) on turnout are negative but small in magnitude for small levels of turnout, more negative for intermediate levels of turnout and less negative for large levels of turnout. So the partial effects of margin are expected to follow a U-shaped pattern as turnout increases.

2.2 Main Data

We employ data for all 435 congressional districts during the period 2000-2008 (5 election cycles).[8] Our response variable, the rate of voter turnout is defined as the percent of voting age population that actually voted.[9] In accordance with the theoretical models we discussed above, two covariates of interest are election margin and spending per voter. Election margin $m_{i,t}$ is defined as the percentage difference by which the winner was elected in office – regardless of political party affiliation, thereby including in our work also elections with independent candidates running for Congress.[10] Campaign expenditure is defined as the total amount per voter spent in election campaigns in any given election cycle (in 2000 US dollars), and allows us to assess the potential effects of spending on the decision to vote.

Further controls include the size of the electorate, education, income, demographic structure of the population and population density. Electorate size is approximated by (the log of) voting age population. In this manner, we avoid having the same covariate in the denominator of our response variable and as an explanatory variable at the same time. We proxy education by the percentage of individuals above 25 years of age that have completed at least four years of high school. Income is proxied by median family income (in 2000 US dollars). The demographic structure in each district is proxied by the fraction of individuals over 65 years of age in voting age population, while population density is measured by the number of persons per square mile.[11] We also control for on-year vs. off-year elections in an extension of our main results, as concurrent Presidential elections have been found to increase turnout strongly.[12] Finally, we also control for other concurrent elections (e.g. Senate and/or

[8] A full account of the data employed is available in a *not for publication appendix*, along with a set of summary statistics. It should be noted that our dataset does not have a panel structure, as the 'distribution' of congressional districts across states has changed at least once in our sample (after the 2000 Census). We discuss the issue of redistricting in our robustness analysis in the next section.

[9] The use of eligible population rather than voting age population is probably more appropriate as pointed out by McDonald and Popkin (2001). Measures of eligible population are not available at the congressional district but rather at the state level, and performing our empirical analysis at the (more aggregate) state level would miss important cross-district variation. Geys (2006) discusses that in many studies this is the most commonly used measure of turnout.

[10] Note that whenever there is an uncontested candidate, margin takes the value one. We should also highlight that in earlier versions, we employed the election margin between Democrats and Republicans. As there are many Independent candidates running for office, we have also included them in our analysis – this choice does not affect any of our conclusions.

[11] Data prior to 2000 at the Congressional District level are not available for demographic structure; and for voting age population, median family income, and education and population density before 1998. As a result it was not possible to extend our sample backwards in time.

[12] For an early theoretical contribution see Campbell (1987) and Knight (2014) for empirical evidence. Data

Gubernatorial) as these might also affect the decision of casting a vote.

2.3 Measuring *Ex Ante* Margin

Recall from our discussion above, that an important variable affecting the decision to participate, is expected margin (or expected closeness) by which a candidate leads prior to the final election outcome.[13] In principle, one could proxy this by employing poll data. Since such data are not easily available at the congressional district level, we make use of the following two political quality variables to obtain measures of *expected* margin.[14] The first (quality_{it}) provides qualitative information about the challenger (whether she held elective office or not, etc.), if there is one; or in the case of open-seat elections about the qualities of the candidates. The second (election_{it}) provides some background information about the specific race. In particular, these variables take the following values:

$$\text{quality}_{it} = \begin{cases} 0 & \text{if} \quad \text{Challenger has not held elective office} \\ 1 & \text{if} \quad \text{Challenger has held elective office} \\ 2 & \text{if} \quad \text{Only Democratic candidate for open seat has held elective office} \\ 3 & \text{if} \quad \text{Only Republican candidate for open seat has held elective office} \\ 4 & \text{if} \quad \text{Both candidates for open seat have held elective office} \\ 5 & \text{if} \quad \text{No challenger} \\ 6 & \text{if} \quad \text{No Democratic candidate (open seat)} \\ 7 & \text{if} \quad \text{No Republican candidate (open seat)} \end{cases} ;$$

and

$$\text{election}_{it} = \begin{cases} 0 & \text{if} \quad \text{Normal (Regular) Election} \\ 1 & \text{if} \quad \text{Unopposed candidate} \; (\,*\,) \\ 2 & \text{if} \quad \text{Incumbent switched parties since last election} \\ 3 & \text{if} \quad \text{Elective office held by candidate} \; (\,*\,*) \end{cases}$$

where ($*$) should be taken to imply that according to available data no other candidate spent a 'significant' amount of money in the race, and ($**$) should be taken to imply: (*i*) challenger was state legislator; or (*ii*) all candidates were state legislators (open seat elections); or (*iii*) challenger is former U.S. Representative.

Note that these variables are *exogenous* with respect to turnout, and known well before the actual election date. We obtain measures of *expected/predicted* margin in a simple way. We first estimate by OLS a model for margin, using only the qualitative variables as explanatory covariates, utilizing data up to the previous election cycle ($t-1$). We then employ the estimated coefficients from this model, use our political 'quality' variables for the current election cycle — which are known well before the elections in November — and obtain a

put together by Michael McDonald available at http://www.electproject.org/national-1789-present, also point in this direction.

[13]Perceptions of electoral competition are also found to lead to higher participation by McDonald and Tolbert (2012).

[14]These are drawing on data provided to us by Gary Jacobson.

prediction for election margin in period t (current election cycle).[15] To make our approach more transparent, we assume that the underlying process for predicting margin is given by

$$m_{i,t} = \gamma_{0,t-1} + \gamma_{1,t-1}\text{quality}_{i,t-1} + \gamma_{2,t-1}\text{election}_{i,t} + \eta_{i,t}, \tag{1}$$

where $\eta_{i,t}$ is a disturbance term. Employing data for 1998, we estimate $\hat{\gamma}_{0,98}, \hat{\gamma}_{1,98}$ and $\hat{\gamma}_{2,98}$ from (1), and using these parameter estimates we obtain *predicted* margin for 2000 in district i as

$$m^*_{i,00} = \hat{\gamma}_{0,98} + \hat{\gamma}_{1,98}\text{quality}_{i,00} + \hat{\gamma}_{298}\text{election}_{i,00}.$$

Then employing data for 1998 and 2000, we estimate $\hat{\gamma}_{0,98-00}, \hat{\gamma}_{1,98-00}$ and $\hat{\gamma}_{2,98-00}$ from (1), and using these parameter estimates *predicted* margin for 2002 in district i is constructed as

$$m^*_{i,02} = \hat{\gamma}_{0,98-00} + \hat{\gamma}_{1,98-00}\text{quality}_{i,02} + \hat{\gamma}_{298-00}\text{election}_{i,02}.$$

Continuing in the same manner, we obtain values of *ex ante* margin for all election rounds, and we avoid having to rely on the actual *ex post* margin, which in turn might be an outcome affected by turnout.[16]

2.4 Empirical Models and Estimators

Let y_{it} denote the observed participation rate (voter turnout) in congressional district i during period t.[17] It is common in the literature (see e.g. Cox and Munger, 1989; Gentzkow, 2006; Washington, 2006) to employ models of the form

$$y_{it} = \alpha + \mathbf{x}'_{it}\beta + u_{it}, \tag{2}$$

which under standard regularity conditions, allow the estimation of the average effect of $\mathbf{x}_{it}$ on y_{it}.[18] While interesting in their own right, focusing on average effects might mask variation in the partial effects of $\mathbf{x}_{it}$ across the turnout distribution. As an example, consider campaign expenditure: an increase in spending raises turnout rate on average, but it might do so more or less depending on whether turnout is high or low. In this paper, we follow an alternative approach and evaluate the effects of $\mathbf{x}_{it}$, at different points (quantiles) of the conditional distribution — hence allowing the effect of these covariates on turnout to be *state–dependent*: these effects depend on whether turnout is actually low or high (see also Bitler, Gelbach and Hoynes (2006) on how focusing on mean effects might miss important

[15] This strategy is similar in spirit to that of Shachar and Nalebuff (1999), who obtain a measure of *predicted closeness* in presidential elections. Their estimates are based on *fitted values* of a regression on variables known before the election date, while ours is an out-of-sample forecast. A different approach is adopted by Arnold (2018) who makes use of the constitutionally prescribed two-round elections to measure electoral closeness instead of using *ex post* margins.

[16] The correlation between the actual and the predicted margin is 0.68 (or the implied R^2 is about 0.47).

[17] We abuse slightly notation here, as while the actual number of congressional districts is the same over time, their distribution across states might change due to redistricting. So district i in 2000 need not be the same with district i in 2002 or later.

[18] If the variables in $\mathbf{x}_{it}$ are strictly exogenous, then estimation by OLS provides consistent estimates of these average effects, while IV or GMM techniques are required when some of the $\mathbf{x}_{it}$ are endogenous.

asymmetries).[19] This is done by employing the quantile regression estimator of Koenker and Bassett (1978) and Koenker (2005).

In particular, let any $\tau \in (0,1)$. We may define the τ-th quantile of y_{it} given $\mathbf{x}_{it}$, $Q_\tau(y_{it}|\mathbf{x}_{it})$, as the function that satisfies

$$\int_{-\infty}^{Q_\tau(y_{it}|\mathbf{x}_{it})} \varphi(y_{it}|\mathbf{x}_{it})dy = \tau,$$

where $\varphi(\cdot|\mathbf{x}_{it})$ is the conditional density of y given $\mathbf{x}$. Similarly to (2), we assume that the τ-th conditional quantile function is given as:

$$Q_\tau(y_{it}|\mathbf{x}_{it}) = \alpha_\tau + \mathbf{x}_{it}'\beta_\tau = \tilde{\mathbf{x}}_{it}'\tilde{\beta}_\tau, \tag{3}$$

where β_τ denotes the relevant slope parameters (partial effects) for the τ-th quantile, and $\tilde{\mathbf{x}}_{it} = [1, \mathbf{x}_{it}']'$. For instance, $\beta_{j,\tau}$ measures the effect of a unit change in $x_{j,it}$ on turnout rate, when turnout lies at it's τ-th conditional quantile.[20] As explained in Koenker and Bassett (1978), estimates of $\tilde{\beta}_\tau$, for any given value of τ, may be obtained as a solution of the following minimization problem:

$$\begin{aligned}
[\hat{\alpha}_\tau, \hat{\beta}_\tau']' &= \arg\min \sum_{\{y_{it} \geq \tilde{\mathbf{x}}_{it}' \tilde{\beta}_\tau\}} \tau|y_{it} - \tilde{\mathbf{x}}_{it}'\tilde{\beta}_\tau| + (1-\tau) \sum_{\{y_{it} < \tilde{\mathbf{x}}_{it}'\tilde{\beta}_\tau\}} |y_{it} - \tilde{\mathbf{x}}_{it}'\tilde{\beta}_\tau| && (4)\\
&= \arg\min \sum_{t=1}^{T}\sum_{i=1}^{N} \rho_\tau(y_{it} - \alpha_\tau - \mathbf{x}_{it}'\beta_\tau) && (5)
\end{aligned}$$

where τ represents the quantile under study and $\rho_\tau(u) = u(\tau - \mathbf{1}(u<0))$ for any $\tau \in (0,1)$ is the so-called "check function". The estimator does not have an explicit form, but the resulting minimization problem can be solved by linear programming techniques.

In order to obtain estimates of the standard errors, we make use of the design matrix bootstrap method (Buchinsky, 1995, 1998). In his Monte Carlo study, Buchinsky (1995) finds that this method performs relatively well, even in small samples; it is robust to changes of the bootstrap sample size relative to the data sample size;[21] and remains valid under many forms of heterogeneity.[22]

---

[19] One could also define *state dependence* if either spending or margin are low or high. We also explore this definition of *state dependence* when we assess the robustness of our main findings.

[20] That is $\beta_{j,\tau} = \partial Q_\tau(y_{it}|\mathbf{x}_{it})/\partial x_{j,it}$ which might vary at different values of τ, whereas the average effect is $\beta_j = \partial \mathrm{E}(y_{it}|\mathbf{x}_{it})/\partial x_{j,it}$. This should not be confused with the effects of margin or spending (and other covariates) when these variable attain high or low values. Focusing on margin, quantile regression allows us to answer what is the effect of an increase in margin by 1%, depending on whether turnout rate is low (say $\tau = 0.10$) or high (e.g. $\tau = 0.9$) – without any reference to the level of margin.

[21] The design matrix bootstrap method works as follows. One samples randomly with replacement from the original observations pairs $(y_j^*, \mathbf{x}_j^*)$ and computes α_τ^* and β_τ^*. Repeating this process L times, we obtain a sample of L $(k+1)$-vectors, whose sample covariance matrix is a valid estimator of the covariance matrix of the original estimator. Here, we use 5000 bootstrap replications to obtain the standard errors.

[22] In fact the design bootstrap matrix performs well even when the errors are homoskedastic (Buchinsky, 1995).

3 Empirical Findings

3.1 Main Empirical Findings

For the sake of comparison, we initially estimate by linear regression the effects of (*ex-post*) electoral margin, spending per voter, (log of) electorate size, education, income, demographic structure and population density on voter turnout. We argued above that such estimates might miss important state dependence inherent even in simple theoretical models; so we also estimate a set of quantile regressions at $\tau = 0.05, 0.25, 0.5, 075$ and 0.95. The results are summarized in Panel A of Table 1 and are also depicted graphically in Figure 1.

[Insert Table 1 and Figure 1 about here.]

Here, we discuss briefly the effects of electoral margin and spending per voter, and focus on the effects of the other covariates below. Starting with the OLS estimates (column 1 in Panel A of Table 1), we note a negative relationship between turnout and electoral margin and a positive relation between spending and election participation — both being significant. Results from quantile regressions (columns 2 to 6 in Panel A of Table 1) present a more complete picture. Margin and spending retain their sign and significance at different quantiles of electoral participation, but the estimated effects differ markedly across quantiles. For instance, the effects of margin differ in 'tail' and 'medium' quantiles: this negative effect is found to be more pronounced when closer to the median ($\tau = 0.5$). The effects of spending, on the other hand, are found to be larger at lower quantiles of turnout, and smaller when turnout is relatively high. These results are also represented graphically in Panels (A) and (B) of Figure 1, where we note a U-shaped pattern of the estimated effect of margin, and a declining effect of spending on voter turnout. These provide prima facie evidence in favor of state dependent effects.

Panel B of Table 1 also presents tests of the null hypothesis that the slope coefficients do not vary across different quantiles.[23] In particular, we perform a series of Wald tests of the null that the slope coefficients of margin ($\mathcal{W}_1$), expenditure ($\mathcal{W}_2$), margin and expenditure together ($\mathcal{W}_3$), and all slope coefficients together ($\mathcal{W}_4$) do not differ at specific quantile pairs τ_1 and τ_2. We also test the null that the slope coefficients are identical in all quantiles considered ($\mathcal{W}_5$). These results show that an important degree of state-dependence (heterogeneity) is present in the data: the partial effects of the covariates included in the model vary, depending on whether turnout is low or high, something which clearly cannot be captured by models focusing only on the conditional mean of the turnout process (OLS), matching the visual evidence from Figure 1. To our knowledge, these results are novel relative to the existing empirical literature on voter turnout.

Our findings thus far are based on a measure of *ex-post* margin, whereas theoretical models use *expected* margin as their integral part. In Table 2 and Figure 3, we report 'structural' estimation results, where we employ our measure of *expected/predicted* margin. We find, as anticipated, a significantly negative effect of expected election margin on turnout, which is not uniform across quantiles; and an asymmetric effect of spending, the higher the turnout.

[23] See Koenker and Bassett (1982) and Koenker (2005).

Panels (A) and (B) of Figure 2 display graphically the effects of *ex-post* margin and spending (from Table 1) and panels (C) and (D) the effects of *ex-ante* margin and spending (from Table 2) on electoral participation rates. From these, one could argue that the estimated effect of both *ex-post* and *expected* margin on election participation display a U-shaped pattern, attaining a minimum close to the median of the (conditional) turnout distribution – when using the *ex-ante* margin the pattern is more pronounced.[24] When electoral margin increases (and the turnout rate is becoming larger towards its median), it is less likely that voters will participate as the probability of (any voter) being pivotal becomes smaller. However, when turnout is large enough (above its median), any increase in the *expected* electoral margin still leads to lower turnout rate, but the resulting reduction is becoming smaller. A possible interpretation of this finding is that whenever the turnout rate is relatively high, there are other factors that influence the decision to participate in an election. To the extent that higher electoral participation is associated with the gravity of the election, voters will be more motivated to participate: so any increase in expected electoral margin will discourage a smaller fraction of voters from participating, as they are already highly motivated to show up and cast a vote.

[Insert Table 2 and Figures 2 & 3 about here.]

On the other hand, we find that the effect of spending is positive and significant across the whole distribution of voter turnout, however it is larger and increasing for quantiles below the 25% of the turnout distribution, and becomes progressively smaller for higher quantiles. Again the pattern of the effect of spending on electorate participation rates is in line with theoretical models discussed above, displaying again a great degree of state-dependence. This result is quite intuitive. Suppose that part of the resources spent by candidates is devoted to mobilizing voters to participate. An extra dollar is more likely to motivate voters to participate when turnout is low, as in this case their vote might be more decisive. The higher turnout is (hence the lower the probability for any voter of being pivotal) the lower is the effect of spending. Our results also indicate that there is a "bliss" level in the turnout distribution: for turnout rates beyond this point (higher quantiles of the turnout distribution), the effect of extra dollar on participation is dwindling down.

Moving away from margin and spending, it is insightful to inspect the effects of the other explanatory variables included in the model. A novel feature of our empirical analysis is that we document that these effects are state-dependent: they are heterogeneous across the distribution of turnout rates (see Panels (C) to (G) in Figures 1 and 3).[25] First, we find that the size of the electorate has a significant negative effect on turnout, in line with existing empirical findings and Downsian models of voting.[26] However, this negative effect becomes smaller

[24]The effect of *ex-post* margin attains its minimum just below the median, while the effect of *ex-ante* margin just below the 60th percentile of the conditional turnout distribution. Additionally, the effect of the former is flatter below the 40th percentile, whereas the effect of the latter is declining more smoothly until reaching its minimum.

[25]As the estimated effects of the other covariates in the two specifications are similar (compare Figures 1 and 3), we focus our discussion only on the latter.

[26]In such models, the larger the size of the electorate, the smaller the probability of being pivotal. So it is more likely that one abstains, as expected utility from voting is lower.

(in absolute value) for quantiles above the 40th percentile of the conditional distribution of turnout.

Turning next to education, we find it exerts a significantly positive effect throughout the distribution of voter turnout. This finding is in line with earlier findings in Dee (2004) and Milligan, Moretti and Oreopoulos (2004), who consider the effect of an individual's education on civic participation (see also Nie, Junn and Stehlik-Barry, 1996).[27,28] Although similar findings were obtained by Cox and Munger (1989) at an aggregate level – focusing on mean effects – our results show a slightly different picture: education not only has a positive effect on election participation, but this (increasing) effect is larger, the higher the turnout. Put differently, increasing the share of high-school graduates in the population, increases turnout rates, but the increase is even larger when the turnout rate is already high.

Examining the effect of median family income, we find that it has a bell-shaped pattern at different quantiles of turnout. In particular, we note that higher income districts are associated with higher turnout only at the middle of turnout distribution (between its 25th and 63th percentile), whereas income is not so important at either too low or too high rates of turnout.

Regarding the effect of demographic structure, we find that turnout is positively influenced by a larger fraction of population above 65 years, but again this effect is state-dependent. It is positive, significant, increasing for low quantiles (below 26th percentile) and decreasing for higher quantiles of the distribution of turnout; and becoming insignificant as a determinant of voter participation when turnout is already high – i.e. when it is above its 95% quantile. To build some intuition behind this result note that to the extent that elderly individuals have higher rates of participation, in low turnout races an increase in the proportion of elderly in the population raises turnout. In high turnout races however, even younger voters are mobilized, so the effect of age-structure diminishes: a change in the distribution of the population does not affect participation in an election in which almost all voters are involved.

Furthermore, we find that while population density affects the decision to vote negatively, its effect becomes essentially zero, when participation is already large (above the 82nd conditional percentile of turnout). This result is compatible with the theoretical argument that social structures are stronger in less densely populated areas, leading to higher turnout rates in these areas. This is because social pressure to participate is higher in areas where interpersonal relationships are stronger and voters are more likely to know the candidates personally (see for example Blank, 1974; Riker and Ordeshook, 1968). The existing literature is inconclusive as to the effect of population density on turnout (see Geys, 2006).[29]

[27]The idea is that an individual's educational development progress leads to an increase in civic skills and knowledge, resulting in greater political interest, mobilization, and involvement (higher turnout).

[28]Campante and Chor (2012) investigate the link between individual schooling and political participation. They employ individual survey data and control for country political institutions and cultural attitudes. They find that political participation is more (less) responsive to schooling in land (human capital) –abundant countries. Furthermore, evidence is provided that political participation is less responsive to schooling in countries with a higher skill premium.

[29]The disparate findings of previous empirical work on the effect of population density on turnout might be due to the fact that many studies at an aggregate level use smaller samples than ours, typically spanning one or two election cycles, hence focusing only on parts (i.e. the cross-sectional variation) and not the entire conditional distribution of turnout.

3.2 Extensions and Robustness Analysis

Our empirical findings are robust to a number of extensions, which we summarize below.[30] First of all, the notion of state dependence we employ pertains to the conditional distribution of turnout, which implies that the effects of *ex ante* election margin and spending per voter are non-linear.[31] On the other hand one could easily imagine situations where expenditure and margin may be used to define "state-dependence" (e.g. they are small or large). Looking at the (conditional) distribution of turnout allows us to examine the effectiveness of each of the control variables in mobilizing voters, when turnout is either small (low quantiles) or large (high quantiles); whereas looking at situations when expenditure or margin are small would allow us to evaluate how their changes affect voter turnout in a *given* (conditional) quantile of turnout.[32] In order to explore this possibility, we have estimated models which include low expenditure and low (predicted) margin as extra covariates. We define expenditure to be low when it lies at the bottom 10% of its distribution and (predicted) margin to be low when it lies at the lower 5% of its distribution.[33]

Results are reported graphically in Figure 4 for the low expenditure experiment and in Figure 5 for the low margin experiment. We first note that the differential effect of low expenditure on the conditional mean of turnout is essentially zero – although the point estimate is negative. We also find that low expenditure has a negative effect on turnout, but only when turnout is above its 85% quantile. This actually tells us that in races when turnout is already very high and expenditure is very low, if expenditure per voter increases by a dollar, turnout would actually be reduced. In all other situations when expenditure is small, increasing it does not seem to affect turnout in a significant way. Similarly, when margin is low, we find no effect on the conditional mean of turnout. Instead, when turnout is between its 49% and its 63% (conditional) quantiles and *ex ante* margin is low, an increase in *ex ante* margin by 1%, leads to a reduction in turnout by about 0.24%. Moreover, apart from the effect of education which is reduced considerably, but its pattern remains the same, the effects of all other variables remains almost identical to our baseline estimates in Table 2 and Figure 3.

[Insert Figures 4 and 5 about here.]

Our second robustness experiment involves controlling for unobserved heterogeneity at the congressional-district level.[34] Doing so, we are faced with two problems. First, our data do not have a panel structure, as redistricting has taken place at least once within the period we look at. While the number of congressional districts is always the same (435), their distribution across states might change – in addition to the geographical definition of each

[30] Most of the results discussed here, are not shown for the sake of brevity. They are detailed in a supplement, which includes also other robustness experiments and is available online.

[31] The same of course applies to all explanatory variables included in our models.

[32] We would like to thank an anonymous referee for raising this issue.

[33] Expenditure per voter is 'low' when it takes values up to \$0.892 in 2000 prices, and predicted margin is low when it attains values up to 13.22% roughly. Note here that minimum values for expenditure per voter and predicted margin are \$0.126 and 10.9% respectively.

[34] We also discuss controlling for state-level fixed effects in the online appendix. The results we obtain are similar to those reported here.

district.[35] Second, incorporating fixed-effects in our estimation practically requires that we have a sample with large enough T (i.e. elections over which districts are observed), which is not the case with our sample – turnout in each district is at best observed only four times. These caveats should be taken into account when one interprets the results that follow.

In order to control for district fixed effects, we follow the two-step method suggested by Canay (2011).[36] The estimated effects of each covariate in our model are shown graphically in Figure 6. Our results regarding margin and expenditure per voter are very similar to our benchmark results (Figure 3), while we also obtain very similar results for demographic structure. The pattern of the effects of the size of the electorate also change slightly, but not in a strong manner. The first notable difference concerns education and median family income: the effects of the former are very close to the estimate of education on the conditional mean to turnout; and the effects of median family income now displays a S-shaped pattern. The second notable difference concerns population density: its effect is significant for high quantiles of turnout, however the effect is now positive. A possible interpretation of this finding is that if population density does not vary much over time, the fixed effects essentially "pick up" the negative correlation between turnout rates and population density.

[Insert Figure 6 about here.]

Our third experiment concerns concurrent elections, since, as it has been argued before, voter participation is significantly higher during on-year elections (Campbell, 1987; Cox and Munger, 1989; Knight, 2014). More generally, voting in House elections might be heavily influenced by the existence of other concurrent elections. In such circumstances, if a voter decides to participate in one type of election, say Presidential, the cost of casting a vote for Congress is reduced considerably — if not driven down to zero. In order to control for the effects of concurrent elections, we include in our specification dummy variables for the coexistence of concurrent Presidential, Gubernatorial and Senate elections, as well as spending per voter in Gubernatorial and Senate elections.[37] Our main findings regarding the existence of state-dependence remain robust to this extension, but there are some notable differences, namely:[38] We find (*i*) that the effect of the electorate size resembles a U-shaped pattern; (*ii*) the effect of education is now monotonically increasing for larger quantiles of the turnout distribution; and (*iii*) the effect of median family income becomes negative for all quantiles of turnout (Blais (2006) further discusses the ambiguity with regard to the effect of socioeconomic factors on turnout). Moreover, (*iv*) we document again the estimated effect of expected margin retains its U-shaped pattern, with the minimum now attained at the 18% quantile. Finally, (*v*) we find that spending per voter is indistinguishable from zero.

[35] We also experimented with eliminating observations that were affected by redistricting from our analysis (about 15.2% of observations). Results shown in the online appendix indicate that none of our results are affected in any substantial manner.

[36] In the first step, one estimates the fixed–effects parameters, which are then subtracted from the turnout rates. The second step is a simple QR on the data obtained from the first step.

[37] Spending per voter in Gubernatorial and Senate elections are defined as total spending divided by voting age population in each state, as spending is available only at the state level. We do not include spending for Presidential elections, as this variable is available only at a national level, hence has no cross-sectional variation.

[38] The estimated effects of demographic structure and population density are almost identical with those reported in our main specification above.

In addition, we find that all concurrent elections increase significantly voter participation in Congressional elections. Concurrent Presidential elections lead to an increase in turnout by about 17% when turnout is below its median, and by about 15.5% when turnout is at is top 90% quantile. Concurrent Gubernatorial and Senate elections increase turnout between 1% and 2%, the latter being significantly above the 42% conditional quantile of turnout. As far as campaign expenditure in Gubernatorial and Senate elections is concerned, we find that it exerts a significant positive effect on voter participation, but for quantiles below the 50% of the turnout distribution. It is interesting to note that these effects are larger for very low quantiles of turnout and diminish towards zero close to the turnout median – resembling the diminishing effect of spending (in Congressional elections) on turnout we uncovered above. Essentially, in this experiment, the positive but declining effect of spending on electoral participation is picked up by spending in Gubernatorial and Senate elections. Based on this last finding one could argue that if spending aims at motivating voter participation, spending in different types of elections can be used as substitutes.

Additionally, one form of population heterogeneity that might influence aggregate voter turnout is the presence of 'minority groups' such as Blacks and Hispanics.[39] To control for this, we include in our specification the percentage of Blacks and Hispanics, as two extra covariates. We find (*i*) that the estimated coefficients in our main model above (Table 2 and Figure 3) remain largely unaffected; and (*ii*) that both measures of 'minority groups' reduce significantly voter turnout at the congressional district level in an asymmetric way. In particular, the percentage of Hispanic population has a significant negative effect on turnout throughout its distribution — having also a U-shape for quantiles below the median, and being more constant for higher quantiles, resembling the effect of electorate size on turnout. Instead, the percentage of black population has a significant negative effect on turnout rate only below its median, this effect being increasing and converging towards zero.[40]

Another form of asymmetry of the electorate relates to gender. We take this into account by including the fraction of males in voting age population for each congressional district in our sample. Once again, we find that our main results remain robust to this extension — the estimated effects of all the covariates in the model are qualitatively equivalent (allowing for sampling error). We also find that an increase in the share of males in the electorate does not influence turnout significantly in low turnout races (below the 18th percentile), but significantly reduces electoral participation as we move towards higher quantiles of turnout (the effect having a V-shape, attaining a minimum at about $\tau = 0.60$).

Finally, our sample of electoral outcomes includes competition in open seat elections, as well as in elections in which the incumbent won the election unopposed, i.e. challenger(s) spent essentially no money in her(their) campaign(s). One could argue that such elections are substantially different (e.g. in terms of the degree of competition between the candidates etc.) from "typical" elections, which could affect our empirical findings. To assess any differential effect of our covariates in "typical" elections, we have eliminated from our sample open-seat elections and elections with unopposed incumbents (reducing the number of observations by

[39]For example Shachar and Nalebuff (1999) find that the percentage of black population reduces significantly voter turnout in presidential elections. Geys (2006) contains a discussion of studies that find similar results.

[40]The findings in Oberholzer-Gee and Waldfogel (2005) indicate that an increased share of Blacks, increases black voter turnout, and at the same time it reduces white voter turnout – with the overall effect being negative.

462, from 2135 to 1673). Our findings — both qualitatively and quantitatively — remain virtually unaffected, with one minor difference regarding the estimated effect of expected margin. In particular, while the effect of expected margin retains its U-shaped pattern, the minimum is now attained at the 33% quantile, relative to the 58% quantile in the full sample.

4 Conclusions

Understanding the behavior of aggregate turnout rates requires detailed assessment of the fact that the response of turnout to different stimuli can be heterogeneous across the distribution of participation rates, in that low participation elections are influenced differently from high participation elections. A prominent feature of our work is that we take this explicitly into account and argue, that even in the context of simple theoretical models, the effects of key variables affecting electoral participation, e.g. expected margin of victory and campaign expenditures, will be state–dependent: they will depend on whether turnout is low or high. Notably, the effects of campaign expenditures and expected margin are — as intuition suggests — positive and negative respectively, but depend on the actual rate of turnout.

We empirically explored these issues using a dataset from the US House elections over the period 2000–2008, within a quantile regression framework. Our empirical results match predictions of existing theory. The effect of campaign spending on turnout rate is positive and significant, it attains its maximum value at about the 25th percentile of the conditional distribution turnout, and is declining for higher participation rates. The effect of margin is negative and also significant throughout the distribution of the rate of turnout, and displays a U-shaped pattern.

A novel feature of our work is to highlight that the effects of important determinants may be heterogeneous across the distribution of turnout rates, providing a more complete picture of the actual turnout process. In particular, while we confirm results of earlier empirical studies regarding the signs of important covariates, such as education, demographic structure, etc., we clearly show that these are far from uniform across the spectrum of turnout rates.

The evidence we present poses a number of new interesting questions, if we are to explain which forces drive political participation. For instance, why does education matter more in elections with high turnout? Or why does an increasing ratio of elderly voters raises turnout a lot only in low turnout elections? At the same time, from a practical, policy perspective, our results provide insights to policies that aim at increasing political participation. For example, if campaign spending is (partly) used to motivate participation, it is more effective in doing so in elections with relatively low participation. Moreover, it would be beneficial to invest resources in mobilizing less educated voters in elections, when turnout is expected to be relatively low, and from a long-run perspective, to increase the percentage of educated people in each congressional district. Similarly, in *low* turnout races, it might be worth addressing younger voters who seem to be more unresponsive than voters in higher age groups.

Overall, our results suggest that a better understanding can be reached by investigating how participation is influenced differently in electoral races where turnout is expected to differ, rather than assuming these patterns to be uniform. Such a research agenda can only lead to richer insights on what drives citizens to exercise their fundamental democratic right of voting. These are questions the existing literature has failed to raise, as "state–dependence" has so far been concealed by focusing on the "average" election.

References

Arnold, F. (2018). Turnout and closeness: Evidence from 60 years of bavarian mayoral elections, *The Scandinavian Journal of Economics* **120**(2): 624–65.

Arzumanyan, M. and Polborn, M. K. (2017). Costly voting with multiple candidates under plurality rule, *Games and Economic Behavior* **106**: 38–50.

Bitler, M. P., Gelbach, J. B. and Hoynes, H. W. (2006). What mean impacts miss: Distributional effects of welfare reform experiments, *American Economic Review* **96**(4): 988–1012.

Blais, A. (2006). What affects voter turnout?, *Annual Review of Political Science* **9**: 111–125.

Blais, A. and St-Vincent, S. L. (2010). Personality traits, political attitudes and the propensity to vote, *European Journal of Political Research* **50**: 395–417.

Blank, R. (1974). Socio-economic determinism of voting turnout: A challenge, *Journal of Politics* **36**(3): 731–752.

Buchinsky, M. (1995). Estimating the asymptotic covariance matrix for quantile regression models: a monte carlo study, *Journal of Econometrics* **68**(2): 303–338.

Buchinsky, M. (1998). Recent advances in quantile regression models: A practical guideline for empirical research, *Journal of Human Resources* **33**(1): 88–126.

Caldeira, G. A. and Patterson, S. C. (1982). Contextual influences on participation in u.s. state legislative election, *Legislative Studies Quarterly* **7**: 359–381.

Campante, F. R. and Chor, D. (2012). Schooling, political participation, and the economy, *Review of Economics and Statistics* **94**(4): 841–859.

Campbell, J. E. (1987). The revised theory of surge and decline, *American Journal of Political Science* **31**: 965–979.

Canay, I. A. (2011). A simple approach to quantile regression for panel data, *The Econometrics Journal* **14**(3): 368–386.

Cox, G. W. and Munger, M. C. (1989). Closeness, expenditures, and turnout in the 1982 U.S. house elections, *American Political Science Review* **83**(1): 217–231.

Crain, M., Leavens, D. R. and Abbot, L. (1987). Voting and not voting at the same time, *Public Choice* **53**: 221–229.

Dawson, P. A. and Zinser, J. E. (1976). Political finance and participation in congressional election, *Annals of the American Academy of Political and Social Science* **425**: 59–73.

Dee, T. S. (2004). Are there civic returns to education?, *Journal of Public Economics* **88**(9-10): 1697–1720.

Dhillon, A. and Peralta, S. (2002). Economic theories of voter turnout, *Economic Journal* **112**(480): F332–F352.

Downs, A. (1957). *An Economic Theory of Democracy*, Addison-Wesley, Boston.

Feddersen, T. (2004). Rational choice theory and the paradox of not voting, *Journal of Economic Perspectives* **18**(1): 99–112.

Gentzkow, M. (2006). Television and voter turnout, *Quarterly Journal of Economics* **121**(3): 931–972.

Gerber, A. S. and Green, D. P. (1999). Does canvassing increase voter turnout? a field experiment, *Proceedings of the National Academy of Sciences of the United States of America* **96**: 10939–10942.

Gerber, A. S. and Green, D. P. (2000). The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment, *American Political Science Review* **94**(3): 653–663.

Gerber, A. S., Green, D. P. and Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment, *American Political Science Review* **102**(1): 33–48.

Geys, B. (2006). Explaining voter turnout: A review of aggregate-level research, *Electoral Studies* **25**(4): 637–663.

Herrera, H., Morelli, M. and Palfrey, T. (2014). Turnout and power sharing, *Economic Journal* **124**: F131–F162.

Knight, B. G. (2014). An econometric evaluation of competing explanations for the midterm gap, *NBER Working Papers* **20311**.

Koenker, R. (2005). *Quantile Regression*, Econometric Society Monographs, Cambridge University Press, Cambridge.

Koenker, R. and Bassett, G. (1978). Regression quantiles, *Econometrica* **46**(1): 33–50.

Koenker, R. and Bassett, G. (1982). Tests of linear hypotheses and l_1 estimation, *Econometrica* **50**(6): 1577–1584.

Koenker, R. and Machado, J. A. F. (1999). Goodness of fit and related inference processes for quantile regression, *Journal of the American Statistical Association* **94**(448): 1296–1310.

Ledyard, J. (1984). The pure theory of large two-candidate elections, *Public Choice* **44**: 7–41.

Matakos, K., Troumpounis, O. and Xefteris, D. (2015). *The Political Economy of Governance*, Springer, chapter Turnout and polarization under alternative electoral systems, pp. 335–362.

Matakos, K., Troumpounis, O. and Xefteris, D. (2016). Electoral rule disproportionality and platform polarization, *American Journal of Political Science* **60**: 1026–1043.

Matsusaka, J. and Palda, F. (1999). Voter turnout: How much can we explain?, *Public Choice* **98**(3-4): 431–446.

McDonald, M. P. and Popkin, S. L. (2001). The myth of the vanishing voter, *American Political Science Review* **95**(4): 963–974.

McDonald, M. P. and Tolbert, C. J. (2012). Perceptions vs. actual exposure to electoral competition and effects on political participation, *Public Opinion Quarterly* **76**: 538–554.

Milligan, K., Moretti, E. and Oreopoulos, P. (2004). Does education improve citizenship? evidence from the united states and the united kingdom, *Journal of Public Economics* **88**(9-10): 1667–1695.

Nie, N. H., Junn, J. and Stehlik-Barry, K. (1996). *Education and Democratic Citizenship in America*, University of Chicago Press, Chicago, IL.

Oberholzer-Gee, F. and Waldfogel, J. (2005). Strength in numbers: Group size and political mobilization, *Journal of Law and Economics* **48**(1): 73–91.

Oberholzer-Gee, F. and Waldfogel, J. (2006). Media markets and localism: Does local news en español boost hispanic voter turnout?, *American Economic Review* **99**(5): 2120–28.

Palfrey, T. and Rosenthal, H. (1985). Voter participation and strategic uncertainty, *American Political Science Review* **79**(1): 62–78.

Patterson, S. C. and Caldeira, G. A. (1983). Getting out the vote: Participation in gubernatorial elections, *American Political Science Review* **77**: 675–89.

Powell, G. B. (1986). American voter turnout in comparative perspective, *The American Political Science Review* **80**(1): 17–43.

Riker, W. H. and Ordeshook, P. C. (1968). A theory of the calculus of voting, *American Political Science Review* **62**(1): 25–42.

Settle, R. F. and Abrams, B. A. (1976). The determinants of voter participation: A more general model, *Public Choice* **27**(1): 81–89.

Shachar, R. and Nalebuff, B. (1999). Follow the leader: Theory and evidence on political participation, *American Economic Review* **89**(3): 525–547.

Smets, K. and van Ham, C. (2013). The embarrassment of riches? a meta-analysis of individual-level research on voter turnout, *Electoral Studies* **32**(2): 344–359.

Uhlaner, C. (1989). Rational turnout: The neglected role of groups, *American Journal of Political Science* **32**(2): 390–422.

Washington, E. (2006). How black candidates affect voter turnout, *Quarterly Journal of Economics* **121**(3): 973–998.

Xefteris, D. (2018). Strategic voting when participation is costly, *Working Paper 12-2018*, University of Cyprus, Department of Economics.

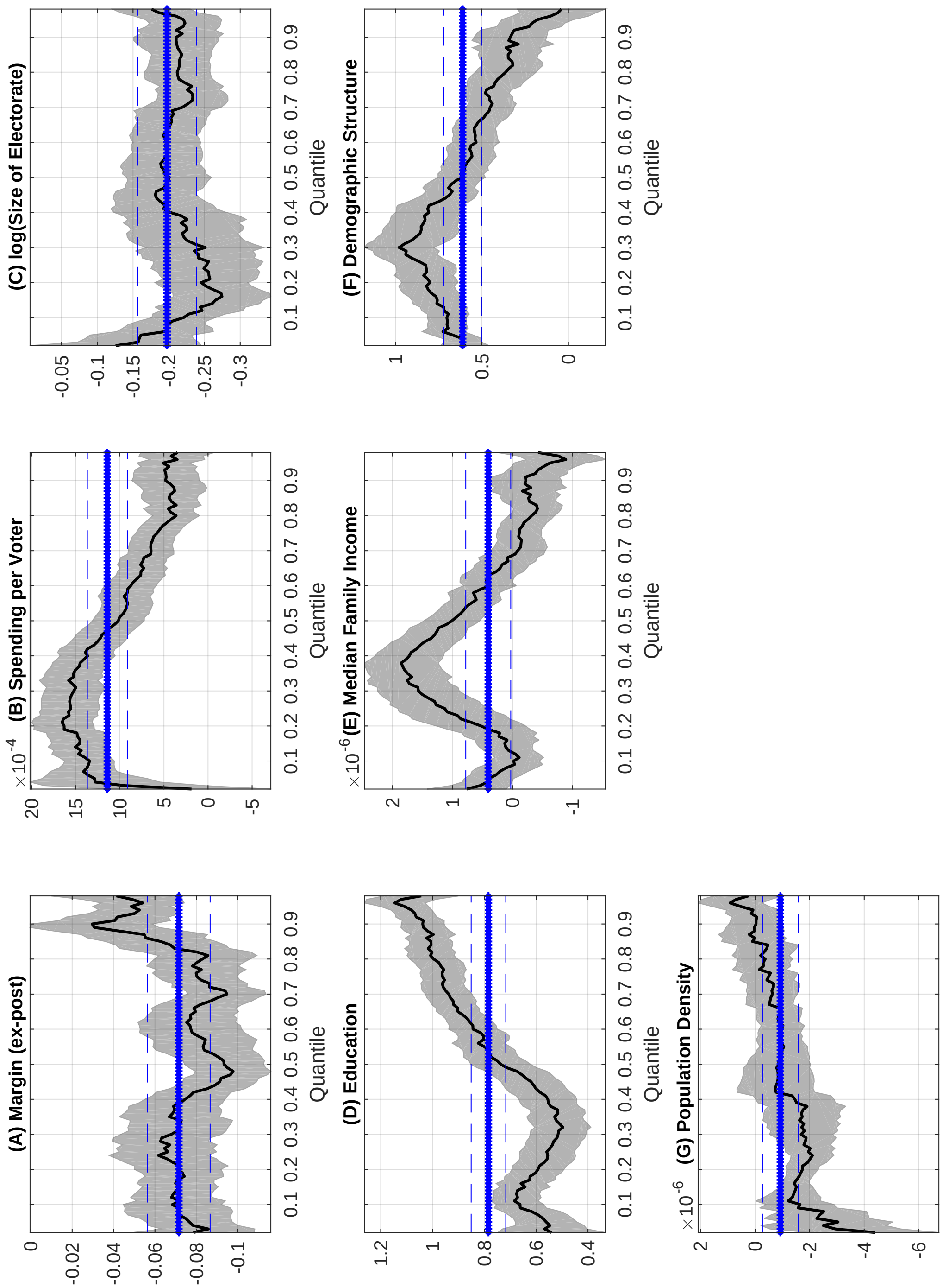

Figure 1: Quantile Regression Estimates of the Partial Effects of All Covariates Included in the Model.

The figure displays the estimated partial effects of all covariates in the model at different quantiles of the turnout rate process (see Table 1). The vertical axis measures the effect of a unit increase in each covariate on turnout rate; the horizontal axis corresponds to the quantiles of the conditional distribution. Shaded areas are the 90 percent confidence intervals of the estimated effects employing bootstrap standard errors, obtained using the design matrix bootstrap method discussed in text (see e.g. Buchinsky, 1995, 1998). Estimates are reported for $\tau \in [0.02, 0.98]$ at 0.1 unit intervals. The blue asterisks show the same partial effects on the conditional mean (estimated by OLS) and the dashed blue lines indicate the 90 percent confidence intervals around these estimates.

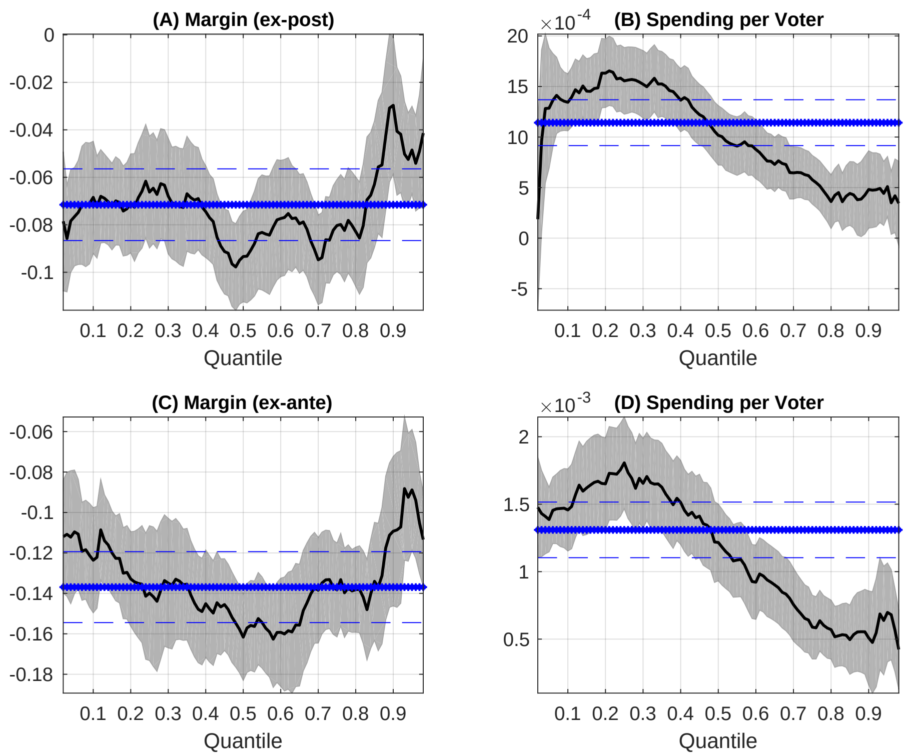

Figure 2: Empirical Quantile Regression Estimates of the Effects of Margin and Campaign Spending on Electoral Participation.

Panel (A) displays the effects of (ex-post) margin and Panel (B) the effects of campaign spending on voter participation, employing the specification estimated in Table 1. Panel (C) shows the effect of (ex-ante) margin and Panel (D) the effects of campaign spending on voter participation, based on the specification estimated in Table 2. See also notes for Figure 1.

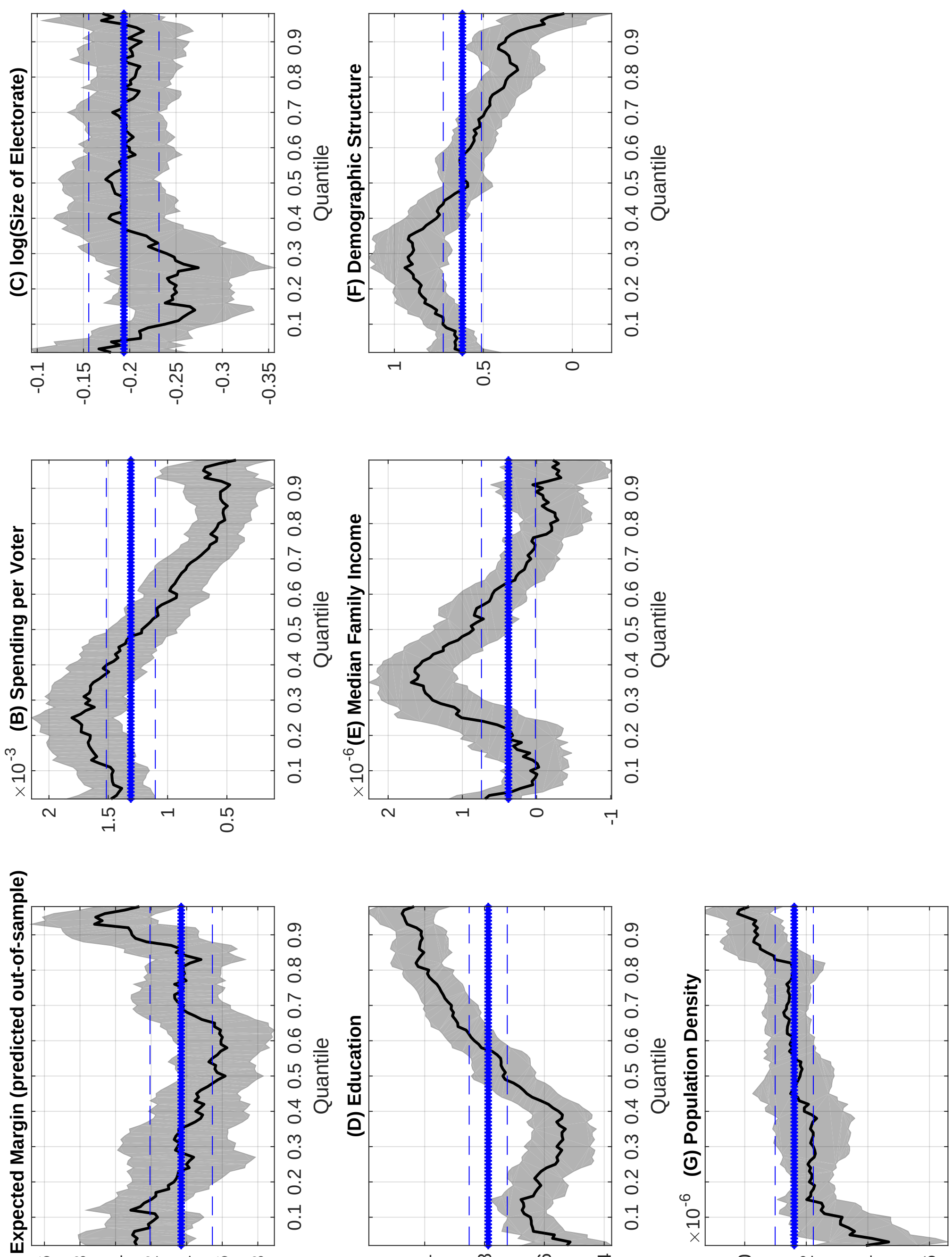

Figure 3: 'Structural' Quantile Regressions Estimates of Partial Effects of All Covariates Included in the Model.

The figure displays the estimated partial effects of all the covariates included in the model, when we employ the predicted/ex-ante measure of election margin described in text (see Table 2). See also notes for Figure 1.

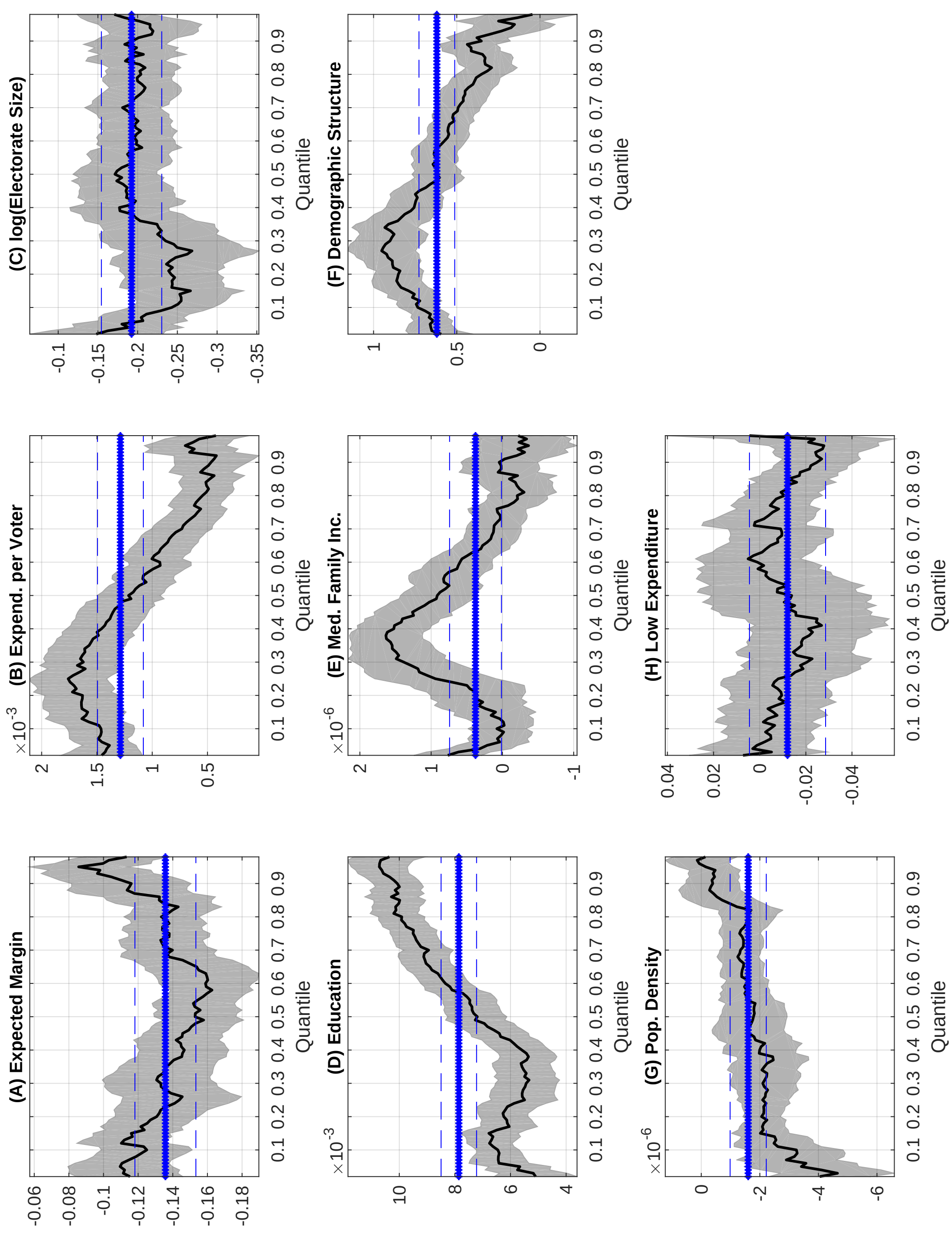

Figure 4: Quantile Regressions Estimates of Partial Effects Controlling for Low Expenditure. The figure displays the estimated partial effects of all the covariates included in the model, when we also control for very low expenditure (i.e. when it takes values at the lower 10% of the distribution of expenditure per voter). See also notes for Figure 1.

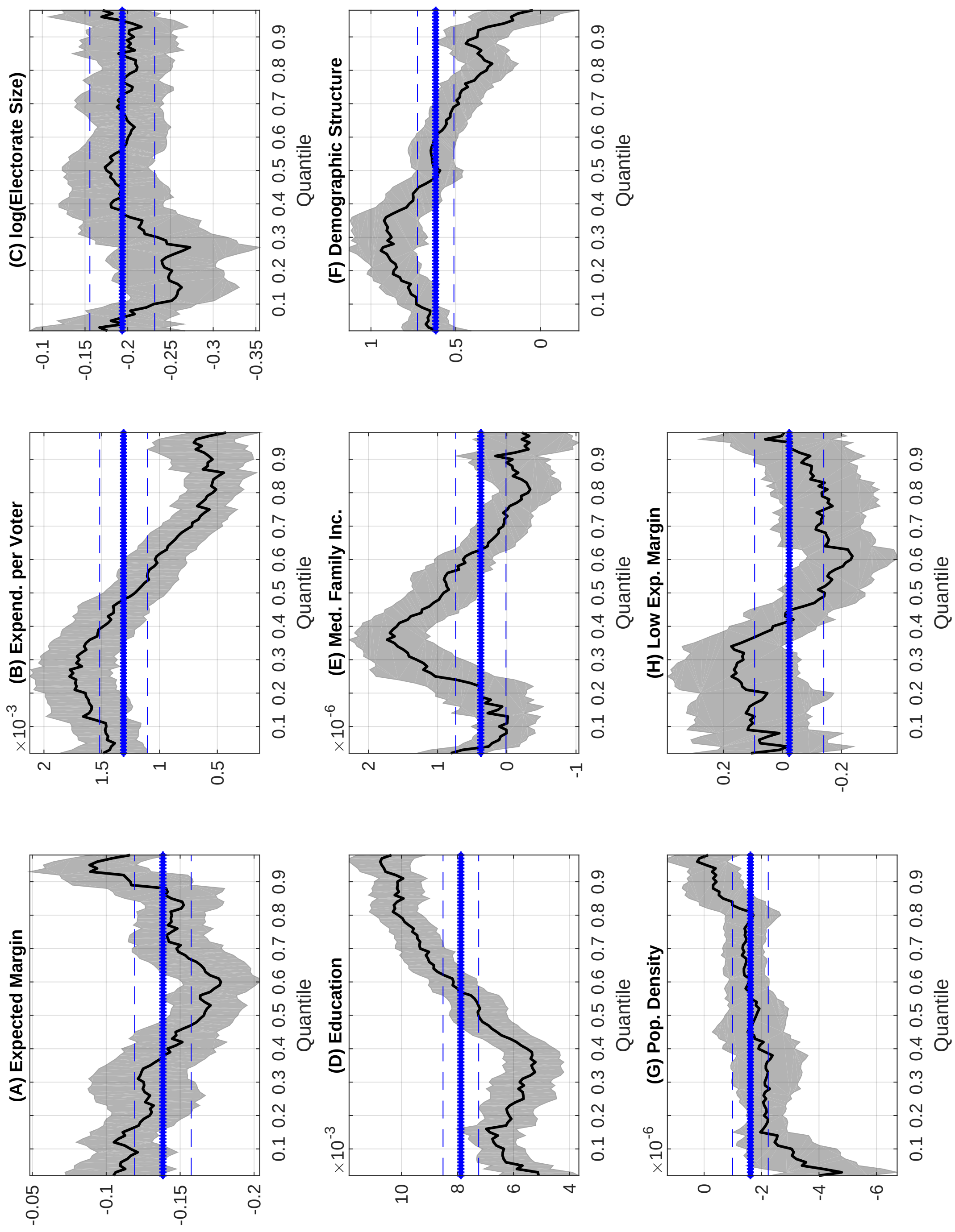

Figure 5: Quantile Regressions Estimates of Partial Effects Controlling for Low Expected Margin.

The figure displays the estimated partial effects of all the covariates included in the model, when we also control for very low predicted margin (i.e. when it takes values at the lower 5% of the distribution of predicted margin). See also notes for Figure 1.

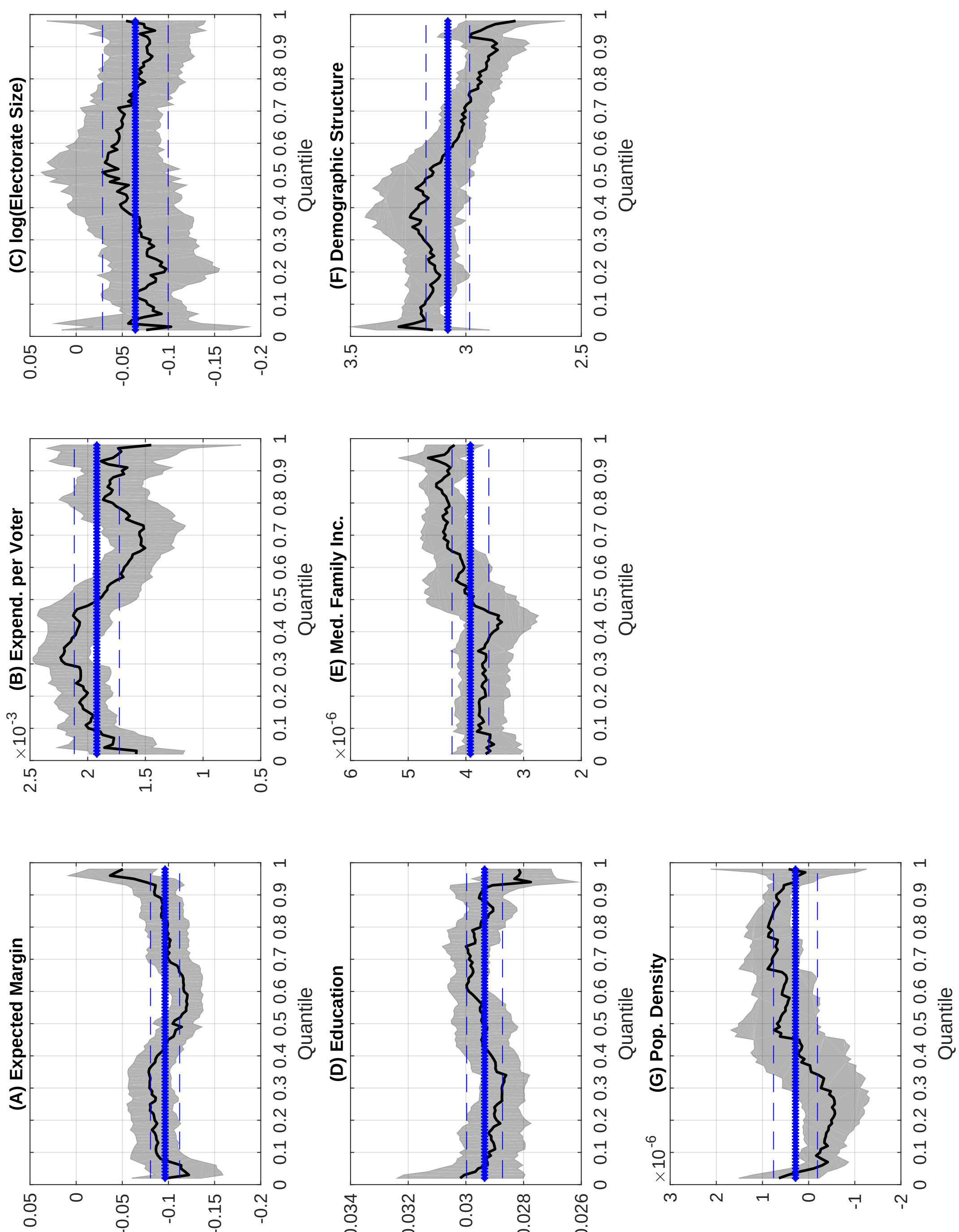

Figure 6: Quantile Regressions Estimates of Partial Effects Controlling for Congressional District FixedEffects.

The figure displays the estimated partial effects of the baseline set of covariates in the model, controlling for state fixed effects following the method suggested by Canay (2011). See also notes for Figure 1.

Table 1: **Main Model Results — OLS and Quantile Regression Estimates**

	Panel A: Parameter Estimates					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	**Quantile Regression**				
		$\tau = 0.05$	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.95$
Covariates	**Dependent variable: Turnout**					
α	2.280***	1.826***	3.045***	2.343***	2.632***	2.568
(s.e.)	(0.322)	(0.535)	(0.600)	(0.449)	(0.354)	(0.426)
Margin_{it}	-0.072***	-0.077***	-0.066***	-0.093***	-0.080***	-0.048***
(s.e.)	(0.009)	(0.012)	(0.014)	(0.011)	(0.012)	(0.011)
$\text{Spending per Voter}_{it} \times 100$	0.114***	0.129***	0.156***	0.102***	0.058***	0.051***
(s.e.$\times$100)	(0.014)	(0.037)	(0.020)	(0.018)	(0.015)	(0.019)
$\ln(\text{Electorate Size}_{it})$	-0.198***	-0.161***	-0.253***	-0.201***	-0.226***	-0.221***
(s.e.)	(0.025)	(0.043)	(0.046)	(0.035)	(0.027)	(0.033)
Education_{it}	0.784***	0.562***	0.559***	0.737***	0.962***	1.108***
(s.e.)	(0.041)	(0.070)	(0.064)	(0.070)	(0.050)	(0.057)
$\text{Median Income}_{it} \times 10^6$	0.404*	0.395	1.130**	1.010**	-0.114	-0.620*
(s.e.$\times 10^6$)	(0.228)	(0.245)	(0.476)	(0.276)	(0.224)	(0.366)
$\text{Demographic Structure}_{it}$	0.611***	0.668***	0.826***	0.617***	0.474***	0.178
(s.e.)	(0.066)	(0.073)	(0.105)	(0.083)	(0.089)	(0.127)
$\text{Population Density}_{it} \times 10^5$	-0.093***	-0.304**	-0.195***	-0.088	-0.050	0.053
(s.e.$\times 10^5$)	(0.040)	(0.124)	(0.050)	(0.070)	(0.060)	(0.064)
(Quasi) R^2	0.382	0.264	0.205	0.211	0.242	0.253

Table 1 Continued

Panel B: Test Statistics			
(I)	(II)	(III)	(IV)
$\mathcal{H}_0 : \beta_{0.05} = \beta_{0.25}$	$\mathcal{H}_0 : \beta_{0.25} = \beta_{0.50}$	$\mathcal{H}_0 : \beta_{0.50} = \beta_{0.75}$	$\mathcal{H}_0 : \beta_{0.75} = \beta_{0.95}$
$\mathcal{W}_1(1) = 0.473$ [0.491]	$W_1(1) = 4.607$ [0.032]	$W_1(1) = 1.295$ [0.255]	$W_1(1) = 6.541$ [0.011]
$\mathcal{W}_2(1) = 0.599$ [0.439]	$W_2(1) = 11.154$ [0.001]	$W_2(1) = 9.812$ [0.002]	$W_2(1) = 0.129$ [0.719]
$\mathcal{W}_3(2) = 1.064$ [0.587]	$W_3(2) = 13.806$ [0.001]	$W_3(2) = 12.614$ [0.002]	$W_3(2) = 6.977$ [0.031]
$\mathcal{W}_4(7) = 9.250$ [0.235]	$W_4(7) = 38.109$ [0.000]	$W_4(7) = 33.767$ [0.000]	$W_4(7) = 21.105$ [0.004]
$\mathcal{H}_0 : \beta_{0.05} = \beta_{0.25} = \beta_{0.50} = \beta_{0.75} = \beta_{0.95}$		$W_5(28) = 166.268$ [0.000]	

Notes for Table 1: Panel A shows OLS and QR estimates of a model of voter turnout which includes the (ex-post) election margin as an explanatory covariate. Estimates are based on $N = 2135$ (election year – congressional district) observations. The numbers in parentheses are standard errors. In the case of OLS these are corrected for heteroscedasticity using White's method. In the QR case, these are have been obtained using the design matrix bootstrap method discussed in text (see e.g. Buchinsky, 1995, 1998). The (quasi) $\bar{R}^2$ are the goodness-of-fit measures for quantile regressions discussed in Koenker and Machado (1999). One, two and three asterisks denote significance at the 10%, 5% and 1% significance level. Panel B reports a series of Wald-type test statistics, to assess whether the estimated slope coefficients obtained by QRs are equal across quantiles. The test statistics $\mathcal{W}_j(\nu)$ are distributed as $\chi^2_{(\nu)}$ variates. In particular, $\mathcal{W}_1$ ($\mathcal{W}_2$) evaluates the null that the slope coefficients of margin (spending) are identical across pairs of quantiles and $\mathcal{W}_3$ assesses whether jointly the slope coefficients of margin and spending are equal for pairs of quantiles; $\mathcal{W}_4$ tests the null that the all slope coefficients are identical for pairs of quantiles; and $\mathcal{W}_5$ tests the null all slope coefficients are identical across *all* quantiles. The numbers in square brackets are the p-values.

Table 2: **'Structural' Results — OLS and Quantile Regression Estimates of Voter Turnout**

Panel A: Parameter Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	**OLS**	**Quantile Regression**				
	(A)	(B)	$\tau = 0.05$	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.95$
Covariates	**Dependent variable: Turnout**						
α	2.280***	2.253***	2.104***	3.056***	2.119***	2.424***	2.243***
(s.e.)	(0.322)	(0.295)	(0.470)	(0.607)	(0.402)	(0.364)	(0.484)
Margin$_{it}$	-0.072***						
(s.e.)	(0.009)						
Margin$^*_{it}$		-0.137***	-0.110***	-0.140***	-0.162***	-0.138***	-0.089***
(s.e.)		(0.011)	(0.018)	(0.020)	(0.013)	(0.014)	(0.018)
Spending per Voter$_{it}$ $\times$ 100	0.114***	0.131***	0.139***	0.181***	0.122***	0.059***	0.070***
(s.e.$\times$100)	(0.014)	(0.013)	(0.015)	(0.020)	(0.017)	(0.014)	(0.022)
ln(Electorate Size$_{it}$)	-0.198***	-0.194***	-0.180***	-0.253***	-0.180***	-0.208***	-0.193***
(s.e.)	(0.025)	(0.023)	(0.038)	(0.047)	(0.031)	(0.029)	(0.038)
Education$_{it}$	0.784***	0.788***	0.566***	0.563***	0.739***	0.951***	1.069***
(s.e.)	(0.041)	(0.039)	(0.068)	(0.072)	(0.064)	(0.052)	(0.061)
Median Income$_{it}$ $\times 10^6$	0.404*	0.379*	0.215	1.010**	0.850***	0.065	-0.303
(s.e.$\times 10^6$)	(0.228)	(0.221)	(0.256)	(0.519)	(0.320)	(0.240)	(0.430)
Demographic Structure$_{it}$	0.611***	0.618***	0.660***	0.913***	0.589***	0.448***	0.157
(s.e.)	(0.066)	(0.065)	(0.073)	(0.109)	(0.083)	(0.093)	(0.139)
Population Density$_{it}$ $\times 10^5$	-0.093**	-0.161***	-0.342***	-0.205***	-0.180***	-0.144***	-0.007
(s.e.$\times 10^5$)	(0.040)	(0.377)	(0.118)	(0.053)	(0.066)	(0.044)	(0.060)
(Quasi) R^2	0.382	0.409	0.292	0.221	0.228	0.255	0.262

Panel B: Test Statistics

(I)	(II)	(III)	(IV)
$\mathcal{H}_0 : \beta_{0.05} = \beta_{0.25}$	$\mathcal{H}_0 : \beta_{0.25} = \beta_{0.50}$	$\mathcal{H}_0 : \beta_{0.50} = \beta_{0.75}$	$\mathcal{H}_0 : \beta_{0.75} = \beta_{0.95}$
$W_1(1) = 1.923$ [0.166]	$W_1(1) = 1.560$ [0.212]	$W_1(1) = 3.147$ [0.076]	$W_1(1) = 7.287$ [0.007]
$W_2(1) = 4.519$ [0.033]	$W_2(1) = 13.044$ [0.000]	$W_2(1) = 20.798$ [0.000]	$W_2(1) = 0.249$ [0.618]
$W_3(2) = 6.152$ [0.046]	$W_3(2) = 15.058$ [0.001]	$W_3(2) = 25.342$ [0.000]	$W_3(2) = 7.862$ [0.020]
$W_4(7) = 15.571$ [0.029]	$W_4(7) = 46.693$ [0.000]	$W_4(7) = 37.841$ [0.000]	$W_4(7) = 20.475$ [0.005]
$H_0 : \beta_{0.05} = \beta_{0.25} = \beta_{0.50} = \beta_{0.75} = \beta_{0.95}$		$W_5(28) = 176.661$ [0.000]	

Notes for Table 2: Panel A shows OLS estimates of models that include ex-post margin (same as in Table 1 for comparison) and predicted margin (Margin$^*_{it}$) in columns 1 and 2 respectively; and QR estimates of a model that includes predicted margin (Margin$^*_{it}$) as an explanatory variable. The number of observations is $N = 2135$. Panel B reports Wald-type test statistics, similar to those in Table 1. See also notes for Table 1.

Online Appendix for: "State-Dependent Effects on Voter Turnout: The Case of US House Elections"

Abstract

This appendix contains results discussed in the robustness analysis but not actually reported in the text. Furthermore additional robustness results are presented.

JEL Classification: D72, C21

Keywords: Voter Turnout, Election Margin, Campaign Expenditure, Quantile Regression.

Current Version: January 25, 2019

A Data Description

A.1 Main Covariates

Total Candidate Spending (in nominal terms): Federal Election Commission data, various files. The variable of interest has been calculated as the total sum of spending by all candidates in each election race.[1]
Nominal spending variables have been converted into real terms, employing the state–level GDP deflator. The GDP deflator has been constructed from data available by the BEA, as the ratio of the GDP by State[2] to state-level Quantity Indices for Real GDP by State (2000=100), again available from the BEA.[3]

The following variables have been obtained from Census' American Community Survey, with missing data having been interpolated using the average growth rates between 2000 and 2008.

• **Education**: We proxy education by the percentage of individuals over 25 years of age, who have completed at least four years of high school.

• **Voting Age Population (VAP)/Size of the electorate**: the number adults (18+) in each congressional district.

• **Over 65 years old**: Individuals with 65 or more years of age.

• **Population**: Total number of individuals in each congressional district.

• **Median Family Income**: Original data are in nominal terms and have been converted into real by employing the state-wide GDP deflator.

The following variables have been obtained from the "*Statistics of the Congressional Election*", available from the Office of the Clerk of the U.S. House of Representatives.

• **Number of Votes:** Total votes cast.

• **Number of Votes in favor of the Winner**: total number of votes received by the winner of the election.

• **Number of Votes in favor of the Runner up**: total number of votes received by the runner up candidate of the election.

Finally, **Land Area (square miles)** data were obtained from the U.S. Census Bureau.

• **Spending per voter**: constructed as total spending in real terms (using the state-wide GDP deflator), normalized by **VAP (Size of electorate)**.

[1]This includes spending by all persons that have either run for office (even in primary elections), even if they have received very few votes in the final election.

[2]Measured in millions of current dollars (All industry total).

[3]These are based on the 2002 the North American Industry Classification System (NAICS), measuring all industry totals.

• (***ex-post***) **Margin**: constructed as (**Nr. of Votes in favor of the Winner — Nr. of Votes in favor of the Runner up**)/(**Nr. of Votes – Scattering/Blank ballots**).[4]

• **Demographic Structure**: constructed as **Over 65 years old/VAP**.

• **Population Density**: constructed as **Population/Land Area**.

Finally, our dependent variable, **Turnout rate,** is defined as the total number of votes cast divided by voting age population (size of the electorate)**.**

A.2 Extra Control Variables used in our Robustness Experiments

Apart from the variables discussed above, we have also used the following variables in our robustness experiments:

• **Percent of Male Voters**: the percent of total males in voting age population, constructed as male voting age population to total voting age population, obtained from Census' American Community Survey.

• **Urbanization Rate**: the percent of population residing in urban areas. This is constructed as the ratio of urban residents (obtained from Census' American Community Survey) to total population.

• **Percent of Government Workers**: the percent of total government workers in Civilian Labor Force, obtained from Census' American Community Survey.

• **Total Spending in Senate Elections** (nominal terms): Federal Election Commission data, various files. This variables is constructed as the grant total of campaign expenditures by all candidates in each state.

• **Total Spending in Gubernatorial Elections** (nominal terms): We use data from the Gubernatorial Campaign Expenditures Database, compiled by Thad Beyle and Jennifer M. Jensen. This variables is constructed as the grant total of campaign expenditures by all candidates in each state.

Nominal spending variables have been converted into real terms employing the state-level GDP deflator, and subsequently expressed in spending per voter using state-wide voting age population (from Census).

Summary statistics of all the variables employed in our work are presented in Table A.1. It is important to highlight here that the correlation between the actual, *ex-post* (Margin_{it}) and the predicted/expected margin (Margin^*_{it}) is quite high, attaining a value 0.685, implying that our measure of expected margin tracks the actual margin quite accurately.

[4] The *Statistics of the Congressional Election* in many cases provide information regarding the number of **Scattering/Blank ballots**. For instance for the 2008 election, there are some Blank/Scattering ballots in various congressional districts in Massachusetts (see e.g. pages 31–32 of the *Statistics of the Presidential and Congressional Election of November 4, 2008*).

B Motivation and First Results

To begin our discussion, it is instructive to give an idea of the difference of the estimates obtained by modelling the conditional mean of the turnout process (estimated by OLS) and by modelling its conditional quantiles. To fix ideas, we estimate two sets of relations. The first set, models electoral participation rates as a function of (*ex-post*) margin only, and the second as a function of spending per voter. These estimated relations are depicted graphically in Figures A.1a and A.1b.

`[Insert Figures A.1 and A.2A.1b about here]`

An important finding is that the slope coefficients seem to differ notably.[5] Take for instance the effect of margin on turnout rates: the slope coefficients estimated for different (conditional) quantiles are quite different from the slope coefficient estimated at the conditional mean of the turnout process. Moreover, these slope coefficients differ also across (conditional) quantiles of turnout. In the same vein, the estimated partial effects of spending per voter across quantiles seem to differ from the slope coefficient estimated at the conditional mean of turnout rate. Based on these preliminary – and admittedly of limited use – results, there seems to be scope in exploring the degree of state dependence of the estimated effects of various determinants of electoral participation. We assess the magnitude of these effects at different (conditional) quantiles, as well as comparing these estimates with those obtained when focusing on the conditional mean (OLS) of the turnout rate process.

The estimated effects of margin (both ex-post and predicted) and spending per voter on electoral participation rates are depicted graphically in Figure A.3, replicating the results reported in Figure 2 in text. What is important is that the estimated partial effects of both spending per voter and margin differ markedly across different (conditional) quantiles of turnout rates. They are clearly different from the OLS estimates and display patterns similar to the one we obtained in our numerical analysis. In particular, the effects of margin on turnout rates display a U-shaped pattern, whereas the effects of spending on turnout are declining up to a certain quantile and become flat thereafter.

`[Insert Figure A.3 about here.]`

C Robustness Experiments

In this section, we provide an overview of our robustness exercises. To facilitate comparisons with our baseline results easier, we start by plotting the estimated partial effects of all the covariates included in our empirical model (see Table 2 in text) in Figure A.4. In what follows we report graphically the estimated partial effects from each of our robustness experiments.

[5]These findings should not necessarily be taken at face value, because of the lack of other controls in the specifications.

[Insert Figure A.4 about here.]

Our first robustness experiment involves the inclusion of 'minority' groups (Blacks and Hispanics) in our specifications. In particular, the baseline empirical specification is augmented by including the percentage of Blacks and Hispanics in voting age population (see Figure A.5). We note that the estimated coefficients in our baseline model (e.g. those in Figure A.4) remain virtually unaffected. In addition, we find that both measures of 'minority groups' reduce significantly voter turnout at the congressional district level, and in an asymmetric way: the percentage of Hispanic population significantly reduces turnout throughout its distribution — displaying a U-shapped for quantiles below the median, and being roughly constant for higher quantiles; while the percentage of black population has a significant negative effect on turnout rate only below its median – this effect increasing towards zero. Finally, including the percentages of Blacks and Hispanics in overall population rather than in the electorate does not make any difference to our findings (see Figure A.6).

[Insert Figures A.5 and A.6 about here.]

In our second robustness experiment, we account for gender. To this end, we include in our baseline specification the percentage of male voters in the electorate. The motivation is that male voters, more often than not, having a less holistic view of the issues are more likely to abstain in any election; as a result a higher fraction of male voters in any district should result in a lower participation rate. We find that our main results remain unchanged by this variation (see Figure A.7). Moreover, while the negative sign of the percentage of males is confirmed, we also find that the estimated partial effect of the percent of male voters on participation rates varies significantly across the distribution of voter turnout rates, being significant for races with turnout rates above the 20th percentile.

[Insert Figure A.7 about here.]

Our third robustness variation involves accounting for the presence of a large fraction of government workers in each congressional district. The idea is that these voters (workers) are more heavily and directly influenced by policies, and hence have more incentives to participate in any given election: so the larger the fraction of government workers, the higher turnout will be. Our results reported in Figure A.8 confirm that this is indeed the case. We note that the effect of the fraction of government workers on electoral participation rates is steadily declining for higher (conditional) quantiles of the turnout process. Moreover, our baseline results remain qualitatively uninfluenced by the inclusion of this extra covariate, the only change being that the effect of (log) electorate monotonically declines until roughly the 30th percentile of turnout, and then remains level.

[Insert Figure A.8 about here.]

Our following two robustness exercises involve first, measuring the size of the electorate using *total* rather than *voting age* population; and second, employing the degree of urbanization as an additional covariate in our empirical work. The idea behind the latter is that by including both urbanization rate and density as covariates in the model, allows us to control and distinguish between 'small city urban' and 'big city urban'. Our estimation results are reported in Figures A.9 and A.10 respectively. We find again that our main conclusions remain unaltered. In addition, we note that a higher rate of urbanization reduces significantly turnout rates (over and above the negative effect of population density), but in an state–dependent manner: the effects of urbanization on turnout display a U-shaped pattern.

[Insert Figures A.9 and A.10 about here.]

Next, in order to examine how much our results may be influenced by the different degree of competition in open seat elections, which may affect turnout rates differently, we performed two sets of experiments. We first eliminated all open seat elections from our sample, and then we eliminated all elections in which the incumbent faced no serious competition (as measured by the spending of challengers) in addition to the open seat elections. Our main conclusions and results remain immune to focusing only on 'regular' elections (see Figures A.11 and A.12.

[Insert Figures A.11 and A.12 about here.]

Perhaps the most important robustness exercise we perform is controlling for concurrent elections: these might increase turnout, as the cost voters face when participating in Congresional elections, depends on whether they have decided already to participate in other concurrent (presidential, senate and/or gubernatorial) elections. We discuss here only two experiments.[6] First, we evaluate whether on-year (presidential) elections differ from off-year elections and find that this is indeed the case: turnout is significantly higher (about 16%) during on-year elections, but more so when turnout rates are below the 65th percentile – the increase is contained to about 15% towards the 95th percentule (see Figures A.13 and A.14).[7] We then control also for concurrent gubernatorial and senate elections (see Figures A.23 and A.24). We find that all three types of elections, lead to higher turnout in Congress elections, with Gubernatorial elections increasing (Congress) turnout at low turnout rates and Senate elections increasing (Congress) turnout when it is above its 42nd percentile. Spending per voter in these election also increases turnout in Congressional elections, but the effect is diminishing for higher turnout elections. In general, as we discuss in the paper, our main conclusions remain qualitatively unaffected with few differences (which are outlined in **Section 5** of the paper).

[6] We do so for space conservation reasons. We provide evidence for more however.

[7] Figures A.14, A.16, A.18, A.20, A.22 and A.24 are identical to figures A.13, A.15. A.17, A.19, A.21 and A.23, the only difference being that the OLS estimates are not reported in the former graphs (A.14, A.16, A.18, A.20, A.22 and A.24). This is done only to facilitate comparisons with previous figures.

[Insert Figures A.13 to A.24 about here.]

Our next robustness experiment involves controlling for time-effects in our estimation, as there might be 'common' factors that influence elections at different districts. Time effects are accounted for by including a set of time dummies in the QR estimation. Our main results remain qualitatively unchanged (see Figure A.25) and are similar to those obtained when controlling for concurrent elections. This should come as no surprise, as the time dummies included coincide – to some extent – with concurrent elections. The only notable difference is that the effect of spending per voter is significant only for electoral races for which the turnout rate is between its 10th and 50th percentile.[8]

[Insert Figures A.25 to A.26 about here.]

When we employ the full set of observations (from 2000 to 2008) we do not explicitly take into account the fact that congressional districts across states (and in some cases within the same state) change because of redistricting. In order to see how much redistricting might impact on our results, we drop from our analysis all the congressional districts that were in fact affected.[9] To start with we note that our sample is reduced from 2142 to 1817 observations, so effectively about 15.2% of observations have been dropped from the sample. Our estimates are reported in Figure A.27, from which we note no difference regarding the pattern and magnitude of the estimated coefficients. The only effect that differs substantially relative to our baseline results is that of education: the pattern remains almost identical, but its effect falls by two orders of magnitude. Therefore, the inclusion of districts that are affected by redistricting in our baseline model, does not alter any of our conclusions in a material way.

[Insert Figure A.27 about here.]

Our final robustness experiment involves controlling for unobserved heterogeneity at the state level – we discuss controlling for district fixed effects in the paper. As we discuss in the paper, incorporating fixed-effects in our estimation practically requires that we have a sample with large enough T (i.e. elections over which states are observed) which is not the case with our sample – each state is at best observed only four times. Here we follow the two-step method suggested by Canay (2011) in which one first estimates the state–fixed effects (we do so by OLS), which are then subtracted from the turnout rates and then we estimate the parameters of interest by simple quantile regression on the data obtained from the first step. The estimated effects of each covariate in our model are shown graphically in

[8]Figure A.26 is identical to figure A.25, the only difference being that the OLS estimates are not reported in the graphs. This is again done only to facilitate comparisons with previous figures.

[9]We would like to thank an anonymous referee for raising this issue.

figures A.28 and A.29.[10] Our results remain virtually unchanged relative to our benchmark results (Figure A.4). The only difference we note concerns education, the effect of which is about two orders of magnitude smaller, but the pattern across the distribution of turnout remains the same.

[Insert Figures A.28 and A.29 about here.]

[10] Figure A.29 is identical to figure A.28, the only difference being that the OLS estimates have been eliminated from the graph. This is done only to facilitate comparisons with previous figures.

References

Buchinsky, M. (1995). Estimating the asymptotic covariance matrix for quantile regression models: a monte carlo study, *Journal of Econometrics* **68**(2): 303–338.

Buchinsky, M. (1998). Recent advances in quantile regression models: A practical guideline for empirical research, *Journal of Human Resources* **33**(1): 88–126.
URL: *www.jstor.org/stable/pdfplus/146316*

Canay, I. A. (2011). A simple approach to quantile regression for panel data, *The Econometrics Journal* **14**(3): 368–386.

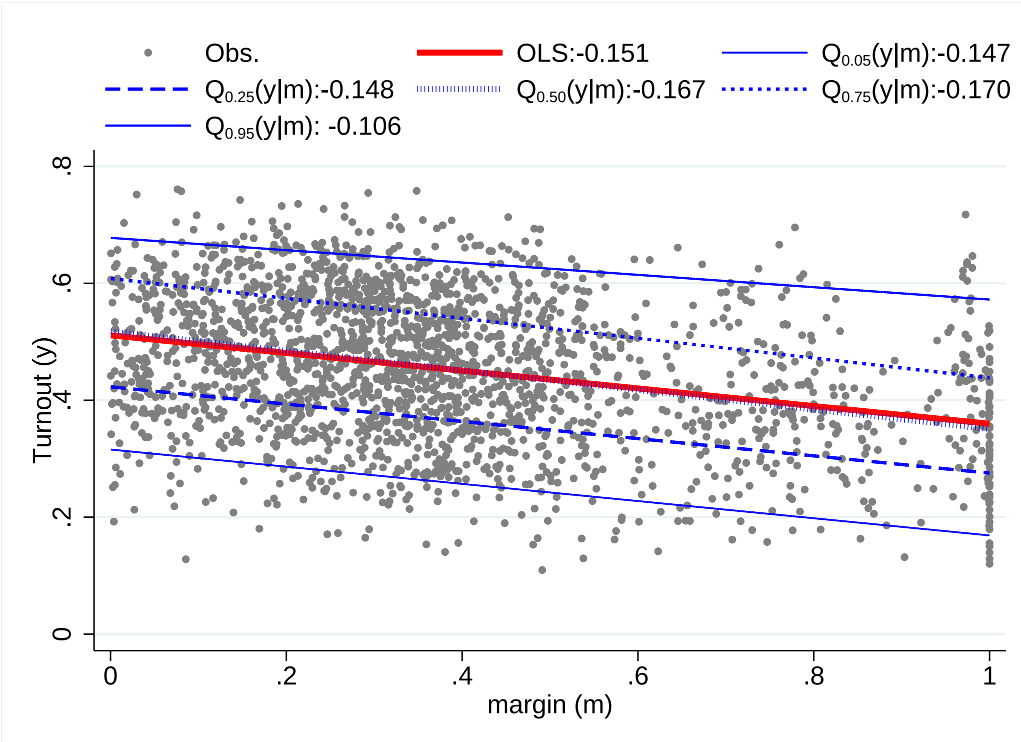

Figure A.1: The effects of margin (m) on turnout (y)

Notes: The figure plots the estimated regression lines of turnout (y) on (*ex-post*) electoral margin (m). In particular it plots the OLS estimates of a relation $y_{it} = b_0 + b_1 m_{it} + \varepsilon_{it}$ and the quantile regression estimates of the same relation, for different (conditional) quantiles, $\tau : Q_\tau(y_{it}|m_{it}) = b_{0,\tau} + b_{1,\tau} m_{it}$. The numbers reported next to the legends are the estimated slope coefficients of the regression at different quantiles. In particular, the estimated relations are:

$\mathrm{E}(y_{it}|m_{it}) = 0.511 - 0.151 m_{it}$

$Q_{0.05}(y_{it}|m_{it}) = 0.316 - 0.147 m_{it}$

$Q_{0.25}(y_{it}|m_{it}) = 0.423 - 0.148 m_{it}$

$Q_{0.50}(y_{it}|m_{it}) = 0.518 - 0.167 m_{it}$

$Q_{0.75}(y_{it}|m_{it}) = 0.608 - 0.170 m_{it}$

$Q_{0.95}(y_{it}|m_{it}) = 0.678 - 0.106 m_{it}$

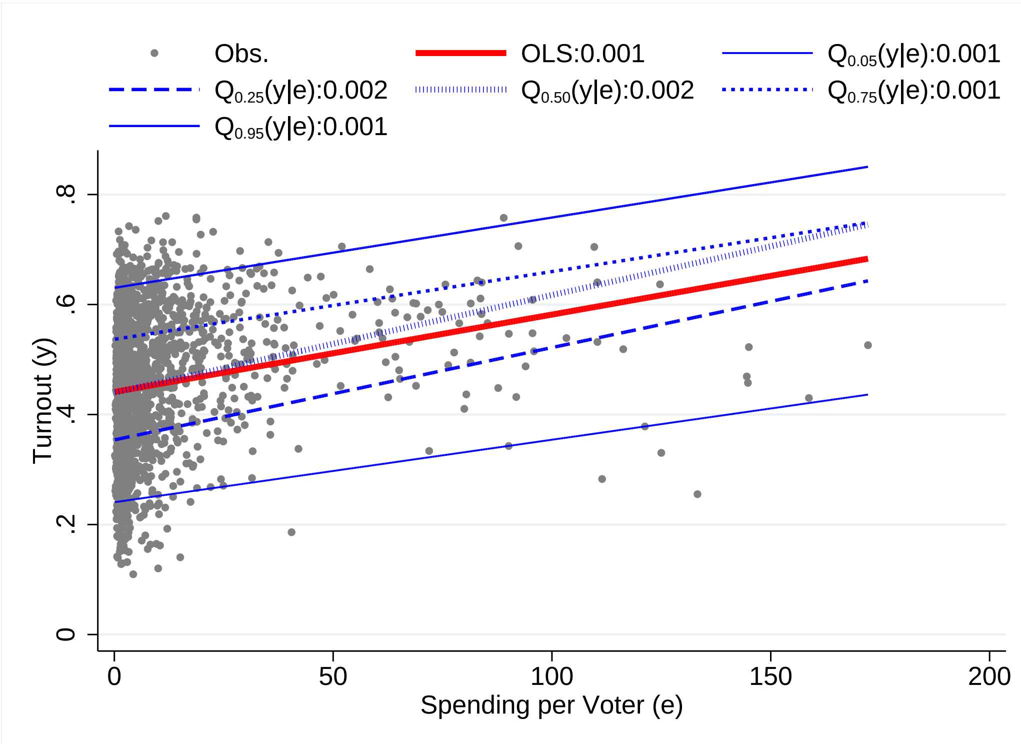

Figure A.2: Quantile Regression Estimates of the Effects of Margin and Campaign Spending on Electoral Participation

Notes: Notes: The figure plots the estimated regression lines of turnout (y) on spending per voter (e). In particular it plots the OLS estimates of a relation $y_{it} = c_0 + c_1 e_{it} + \varepsilon_{it}$ and the quantile regression estimates of the same relation, for different (conditional) quantiles, $\tau : Q_\tau(y_{it}|e_{it}) = c_{0,\tau} + c_{1,\tau} e_{it}$. The numbers reported next to the legends are the estimated slope coefficients of the regression at different quantiles. In particular, the estimated relations are:

$\mathrm{E}(y_{it}|e_{it}) = 0.441 + 0.0014 e_{it}$

$Q_{0.05}(y_{it}|e_{it}) = 0.241 + 0.0011 e_{it}$

$Q_{0.25}(y_{it}|e_{it}) = 0.354 + 0.0017 e_{it}$

$Q_{0.50}(y_{it}|e_{it}) = 0.440 + 0.0018 e_{it}$

$Q_{0.75}(y_{it}|e_{it}) = 0.537 + 0.0012 e_{it}$

$Q_{0.95}(y_{it}|e_{it}) = 0.630 + 0.0013 e_{it}$

See also notes for Figure **??**

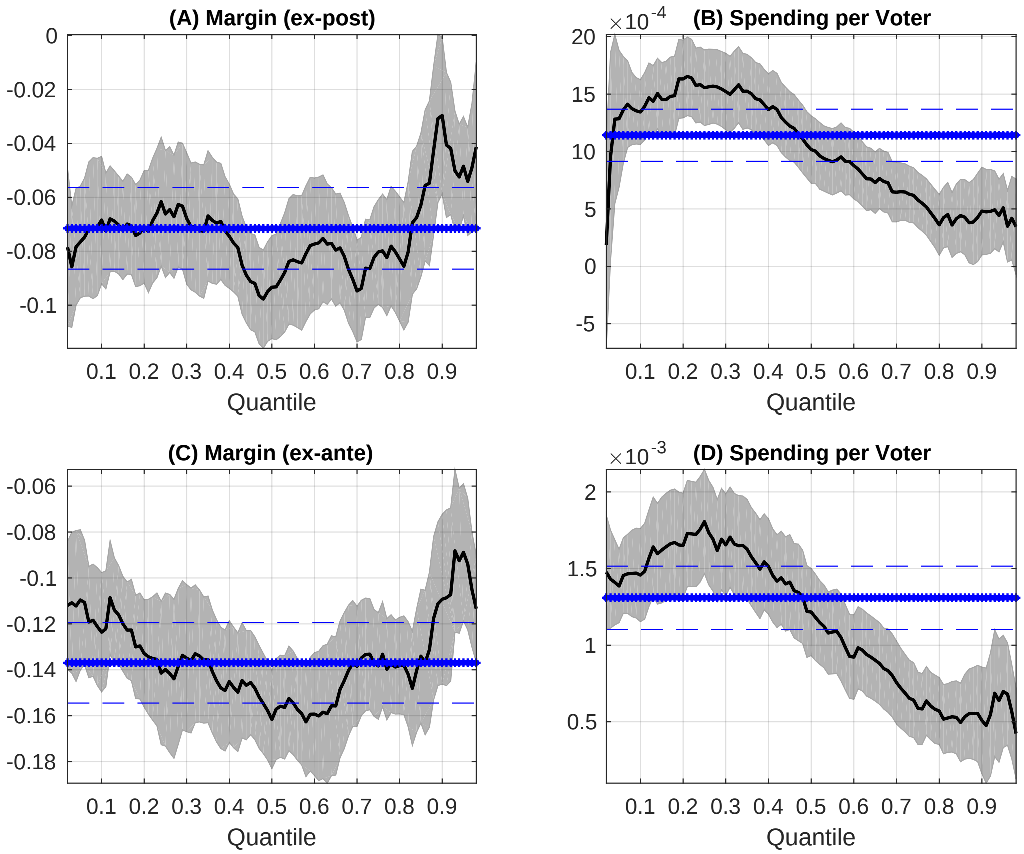

Figure A.3: Quantile Regression Estimates of the Effects of Margin and Campaign Spending on Electoral Participation

Notes: This figure is identical to Figure 2 in text. Panel (A) displays the effects of (*ex-post*) margin, Panel (B) the effects of campaign spending on voter participation, Panel (C) shows the effect of (*ex-ante*) margin and Panel (D) the effects of campaign spending on voter participation. The vertical axis measures the effect of a unit increase in each covariate on turnout rate; the horizontal axis corresponds to the quantiles of the conditional distribution. Shaded areas are the 90 percent confidence intervals of the estimated effects employing bootstrap standard errors, obtained using the design matrix bootstrap method discussed in text (see e.g. Buchinsky, 1995, 1998). Estimates are reported for $\tau \in [0.02, 0.98]$ at 0.01 unit intervals. The blue asterisks show the same partial effects on the conditional mean (estimated by OLS) and the dashed blue lines indicate the 90 percent confidence intervals around these estimates.

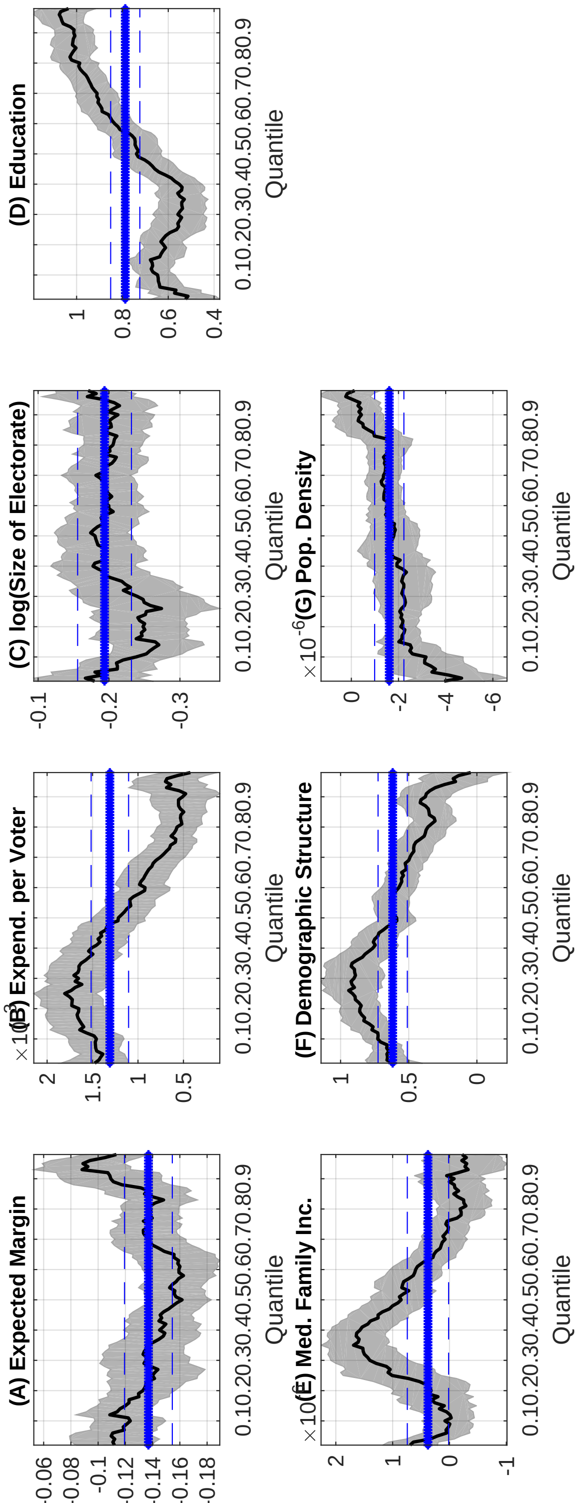

Figure A.4: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation.

Notes: The figure displays the estimated partial effects of all covariates in the model at different quantiles of the turnout rate process (see Table 2 in text). See also notes for Figure A.3.

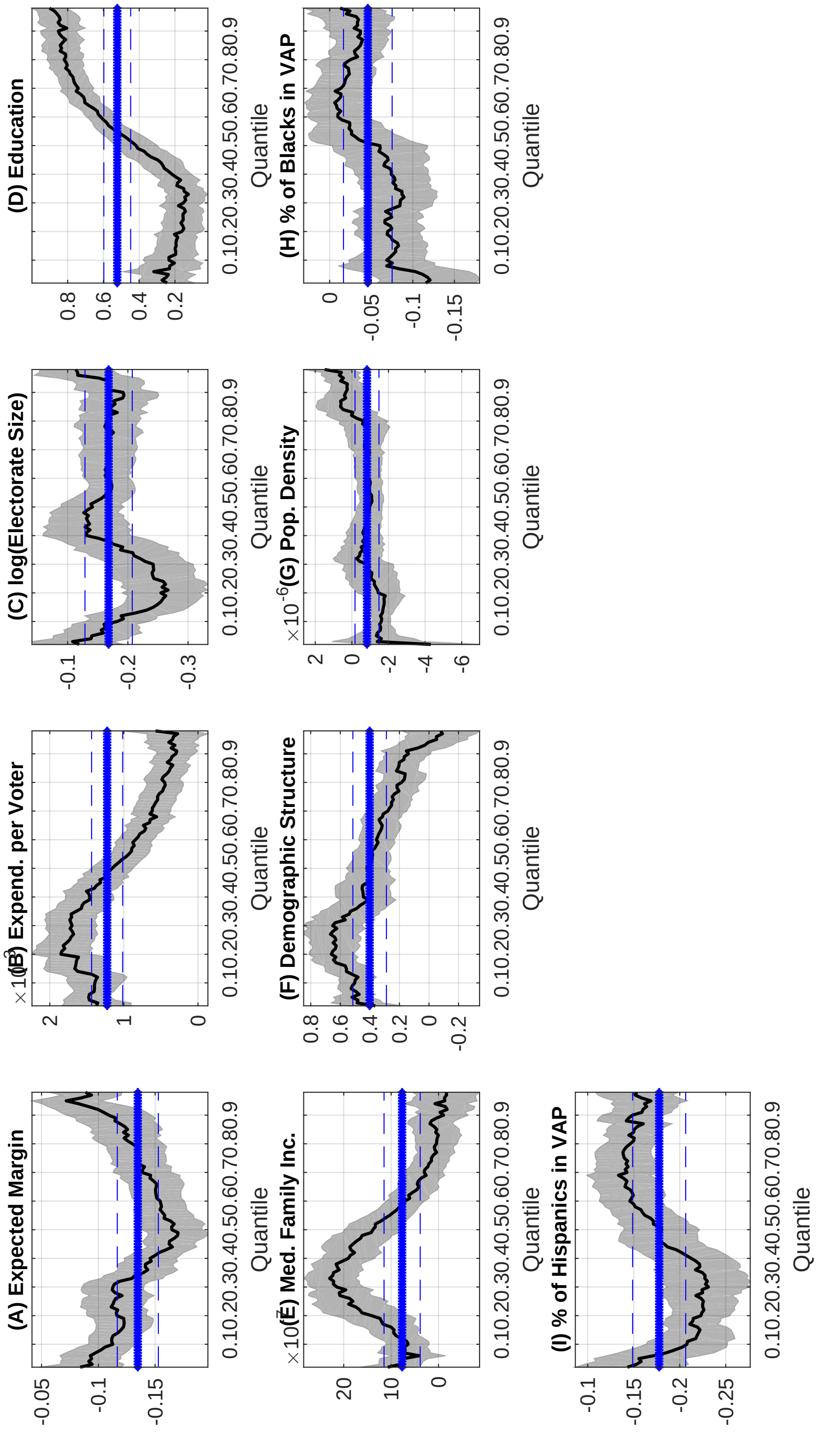

Figure A.5: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling 'minority' groups in the electorate.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, as well as the % of Blacks and Hispanics in voting age population at different quantiles of the turnout rate process. See also notes for Figure A.3.

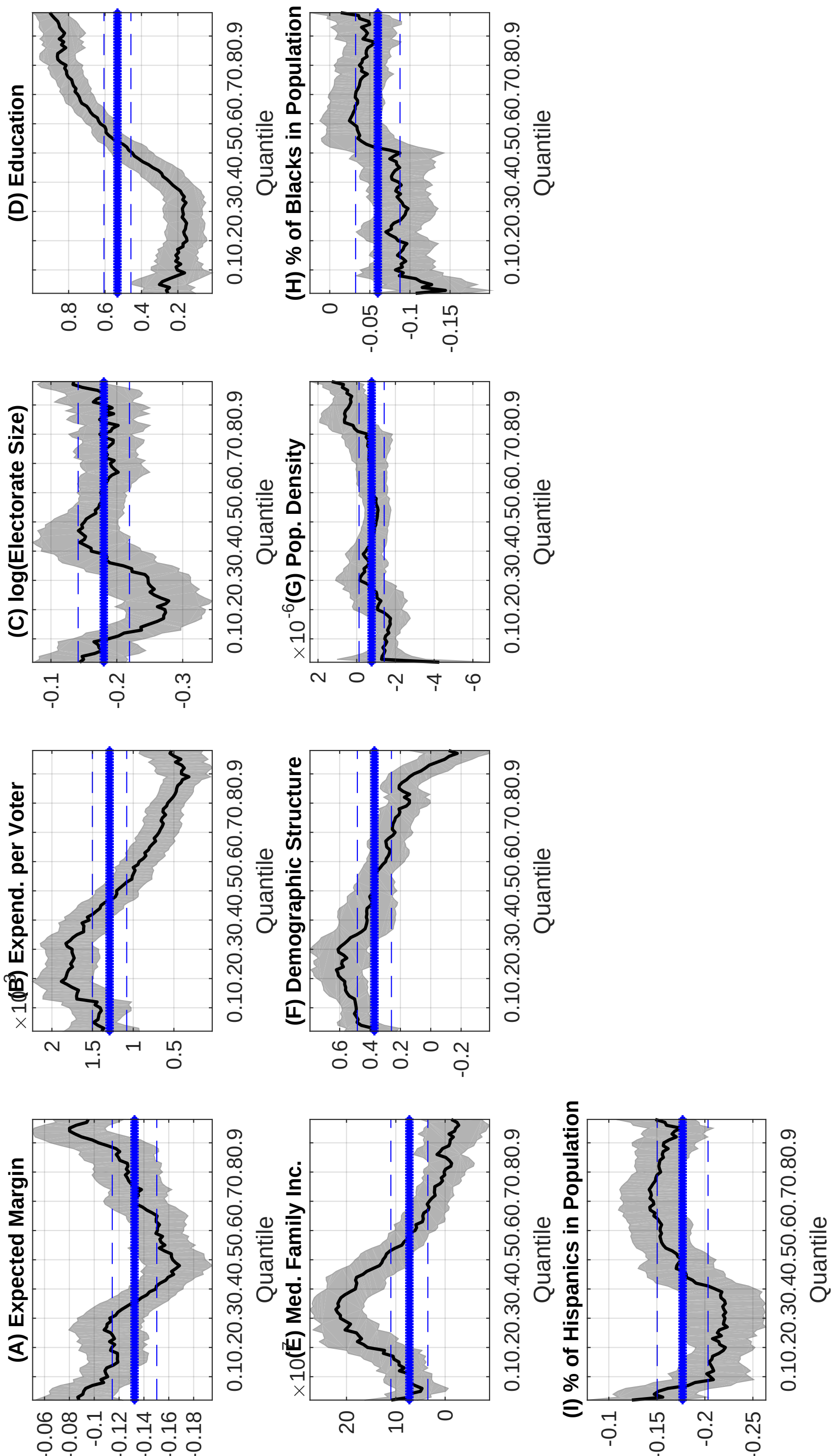

Figure A.6: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling 'minority' groups in overall population.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, as well as the % of Blacks and Hispanics in total population at different quantiles of the turnout rate process. See also notes for Figure A.3.

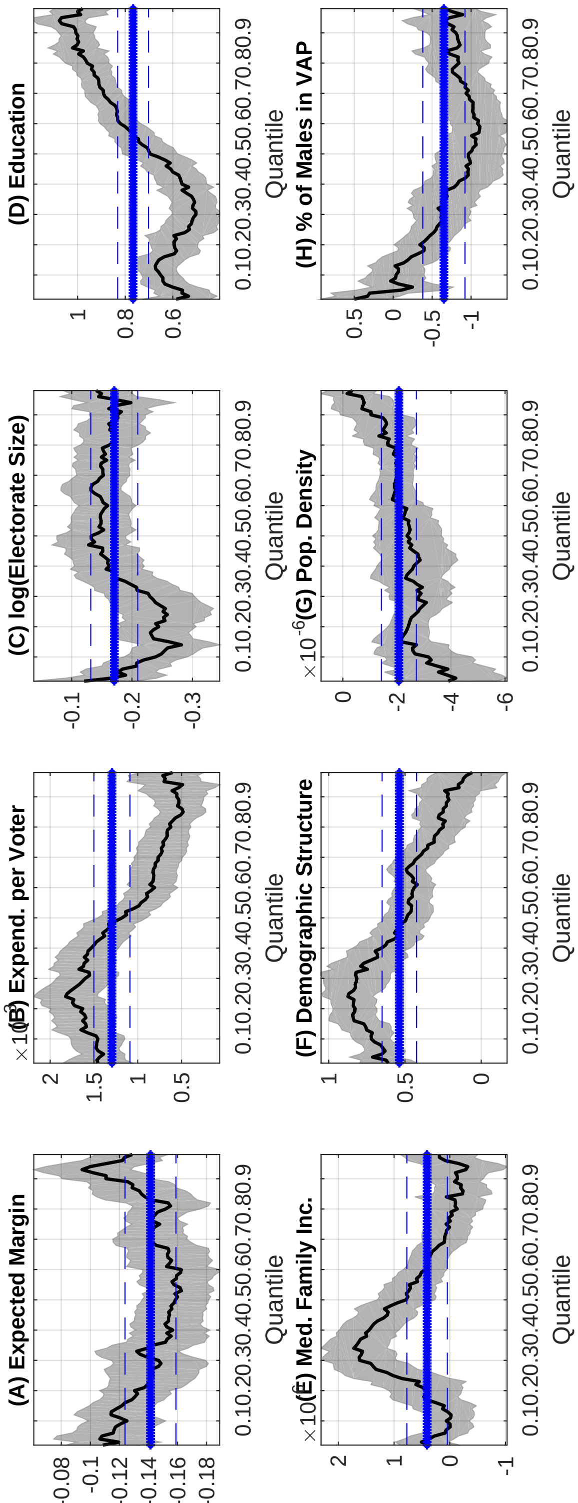

Figure A.7: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling for gender in the electorate.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, as well as the % of males in voting age population, at different quantiles of the turnout rate process. See also notes for Figure A.3.

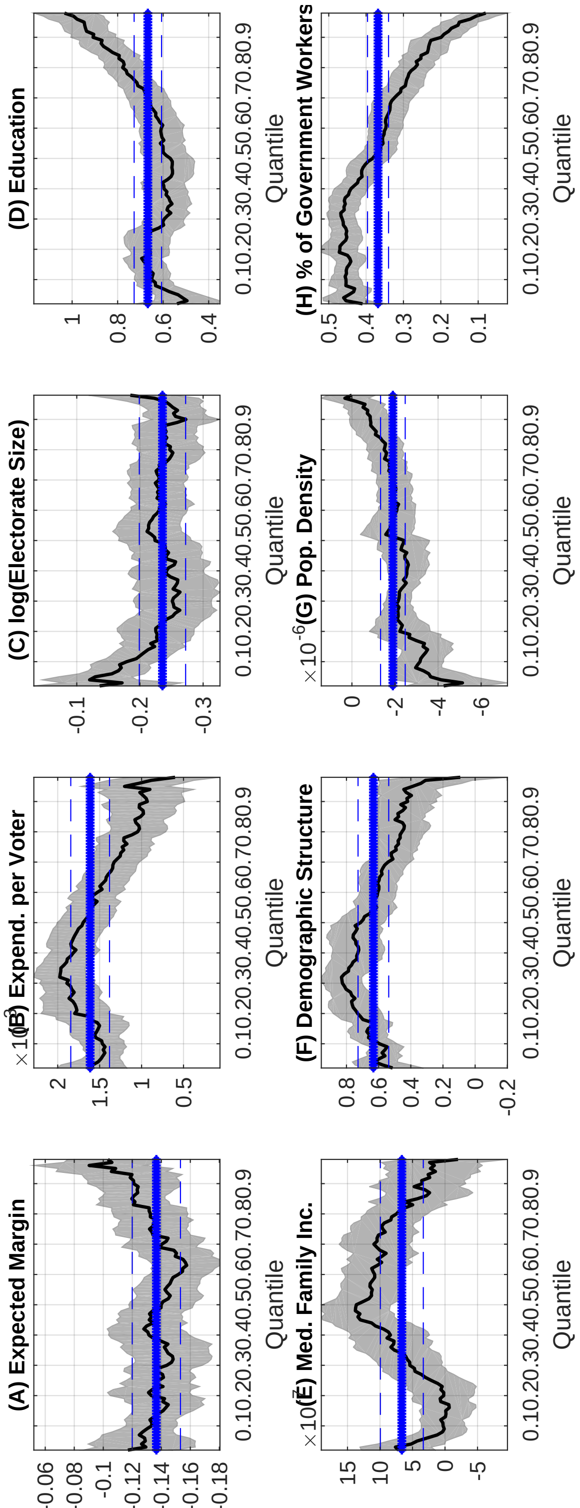

Figure A.8: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling for government employment.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, as well as the % of government workers in civilian labor force, at different quantiles of the turnout rate process. See also notes for Figure A.3.

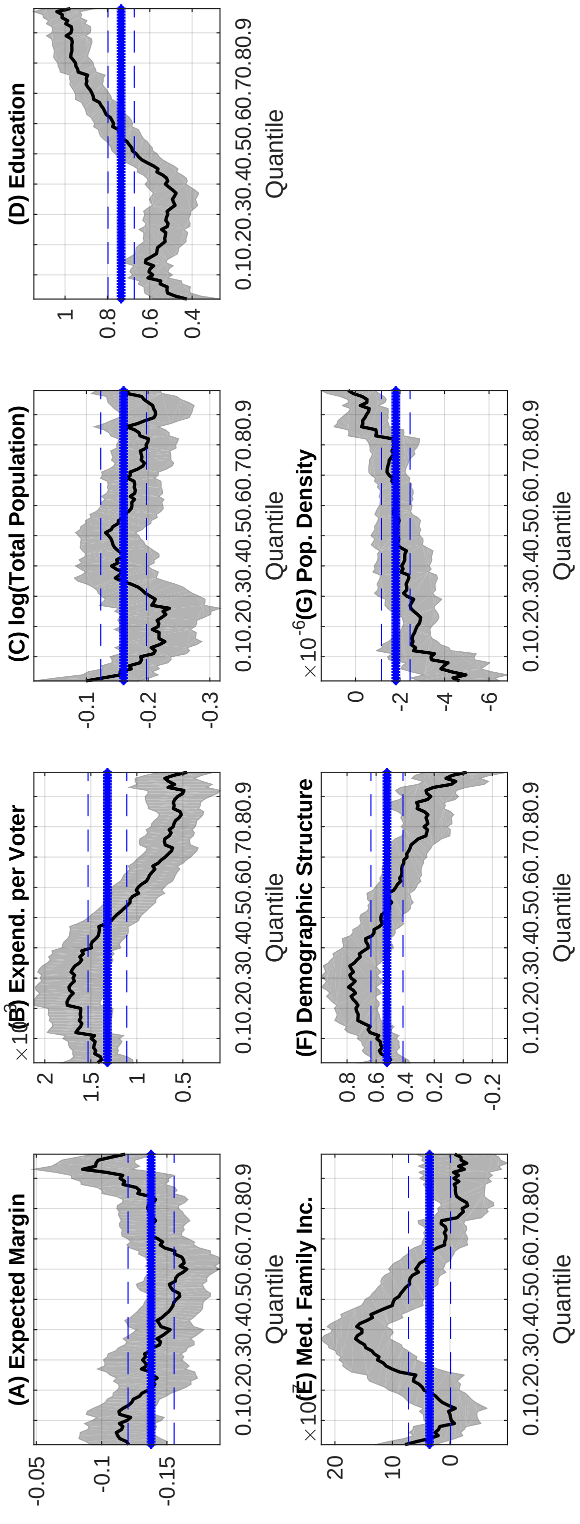

Figure A.9: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, replacing the size of the electorate with the size of total population.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, apart from the size of the electorate which is replaced by the size of total population in each congressional district. See also notes for Figure A.3.

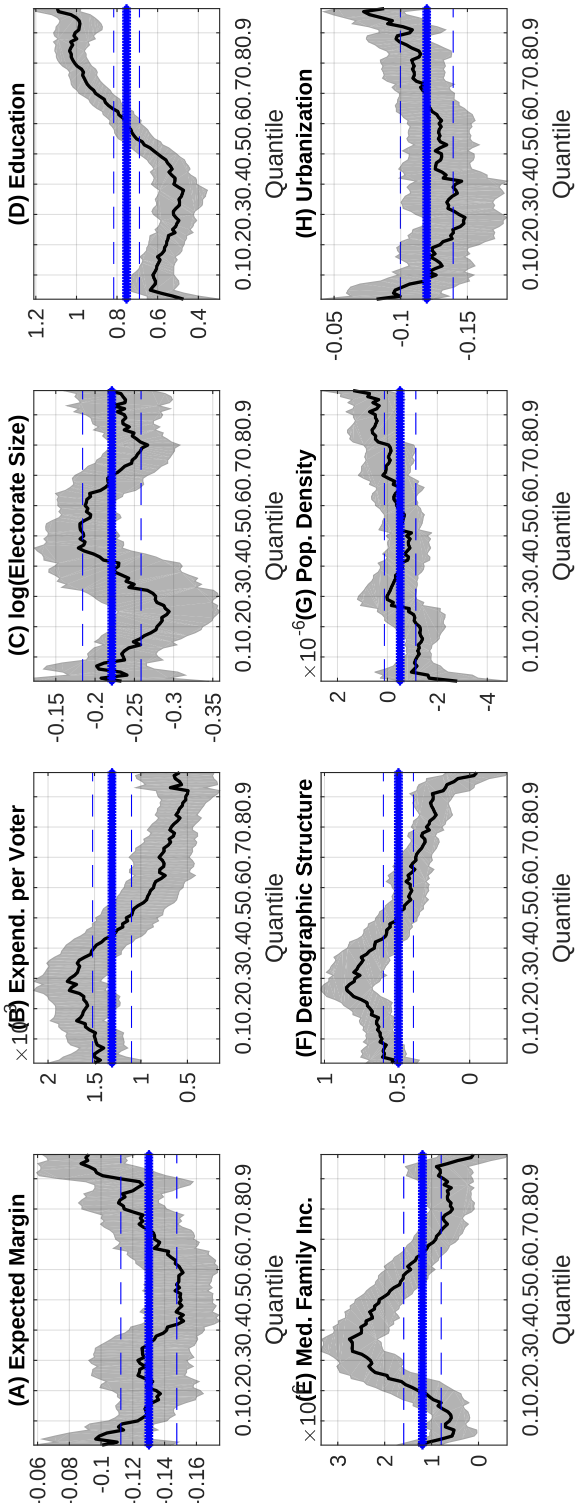

Figure A.10: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling for urbanization.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, as well as the % of urban population, at different quantiles of the turnout rate process. See also notes for Figure A.3.

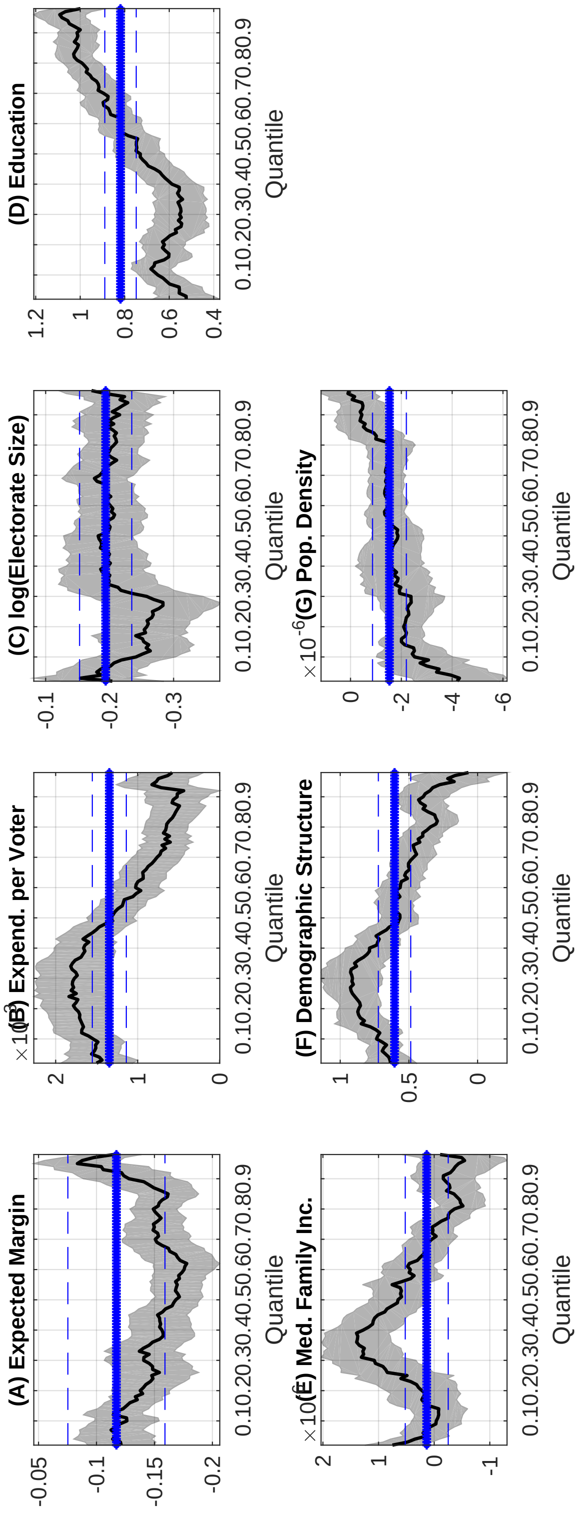

Figure A.11: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, eliminating open seat elections.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, when open seat elections have been eliminated from the sample. The sample is reduced from 2142 to 1960 observations (about 8.5% of observations have been dropped from the sample). See also notes for Figure A.3.

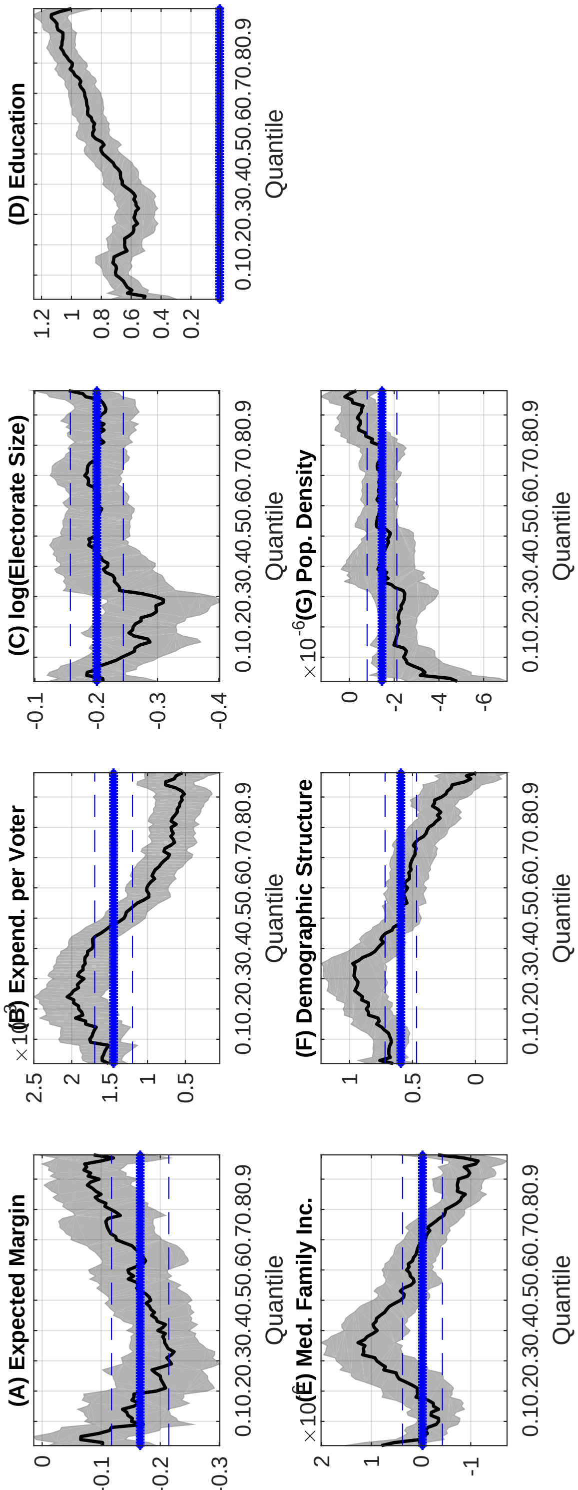

Figure A.12: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, eliminating open seat elections and elections with low competition.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, when open seat elections and elections in which the challenger had zero spending have been eliminated from the sample. The sample is reduced from 2142 to 1677 observations (21.71% of observations have been dropped from the sample). See also notes for Figure A.3.

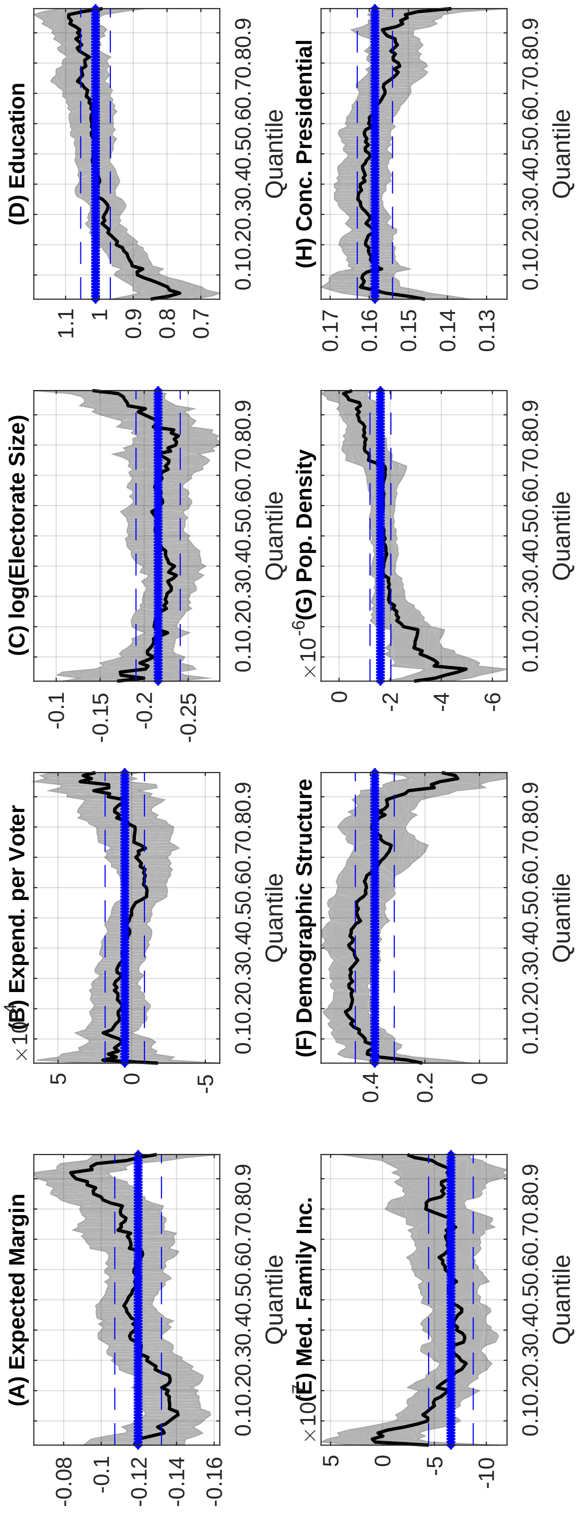

Figure A.13: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling for on-yeal elections (concurrent presidential elections).

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, controlling for on-year elections (concurrent presidential elections). See also notes for Figure A.3.

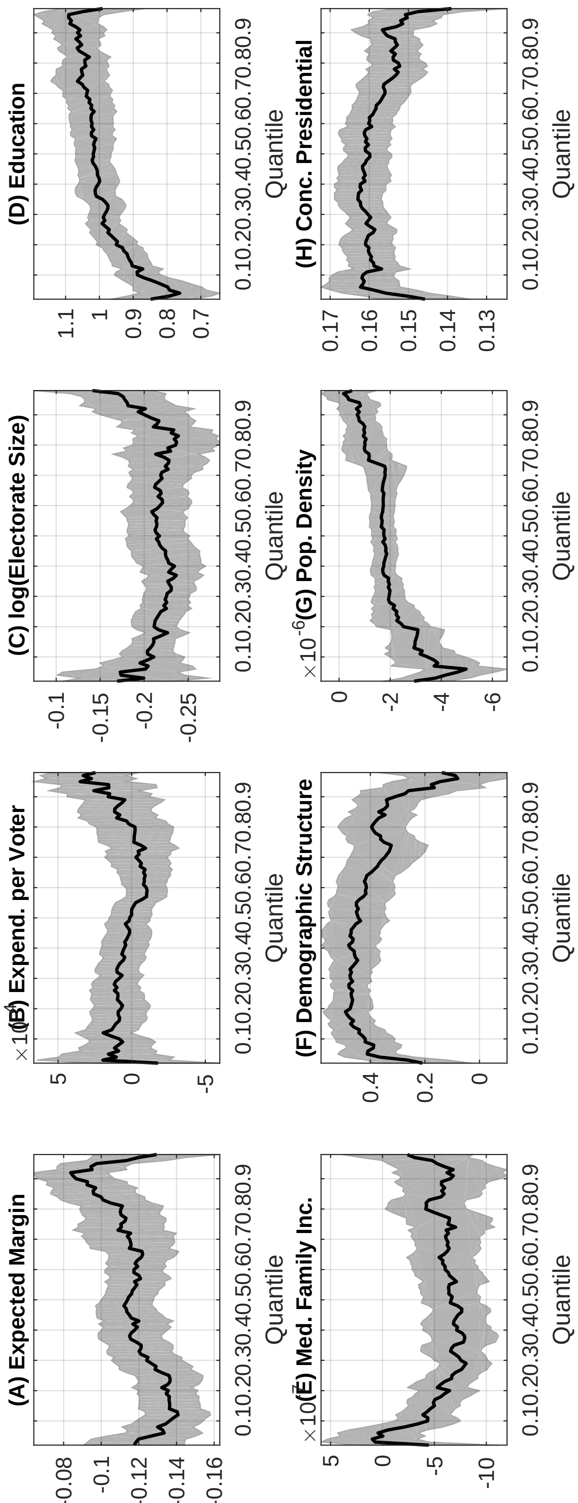

Figure A.14: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling for on-yeal elections (concurrent presidential elections).

Notes: Notes: The figure displays the same estimates with those reported in Figure A.13, excluding the estimates of the conditional mean model (OLS). See also notes for Figure A.3.

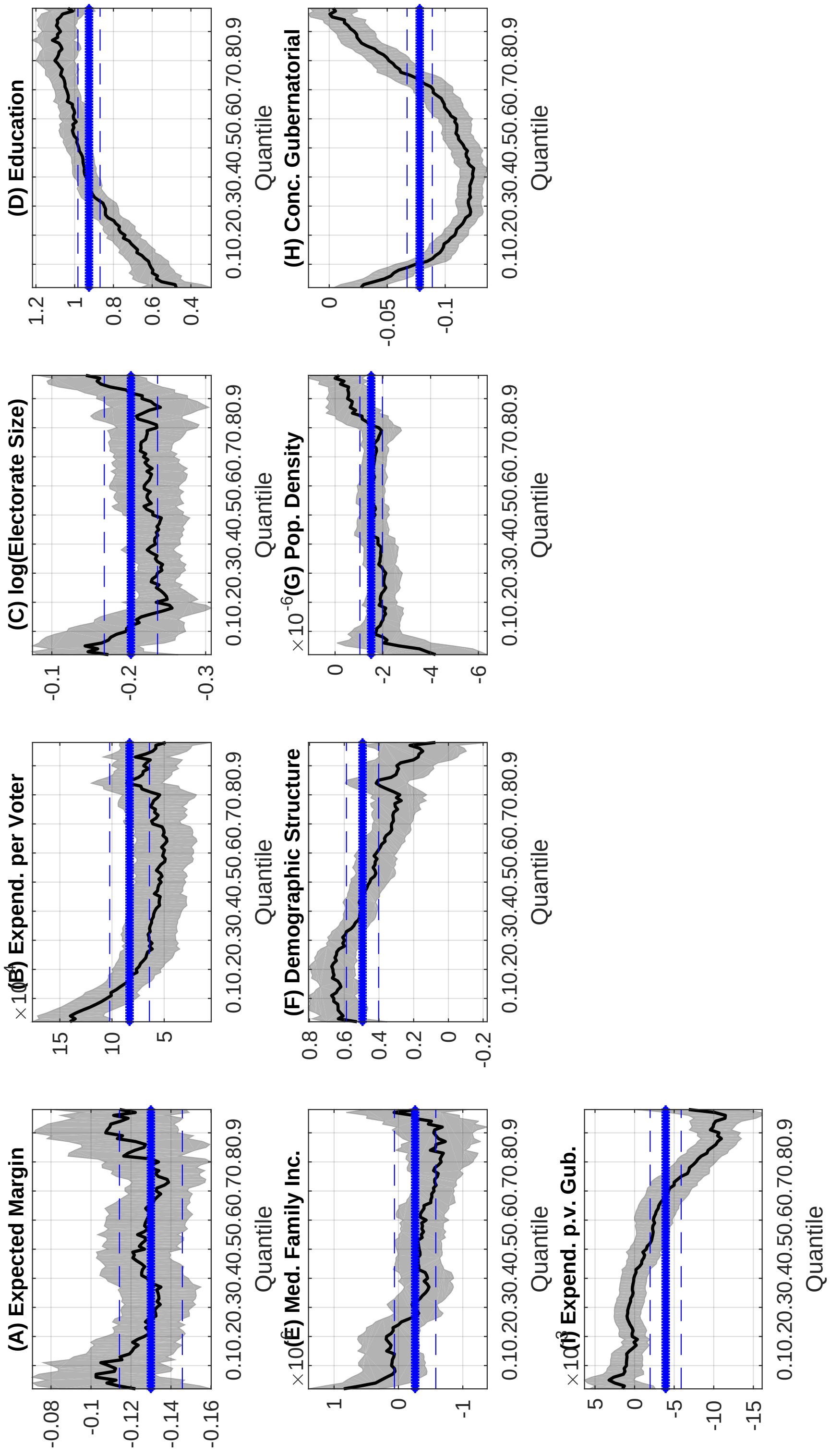

Figure A.15: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling for concurrent gubernatorial elections.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, controlling for concurrent gubernatorial elections, as well as for campaign expenditures in gubernatorial elections. See also notes for Figure A.3.

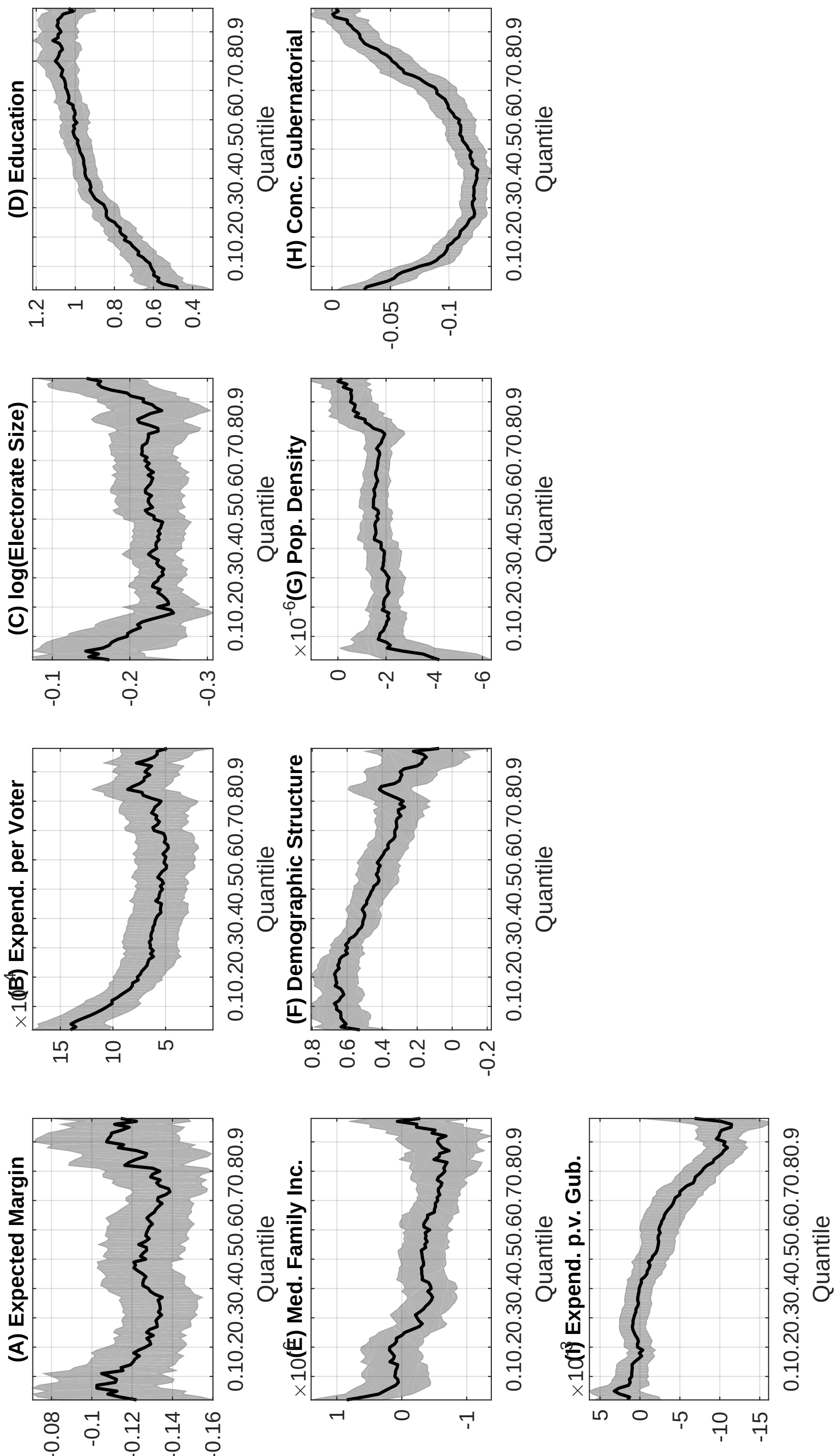

Figure A.16: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling for concurrent gubernatorial elections.

Notes: The figure displays the same estimates with those reported in Figure A.15, excluding the estimates of the conditional mean model (OLS). See also notes for Figure A.3.

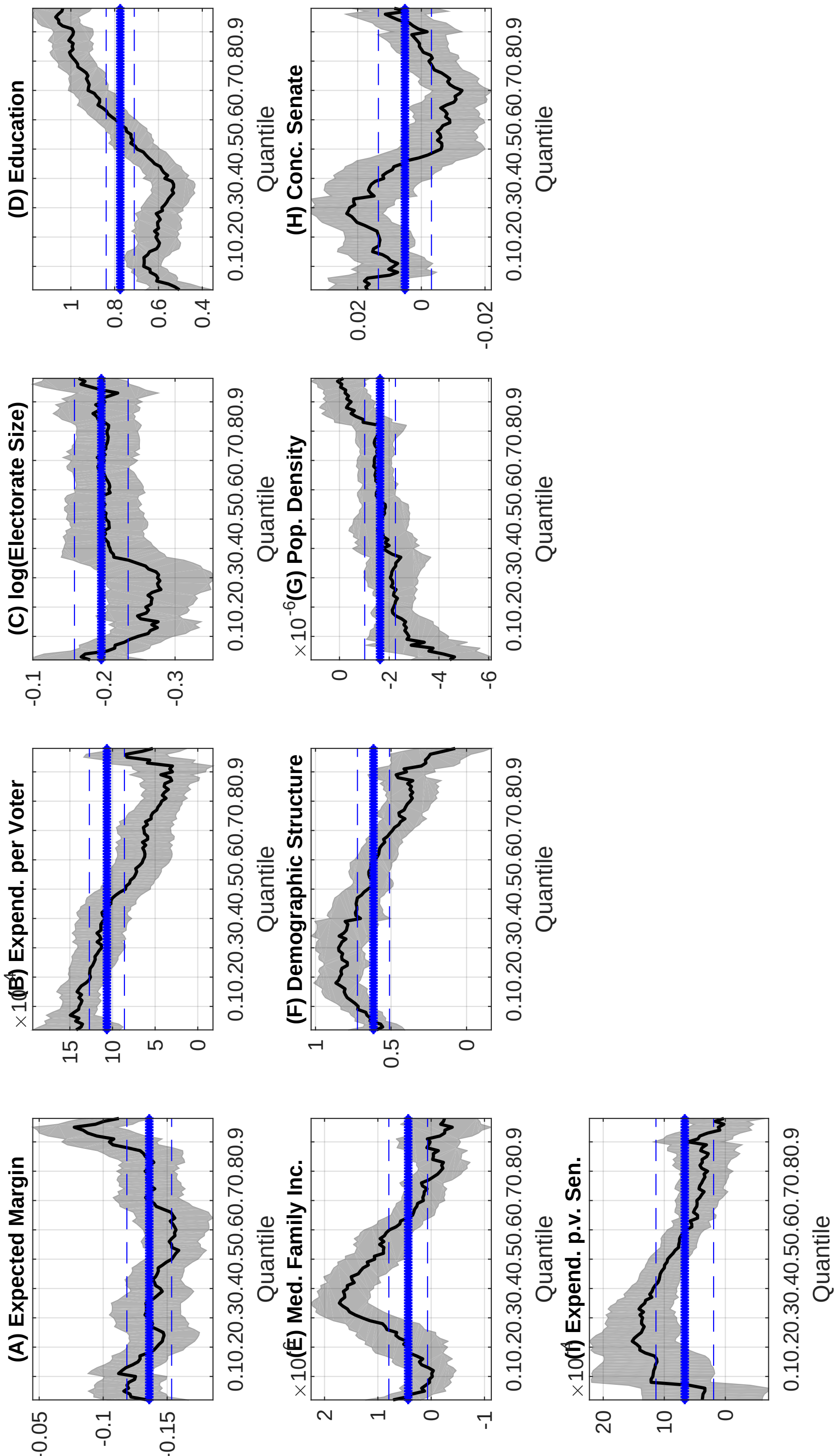

Figure A.17: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling for concurrent Senate elections.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, controlling for concurrent Senate elections, as well as for campaign expenditures in Senate elections. See also notes for Figure A.3.

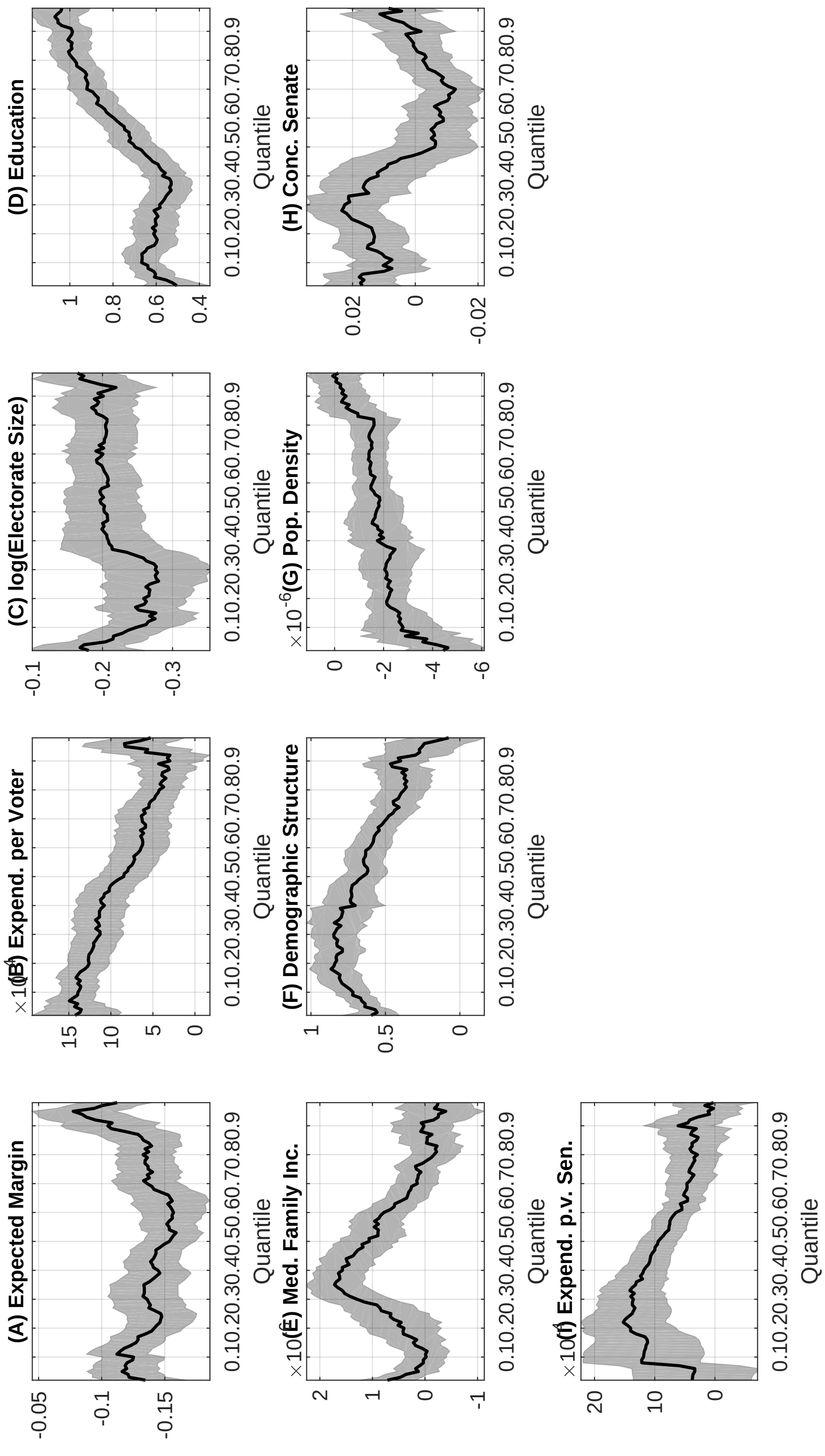

Figure A.18: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling for concurrent Senate elections.

Notes: The figure displays the same estimates with those reported in Figure A.17, excluding the estimates of the conditional mean model (OLS). See also notes for Figure A.3.

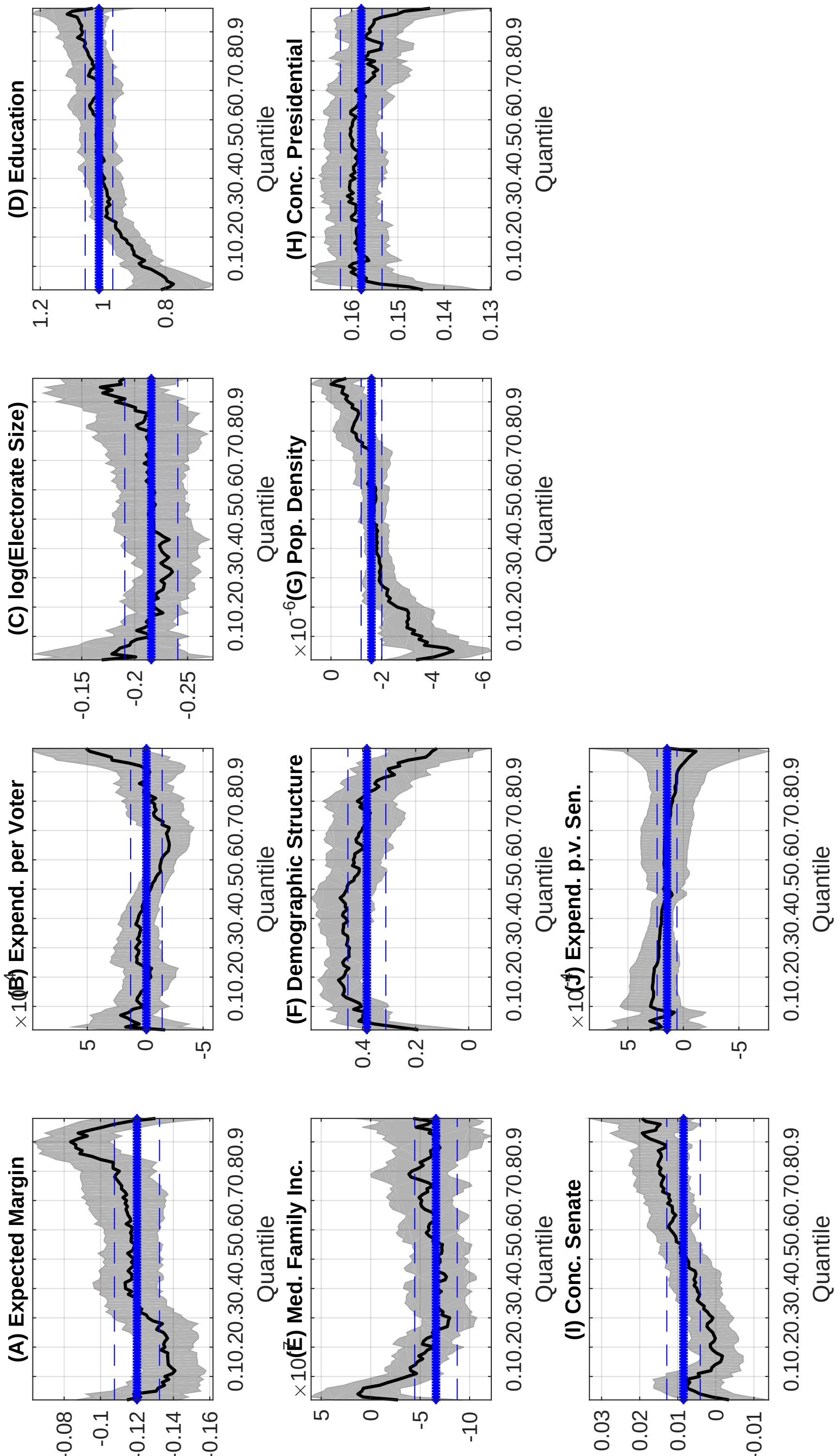

Figure A.19: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling for concurrent Presidential (on-year) and Senate elections.

Notes: figure displays the estimated partial effects of the baseline set of covariates in the model, controlling for concurrent presidential (on-year) and Senate elections, as well as for campaign expenditures in Senate elections. See also notes for Figure A.3.

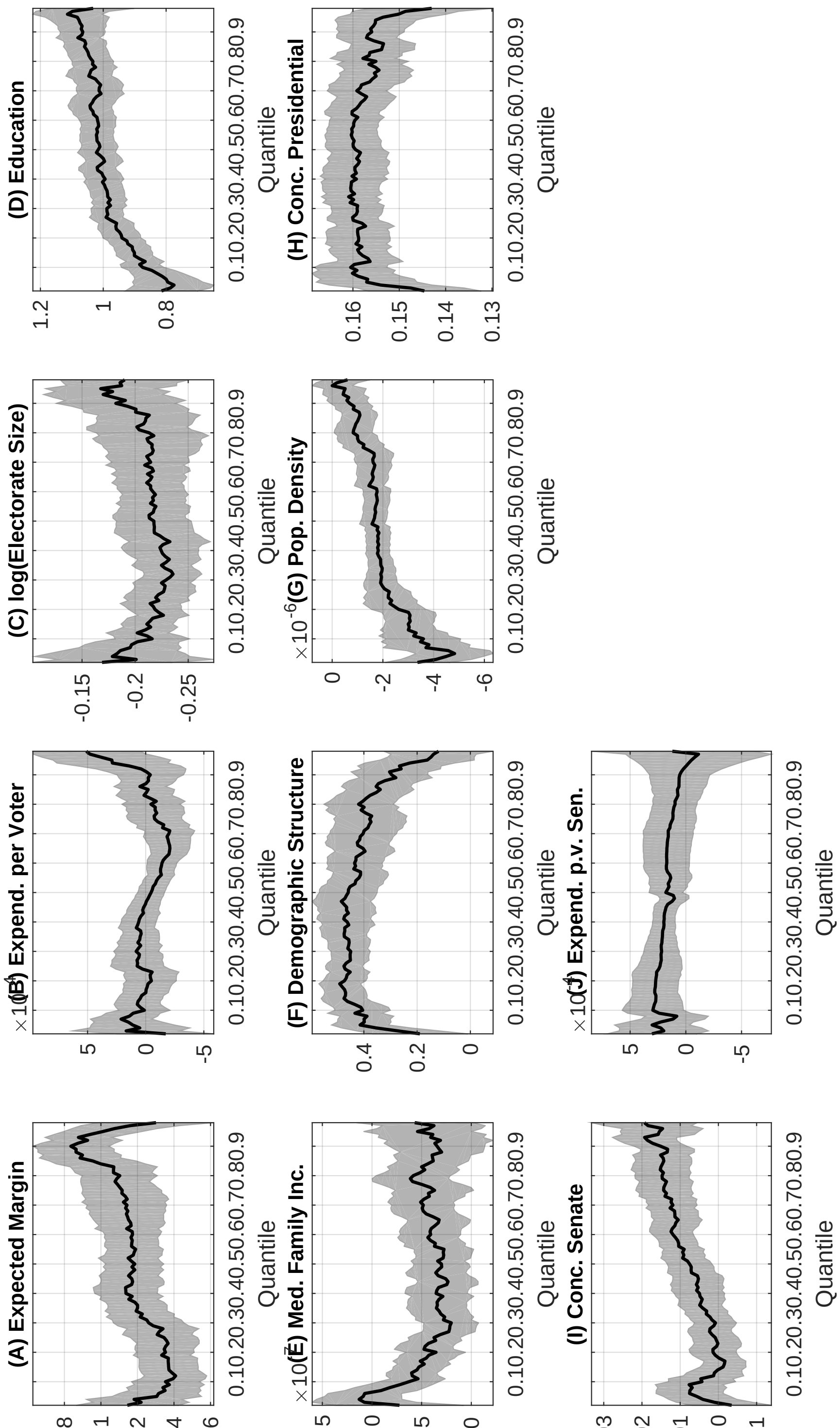

Figure A.20: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling for concurrent Presidential (on-year) and Senate elections.

Notes: The figure displays the same estimates with those reported in Figure A.19, excluding the estimates of the conditional mean model (OLS). See also notes for Figure A.3.

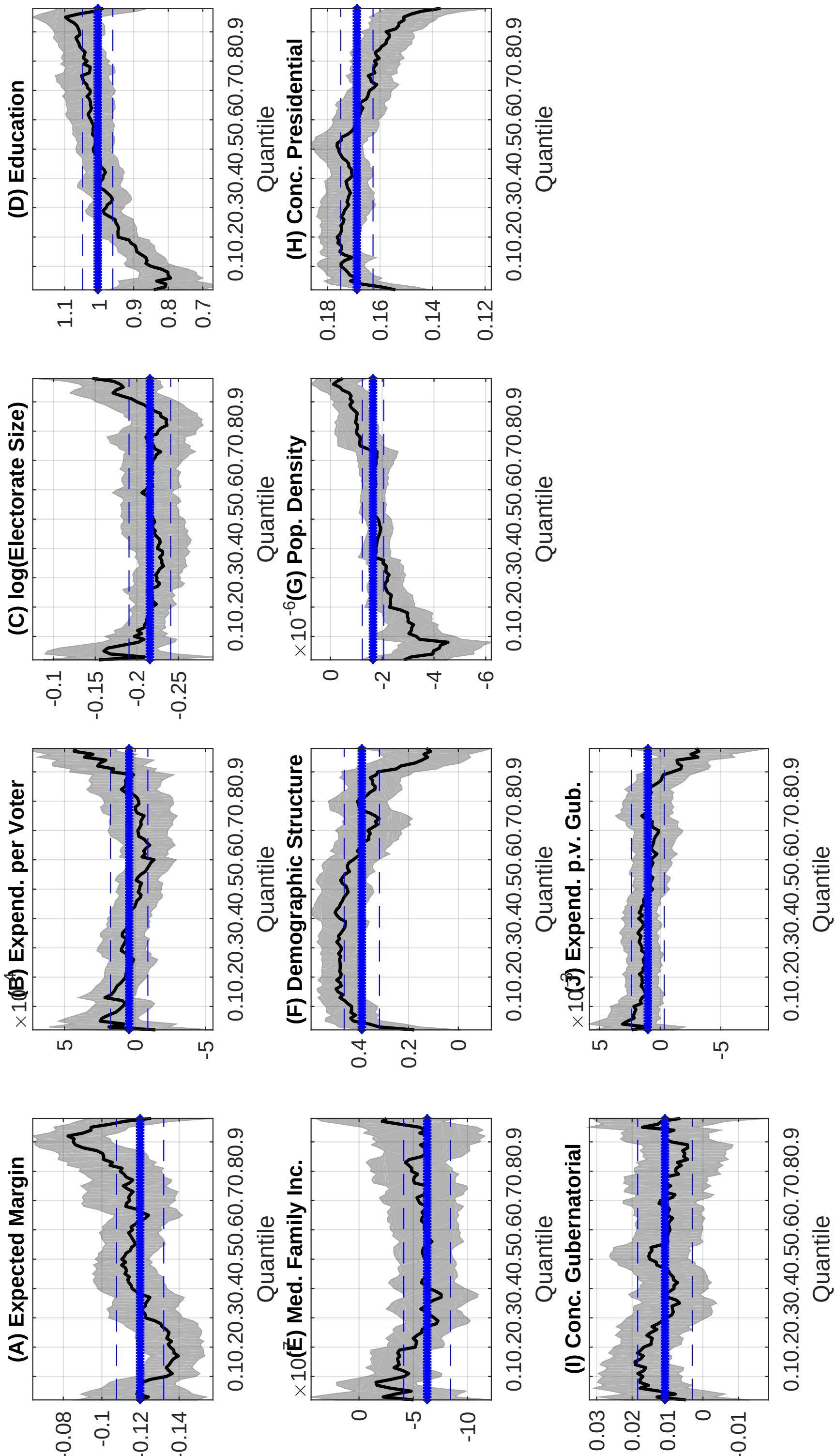

Figure A.21: Quantile regression estimates controlling for concurrent Presidential (on-year) and Gubernatorial elections.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, controlling for concurrent presidential (on-year) and gubernatorial elections, as well as for campaign expenditures in gubernatorial elections. See also notes for Figure A.3.

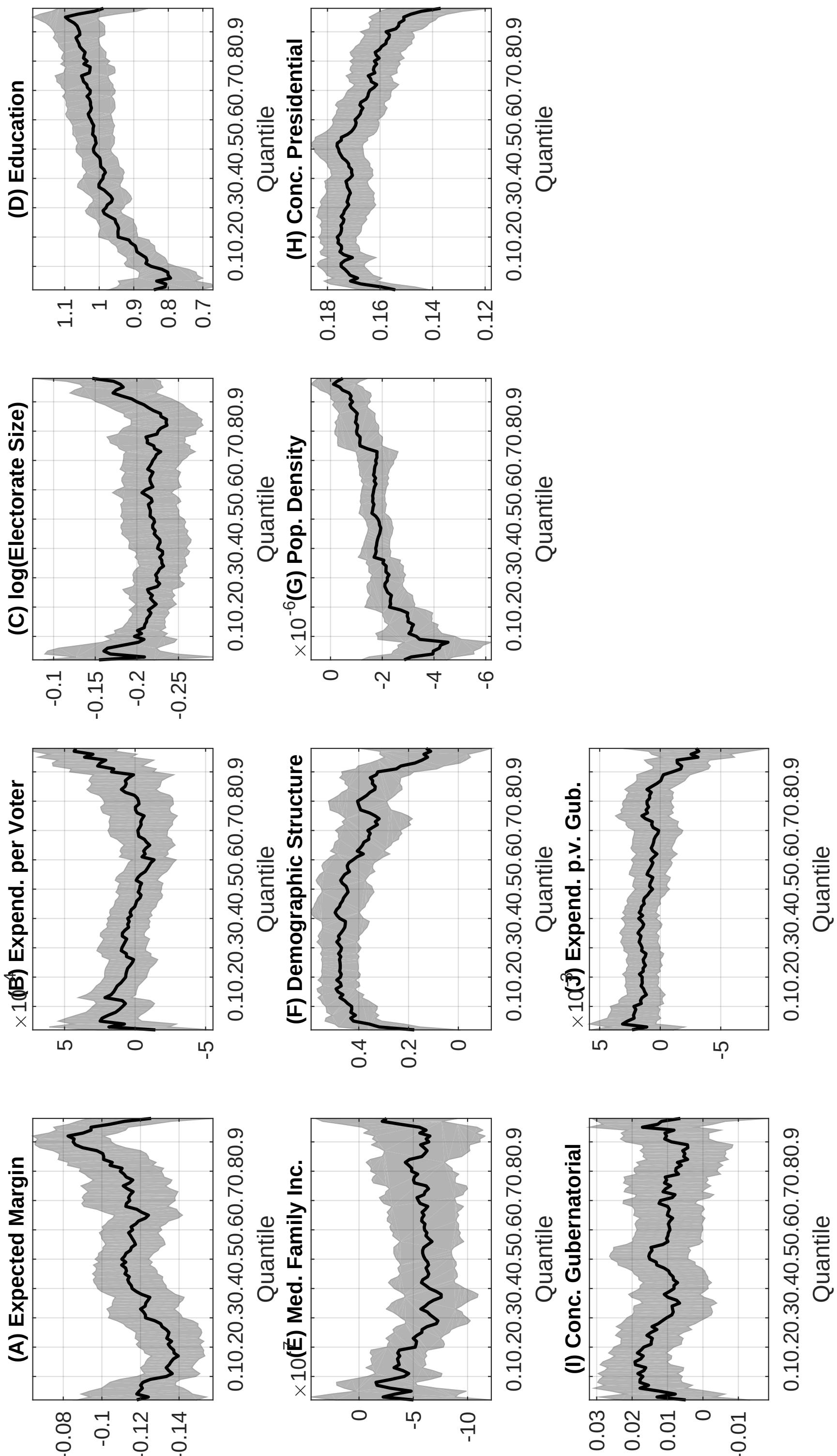

Figure A.22: Quantile regression estimates controlling for concurrent Presidential (on-year) and Gubernatorial elections.

Notes: The figure displays the same estimates with those reported in Figure A.21, excluding the estimates of the conditional mean model (OLS). See also notes for Figure A.3.

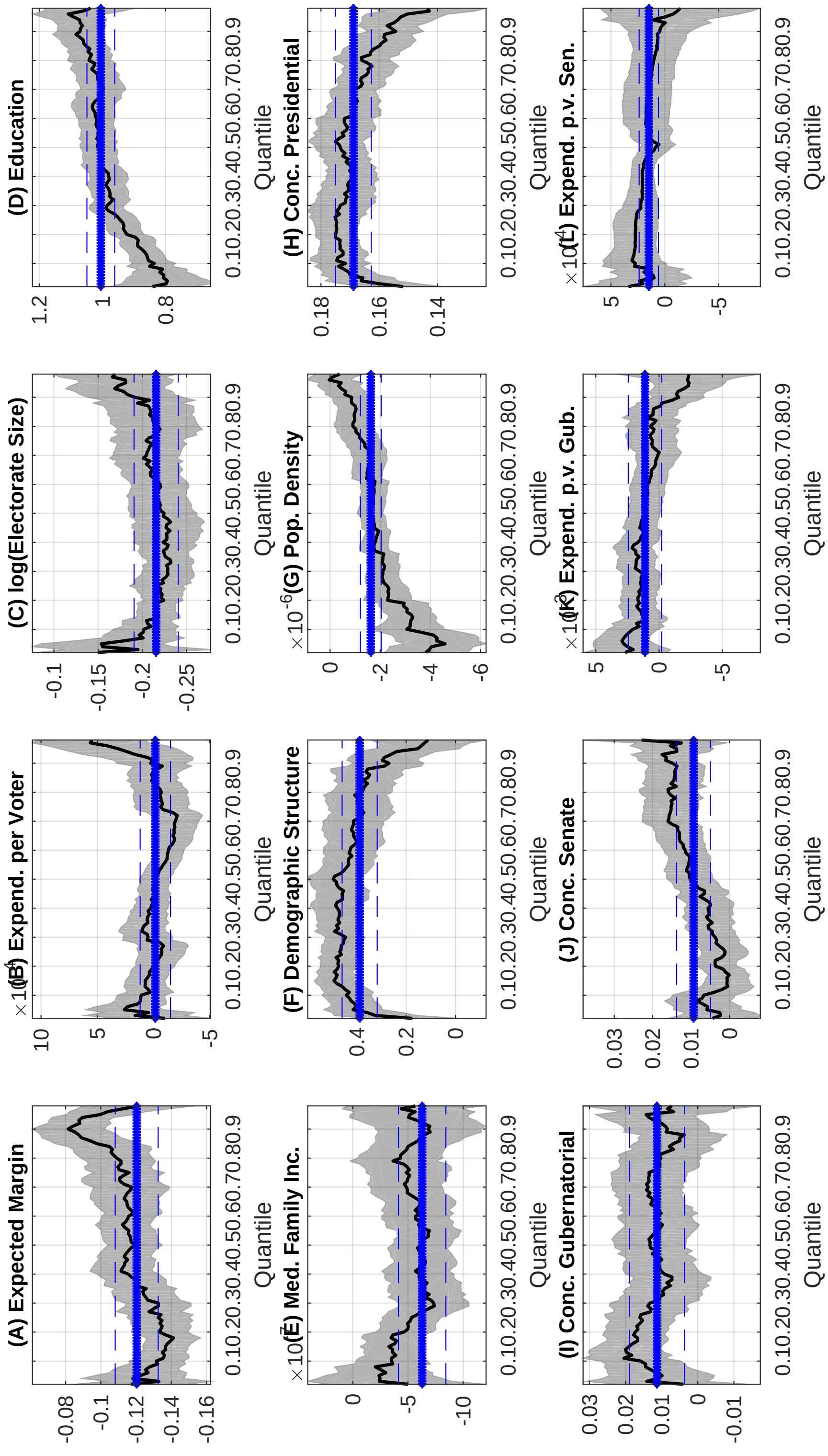

Figure A.23: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling for all types of concurrent elections.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, controlling for concurrent presidential, gubernatorial and senate elections, as well as for campaign expenditures in gubernatorial and senate elections. See also notes for Figure A.3.

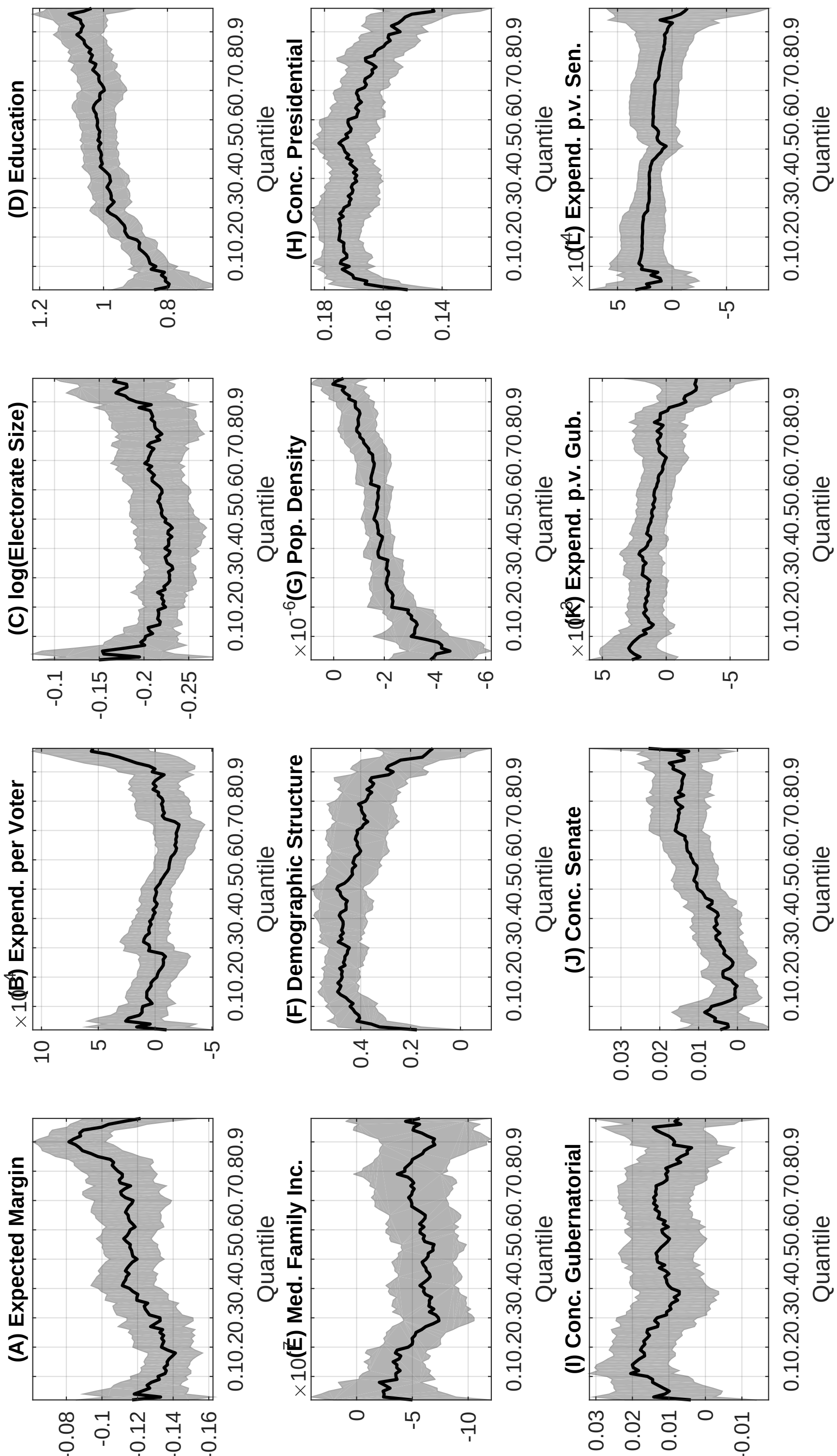

Figure A.24: Quantile regression estimates of the effects of the baseline set of covariates on electoral participation, controlling for all types of concurrent elections.

Notes: The figure displays the same estimates with those reported in Figure A.23, excluding the estimates of the conditional mean model (OLS). See also notes for Figure A.3.

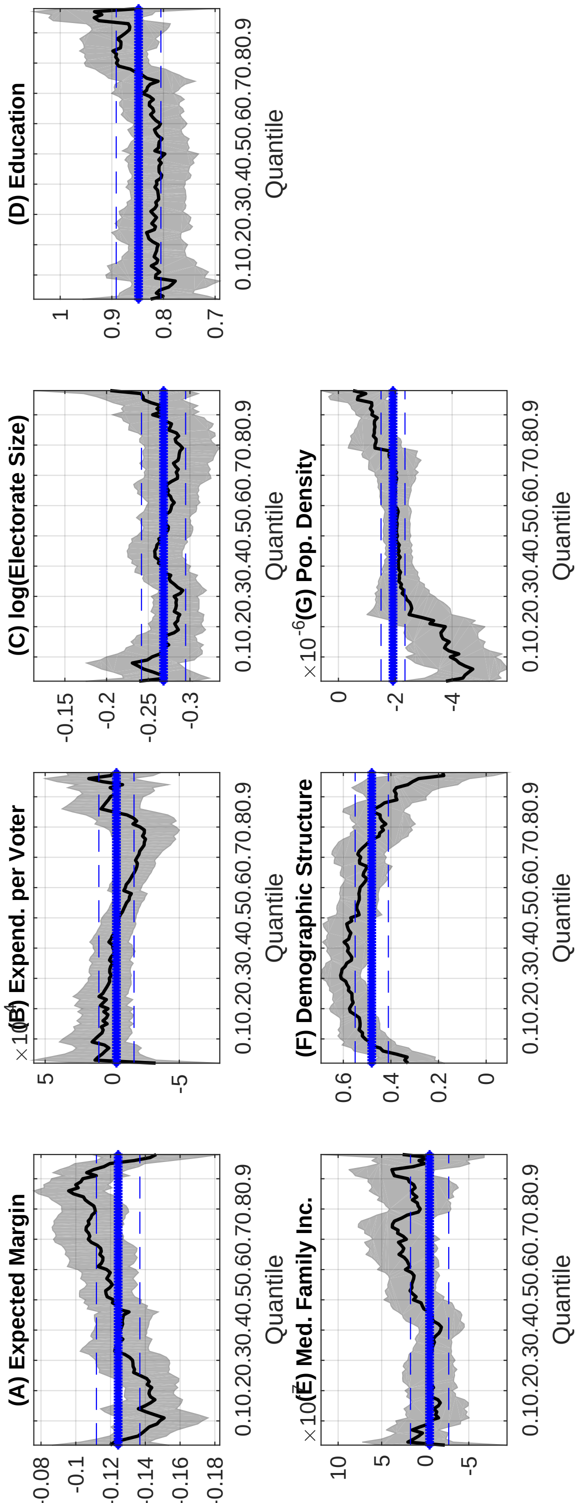

Figure A.25: Quantile regression estimates of the effects of baseline set of covariates on electoral participation, controlling for time effects.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, controlling for time effects by including a set of time dummies in the QR estimation (for 2002, 2004, 2006, and 2008). See also notes for Figure A.3.

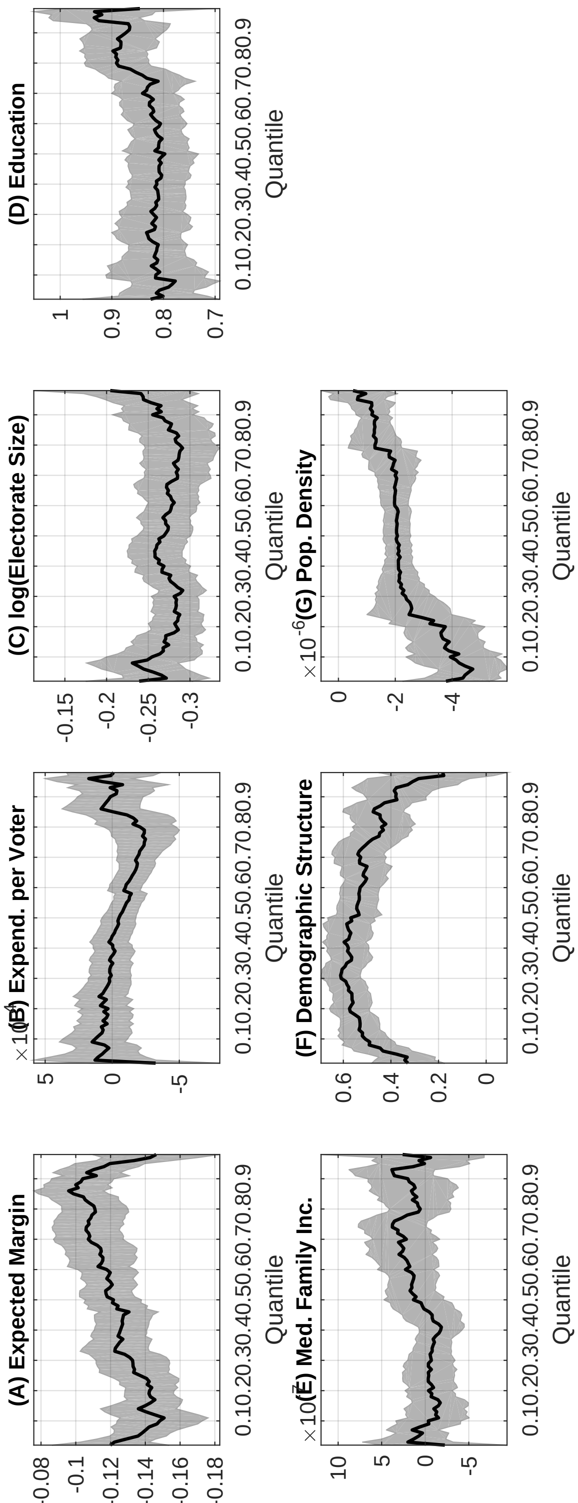

Figure A.26: Quantile regression estimates of the effects of baseline set of covariates on electoral participation, controlling for time effects.

Notes: The figure displays the same estimates with those reported in Figure A.25, excluding the estimates of the conditional mean model (OLS). See also notes for Figure A.3.

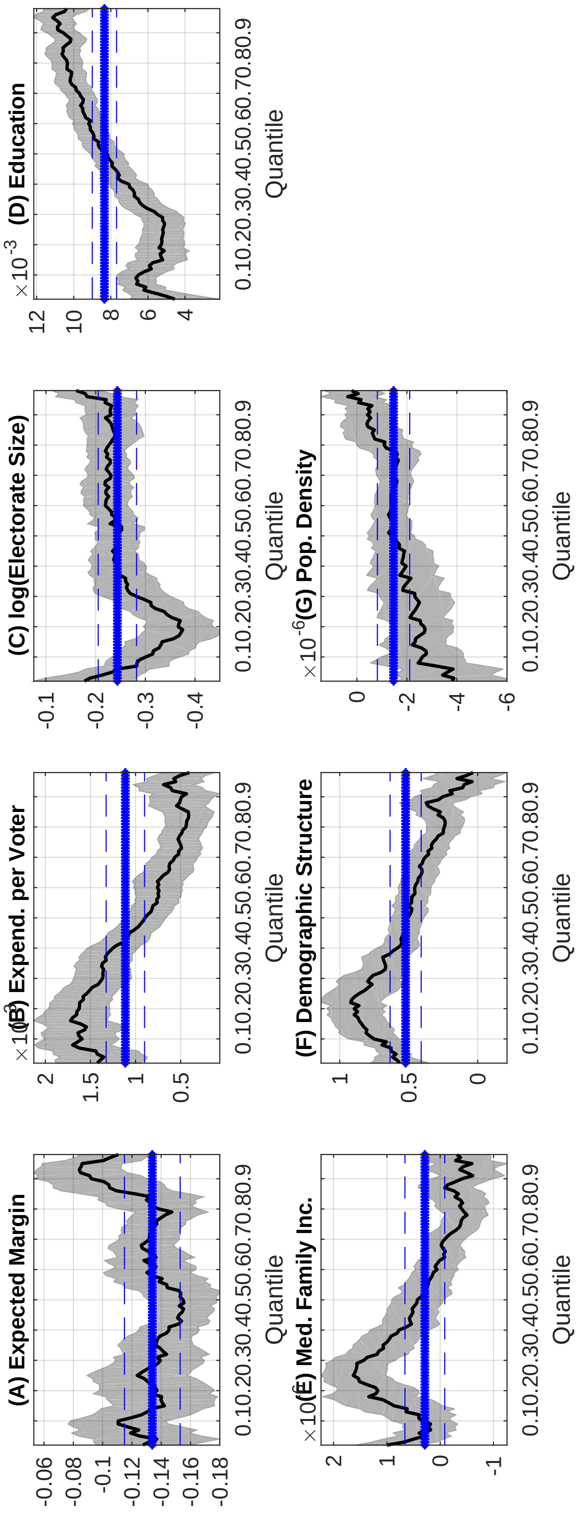

Figure A.27: Quantile regression estimates of the effects of baseline set of covariates on electoral participation, eliminating the observations that are affected by redistricting.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, when we eliminate from our sample all observations that are affected by redistricting. The sample is reduced from 2142 to 1817 observations (about 15.2% of observations have been dropped from the sample). See also notes for Figure A.3.

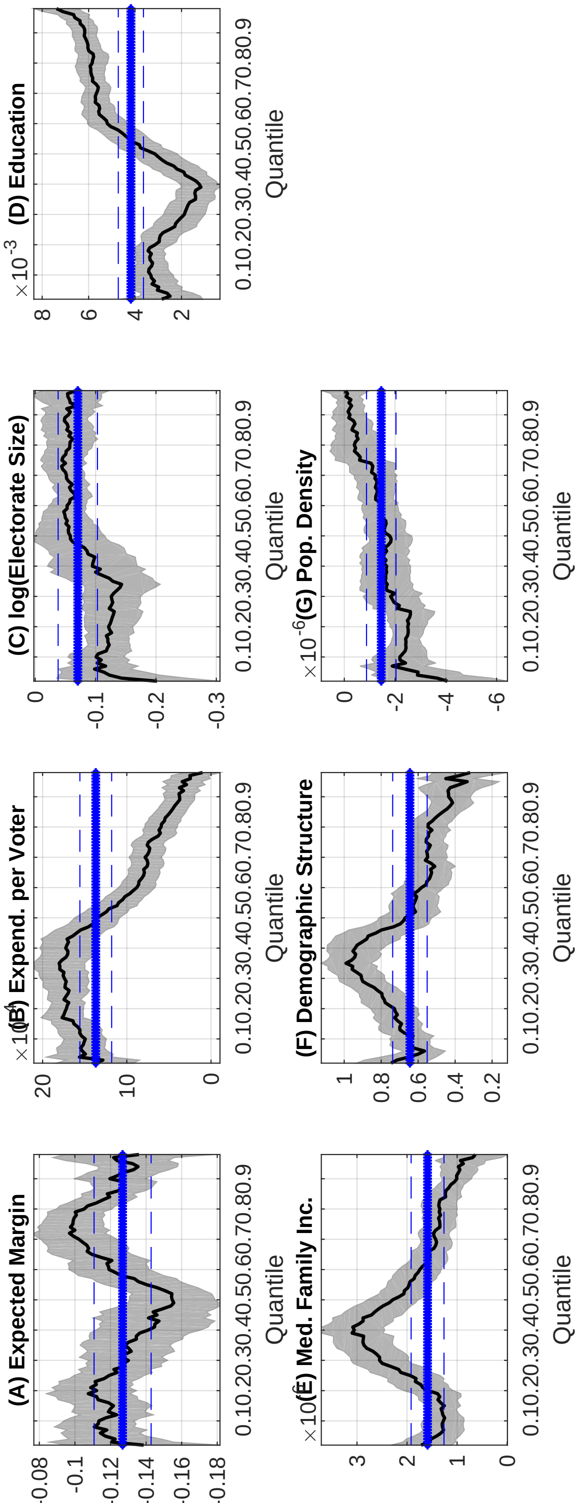

Figure A.28: Quantile regression estimates of the effects of baseline set of covariates on electoral participation, controlling for state fixed effects.

Notes: The figure displays the estimated partial effects of the baseline set of covariates in the model, controlling for state fixed effects following the method suggested by Canay (2011). See also notes for Figure A.3.

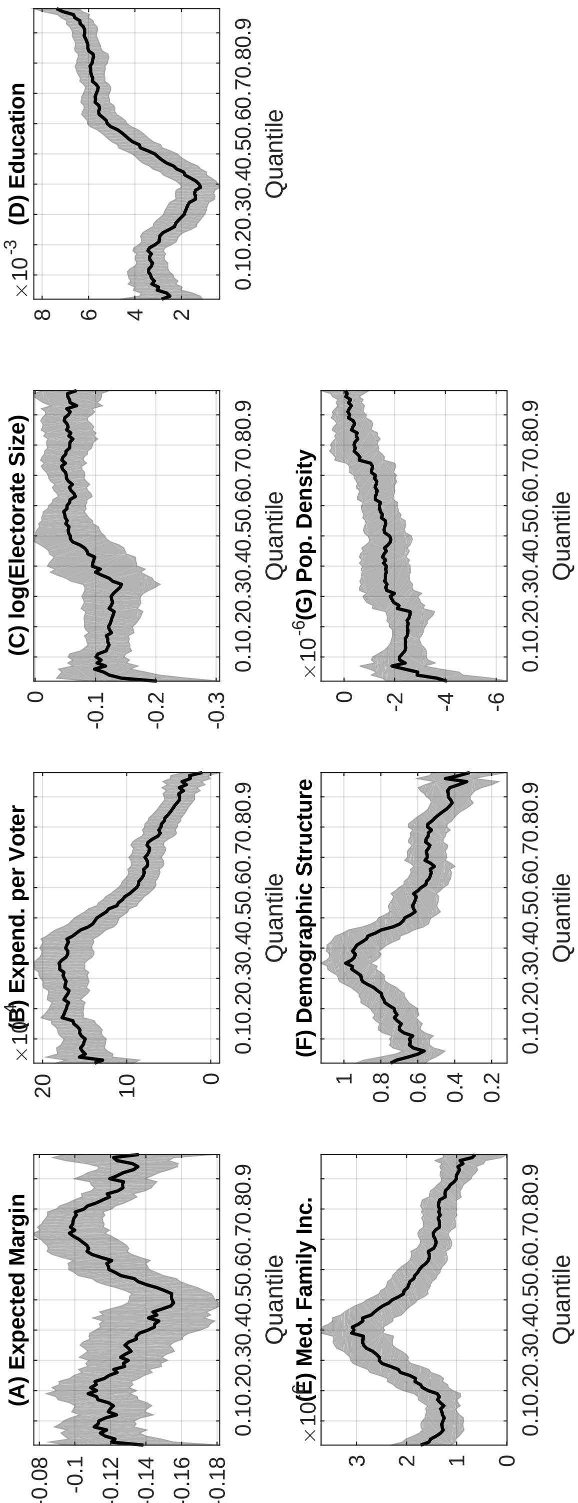

Figure A.29: Quantile regression estimates of the effects of baseline set of covariates on electoral participation, controlling for state fixed effects.

Notes: The figure displays the same estimates with those reported in Figure A.28, excluding the estimates of the conditional mean model (OLS). See also notes for Figure A.3.

Table A.1: **Summary Statistics**

	N	Mean	St. Deviation	Min	Max	$Q_{0.05}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.95}$
Turnout$_{it}$	2,150	0.453	0.123	0.110	0.761	0.245	0.364	0.455	0.550	0.646
Margin$_{it}$	2,175	0.391	0.255	0.342×10^{-3}	1	0.0499	0.212	0.342	0.497	0.980
Margin$^*_{it}$	2,167	0.410	0.190	0.109	1	0.132	0.333	0.341	0.441	0.794
Spending per Voter$_{it}$	2,175	8.426	16.57	0.126	172.2	0.692	1.370	2.586	8.299	32.66
Education$_{it}$	2,175	0.848	0.074	0.361	0.974	0.702	0.816	0.865	0.901	0.933
(Real) Median Family Income$_{it}$	2,175	52,233.190	13,569.470	19,390.39	98,788.29	33,826.700	42,651.770	49,657.320	59,457.040	79,593.230
ln(Population$_{it}$)	2,175	13.409	0.093	12.881	13.876	13.276	13.357	13.399	13.454	13.573
ln(Voting Age Population$_{it}$)	2,175	13.122	0.094	12.586	13.582	12.975	13.066	13.121	13.176	13.278
Demographic Structure$_{it}$	2,175	0.165	0.035	0.0678	0.359	0.106	0.143	0.165	0.184	0.218
Population Density$_{it}$	2,175	2,423	6,799	1.091	65,385	22.284	95.018	305.033	1,938.986	10,860.240
Urbanization$_{it}$	2,175	0.789	0.198	0.212	1	0.429	0.642	0.845	0.979	1
Quality$_{it}$	2,174	1.098	1.831	0	7	0	0	0	1	5
Election$_{it}$	2,168	0.501	0.982	0	3	0	0	0	1	3
Blacks Percentage$_{it}$ (VAP)	2,034	0.124	0.147	0.002	0.765	0.012	0.032	0.064	0.153	0.515
Hispanics Percentage$_{it}$ (VAP)	2,123	0.126	0.159	0.005	0.819	0.011	0.026	0.063	0.157	0.513
Blacks Percentage$_{it}$ (Total Pop.)	2,073	0.130	0.153	0.002	0.789	0.011	0.033	0.067	0.162	0.548
Hispanics Percentage$_{it}$ (Total Pop.)	2,140	0.141	0.171	0.005	0.860	0.012	0.029	0.072	0.179	0.573
Percent of Government Workers$_{it}$	2,154	0.185	0.104	0.060	0.543	0.088	0.115	0.146	0.197	0.413
Percent of Male Voters$_{it}$	2,175	0.485	0.014	0.423	0.561	0.463	0.478	0.486	0.494	0.506
Concurrent Gubernatorial Elections$_{it}$	2,175	0.389	0.488	0	1	0	0	0	1	1
Concurrent Senate Elections$_{it}$	2,175	0.670	0.470	0	1	0	0	1	1	1
Concurrent Presidential Elections$_{it}$	2,175	0.600	0.490	0	1	0	0	1	1	1
Spending per voter in Gubernatorial$_{it}$	2,175	1.549	2.574	0	19.254	0	0	0	2.781	6.584
Spending per voter in Senate$_{it}$	2,175	5.775	17.612	0	555.333	0	0	1.396	4.041	30.684

Notes for Table A.1: The table reports descriptive statistics of the variables employed in our analysis. Margin$^*_{it}$ denotes the expected/predicted (out of sample) margin obtained as outlined in text.