CONTRIBUTI DI RICERCA CRENOS

UNIVERSITÀ DI CAGLIARI UNIVERSITÀ DI SASSARI
CENTRE FOR NORTH SOUTH ECONOMIC RESEARCH
CRENoS
CENTRO RICERCHE ECONOMICHE NORD SUD

EDUCATIONAL DIVIDE BETWEEN VOTERS: A NATIONWIDE TREND?

Maria Grazia Pittau
Francesco Politano
Roberto Zelli

WORKING PAPERS

2024/08

arkadia

Centro Ricerche Economiche Nord Sud
(CRENoS)
Università di Cagliari
Università di Sassari

CRENOS was set up in 1993 with the purpose of organising the joint research effort of economists from the two Sardinian universities (Cagliari and Sassari) investigating dualism at the international and regional level. CRENoS' primary aim is to improve knowledge on the economic gap between areas and to provide useful information for policy intervention. Particular attention is paid to the role of institutions, technological progress and diffusion of innovation in the process of convergence or divergence between economic areas. To carry out its research, CRENoS collaborates with research centres and universities at both national and international level. The centre is also active in the field of scientific dissemination, organizing conferences and workshops along with other activities such as seminars and summer schools.
CRENoS creates and manages several databases of various socio-economic variables on Italy and Sardinia. At the local level, CRENoS promotes and participates to projects impacting on the most relevant issues in the Sardinian economy, such as tourism, environment, transports and macroeconomic forecasts.

www.crenos.unica.it
crenos@unica.it

CRENoS – Cagliari
Via San Giorgio 12, I-09124 Cagliari, Italia
tel. +39-070-6756397

CRENoS - Sassari
Via MURONI 25, I-07100 Sassari, Italia
tel. +39-079-213511

Title: EDUCATIONAL DIVIDE BETWEEN VOTERS: A NATIONWIDE TREND?

Prima Edizione: Maggio 2024
ISBN: 978 88 68515 188

Viale Bonaria 98 - 09125 Cagliari
Tel. 070/6848663 - info@arkadiaeditore.it
www.arkadiaeditore.it

Educational divide between voters: A nationwide trend?

Maria Grazia Pittau
Sapienza University of Rome & CRENoS

Francesco Politano and **Roberto Zelli**
Sapienza University of Rome

Abstract

Objectives. The increasing of higher education in almost all the Western democracies, has driven the growth of a mass graduate class. Has this produced an increase in partisan voting differences between lower-educated and high-educated? Does education affect in the same way low-income and high-income voters? *Methods*. We examine 2020 post-election data in the United States as a whole and in the states and we allow interaction between education and income at both individually and state level. *Results*. We find no clear pattern in educational attainment when associated with income. Education matters differently between low-income and high income voters. After controlling for individual characteristics and state-level of wealth, interaction between education and income results in a more complicated pattern of class-based voting than we might expect based on education and income alone.

Keywords and phrases: Income, education and voting, Inequality, Multilevel models, Hierarchical Bayes methods
Jel Classification: D72; C25; D31

Introduction

According to Gethin *et al.* (2022), the conflict between left-wing and right-wing parties has always been a fundamental issue in politics, power and elections. In the post-World War II period in 21 Western countries, conservative and Christian Democratic parties tended to draw more support from the more affluent social classes, while social democratic and left-wing parties tended to gain more support among low-income classes. The most important result of this study is a gradual process of disconnection between the effects of income and education on individuals' vote. While the household income effect seems to be stable over time, education has a reverse effect on voting behavior. In the United States during the 1950s–1960s, lower-income and lower-educated electorate was more prone to vote for social democratic and affiliated parties, and rich and high educated voters were supporting right-wing parties. Starting from the 1970s, this evidence gradually becomes less clear, with a disconnection between income and education. The top-income voters continue to prefer the right-wing parties, and the poor the social democratic parties, but the distinction between lower-educated and high-educated voters becomes blurred. The disadvantage of Democrats among college graduates compared to non-college graduates (after controlling for individual characteristics include income) narrows as time unfolds and the situation reverses in the 1990s. In the 2010, the disconnection between income and education becomes clear: high-income individuals continue to vote for the right, while higher-educated voters shift their preferences toward the left. This trend continues, reaching a record high in 2016 in the United States during Hillary Clinton's candidacy against Donald Trump. United States have transitioned from a single-elite system in the 1970s to a multiple-elite system in the 2010s. Specifically, the American right gets more support from the wealthy, while the left receives more support from the better educated, all other factors being equal.

This tendency is not isolated among Western countries. Stubager (2010) analyzed Denmark as a critical case for demonstrating the existence of an educational cleavage at the electoral level. Ford and Jennings (2020) found that the expansion of higher education and the rise of a mass graduate class, among other factors, has a major role in generating new cleavages in many Western European democracies. Class and education still show a massive impact on the political preferences and electoral behavior in nine Western countries, even when the individual level of "globalization", like the degree of cosmopolitan identification, is considered (Langsæther *et* al., 2019).

Starting from this evidence, this study aims at understanding whether this disconnection between education and income is always true. How individual's education level matter in shaping preferences of political attitudes? Lower-educated people are loyal to the right, regardless their household income and the place in which they reside and vote? Do the higher-educated poor and low-educated poor vote differently? Or it could be that geography, the place where people live, matters more than household income experiences?

These scenarios, or even their combination, could explain the puzzling patterns of voting behavior during the recent years in the United States. The U.S. provides an opportunity to study a Western democracy where the distinction between right-wing conservative and left-wing liberal parties is evident and dominates the total votes cast in each election.

In the U.S., Republican candidates have traditionally performed better in poor states, while Democrats candidates were favorite in the wealthiest states and among working class voters. However, a peculiar event occurred in the 2016 presidential elections. The "Blue Wall", a set of states (i.e. Pennsylvania, Michigan, and Wisconsin) characterized by a strong working class and consecutively won by the Democratic Party from 1992 to 2012, was breached by opposing candidate Donald Trump. By winning in states with a large working-class vote, 2016 Trump election has raised expectations about whether Republicans can gain support from poorer social classes that have been historically hostile towards them. In his inaugural address, Trump also acknowledged the 'forgotten ones', citizens who have been most disadvantaged by recent social transformations. The Republican Trump lost the presidency in 2020, although preserving a large part of the working class support.

On average, the progressive Democrats received the most votes in affluent states on the West and East coasts, while conservative Republicans performed better in less affluent states, particularly in the Southeast. Democratic candidate Biden won his way into the White House due to his success in 10 of the 11 richest states in the country, while Trump won 10 of the 11 poorest. It appears that there has been a shift in class support for the two major American political parties. This is a simple state-by-state evidence, however, at personal level it is still not very clear whether the poor still tend to lean liberal and the rich to the conservative.

Assuming that the research question pertains to individual-level occurrences but also geography matters for sharpening lines of political views, Gelman *et al.* (2007; 2010a) have already explained the inconsistency between individual patterns and those present in the aggregate data, showing different patterns of partisan voting by income. They found that in poor states such as Mississippi, richer people are much more likely than poor people to vote Republican, whereas in rich states such as Connecticut, there is very little difference in vote choice between the rich and the poor.

We contribute to this literature by understanding the reversal role of education not separately but interacted with household income. Their interaction along with the level of prosperity of the state in which voters reside could contribute to explain the strength and the influence of education. We also contribute to this evidence by adopting a very large dataset with a sample size of over 60,000 voters, that ensures a reasonable representativeness at state level. We use the most recent Cooperative Election Study (CES) post-election survey, conducted in 2020 on the United States presidential elections between Trump and Biden. Our analysis confirms many of the findings of the existing literature, but it also provides new insights into the transformation of political cleavages in the United States.

The rest of the paper is structured as follows. Section 1 presents the dataset exploited in the analysis and highlights the advantages of using these data with respect to more traditional data sources. Section 2 documents the empirical strategy and the econometric model estimated to explore the patterns of income, education and voting within and between states. Section 3 discusses the main results and suggests possible explanation and distinctness of the reversal of education cleavage. Section 3 concludes.

1 Data Sources

At micro level our analysis relies on survey data from the 2020 Cooperative Election Study (CES). The CES is a national stratified online survey sample administered by YouGov. Data are collected in two waves: the pre-election survey data were collected from late September to late October, and the post-election survey data were collected in November after the presidential election (Schaffner *et al.*, 2021). The CES survey, conducted since 2006 and originally known as the Cooperative Congressional Election Study has the purpose of studying and understanding the political behavior and attitudes of Americans citizens (Ansolabehere *et al.*, 2021). The 2020 edition, which released data in two parts during 2021 and is available in the Harvard University

database, surveyed a sample of 61,000 people. The availability of data on such a large sample, which for California goes as high as 5035 voters, made it possible to analyze variation across the states, as well as the interaction between individuals effect in relation to the average level of wealth in each state. This feature makes CES post-election micro data preferable to American National Election Study (ANES) when looking for hierarchical interaction between individual and state-level predictors. In fact in 2020 the ANES post-election interviews were around 8,600 in the whole country.[1] The 2020 CES edition involved 60 research groups, each of whom commissioned YouGov to conduct a survey of approximately 1,000 interviews through CAWI (Computer Assisted Web Interviewing). Rather than using a traditional probabilistic selection method, matched random sampling was used. The researchers utilized volunteer panel members who had already responded to pre-election interviews from YouGov, Dynata, Critical Mix, and Prodege (74,099 people after quality controls). These panel members were reinvited to respond to post-election interviews as well.
A target sample was selected from the 2019 American Community Survey list of citizens who were not contacted directly. Instead, the CES 2020 data available come from associated sample, which is the sample of panel respondents who were associated with the target sample through association by proximity, which was based on a weighted Euclidean distance metric conditioning on registration status, age, race, gender and education. This association is considered reliable since it complies with three properties: ignorability, smoothness, and common support. It is important to note that the sampling process was validated by comparing election results among sample members state by state and actual voting results in the presidential election. This proves the quality of the sample and the goodness of the weights for post-stratification.

The variable indicating the candidate for U.S. president voted for by the respondent (*For whom did you vote for President of the United States?*, `CC20_410`) was recoded into a binary variable with value equal to 1 if the respondent voted for Trump, 0 otherwise.

At macro level, data information is derived from the Bureau of Economic Analysis (BEA) and the Census Bureau. Potential state-level predictors are the average personal income in each state for the year 2020 and income inequality, as measured by the Gini index, for the year 2019. Starting from population estimates provided by the Census Bureau, BEA estimates the average personal income as the ratio of total personal

[1] Because of substantial missingness in some variables, the original CES sample size reduced to around 40,000 voters, but still remains larger than the 2020 ANES sample.

income to the total population at midyear.[2]

Estimations of Gini index of income inequality at sub-national level are based on American Community Surveys data and they are officially released by the Census Bureau's Population Estimates Program.[3]

2 Empirical Strategy

The Divergence of Income and Education

To document how dissimilar is the influence of education and income on partisan preferences, we start by estimating the difference in supporting Trump between highly educated and low educated individuals for different classes of annual household income. The results from a simple linear logistic model of the form:

$$\widehat{\Pr(\texttt{Trump})} = logit^{-1}\left(\alpha + \beta\ \texttt{income} + \gamma\ \texttt{education}\right) \tag{1}$$

are shown in Figure 1 that simultaneously displays the probability of support for Trump in 2020 for voters with at least a bachelor's degree (blue line) and voters with a lower level of education (red line) and for each of the 16 income categories.[4] Highly educated voters are here defined as those with at least a Bachelor of Arts (B.A.) or Bachelor of Science (B.S.), corresponding to a four-year college. In either programs, students typically take general education courses for their first two years and major-specific courses for the final two years. Going to college has in fact numerous short and long-term benefits that are fundamental in delineating voting behavior. The possibility of attending a wider range of intellectual discourse on campus results in a more open-minded way of considering differing opinions and understanding the intellectual value and democratic significance of political tolerance. In the long term, going to college implies, on average, better job opportunity and future higher earnings.

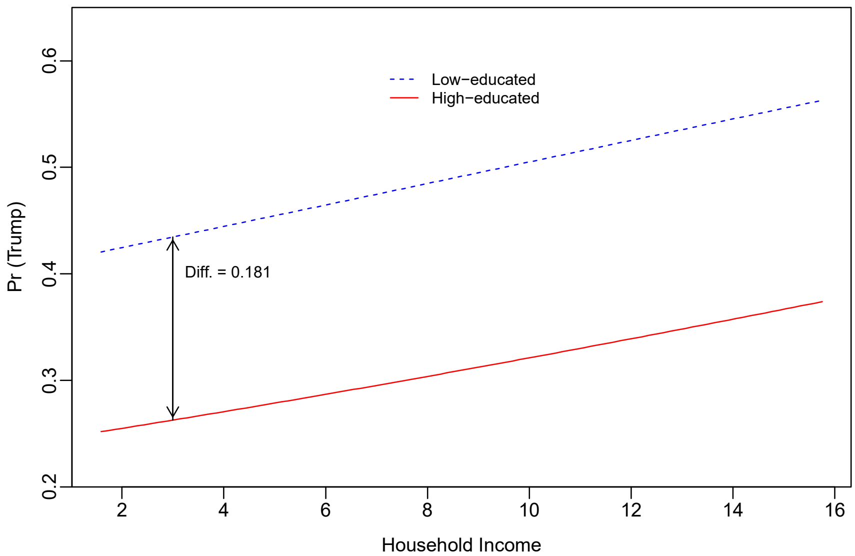
Graduates are less likely to vote Republican with respect to non graduates with an estimated difference in probability of around 18% (Figure 1). At the same time,

[2] https://apps.bea.gov/itablehttps://apps.bea.gov/itable.

[3] Datacanbedownloadedathttps://https://data.census.gov/all/tables?q=B19083:\%20GINI\%20INDEX\%20OF\%20INCOME\%20INEQUALITY. The margin of errors representing the degree of uncertainty around the estimated figures are also available.

[4] Household income is actually ordinal in 16 classes, but it is included in the model as a numerical variable. The categories correspond to annual family income below \$10,000, \$10,000 –\$19,999 and so on. Above \$100,000 classes are: \$100,000 – \$119,999; \$120,000 – \$149,999; \$150,000 –\$199,999; \$200,000 –\$249, 999; \$250,000 –\$349, 999; \$350,000 – \$499, 999; and \$500,000 or more. On the use of ordinal variables with a reasonable number of categories as a continuous variable see, among others, Gelman and Hill, (2006).

Figure 1: Estimated difference in supporting Trump between high-educated and low-educated as a function of household income.



high-income individuals tend to vote for the "right", with an expected difference in the likelihood of supporting Republicans between "rich" (household income in the class 120,000–149,000, corresoponding to the 80th percentile) and "poor" (household income in the class 20,000–30,000, 20th percentile) of around 8% in line with what already found in the literature (Ghetin *al.*, 2022 among others).

A more insightful and interesting question is the following: do voting patterns of the rich and the poor differ according to individuals' level of education? To compare low-educated and high-educated poor together, or low-educated and high-educated together, we allow for a simple interaction between income and education (see equation 3).

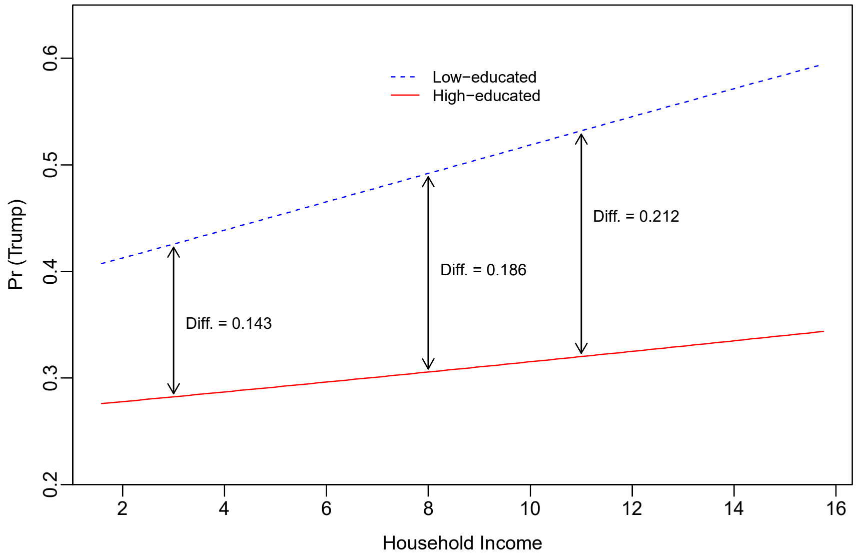
$$\Pr(\widehat{\texttt{Trump}}) = logit^{-1}\left(\alpha + \beta\ \texttt{income} + \gamma\ \texttt{education} + \delta\ \texttt{income} * \texttt{education}\right) \qquad (2)$$

In the absence of controls, the income effect is positive and the education effect negative, as expected. The estimated coefficient of the interaction is negative (-0.22) and statistically significant: the difference between graduates and not graduates increases as household income rises.[5] The income gradient of support to the Republican Party is steeper for low-educated voters than for graduates voters. Figure 2 shows this empirical evidence: poorly educated voters (20th income percentile) have a probability

[5]Model estimations are reported in Appendix A, table A.1.

of supporting Trump less than 14.3% compared to their graduated counterparts. This difference becomes equal to 18.6% for median income voters and is equal to 21.2% for upper-income individuals (80th income percentile). This gap between high-educated and low-educated gradually enlarges as income increases and reaches its highest level for individuals with income exceeding $500,000.

Figure 2: Estimated difference in supporting Trump between high-educated and low-educated as a function of household income along their interaction.



The divergent pattern is robust to additional socio-demographic individual controls.[6] The signs of the control variables are as expected, income coefficient is still positive and education negative. The interaction becomes stronger with an estimated coefficient equal to -0.266 (see Appendix A), meaning that for two individuals with annual income exceeding $350,000 with same characteristics except education the expected difference in supporting Trump is equal to 22%. The difference drops to 9% when considering two individuals with different level of education (graduated and not graduated) with annual income less than $20,000, other things equal.

Party affiliation and education in American states

Considering the country as a whole, we have found that the income gradient is most steep–indicating differences in preferences across income groups–for low-educated individuals than for graduated voters. As a matter of fact, in almost all the recent elections

[6] We control for gender, birth generation, ethnic group, importance of religion and for the place of residence: urban/rural areas.

Democrats performed well in the richer blue states in the Northeast and Coasts, while Republicans dominate in the red states in the Midwest and the South. Gelman *et al.* (2010) have already discussed that in poor states, rich people are much more likely than poor people to vote for the Republican presidential candidate, but in rich states income has a very low correlation with vote preference.

Comparing to these previous studies, our key contribution is to understand if the effect of education between income groups, has the same influence in wealthy states, as compared to poor U.S. states. In other words, how much does context matter? The wealth of U.S. states is approximated by the average personal income. Within this framework we have two levels of variation: individuals clustered within states with potential for predictors at different levels.

In summary, beyond household income, there are seven categorical variables: State[7], gender, race, age, religiosity, area of residence (urban/rural) and education. Age and education are actually ordinal, but we include them as unordered categories in our model. Of the 51 x 2 x 4 x 5 x 2 x 2 x 2=16320 distinct categories, we have at least one observation from only 7528 of them. The high level interaction effects between sets of these categorical variables requires a high ratio of parameters to data points, motivating the use of a regression model that allows for shrinkage, or regularization, such as a multilevel Bayesian model.

Modeling relations between Income and Education across States

To study the relation of education and income to individual vote preferences, controlling for state, we fit a multilevel logistic regression of vote preferences on household income (using the sixteen-point scale in footnote 3 and state variables.

The model uses a dichotomous response variable in the analysis of the 2020 election data, indicating whether the preference was for Trump or Biden. Let $\pi_{i[j]} = P(Y_{i[j]=1})$ be the probability that individuals i resident in state j support Trump, where Y is a binary variable indicating whether in the 2020 election the vote was for Trump $(Y = 1)$ or Biden $(Y = 0)$.

The probability of supporting Trump is directly estimated by the following *varying-intercepts* and *varying-slopes* multilevel logistic model:

$$\pi_i = logit^{-1}(\underbrace{\alpha_s + \beta_s z_{(i,s)})}_{\substack{\text{varying}\\ \text{across states}}} + \underbrace{\beta x_i}_{\substack{\text{invariant}\\ \text{across states}}} + error) \tag{3}$$

[7]Including the District of Columbia

where

- $logit^{-1}$ is the inverse-logistic function, s indexes the area s where household i resides,
- z are individual-level predictors with states varying coefficients,
- x are individual-level predictors with states invariant coefficients,
- α_s (intercept), β_s (slopes) are the varying-parameters of the model.

What makes model (3) multilevel is the inclusion of contextual variables and the modelling of α and β.

The first level of our main model states:

$$\begin{aligned} &\Pr(Y_i = \text{Trump}) = logit^{-1}\big(\alpha_{s[i]} + \beta_{1(s[i])}\,\texttt{income}_i + \\ &\beta_{2,(s_[i])}\,\texttt{education}_i + \beta_{3,(s[i])}\,\texttt{income}_i \times \texttt{education}_i + \\ &\boldsymbol{\beta}'_{0,([i])}\,\mathbf{X}_i + error\big) \end{aligned} \tag{4}$$

for $n = 39664$ individual responses, $s = 1, \cdots, 51$ states and $\mathbf{X}$ is the matrix of invariant control variables. Family income, treated a numerical variable, is scaled in order to have mean 0 and standard deviation 1, education is coded 1 if the respondent has at least a college degree and 0 otherwise. The subscript notation $s[i]$ denotes the state of residence for individual i.

The main feature of the higher level models for the parameters in equation (4) is that every group of state level parameters is normally distributed around a national mean. We therefore allow α_s and β_s to vary across U.S. states, modeling as a function of contextual predictors (the average personal income and the level of inequality), to capture cross-sectional contextual effects (see among others Massari *et al.*, 2013; Shirley and Gelman, 2015; Farcomeni *et al.*, 2016)

$$\alpha_s \sim \mathrm{N}(\alpha + \delta_1\,\mathbf{Z}_{1,s} + \delta_2\,\mathbf{Z}_{2,s}, \sigma^2_\alpha), \tag{5}$$

$$\beta_{1,s} \sim \mathrm{N}(\beta_1 + \gamma_1\,\mathbf{Z}_{1,s}, \sigma^2_{\beta_1}), \tag{6}$$

$$\beta_{2,s} \sim \mathrm{N}(\beta_2 + \gamma_2\,\mathbf{Z}_{1,s}, \sigma^2_{\beta_2}), \tag{7}$$

$$\beta_{3,s} \sim \mathrm{N}(\beta_3 + \gamma_3\,\mathbf{Z}_{1,s}, \sigma^2_{\beta_3}), \tag{8}$$

where $\mathbf{Z}_{1,s}$ and $\mathbf{Z}_{2,s}$ denote the 51 matrices of state level covariates that affect the state intercepts and slopes respectively, $\mathbf{Z}_{1,s}$ is average personal income, and $\mathbf{Z}_{2,s}$ is Gini income inequality, for $s = 1, \cdots, 51$.

Since the challenge in fitting a hierarchical model is estimating simultaneously all the data-level regressions, Bayesian inference treats the macro-level model as "prior information" in estimating the micro-level coefficients. The resulting posterior distribution is proportional to the likelihood multiplied by the prior distribution of the parameters, and inferences are typically summarized by random draws from this product. In vast majority of cases posterior distribution is not directly available and simulation techniques (as Markov chain Monte Carlo, MCMC) are required to obtain a sample consisting of many draws from the target posterior distribution. MCMC generates samples from the posterior distribution by constructing a reversible Markov-chain that has as its equilibrium distribution the target posterior distribution. The final estimates will depend on the information that comes from the data and from the priors; the more information is contained in the data, the less influential are the priors. Point estimates are the medians computed from simulations. The standard deviations are computed from the same set of draws and are proportional to the median absolute deviation (MAD). We fit the model by using the function `stan_glmer` (Goodrich *et al*, 2020), which allows for weakly informative default prior distributions for estimation to enhance calculation stability. The default options provided by the software are often more practical for many models, as a priori knowledge is restricted to the variables' order of magnitude, despite being technically reliant on data. We run the MCMC algorithm for 6000 iterations, half of which were warm-up or burn-in, for each of the four required Markov chains. We performed all pre-processing and post-processing in R (R Core Team, 2023).

3 Main Results

We discuss our results in three sections: first we re-examine the relation of education and income to individual vote preferences, controlling for gender, religion, church attendance, rural/urban, state, race/ethnicity, and birth generation at national level.

Second, we investigate how the variability between the estimated 51 state-level coefficients (see equation 5) – representing the probability of support for the Republican candidate (on the logistic scale) – can be explained by the state-level of wealth and inequality. Finally, we look at the interaction between individuals and state effects.

Average Individual Effect

After controlling for individual demographic variables and for variation between state, the estimated effects of income, education and their interaction (average of β_1, β_2 and β_3 from equation 4) are still big in size (0.37, -0.69 and -0.26) and statistically significant (with standard error of 0.04, 0.04 and 0.06 respectively). The first estimate confirms that higher levels of household income are associated with higher levels of support for Republican presidential candidates. In contrast, highly educated voters are less prone to support Republicans and the results confirm that the role of income for less educated voters is more than three times higher than for graduated individuals.

State Wealth Effect

For a baseline voter (a not religious Caucasian man, baby boomer, with average income, not educated, living in an urban area), the probability of supporting Trump in the 2020 presidential elections was equal to around 35%, with great variation among states, with posterior standard deviation of 0.19. These varying intercepts estimate the predictive effects[8] of residing in a given state that are not already explained by whatever state level variables are included in the model. For example, a Washington DC resident has an expected probability of supporting Trump around to 28%, while for a resident of Wyoming the likelihood of voting Trump rises up to 41%, other things equal. This variability is partially explained by state's average income and, to a lesser extent, state level of inequality measured by Gini coefficient. The posterior mean of the state level average income $\delta_1 = -0.36$ with posterior standard deviation of 0.06 in equation (5) has a negative significant geographical effect: Republicans do better in the poorer states while richer states support Democrats, in line with the recent literature (see among others Gelman *et al.*, 2007). State-level inequality is positively, although not strongly, related to the probability of supporting Trump. The posterior mean of the state level Gini inequality coefficient in equation (5) is equal to $\delta_2 = 0.18$ with posterior standard deviation of 0.06 indicating that, all else equal, the greater the inequality, the more one would expect to find more support for republican candidates.

Economic inequality in all the U.S. states is particularly high and rising, according to the 2020 data from the Census Bureau, Gini coefficient ranges from 0.43 in Utah to

[8] This is not a *causual* effect but merely an association between Trump vote share and state of residence.

0.51 in New York State. Americans citizens seem however more willing than citizens of other rich countries to tolerate inequality and the government does comparatively little to redress the balance. Economic theories suggest that in countries with high inequality the majority of the population will support democracy as a potential mechanism for redistribution. In line with that, Galbraith and Hale (2008) found evidence of a negative correlation between state-level inequality and the Republican vote share, at least until 2004. The authors suggest however caution in interpreting this results: citizens do not necessarily have the right perception of inequality and accordingly they do not consider local economic disparities when they cast their vote.

In any case American citizens do not ask for government action to address inequality, they are not clamoring for demand for redistribution. In fact, more recent studies (McCarty *et al.* 2006; Gelman *et al.*, 2010b; Condon and Wichowsky, 2020; Kelly, 2020) show that economic inequality helps Republicans win elections, rather than making the public favor redistribution, legitimating the positive sign of our estimate.

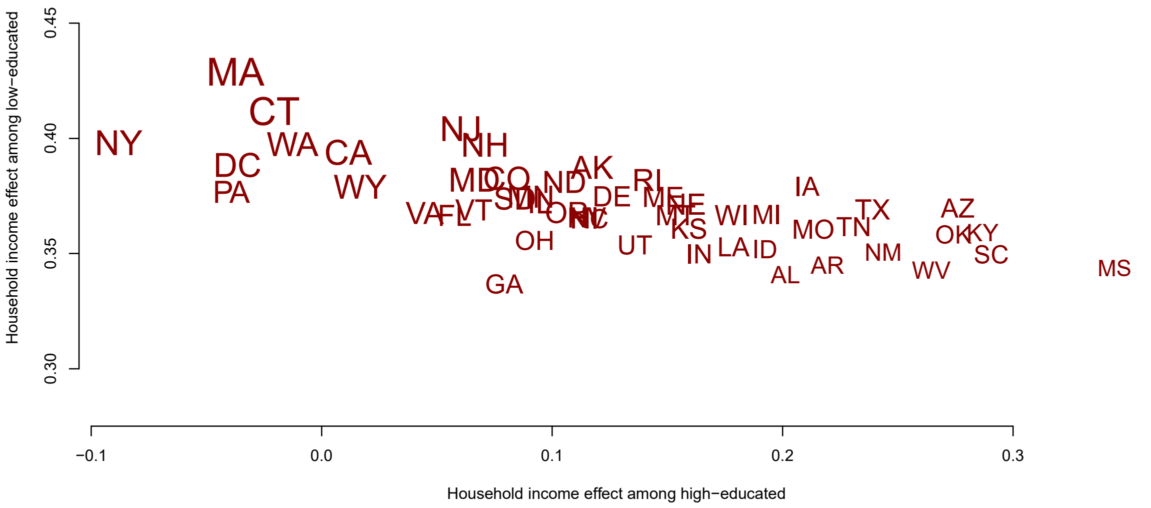
The state-level Divergence of Income and Education

We have already found a quite important gap in the income effect on voting support between low-educated and high-educated individuals considering the country as a whole. This evidence assumes different characteristics if we allow the coefficients β_1, β_2 and β_3 in equations (6 –8) to vary across American states. Within states, the difference in the income coefficients between low-educated and high-educated voters is far from being uniform. The divergence between the two income effects becomes larger in wealthier states than in poor states. Figure 3 shows the difference in estimated income slopes for all 51 states[9], revealing a clear pattern with high differences in rich states–bigger abbreviation size–and low differences in poor states–smaller abbreviation letters. Part of this variability is explained by the level of wealth of the state, measured by the average state income and included as group-level predictor. The inclusion of this predictor reduces the unexplained variability among American states and increases the precision of the multilevel model fit. Particularly, it reduces the residual error of the interaction coefficient $\beta_{3,s}$ in equation (8), indicating the the variation of the interaction effect across states is partially explained by the different state's general level of wealth.

The corresponding estimated curves are shown in Figure 4, in which the probability of supporting Trump is plotted as a function of income category. For a better understanding, we reported only a selection of American states, say two rich states

[9]Coefficients estimates are reported in Table in Appendix A.2.

Figure 3: Income effect estimated by the multilevel model among high-educated (x-axis) and among low-educated (y-axis), by state. The size of the abbreviation of each state is based on its average personal income

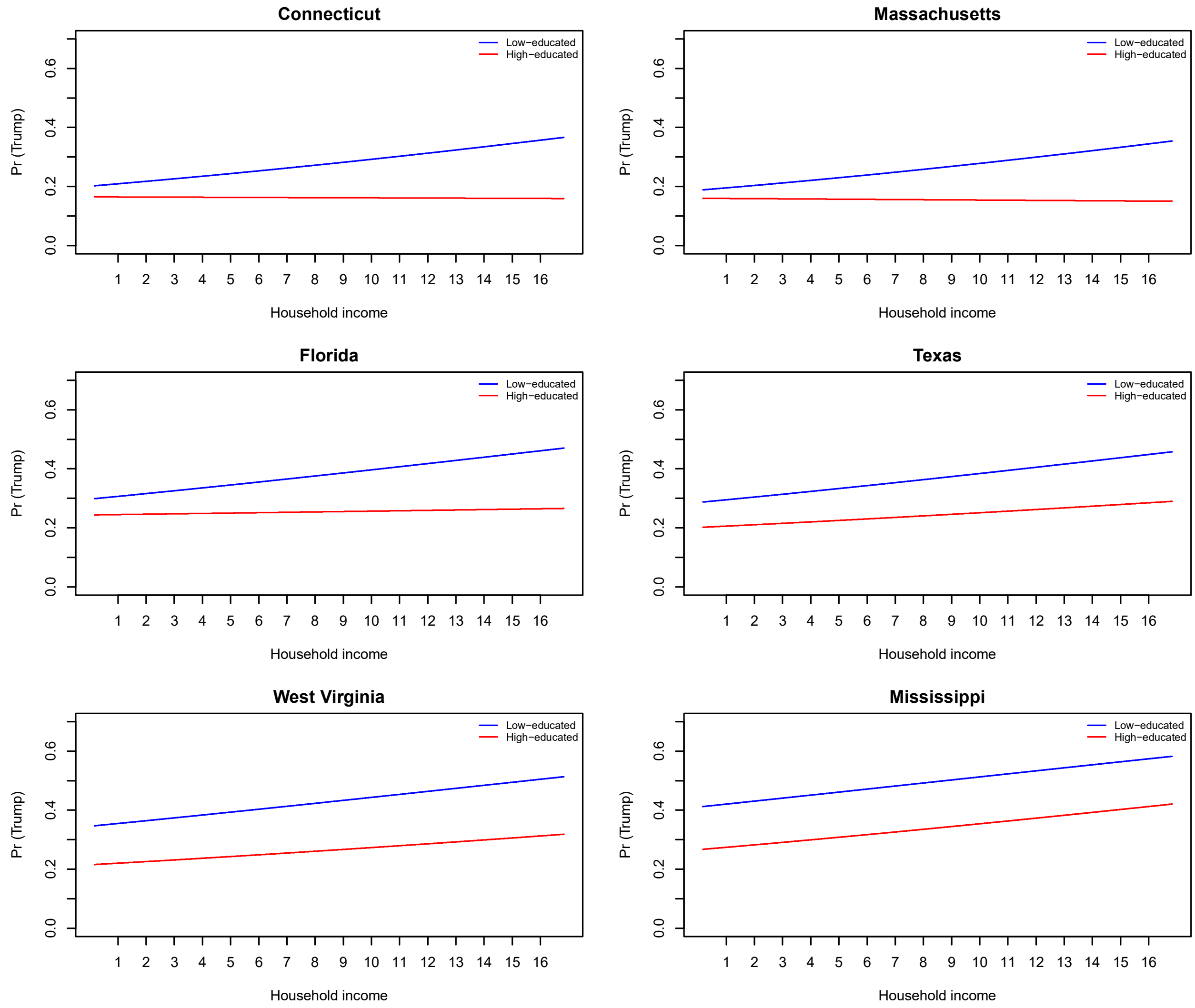


(Connecticut and Massachusetts), two middle-income states (Florida and Texas) and two poor states (West Virginia and Mississippi), respectively.

Overall, Figure 4 reveals three patterns. First, the probability of supporting Republicans is always higher for uneducated voters (blue curve in the graph) than for college-educated voters (red curve). Second, there is evidence of a strong correlation between household income and preferences for Republican candidates and support for Republicans increases with income. What is new is that these two curves are not necessarily parallel. The difference in household income effect among low- and high-educated voters depends on the state of residence, with the largest gap in the richest states – Connecticut and Massachusetts – and the smallest gap in the poorest states – West Virginia and Mississippi.

In more details, among low-educated voters, ho usehold income is always a positive predictor of Republican presidential vote support: within states richer low-educated voters are more likely to support the Republicans, and this pattern is consistent across all the 51 states. More intriguing, among the college-educated voters household income is not always a positive predictor of Republican vote support. The income effect among college-educated individuals in fact vary from rich state (i.e. Connecticut, Massachusetts, New York, Washington DC, Pennsylvania) in which is negative or near

Figure 4: Difference between high-educated and low-educated estimates of Trump support (where the national average has been subtracted out) within six states that are rich (Connecticut and Massachusetts), middle-income (Florida and Texas), and poor (West Virginia and Mississippi). The curves show the probability of supporting Trump as a function of income category. The blue curve is for low-educated voters and the red curve for college-graduated individuals.



zero, to middle and low-income states (i.e. Florida, Texas, Nevada, Maine, Michigan, Alabama, South Carolina, West Virginia and Mississippi) in which the coefficients for income are positive and tend to be higher in poorer states.[10] These results translate in bigger differences among income effects between low-educated and college-educated voters on Republican support in rich states than in states with lower average incomes.

Specifically, in rich states, (Connecticut and Massachusetts, but also in other states like New York States, New Jersey, California, District of Columbia ecc.) among the low-educated voters only a small proportion (around 23%) of low-income voters support Trump, *ceteris paribus*. However, this probability increases among higher income voters, reaching the value of 39% for very wealthy people.

In rich states individuals with college degrees tend to lean more towards supporting Democratic candidates: the probability of voting Trump in Connecticut for poor individuals, other things equal, is around 16% and it remains almost stable as household income increases. The difference in supporting Trump between two poor individuals (annual household income between 10,000 and 20,000 US$) one with low education and the other with at least a college degree is equal to 6% but, when we compare two individuals with household income in the class 350,00–500,000 US$ this gap rises to 23%. Similarly if we compare two individuals residing in Massachusetts or in similar rich states. There is a clear education gap in voting patterns among the wealthy people, with Trump performing better among voters with lower levels of education compared to those with higher levels of education.

In Florida and Texas, and similarly in other medium-rich states, the probability of support for Republicans is, on average, higher than in the richest states, independently of household income. Although less pronounced with respect to rich states, the difference between voters with and without college is still significant, especially among rich individuals. Among poor voters we estimate an around 8% gap in Florida and a 9% gap in Texas, but these gaps increase to 22% in Florida and 28% in Texas among upper-income voters, indicating again that education level still matters and matters more for wealthy individuals than for poor voters. Republicans have in fact a substantial advantage over Democrats among educated voters in the highest income tier, but only a modest advantage among them at the lowest income tier. Among voters without a bachelor's degree, higher income is associated with being more Republican, but this difference between low and high educated becomes unremarkable as we move to less

[10]See Table A.2 in Appendix A.

wealthier states. Although not all residents of lower-income states voted for Trump, support for Republican candidates is, on average, higher than in medium-rich states. After controlling for demographic and individual characteristics, in West Virginia 38% majority of voters without a college degree who are lower income are Republicans and around 44% in Mississippi. In West Virginia 41% of voters without a college degree who are middle income are associate with the Republican Party and this percentage rises to 48% in Mississippi. Voters without a college degree at the highest income levels affiliate with the Republican Party are in 54% West Virginia and 61% in Mississippi.[11] The preference for Republican party in poor states, like West Virginia and Mississippi, increases with income and in almost the same way for voters with and without college degree. The the gap in voting patterns between high and low educated is almost stable regardless the level of income, other things equal ((lower panel in Figura 4). In West Virginia, among low-income voters, the gap between low and high educated in supporting Donald Trump is around 19% and equal to 14% for voters in the upper income tier. In Mississippi the educational gap is even smaller: among voters in the lower income tier is 15% and equal to 16% for voters in the upper income tier. In poor states the educational gap does not significantly change with income.

This is to say that for poor voters residing in rich states education does not really matter in supporting Trump, but Republicans have a modest edge among low-income voters in poor states. As household income increases, education counts more and more in shaping voting preferences, even after controlling for basic demographic variables. In addition, this correlation is stronger in rich states (i.e Connecticut), than in middle-income states (i.e. Texas) and almost negligible in poor states (i.e. Mississippi).

Diagnostic

Can we trust these results? Concerning the diagnostics of the logistic model, the graphical analysis of the binned residual proposed by Gelman and Hill (2006), divided into 700 classes, did not reveal any anomalies. In fact, 92% of the points fall within the 95% confidence intervals, and there are no clear patterns that suggest the omission of a relevant variable for the analysis or the need for algebraic transformations of the predictors.

The necessary checks, given that the model estimation was carried out with a Bayesian approach using the Monte Carlo Markov chain method, concern the convergence of the Markov chains. The trace plots of the main predictors show how the

[11] A similar pattern is found in the other poor states. All the estimates are available upon request.

chains, although starting from different points, arrived at the same area. The R statistic, which compares the between–and within–chain estimates for model parameters, never exceed the limit equal to 1.1. The Effective Sample Size of the estimates, higher than 1,000 as recommended by Muth *et al.* (2018) for social science problems, highlights that the autocorrelation present in the Markov chains does not negatively affect the posterior distribution of any of the parameters. Furthermore, except for the posterior log density, which has a value of 0.24, all the estimated parameters have a Monte Carlo standard error less than 0.0015.
Finally, the "Posterior Predictive Checking" technique verified the good fit of the model to the data by simulating from the posterior distribution according to its specification: the simulated results reproduce quite faithfully the characteristics of the distribution of the outcome variable.[12]

Concluding Remarks

Our multilevel analysis reveals the puzzling role of education as a potential cleavage between Republicans and Democrats, whose role is not uniform within the 51 American states. To explore the role of education from a different perspective, we estimated a multilevel Bayesian model to 39,664 individual responses to the question 'For whom did you vote for President of the United States?' by using individual data from the 2020 released of the Cooperative Election Study and state level variables from the Census Bureau. The use of a structured prior distribution on the state effects allowed us to estimate simultaneously their variation, while also estimating interaction effects that shed new light on certain relationships between education and income and Trump support, and on varying patterns in Republican support among states.

We start with the basic facts that both income and education are strongly associated with vote preferences. Within states the traditional rich-poor divide remains and to some extent, also the role of education. The percentage of voters without a bachelor's degree, associated with the Republican Party is higher compared to graduates voters, confirming the "education gap" in political preferences.

We also explore the way in which education interacts with income and we found that the income gradient of support to Republican Party is steeper for low-educated individuals compared to voters with a college degree or more. Household income pre-

[12]Details of the deviance, effective number of parameters and deviance information criterion of the multilevel Bayesian models that we fitted are available upon request.

dicts vote choice much more for voters without a bachelor's degree, than for college graduates, and this pattern is not equal across all the American states. States differs in many ways, both socially and economically. Rich, educated people in rich states are socially and economically more liberal than rich educated people in poor states. Affluent individuals residing in a poor state, may still be part of communities or social circles where conservative values are prevalent, and this translates into Republican dominance among upper-income voters, with and without a college degree. For analyzing this variation in the role of income and education within and between states we allow for interactions between income, education and level of state wealth.

The most important finding in our research is that the education gap in conservative voting is greater in rich states than in poor states, also after controlling for economic, demographic and social indicators. In addition, educational division becomes extreme significant for upper-income voters. In poor states the role of education is almost negligible: Republicans have always the support of the rich voters, independently on their level of education. Low-educated people are mostly loyal to the right and their loyalty increases as income rises. In wealthier states, education level often becomes a significant cleavage in political preferences. This is partially due to access to information and critical thinking skills: individuals with higher levels of education are more likely to critically evaluate political messages and policies. Furthermore, social concerns such as health care, inequality, discrimination, environmental sustainability and so on are issues that college degrees voters care about which align more closely with progressive or Democratic agendas. And this is particularly true in wealthier states, where higher education levels may be more closely associated with economic security rather than just economic advancement.

Bibliography

Ansolabehere, S., Schaffner, B. F. and Luks, S. (2021). Cooperative Election Study 2020, Common Content. Release 2: August 4th, 2021. *Cambridge, MA: Harvard University.* `http://cces.gov.harvard.edu`

Condon M., and Wichowsky A. (2020). *The Economic Other Inequality in the American Political Imagination.* The University of Chicago Press.

Farcomeni A., Pittau M.G. , Zelli R. (2016). Has the attitude of US citizens towards redistribution changed over time?, *Economic Modelling*, 52 (B), 714–724.

Ford, R. and Jennings W. (2020). "The Changing Cleavage Politics of Western Europe." *Annual Review of Political Science*, 23, pp. 295–314.

Galbraith J. K. and T. Hale (2008). State Income Inequality and Presidential Election Turnout and Outcomes, *Social Science Quarterly*, Vol. 89 (4), pp. 887-901.

Gelman, A. and J. Hill (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gelman, A., Park, D., Shor, B., and Bafumi J. (2007). Rich State, Poor State, Red State, Blue State: What's the Matter with Connecticut?, *Quarterly Journal of Political Science*, 2, pp. 345-–67.

Gelman, A., Park, D., Shor, B., and Jeronimo C. (2010a). *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do.* Princeton University Press.

Gelman A., Kenworthy L. and Su Y-S (2010b). Income Inequality and Partisan Voting in the United States, *Social Science Quarterly*, Vol. 91, No. 5, Special Issue on Inequality and Poverty: American and International Perspectives, pp. 1203–121.

Gethin, A. , Martínez-Toledano, C., and Piketty, T. (2022). Brahmin left versus merchant right: Changing political cleavages in 21 Western Democracies, 1948–2020, *The Quarterly Journal of Economics*, 137, 1, pp. 1–48.

Goodrich B., Gabry J., Ali I. and & Brilleman S. (2020). `rstanarm`: *Bayesian applied regression modeling via Stan.* R package version 2.21.1

Kelly N. J. (2020). *America's Inequality Trap*, Chicago Studies in American Politics. The University of Chicago Press.

Langsæther, P. E., and Stubager R. (2010). Old Wine in New Bottles? Reassessing the Effects of Globalisation on Political Preferences in Western Europe, *European Journal of Political Science Research*, 58, pp. 1213–1233.

Massari R., Pittau M. G. and R. Zelli (2013). Hierarchical modelling of disparities in preferences for redistribution. *Oxford Bulletin of Economics and Statistics*, 75, 4.

McCarty, N. M., Poole K. T., and Rosenthal H. 2006. *Polarized America: The Dance of Ideology and Unequal Riches.* Cambridge, MA: MIT Press.

Muth, C. , Oravecz, Z. Gabry, J. (2018). User-friendly Bayesian regression modeling: A tutorial with rstanarm and shinystan. *The Quantitative Methods for Psychology*, 99–119.

R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

https://www.R-project.org/

Schaffner B, Ansolabehere S and Luks S (2021). Cooperative election study common content, 2020. *Harvard Dataverse*, V4, 1.

Stubager, R. (2010). The Development of the Education Cleavage: Denmark as a Critical Case, *West European Politics*, 33, pp. 505–533.

Shirley K. E. and A. Gelman (2015). Hierarchical models for estimating state and demographic trends in US death penalty public opinion. *Journal of Royal Statististical Society*, series A, 178, pp. 1–28.

Appendix A

Table A1: Estimated coefficients with relative standard errors of individual characteristics in the fitted varying-intercept and varying-slope multilevel logistic regression model in the US. From basic model to hierarchical core model.

	Basic model		Basic model with interaction			Interaction Inc:educ and controls			Hierarchical core model		
	Estimates	S.E.	Estimates	S.E.		Estimates	S.E.		Estimates	S.E.	
Intercept	-0.126	0.013	-0.119	0.014	***	-0.656	0.029	***	-0.623	0.043	***
Household income	0.285	0.023	0.376	0.030	***	0.351	0.035	***	0.376	0.037	***
Education	-0.769	0.023	-0.738	0.029	***	-0.676	0.027	***	-0.694	0.035	***
Household income:Education			-0.218	0.047	***	-0.266	0.052	***	-0.261	0.059	***
Importance Religion						1.555	0.025	***	1.536	0.025	***
Silent Generation						0.210	0.042	***	0.212	0.043	***
X Generation						0.003	0.030		0.000	0.031	
Y Generation						-0.299	0.032	***	-0.288	0.032	***
Z Generation						-0.759	0.063	***	-0.758	0.064	***
Ethnic group: afro-american						-2.771	0.066	***	-2.863	0.068	***
Ethnic group: hispanic						-0.591	0.045	***	-0.604	0.048	***
Ethnic group: other						-0.169	0.046	***	-0.152	0.047	***
Female						-0.497	0.024	***	-0.505	0.024	***
Area of residence: rural						0.633	0.029	***	0.610	0.030	***
State income									-0.365	0.064	***
State Gini inequality									0.176	0.060	***
Household income: state income									0.028	0.076	
Education: state income									0.038	0.069	
Education: Household income: state income									-0.198	0.121	*
σ_α									0.188		
σ_{β_1}									0.067		
σ_{β_2}									0.123		
σ_{β_3}									0.152		
n	39672		39672			39664			39664		
States									51		

Note1: *** significance level at 0.01: ** significance level at 0.05: * significance level at 0.10.

Note2: Boomer is the baseline group for Generation and Caucasian for Ethnic group.

Table A2: Varying income effects between high and low educated voters by state wealth.

State	*Avg inc*	$\beta_{IncL.E.}$	$\beta_{IncH.E.}$
Connecticut	78463	0.41	-0.02
Massachusetts	78388	0.43	-0.04
New York	71577	0.40	-0.09
New Jersey	71505	0.40	0.06
California	70647	0.39	0.01
New Hampshire	68542	0.40	0.07
District of Columbia	68350	0.39	-0.04
Washington	68350	0.40	-0.01
Wyoming	65782	0.38	0.02
Maryland	65685	0.38	0.07
Colorado	65358	0.38	0.08
Alaska	62756	0.39	0.12
Minnesota	62240	0.37	0.09
Virginia	62189	0.37	0.04
Illinois	62139	0.37	0.09
North Dakota	60864	0.38	0.11
Pennsylvania	60685	0.38	-0.04
South Dakota	60446	0.37	0.08
Rhode Island	59941	0.38	0.14
Vermont	59296	0.37	0.07
Nebraska	57421	0.37	0.16
Florida	57292	0.37	0.06
Hawaii	57241	0.37	0.11
Oregon	57005	0.37	0.11
Delaware	56324	0.37	0.13
Kansas	55974	0.36	0.16
Wisconsin	55941	0.37	0.18
Texas	55601	0.37	0.24
Nevada	55406	0.37	0.12
Maine	54912	0.37	0.15
Montana	54106	0.37	0.15
Ohio	53545	0.36	0.09
Michigan	53388	0.37	0.19
Iowa	53312	0.38	0.21
Tennessee	52351	0.36	0.23
Arizona	52327	0.37	0.28
Utah	52225	0.35	0.14
Indiana	52219	0.35	0.16
Missouri	52108	0.36	0.21
Georgia	51987	0.34	0.08
North Carolina	51900	0.37	0.12
Louisiana	50809	0.35	0.18
Oklahoma	50518	0.36	0.27
Idaho	49491	0.35	0.19
South Carolina	49105	0.35	0.29
Kentucky	47525	0.36	0.29
Arkansas	47154	0.34	0.22
New Mexico	46760	0.35	0.24
Alabama	46179	0.34	0.20
West Virginia	45240	0.34	0.26
Mississippi	42716	0.34	0.34

Ultimi Contributi di Ricerca CRENoS

I Paper sono disponibili in: http://www.crenos.unica.it

24/07 *Federico Atzori, Vittorio Pelligra,* "Framed norms. The effect of choice-belief information on tax compliance"

24/06 *Simone Nobili,* "Concentration, Market Power and International Tax Competition"

24/05 *Maria Grazia Pittau, Pier Luigi Conti, Roberto Zelli,* "Inference for deprivation profiles in a binary setting"

24/04 *Edoardo Otranto,* "A Vector Multiplicative Error Model with Spillover Effects and Co-movements"

24/03 *Luca Serafini, Emanuela Marrocu, Raffaele Paci,* "Smart Strategies, Smarter Performance: the Impact of S3 and Industry 4.0 on Firms' Outcomes"

24/02 *Hugues Champeaux, Elsa Gautrain, Karine Marazyan,* "Men's premarital migration and marriage payments: Evidence from Indonesia"

24/01 *Pablo Álvarez-Aragón, Hugues Champeaux,* "Measuring Norms and Enumerator Effects: Survey Method Matters"

23/15 *Fabio Cerina, Xavier Raurich,* "Saving Behaviour and the Intergenerational Allocation of Leisure Time"

23/14 *Leonardo Vargiu, Bianca Biagi, Maria Giovanna Brandano, Paolo Postiglione,* "Research Infrastructures and Regional Growth: the case of Europe"

23/13 *Alberto Tidu,* "Dissecting inequality: conceptual problems, trends and drivers"

23/12 *Luciano Mauro, Francesco Pigliaru,* "Italy's National Recovery and Resilient Plan: Will it Narrow the North-South Productivity Gap?"

23/11 *Fabrizio Antolini, Samuele Cesarini, Giorgio Garau,* "The economic impact of the tourism sector on the overall Italian economy: An Input-Output Approach"

23/10 *Giorgio Garau, Andrea Karim El Meligi,* "The Impact of the Pandemic and War on Surplus Redistribution Mechanisms: A Sectoral Analysis of France and Italy"

23/09 *Maria Giovanna Brandano, Alessandra Faggian, Adriana C. Pinate,* "The impact of COVID-19 on the tourism sector in Italy: a regional spatial perspective"

23/08 *Fabrizio Cipollini, Giampiero M. Gallo, Alessandro Palandri,* "Modeling and evaluating conditional quantile dynamics in VaR forecasts"

23/07 *Ugo M. Gragnolati, Luigi Moretti, Roberto Ricciuti,* "Early railways and industrial development: Local evidence from Sardinia in 1871–1911"

23/06 *Giampiero M. Gallo, Demetrio Lacava, Edoardo Otranto* "Volatility jumps and the classification of monetary policy announcements"

23/05 *Diego Dessì, Raffaele Paci,* "The impact of Global Value Chains participation on countries' productivity"

23/04 *Edoardo Otranto, Luca Scaffidi Domianello,* "On the relationship between Markov Switching inference and Fuzzy Clustering: A Monte Carlo evidence"

23/03 *Mario Agostino Maggioni, Emanuela Marrocu, Teodora Erika Uberti, Stefano Usai,* "The role of localised, recombinant and exogenous technological change in European regions"

23/02 *Bianca Biagi, Laura Ciucci, Claudio Detotto, Manuela Pulina,* "University study programmes and students dynamics"

23/01 *Giovanni B. Concu, Claudio Detotto, Marco Vannini,* "Drivers of intentions and drivers of actions: willingness to participate versus actual participation in fire management in Sardinia, Italy"

22/06 *Alberto Tidu, Stefano Usai, Frederick Guy,* Measuring spatial dispersion: an experimental test on the M-index www.crenos.unica.it

ISBN 9788868515188