

“Compensate the Losers?”

Economic policy and partisan realignment in the US

Ilyana Kuziemko, Nicolas Longuet Marx, Suresh Naidu*

September 26, 2023

Abstract

We argue that the Democratic Party’s evolution on economic policy helps explain partisan realignment by education. We show that less-educated Americans differentially demand “predistribution” policies (e.g., a federal jobs guarantee, higher minimum wages, protectionism, and stronger unions), while more-educated Americans differentially favor redistribution (taxes and transfers). This educational gradient in policy preferences is largely unchanged since the 1940s. We then show the Democrats’ *supply* of predistribution has declined since the 1970s. First, Congressional roll-call votes exhibit a decline of Democrat-introduced predistribution legislation. Second, we document the rise of new conservative coalitions in the party particularly hostile to New-Deal-style predistribution (e.g., the Democratic Leadership Coalition, or DLC). Consistent with these developments, educated donors play an increasingly important roll in Democratic (especially DLC) primary campaigns relative to Republican primaries. Democratic candidates hail from Ivy League universities to a significantly greater extent than Republicans, whereas the reverse was true in the immediate postwar decades. We then examine voter reaction to these supply changes. Less-educated Americans begin to leave the Democratic Party in the 1970s, after decades of serving as the party’s base. Roughly half of the total shift can be explained by their changing views of the parties’ economic policies. We also show that in the crucial transition period of the 1970s and 1980s, DLC-aligned candidates draw disproportionately from more-educated voters in both hypothetical presidential match-ups and actual Congressional elections.

Key words: Predistribution, Political Competition, Voting, Neoliberalism.

JEL codes: D72, D78, H1.

*We thank Jessica Fuchs-Shafer, Clément Herman, Madhavi Jha, Donato Onorato, Noah Simon and especially Diva Baristone for excellent research assistance. We thank Stephen Ansolabehere, David Autor, Anne Case, Angus Deaton, Lily Geismer, Amory Gethin, Jacob Hacker, Alex Hertel-Fernandez, Shigeo Hirano, Ethan Kaplan, Thomas Piketty, Robert Shapiro, Eric Schickler, Gavin Wright, as well as seminar participants at Berkeley, Boston University, UBC, Columbia, LSE, University of Maryland, NBER Political Economy, Princeton, and Sciences Po for helpful comments. Kuziemko: kuziemko@princeton.edu; Longuet Marx: nkd2120@columbia.edu; Naidu: sn2430@columbia.edu.

1 Introduction

Over the past several decades, less-educated voters in rich democracies have abandoned center-left parties, their political home for generations (Kitschelt and Rehm, 2019; Gethin *et al.*, 2021).¹ In the US context, many pundits have argued that less-educated voters who leave the pro-redistribution Democrats are voting against their economic self-interest. Scholars have proposed social issues (Lee and Roemer, 2006; Gennaioli and Tabellini, 2019; Enke *et al.*, 2021), misinformation (DellaVigna and Kaplan, 2007; Martin and Yurukoglu, 2017; Cruces *et al.*, 2013), or distrust in government (Kuziemko *et al.*, 2015) to help explain this seeming puzzle. A recent review of this evidence by a labor historian asked: “Is America too Rich for Class Politics?”²

In this paper we focus on the American case and argue that reports of the death of class politics have been exaggerated. Our argument has three main steps. First, we focus on *demand* for economic policies by education. We show that less-educated Americans differentially support *predistribution* (e.g., a federal jobs guarantee, higher minimum wages, pro-union industrial relations, and protectionist trade arrangements) and always have, at least since polling became widely available in the 1940s.³ By contrast, more-educated Americans are stronger supporters of *redistribution* (e.g., higher taxes on the rich), though this gradient is smaller than their relative distaste for predistribution. We thus uncover a stable, long-term educational cleavage over pre-fiscal vs post-fiscal dimensions of egalitarian policies.

Second, we focus on the parties’ *supply* of economic policies and show that since the mid-1970s, the Democrats have moved away from the preferences of less-educated voters. We begin by using Congressional votes as our main source of data. Before the 1970s, predistribution-related topics accounted for nearly twenty percent of House votes in years the Democrats controlled the Speakership, but since then for only ten percent. The redistribution share holds steady. Using natural language processing techniques, we find a similar mid-1970s decline in Democratic platform semantic similarity to pre- versus redistribution survey question wordings.

This shift against predistribution coincides with the rise of a faction within the Demo-

¹Two salient examples from 2016 include less-educated voters breaking with Labour and supporting Brexit in the UK (Hobolt, 2016) and fueling the successful presidential campaign of Donald Trump in the US (Sances, 2019).

²See <https://nymag.com/intelligencer/2021/09/is-america-too-rich-for-class-politics.html>.

³Gallup, from which we draw much of our historical data, begins fielding regular surveys in 1935, but only in the early 1940s does it consistently ask respondents their education level. We provide more information about Gallup and our other surveys sources in Section 3.

cratic party that called themselves the “New Democrats” and would eventually form a more official organization called the Democratic Leadership Council (DLC). We analyze vote patterns to show that this faction was generally more conservative than other Democrats but especially so on predistribution topics, consistent with their stated positions on these issues.

To further our claim that Democratic politicians begin to reflect the preferences of the educated, we provide evidence that educated donors play an increasingly important role in Democratic primary races relative to Republican primaries. Since the mid 1980s, Democratic primary candidates have relied more on educated census tracts for their donations than have Republican candidates, suggesting that educated neighborhoods play a greater role in candidate selection for the Democrats. Importantly, this trend has been driven by out-of-district (or out-of-state in the case of Senate races) donors, so cannot be explained by changes in local voter preferences as these donors cannot even vote for the candidates they support financially. When we look at within-party factions, we find that, as we would expect given their economic policy preferences, educated donors differentially support DLC Democratic primary candidates over other Democrats.

While consistent donations data are only available since 1980, we can take a longer-run view via Congressional biographies. In the decades immediately after World War II, Republican legislators were significantly more likely to hail from Ivy League universities. The reverse is true today, with the inflection point again occurring in the 1970s. In the 1970s Democrats also begin to speak in a manner requiring higher levels of education to understand relative to Republicans.

Finally, we trace out voters’ reaction to this change in the Democratic party’s policies and candidates. First, we clean and harmonize hundreds of surveys ($N \approx 2$ million) to precisely estimate the year-by-year evolution of party identification with respect to education since the early 1940s. In the 1940s, every additional year of education predicts a three percentage-point decrease in the likelihood of identifying as a Democrat. This relationship holds with little change until an inflection point, which we estimate as occurring in 1976. Since then, the pace of realignment remains relatively steady: by 2000 the gradient reaches zero and today it is essentially reversed from its 1940s-1960s baseline, with each additional year of education predicting you are three percentage points more likely to identify as a Democrat. While this evidence does not directly tie realignment to economic policy, it is reassuring that realignment begins in the mid-1970s, exactly when we identify shifts in economic policy supply.

We then show that roughly half of this shift in the gradient can be explained by voters' assessment of the parties' economic policies. Since the 1940s, Gallup has asked which party is better to keep the country prosperous. The less-educated were consistently more likely to answer "Democrats" than were more-educated respondents in the immediate decades following World War II. Today, the less-educated are more likely to name the Republicans as the party better for prosperity (a result that is robust to many controls, including own party ID). Moreover, half of the total educational realignment is erased when we control for respondents' answer to this question.

While the above analysis takes an eighty-year view of the realignment, we then focus on the pivotal years of the 1970s through early 1990s. Most of the Democratic presidential primaries in these years came down to a "new" versus an "old" Democrat (e.g., in 1984 "new" Gary Hart versus "old" Walter Mondale) as the top two candidates in terms of total delegates. We use survey questions that ask respondents for whom they would vote in hypothetical presidential elections (e.g., again using 1984, we compare answers to a hypothetical Hart-versus-Reagan election and a Mondale-versus-Reagan election). More-educated voters differentially vote for the Democrat in these hypothetical match-ups when he is a "New Democrat" as opposed to an old-style Democrat.

We then move to actual 1980s House election results and thus an ecological approach (as we cannot observe actual votes of individuals). Using returns data disaggregated into granular geographic units (about sixty units per Congressional District) we again show that more-educated areas differentially vote for the Democrat when he is a DLC Democrat.

In summary, we argue that changes in the Democratic party's economic policy played a key role in partisan realignment by education. Voters' economic preferences by education have in fact changed very little since the 1940s—today as then, less-educated voters appear to prefer a less market-based and more interventionist economic program that aims to promote domestic employment and wages. Beginning in 1970s, influential factions within the Democratic party began to move away from those policies, and we show it is indeed more-educated voters who disproportionately find these candidates attractive. We show this timing is consistent with party reforms that reduced the within-party financial influence of labor unions, a traditionally low-education constituency within the Democratic party. By the 1990s, DLC-aligned politicians and advisors controlled the Democratic agenda: a former president of the DLC (Bill Clinton) was in the White House most of the decade and by the late 1990s near-majorities of Democratic legislators joined DLC-aligned Congressional cau-

cuses. A final point to emphasize here is that DLC or “New” Democrats are significantly more socially conservative than other Democrats, so the differential support they enjoy among the educated (who we show have always been more socially liberal) is in *spite* of social issues, thus suggesting a particularly strong role for economic policy in explaining realignment.

Our paper is most directly related to the recent work on the changing relationship between education and partisan identity (what Thomas Piketty has termed the “Brahmification of the left”). As noted earlier in the introduction, economists and political scientists have explored many hypotheses, though none to our knowledge focuses exclusively on how well a purely economic-policy-based mechanism performs. Historians, however, have provided more qualitative evidence for this hypothesis (we benefited in particular from reading Stein, 2010 and Geismer, 2022).⁴ The qualitative historical work often identifies the 1970s as a key moment of transition (which we also find in our quantitative analysis). In this paper, we do *not* seek to answer (a) *why* Democratic legislators moved away from their traditional support of predistribution in the 1970s or (b) whether this decision was electorally optimal. In the Conclusion we briefly return to these questions and believe them to be natural questions for future work.

We also join a large literature in economics and political science that attempts to measure voters’ preferences, in particular on economic issues. For the most part, past papers group together pre- and re-distribution in a single “economic preferences” index. Other papers focus on redistribution explicitly.⁵ We find that breaking up economic policies into pre- and re-distribution in fact creates a stable educational cleavage over the past eight decades.

We are also related to a smaller economics and political science literature on predistribution versus redistribution. Hacker (2011) defines predistribution as a “focus on market reforms that encourage a more equal distribution of economic power and rewards even before government collects taxes or pays out benefits.” Bozio *et al.* (2020) and Blanchet *et al.* (2020) argue that differences in predistribution better explain differences in inequality across countries than do differences in redistribution.

⁴While to the best of our knowledge economists have not examined the political demand for predistribution more generally, there are several papers on the political effects of trade and immigration. See Choi *et al.* (2021) and Dorn *et al.* (2020) on trade and Mayda *et al.* (2022) on immigration.

⁵For examples of papers that aggregate preferences into a single economic index, see Ansolabehere *et al.* (2006) and Enke *et al.* (2021). The first component of the widely-used DW-nominate is another measure often used to capture the economic ideology of legislators. For papers focusing on preferences for *redistribution*, see Kuziemko *et al.* (2015), Ashok *et al.* (2015), and Alesina *et al.* (2018). Recent and important exceptions are Stantcheva (2022) and Dechezleprêtre *et al.* (2022) examining how voters form views on policies not directly related to taxes and transfers, such as trade and climate change.

The exceptions noted above notwithstanding, predistribution has fallen out of favor among economists in recent decades. A distinguished theoretical literature in public finance (Diamond and Mirrlees, 1971) and law and economics (Kaplow and Shavell, 1994) has argued that *ex-post* redistribution via taxes and transfers best implements egalitarian goals. This theoretical argument has been brought into public policy recommendations by both conservative and liberal economists, especially in the US. For example, Greg Mankiw writes that “policymakers do not have the tools to exert such a strong influence over pretax earnings, even if they wanted to do so.” Similarly, Brad DeLong writes: “I can’t see the mechanism by which changes in government policies bring about such huge swings in pre-tax income distribution.”⁶ Instead of the *normative* question of whether predistribution is an efficacious and desirable policy tool, in this paper we focus on the *positive* questions of which voters support predistribution and whether the changes in partisan positions on predistribution can explain observed shifts in partisan identity.

Methodologically, we develop new approaches for comparing political opinions and policy positions over time. For example, the changing selectivity of educational categories complicates relating education to opinions over long periods. Our survey data are in fact quite detailed (typically five or more educational categories) and we then use Census data to convert intervals into year of schooling based on additional demographic information from the surveys, creating a common metric over our eighty-year analysis period. Another challenge is the sensitivity of survey responses to even small changes in wording. As we are working with survey questions from diverse sources over eight decades, we use text embeddings to confirm that our results are robust to wording changes (i.e., controlling flexibly for principal components of the embedding vectors), and we show questions on the same issue are semantically similar despite being syntactically different. We also use survey question embeddings to measure the distance between national party platforms and the specific wordings of opinion questions, and show that the similarity between Democratic platforms and the predistribution policy questions decrease after the 1970s.

One point to emphasize is that predistribution policies fit awkwardly with work-horse models in political economy, which typically have voters maximize preferences over leisure and consumption (which in most static models is given by after-tax-and-transfer earnings). Under these assumptions, low-wage voters would never prefer predistribution over a standard

⁶See https://scholar.harvard.edu/mankiw/files/inequality_final.pdf for the Mankiw quote and https://delong.typepad.com/sdj/2006/08/the_primacy_of_.html for the DeLong quote. Mankiw chaired the CEA under Republican George W. Bush while DeLong worked at the Treasury under Bill Clinton.

progressive tax-and-transfer scheme, a result at odds with our empirical findings. For this reason, we sketch a simple model in Section 2, formalized in the Appendix, that specifies differences between pre- and re-distribution and how they relate to classical models in political economy.

After laying out this theoretical groundwork, Section 3 then briefly introduces our data sources, focusing on those that are novel. In Section 4, we document the strong, robust support among less-educated voters for “predistribution” policies and contrast this pattern with that between education and support for redistribution. Section 5 provides some history on the rise of anti-predistribution factions (e.g., the DLC) in the Democratic party and then, relying largely on roll-call vote data, documents the declining supply of predistributionist policies supplied by Democrats, a trend that first appears in the 1970s and is driven by DLC legislators. Section 6 focuses on candidate selection and biographies, finding that Democrats rely increasingly on educated donors and candidates, relative to Republicans. Section 7 documents the reaction of voters to these supply shifts, and again the 1970s emerges as a key inflection point. Section 8 briefly considers alternative stories, though we mostly leave this work to future authors. Section 9 concludes.

2 Predistribution vs Redistribution in Electoral Competition

Standard models of distributive politics assume that voters only care about consumption (in static models, equivalent to after-tax income) and leisure. If voters maximize utility based exclusively on consumption and leisure, then demand for predistribution should be low. Under such assumptions, a coal-miner, say, choosing between industrial policy that would allow him to keep working or a monthly check that would cover his lost salary but not provide future work would always choose the latter. Economists often suggest that efficiency-enhancing policies such as free trade can be made Pareto efficient by “compensating the losers” and under a standard model with sufficiently large surplus majorities would support such a policy.

But scholars have also raised plausible reasons why individuals may care about their *pre-tax-and-transfer* income independently of their post-tax income, and thus why they might favor predistribution-type policies over redistribution. First, voters might be making relative comparisons with their social reference group (Killian *et al.*, 2008), and the metric of the comparison could be pre-tax income (perhaps proxying for talent, human capital, or contribution to society). Second, voters might have direct preferences over their beliefs that they

are productive, as in Bénabou and Tirole (2016), and pre-tax income might be a signal of that productivity. Obscuring that signal with predistribution policy could be preferred by voters with strong priors that they are low productivity (whereas redistribution in modern tax systems tends to be rank-preserving). Third, voters may believe that the tax and transfer system is more opaque, corrupt, or inefficient than more transparent policy interventions (Kuziemko *et al.*, 2015). These various theories suggest heterogeneity in support for pre- versus re-distribution by socioeconomic status, productivity or trust in government, all of which are plausibly proxied by education.

If voters have preferences over their *pre-tax* income, then electoral competition over tax rates alone will miss dimensions of economic policy that are salient to voters. Indeed, while influential, the Meltzer-Richards framework has not met with overwhelming empirical success. For example, the model predicts that demand for redistribution should increase with pre-tax inequality, a result which is remarkably difficult to find in the data (see Bonica *et al.*, 2013, Acemoglu *et al.*, 2015 and Ashok *et al.*, 2015).

A key claim of our paper is that economic policy is multi-dimensional and thus cannot be captured by a single tax rate (as in Meltzer-Richard) or a single index. But issues of equilibrium existence and uniqueness often arise in standard political competition models with multi-dimensional policies, often requiring specific assumptions about timing to resolve. We further document the presence and influence of *factions* within parties, and show empirically that changing faction influence changes a party’s policy priorities. Our model captures both of these phenomena, building on ideas in Roemer (1998) and Besley (2007) that model political competition along two dimensions with intra-party bargaining, but unlike these papers (which typically assume an “economic” and a “social,” often racial equality in the U.S. context, dimension), we break economic policy into “predistribution” and “redistribution” dimensions. Appendix C provides a formal model, but here we highlight the key assumptions and predictions.

In particular, *we assume that voters have preferences over redistribution τ and predistribution θ and that more educated voters differentially oppose predistribution relative to redistribution* (i.e. either monotonic single-crossing preferences, or single-peaked preferences with ideal points ordered by education). We provide extensive evidence for this assumption in Section 4.

Following (Besley, 2007), we suppose that parties compete on providing a given level of utility to a swing voter, but then within each party, factions bargain, with the more

educated faction having bargaining power λ^j , $j \in \{D, R\}$ over the specific policy mix that would implement the swing voter’s utility. Parties commit to providing a level of swing-voter utility, but cannot commit to policies.

The model shows that demand matters via either party factions or swing voters’ changing preferences over predistribution vs redistribution, both of which we rule out empirically below. It also suggests that supply of policies by a party is partly determined by the relative bargaining power of factions. We argue that in the 1970s, λ^D increases. As we discuss later in the paper, changes in federal campaign finance laws as well as Democratic-specific party reforms both substantially increased the relative power of the educated within the Democratic party (for example by diminishing the influence of labor unions which had previously had substantial power in the Democratic party to represent the political interests of less-educated voters). The model thus predicts : (a) a decline in θ relative to τ in the Democratic agenda; (b) a relative movement of less-educated voters away from the Democrats. Both of these changes should be observed starting in the 1970s as well.

Finally the model predicts that, owing to the swing voter constraint, changes in policy supply result in changes in the composition of voter support, without changing the overall vote share of the party. We show that there was a sharp break in Democratic partisan affiliation by education beginning in the 1970s, without a simultaneous change in overall Democratic support. All of these results would be difficult to generate in a model of sharp political competition with a single policy dimension or without party factions.

3 Data and methodology

3.1 Data

This section describes the survey data used in this paper (in particular in Sections 4 and 7), which we believe in many cases we are the first to use. We only briefly describe the other data used in the paper, which are less novel, when we actually analyze them. We provide greater detail on all data sources in the Appendix B.

3.1.1 Survey sources

While we include standard data sources such as American National Election Study (ANES), the General Social Survey (GSS), and the Cooperative Congressional Election Study (CCES), we also make heavy use of less familiar historical data from survey corporations, for the most

part housed by iPoll at Cornell. The majority of these data come from Gallup, which beginning in 1942 asked respondents both their educational attainment and their self-reported partisan identification (Gallup surveys begin in 1935 and since then have always asked age, race, sex and state of residence). Many datasets require manual construction of dictionary files from codebooks to convert raw data files into usable datasets and we contracted with iPoll to perform these conversions. Altogether, we have more than 2 million observations that include education and partisan identity, from 1942 until 2020. While education and partisan identity is asked in essentially all Gallup surveys from 1942 onward, Gallup would often also ask views on specific policy questions (e.g., the minimum wage, tariffs, taxes), which we also collect.

To be consistent across datasets, we drop any respondent below the age of 21. We also end most of our analysis in 2015 (or earlier, depending on the sample period of various datasets) to avoid any of our results being driven by the Trump election, though in most cases we find that voting patterns during the Trump era was a continuation of past trends that date to the 1970s. Otherwise, we impose no sample restrictions. We provide summary statistics by data source and time period in Appendix Table G.7.

3.1.2 Key survey questions

We include all survey questions on pre- or redistribution topics so long as they are comparably worded across time.⁷ The data appendix provides the exact wording of each question so readers can judge comparability themselves. We provide illustrative examples below.

Predistribution questions

—**Minimum wage.** We include questions that ask whether the respondent approves increasing the minimum wage. Typically, a new, higher level is proposed. An example from Gallup in 2013: “Would you vote for a law that would raise the federal minimum wage to nine dollars an hour?”

—**Government job guarantee.** We include questions that ask respondents whether the government has the responsibility to provide a job for anyone who wants to work. An example from *Time* magazine in 1976: “Do you favor or oppose the passage of a full employment bill in which the government guarantees a job to everyone who wants to work?”

⁷For example, we would *not* include a question about the “Bush income tax cuts” because it is not directly comparable to other policies across time. Instead, we would include questions about income tax cuts in general.

—**Support for unions.** We rely largely on Gallup for this topic area, which since the 1940s has asked individuals whether unions should have more, less or the same amount of influence than they currently enjoy.

—**Trade policy.** In earlier years, Gallup typically asked about tariff levels. For example, in a survey from 1953 they ask: “By and large, do you favor higher or lower tariffs than we have at present?” In more modern datasets, the term “tariffs” is not often used, and instead respondents are asked whether there should be more or fewer “limits” on imports. For example, since 1986 ANES has asked a survey question of the form: “Some people have suggested placing new limits on imports in order to protect American jobs. Others say that such limits would raise consumer prices and hurt American exports. Do you favor placing new limits on imports, or not?”

Redistribution questions

—**Tax the Rich More.** We begin with Gallup questions going back to the 1940s on whether there should be limits so that even the very rich do not pay more than half their income in taxes. We then add GSS questions that ask whether the respondent considers that the amount high income earners are paying in taxes is too low.

—**Views on own taxes.** Since the 1950s, Gallup has asked respondents whether they view their own federal income taxes as fair, too high or too low. We also rely on a very similarly worded question from the GSS.

—**Prioritizing tax cuts.** A variety of pollsters have asked respondents how strongly they prioritize tax cuts, even if the government would have to put off other goals.

—**Transfers to the poor.** An example comes from the GSS: “Are we spending too much, too little, or about the right amount on welfare.”

3.2 Estimating educational gradients over long periods

One challenge faced by any long-run analysis involving educational attainment is that a given educational category can represent very different levels of selectivity at different points in time. As just one example, in 1940 only one-fourth of U.S. adults over age 25 had completed twelfth grade, whereas today nearly ninety percent have.

Our goal is to translate our education variable into the same units—years of schooling—across all of our hundreds of datasets. Most of our datasets include detailed *categories* of education, typically five to seven categories. The categories naturally shift over time, so a

1940s-era Gallup survey will often provide “less than fourth grade” as a category, whereas such a category would not exist today. We then use the Census (which records education in terms of *years of schooling*) to estimate years of schooling conditional on self-reporting a given educational category. Our preferred method (though we show robustness to alternatives) uses sex, race, year of observation (interpolated between Census years), and birth cohort in ten-year bins to predict years of schooling conditional on a given category of education. In this manner, we can combine hundreds of datasets with slightly different educational categories via a common metric.

We then estimate the relationship between our main outcome variables—e.g., economic-policy preferences, partisan self-identification—and this *AdjYearsEduc* variable, separately by time period p (where p is often a single year when data permit or a four- or five-year period when data are more sparse). That is, for each time period p , we estimate:

$$y_i = \beta^p \text{AdjYearsEduc}_i + f(a_i) + \mu_{s(i)} + e_i, \quad (1)$$

where y_i is the outcome of interest, $f(a_i)$ are flexible controls for age (typically age fixed effects in five-year bins), and $\mu_{s(i)}$ are survey fixed effects (e.g., if in $p = 1947$ we had three surveys for a given analysis sample, say a Gallup survey in May and December and a Roper survey in January, each would get its own fixed effect), which therefore subsume date fixed effects. In general, we view the unconditional covariance between education and preferences as our target parameter, with the exception of age, which we treat as a nuisance variable and thus try to absorb. Many of our main results plot the β^p values over time periods p to display long-run trends.

4 Demand for economic policies by education

We now begin to present evidence related to the model in Section 2. This section empirically validates a key assumption of the model: that higher levels of education differentially reduce support for predistribution compared to redistribution.

4.1 Educational gradient of support for economic policies

Figure 1 displays the coefficients from equation (1), separately for each of our eight topic areas. To better compare the various outcomes in one graph, we standardize all outcomes to have a mean of zero and a standard deviation of one. If needed, a question is re-oriented

to be increasing in the left-wing (i.e., pro-predistribution, pro-redistribution) answer. To avoid clutter we suppress confidence intervals but Appendix Figure A.1 shows each outcome separately with confidence intervals.

4.1.1 Predistribution

Figure 1 shows a strong, negative association between support for predistribution policies and years of education (predistribution coefficients are depicted with solid markers, connected with solid lines).

In the first series, we show the educational gradient in support for a government job guarantee. Of all of our economic policy questions, this is perhaps the most striking result in that the magnitude is both large and relatively steady over our eight-decade sample period. From the 1940s until today, an additional year of education reduces support by between 0.05 and 0.1 standard deviations. As we show in Appendix Figure A.1, in each period this relationship is highly statistically significant.

The second series in Figure 1 shows a consistent, negative association between years of education and support for the minimum wage. Since the 1970s, the educational gradient has been relatively steady: each year of additional education reduces support for the minimum wage by roughly 0.04 standard deviations. The educational gradient was somewhat larger in magnitude in the 1940s and 1950s. Again, each coefficient is statistically significant.

A similarly negative and robust pattern of coefficients obtains when support for greater union influence serves as the outcome variable (third series).

The final predistribution series documents the relationship between education and support for protectionist trade policies. Of all of our predistribution outcomes, support for protectionism has the least stable relationship with education over time (though like our other outcomes is on average negative over the long sample period). In the 1940s and 1950s, there is little educational gradient in views toward trade. In general, trade policy was not politically salient in the U.S. during this period—the country had few industrial rivals in the immediate post-war decades and a Cold-War, bi-partisan consensus held that the US should provide favorable terms of trade to countries at risk of falling to the communists (Stein, 2010). But beginning in the 1970s, a strong, negative educational gradient emerges and by the 1990s it is the predistribution policy that has the largest educational gradient in magnitude.

4.1.2 Redistribution

Whereas the educational gradient for predistribution policies is large and negative, the gradient for redistribution is generally positive and in all cases smaller (redistribution coefficients are depicted with hollow markers and connected by dashed lines).

Both support for more progressive taxation and views that one's own taxes are not unfairly high are generally correlated with higher levels of education, though this relationship is not very large in magnitude. There is a somewhat stronger, positive gradient for willingness to delay tax cuts for other priorities, but, again, the magnitude of this coefficient is smaller than for the predistribution questions.

On the spending side, we examine the educational gradient on support for transfers to the poor. Unlike our three tax questions, the pro-redistribution response to this question is negatively associated with education. But the magnitude of the coefficients is always smaller than the coefficients on the predistribution questions, and there is a slight trend towards zero.

In sum, the questions on redistributive policy generally show an opposite-signed and much more attenuated correlation with education compared to the questions on predistribution, suggesting that important information may be lost by aggregating all economic policies into a single index of egalitarianism, without disaggregating into pre- versus re-distribution.

4.2 Overall trends in support for pre- and redistribution

Given that educational attainment has increased over time, one possibility is that overall support for predistribution versus redistribution has declined even if the educational gradient is stable. Moreover, stepping outside the model, support for predistribution across dimensions unrelated to education might be declining and thus the Democrats might be moving away from a predistributionist agenda merely as a result of its diminishing popularity.

In Appendix Table A.1, we show that the overall support to predistribution relative to redistribution is stable across time (if anything, predistribution is becoming more popular over time, relative to redistribution, though this trend is not statistically significant). Of course this exercise involves pooling our four pre- and re-distribution questions into two (pre- and re-distribution) categories, and we thus show this result is stable after flexibly controlling for question type.

4.3 Robustness to question wording

Researchers have long cautioned that even small differences in question wording can affect the share of respondents who agree with a statement (Schuman and Presser, 1996). To make this concern concrete, suppose support for the minimum wage is in fact *decreasing* over time, but we might fail to identify this trend because simultaneously the wording of minimum-wage questions changed so as to elicit more positive responses.

Appendix G addresses this concern using recent advances in natural language processing (NLP). We calculate embedding vectors (Devlin *et al.*, 2018) and dictionary-based text measures for each survey question. We consider as ground truth that our 95 survey questions on economic policy can be meaningfully separated into eight categories (minimum wage, trade policy, etc.). We show that eight hierarchical clusters induced by the embeddings reproduce this grouping as well, even though the exact wording of two questions about the same policy might use very different words. In a more demanding test, we show that the embeddings in fact can separate predistribution and redistribution question texts.

Having shown that the numeric text features can distinguish meaningful differences *across* topic areas, we then use them to identify meaningful differences in wording *within* topic areas. Appendix Figure A.6 reproduces our main Figure 1 after flexibly controlling for the eight hierarchical clusters and five top principal components of the embedding vectors, as well as the dictionary-based measures. The result is largely unchanged. We also show that the stability over time of overall support for predistribution relative to redistribution, which we introduced in the previous subsection, is robust to controlling flexibly for text features. While we are somewhat hesitant to compare aggregate levels of support for redistribution and predistribution (is the minimum wage more popular than tax hikes?), when we perform this exercise after controlling for text features (in case tax questions are asked in a more negative manner than minimum wage questions, e.g.) we find that predistribution is more popular.⁸

⁸We in fact find that the within-topic-area differences in wordings are only weakly related to overall support for a given topic area. This result is consistent with historical accounts that polling companies such as Gallup were sensitive since at least the 1950s to the possibility of biased question wording. They would often randomly assign respondents within a survey to small differences in wording to test for bias.

4.4 Other robustness and related results

We show in the appendix that the educational gradients depicted in Figure 1 are robust to various modifications: controlling for income, gender and age (Appendix Figure A.4, though we lose observations as income is not always included in our surveys, especially for earlier years); and re-weighting observations so as to hold fixed *race* \times *education* composition to its late-1940s distribution (Appendix Figure A.5).

In the interest of space, Appendix Table A.2 performs a number of robustness exercises and sub-sample analysis using more concise, parametric approach: pooling all questions and regressing the response on a *Years educ.* \times *Predis* interaction (along with all lower-order terms). We document a precise -5 to -7 percentage-point effect of an additional year of education on support for predistribution relative to redistribution (which one can roughly eye-ball from Figure 1). We further show that there is no significant trend in the overall educational difference in support for predistribution vs redistribution (again, this result can be gleaned from the figure). This result is robust to controlling for survey question embedding variables interacted with age and year fixed effects, as well as a variety of other demographic controls interacted with question fixed effects. As much of the realignment discourse focuses on whites, we show that the pre-versus-redistributional educational gradient holds for a whites-only sample (in fact, it is slightly more negative and the trend over time even closer to zero).

Finally, we also examine views toward pre- and re-distribution *by race*, setting aside for the moment differences by education. Appendix Figure A.3 shows similar patterns when an indicator variables for white is used instead of years of education to capture socioeconomic advantage (with the exception of trade, which does not have a marked gradient by self-reported race). As such, predistribution policies appear differentially supported by non-whites and less-educated whites.⁹

In this section we have shown that grouping all economic policy into a single index obscures important heterogeneity in the relationship between economic preferences and educational attainment. In particular, we have shown that education has, at least since the 1940s, predicted substantially less support for policies aimed at reducing differences in pre-tax-and-transfer income (“predistribution”) but, if anything, has predicted somewhat greater

⁹It is interesting to note that the organizers of the famed 1963 March on Washington for Jobs and Freedom listed many predistributionist policies in their official list of goals (including a federal jobs guarantee and an increased minimum wage).

support for tax and transfer policies aimed at reducing differences in disposable income (“re-distribution”). Of course, past work has shown correlations between education and certain specific policy questions (e.g., Choi *et al.*, 2021 shows a strong, negative relationship between education and support for NAFTA in particular and free trade more generally), but we believe our analysis in this section is novel in terms of documenting these preferences over several policy domains and over the last eight decades.

5 Democrats’ changing supply of economic policies

Historians and political scientists have argued that Democrats who entered the political scene in the 1970s were distinct from their predecessors, especially on economic issues. “New Democrats, often from suburban, affluent districts, made it a badge of honor that they were not New Dealers (Stein, 2010).” Many entered Congress in the first-post-Watergate election of 1974 (and were known as the “Watergate Babies”) and had particular interest in the technology sector and other high-skill industries (and gained the nickname “Atari Democrats”). “The freshman Democrat today is likely to be an upper-income type and that causes some problems with economic issues” noted the AFL-CIO counsel at the time.¹⁰ “[New Democrats’] efforts were backed by connections to finance and corporations rather than organized labor.” (Mudge, 2018, pp. 291).

In this section we attempt to quantify some of the arguments made more qualitatively by previous authors. We begin with analysis that examines the policies of the party as a whole and then move toward identifying the rise of within-party factions who favored more conservative positions (especially on predistribution issues).

5.1 Decline of predistribution in the Democratic agenda

5.1.1 Data

We make heavy use of the Comparative Agendas Project (CAP) dataset. CAP groups all Congressional votes since 1947 into policy-related categories and sub-categories. We take these categories as given and then place them into pre- and re-distribution groups.¹¹ We

¹⁰As quoted in Stein (2010).

¹¹See <https://www.comparativeagendas.net/> for the CAP website, which includes the dataset we use in this section as well as many others. All CAP categories and sub-categories, including their description and examples, can be found here: https://comparativeagendas.s3.amazonaws.com/codebookfiles/Codebook_PAP_2019.pdf.

provide the exact groupings in Appendix D and only briefly summarize here.

In general, we define a category of bills as “predistribution” if it relates to labor market regulation (e.g., minimum wage, working conditions, union organizing), industrial policy and public works, and trade. We err on the side of choosing topics that match the survey questions in Section 4. For redistribution, we include tax and budget topics and transfer programs.

We can also use an alternative bill classification dataset from Bateman *et al.* (2018). In the interest of space we show in the main part of the paper results using the CAP dataset and replicate the analysis in Appendix D using the Bateman et al. classification data.

5.1.2 Decline of predistribution bills brought to a vote

As a first step towards documenting the changing supply of Democratic policies, we look at the composition of roll-call votes in the House. We focus on the House of Representatives because Democrats were in near-constant control of this chamber in the post-war period until 1995 (and then again in the mid 2000s), allowing us a long time-series to study any changes in the composition of roll-call votes while Democrats controlled the chamber. Throughout the analysis in this subsection, we make the assumption that the party in control of the chamber of Congress controls the supply of roll-call votes in that chamber.¹²

Figure 3 shows the share of pre- and re-distribution roll-call votes every year that the Democrats are in control of the House from 1947 until 2009. The (infrequent) gaps in the series occur when Republicans gain control of the chamber.

The figure shows that while predistribution figured prominently in the Democrats House agenda in the decades following World War II, the share of roll-call votes in predistribution-related policy categories begins to decline in the late 1970s. Following historians that date Jimmy Carter’s presidency as marking the start of the “New Democrat” era, we show that the share of predistribution votes declines significantly before and after 1977 in both datasets.

The pattern for the share of votes concerning redistribution is more stable over time, at roughly five to ten percent a year (with a statistically significant but economically small increase since 1977).

¹²Cox and McCubbins (2005) argue that the structure of House rules strongly favored the majority party, since Thomas Reed systematized House procedures in 1894, writing that “the rules of the legislative game have been heavily stacked in the majority party’s favor since the re-adoption of Reed’s rules in 1894.”. The Republicans explicitly formalized this practice with the so-called “Hastert Rule,” and since Dennis Hastert’s speakership in 1999 do not bring to the floor of the House any measure that a majority of Republican House members do not support. Democrats do not publicly follow such a rule but the vast majority of roll-call votes brought by Democratic Speakers enjoy the support of the majority of Democratic House members.

To bolster our claim that Democrats moved away from predistribution policies at some point in the 1970s, we use both word counts and text embeddings to analyze the language of party platforms in Appendix G.1. In contrast to bills, the platforms allow us to examine the parties’ public-facing language when unconstrained by parliamentary procedure. Appendix G.1 shows that Democratic party platforms, at both the national and state level, become more semantically similar to predistribution over redistribution after the 1970s.

5.2 The “New Democrats” and the decline of predistribution

There is no official list of “New Democrats” so to facilitate quantitative hypothesis testing, we will often focus on the largest and most influential organization associated with the movement, the Democratic Leadership Council (DLC).

5.2.1 The rise of the DLC

The DLC officially launches in 1985. We have been able to obtain official membership lists for 1985, 1986, 1991, 1993 and then 1997 onward. We define a member of Congress as “DLC” for their entire career if they ever appear on any of these lists.¹³ Note that the gaps in our membership lists mean we will have some classification errors in the late 1980s and early 1990s.

The DLC was explicitly skeptical of predistribution policies. In fact, in his memoirs, DLC founder Al From specifically criticizes all four of the predistribution policies we examined in Section 4. As just one example, he emphasizes his group’s break with the “old” Democrats’ tolerance of protectionism: “Our [the DLC’s] pro-trade stance clearly reinforced our message that we were different from the old Democrats.”¹⁴

Figure 2 documents the evolution of the DLC as a share of Democratic House members. The DLC enjoyed robust growth since its founding in 1985, notwithstanding a short-term dip in the late 1980s and early 1990s (which we attribute to our lists being incomplete during

¹³Please see Appendix E for greater detail on our definition of a DLC member and our methodology for acquiring membership lists. While the DLC officially closed in 2011, in 1997 both House and Senate Democrats associated with the DLC formed “New Democrat” caucuses, which existed until the end of our sample period. We thus call anyone who is on an official DLC membership list or a member of the New Democrat Caucus a DLC member.

¹⁴On the minimum wage, he writes: “The bottom line was that the country and economy had changed, and the minimum wage, so important in the New Deal, had ceased to be an effective way to help the working poor.” From often describes unions as a major obstacle to passing DLC-favored policies: “But we needed to change our policy, and I wasn’t about to give the unions or any other constituency group a sign-off on what we recommended (p. 39).” The DLC was also opposed to public jobs programs and in fact called for the elimination of over 250,000 federal government jobs.

this period). By the 1990s and early 2000s, almost half of Democrats in the House were DLC members. We show a similar evolution in the Senate (Appendix Figure A.7) and note as well that official DLC lists included governors and local politicians.

When did the types of politicians who would become DLC members first appear in Congress? While the DLC did not officially exist until 1985, we can see the share of Representatives who *would later become* DLC members before this date, and we track the “eventual DLC” share going back to the 1960s in Figure 2. We see the DLCers begin to enter Congress in greater numbers in the 1970s, consistent with historical accounts. Of course, the large entry in the 1970s may simply be an artifact of our definition of DLC requiring membership in Congress in 1985 (or later). In the second series, we compare the DLC share among all representatives who serve in 1985 or later. We again find that future DLCers markedly increase their share of all Democratic House members in the 1970s.

While the next subsection ties the DLC to the decline of predistribution more directly, it is reassuring to see that DLC politicians differentially come to power just as Democrats are moving away from predistribution.

5.2.2 DLC votes on predistribution

We examine voting patterns of the DLC using two complimentary approaches.

Issue-specific ideal points. A large literature in political science is devoted to inferring legislators ideology by estimating “ideal points” (McCarty *et al.*, 2006) based on roll-call voting patterns.¹⁵ Traditionally, these models have been estimated on all votes a legislator takes, but more recently some papers have produced “issue-specific ideal points,” which we use below. As we are simply applying the methodology of past authors to the subset of votes related to predistribution and other topic areas, we summarize the ideal-point estimation procedure in Appendix F but do not elaborate here.

In Table 1 we present results from a regression testing whether the DLC is more conservative relative to other Democrats and in particular on votes related to predistribution. In col. (1) we present results when the familiar DW-Nominate ideal point is regressed on the DLC dummy in a bivariate regression. We use the “first dimension” of the DW-Nominate ideal point, which proxies general ideology (increasing, as are all our ideal points, in conservativeness). By this general measure, DLC members are significantly more conservative than other Democrats.

¹⁵Recent papers have extended the basic logic to other observable behaviors such as speech and fundraising sources.

The rest of the table shows that the DLC are particularly more conservative on economic policy and even more for policies involving predistribution. Col. (2) shows that the DLC are significantly more conservative than other Democrats on predistribution votes, with this gap equal to one-fifth the Democratic-Republican gap (reported in the rows at the bottom of the table). Note that in some years we are missing DLC membership lists, so the coefficient on DLC is likely biased toward zero due to this measurement errors, but even so it appears to be economically meaningful. The next column shows that this gap is robust to adding *State* \times *Congressional term* fixed effects, which isolates comparisons to DLC versus non-DLC members in the same state and time period. Col. (4) shows that the DLC is more conservative on redistribution votes as well (and indeed on social issues, too, in col. 5, a result we return to in Section 8).

The final column shows that while the DLC is in general more conservative than are other Democrats, the opposition against left-wing economic policy and in particular predistribution still stands out. In this regression, each *Congressperson* \times *Congressional term* has three observations: an ideal point for predistribution, an ideal point for redistribution, and the generic DW-nominate ideal point (the omitted group in the regression). The positive coefficient on the DLC dummy variable reflects the overall conservativeness of the DLC captured in col. (1). The positive and significant interactions show that the DLC is particularly conservative on economic policies (whether pre- or re-distribution) in a manner that can be separated from their generic conservativeness. And in the rows beneath the column we show that the larger (i.e., more conservative) coefficient on predistribution than redistribution is indeed statistically significant. In sum, the DLC is more conservative than other Democrats, especially on economic policy, and especially on *predistribution* topics within economic policy.

Concordance with Republicans. One virtue of studying within-party factions (in our case, the DLC versus other Democrats) is that differences with the opposition party (in our case, the Republicans) can be used to measure ideology in a transparent manner, without parametrically constructing ideal points. We estimate the following baseline regression to assess whether the DLC tends to vote more often with Republicans than do other Democrats:

$$Pr(Yea)_{ir} = \beta DLC_i \times Mean\ Repub_r + \eta_r + \gamma X_{ir} + e_{ir}, \quad (2)$$

where the outcome is Representative i voting “yea” on roll-call bill r ; DLC_i is a dummy for whether i belongs to the DLC; $Mean\ Repub_r$ is the share of GOP Representatives who voted for bill r ; η_r are roll-call-vote fixed effects (which subsume year or Congressional session

fixed effects); X_{ir} are controls (which always include the lower-order terms of the interaction or fixed effects that subsume them) that we vary to probe robustness, and e_{ir} is the error term. We estimate this regression only for Democrats, so the DLC dummy represents the difference between DLC Democrats and other Democrats. We include the vector η_r of roll-call-vote fixed effects (so that each vote receives its own dummy) given the evidence in Figure 3 that the composition of votes has changed across time.

Col. (1) of Table 2 estimates equation (2). Relative to other Democrats, when the share of Republicans voting for a bill increases by ten percentage points, the probability a DLCer votes for a bill increases by 1.48 percentage points. This result is the rough analogue to the first column of Table 1 which uses the DW-nominate ideal point as the outcome.

Our main hypothesis is that DLC-affiliated representatives are differentially more conservative on predistribution issues. Using the same categorization of bills as in the previous section, we augment specification (2) with a triple-interaction term, which allows us to separate the DLC's conservatism on predistribution from votes on other issues:

$$\begin{aligned} Pr(Yea)_{ir} = & \beta^1 Predist_r \times DLC_i \times Mean\ Repub_r + \beta^2 DLC_i \times Mean\ Repub_r \\ & + \beta^3 DLC_i \times Predist_r + \eta_r + \gamma X_{ir} + e_{ir}. \end{aligned} \quad (3)$$

Col. (2) estimates specification (3). In this specification, the coefficient on $DLC \times Predist$ captures the likelihood a DLCer (relative to other Democrats) is likely to vote for a predistribution bill *with no Republican support* (so, likely, a left-wing bill). Consistent with historical accounts, the coefficient is negative. And as Republican support rises for a predistribution bill, so does DLC support (and this effect can be separated in a statistically significant manner from this tendency on all other votes). For predistribution votes, a ten percentage-point increase in Republican support increases DLC support by $(1.41 + 0.51)$ 1.92 percentage points compared to the non-DLC Democratic baseline.

In col. (3) we show that the DLC's concordance with Republicans on predistribution votes is significantly greater than that for votes on redistribution or on all other votes (note that the significance of the predistribution minus redistribution concordance is shown in the bottom rows of the table). This result holds when in col. (4) we include $State \times Congressional\ session$ fixed effects, which restricts comparisons to DLC versus non-DLC Democrats in the same state at the same time. Note that in the ideal-points analysis, DLCers were also more conservative (relative to the non-economic votes omitted group) on redistribution compared

to other Democrats. In this analysis, their conservativeness on redistribution cannot be distinguished statistically from their conservativeness on the omitted-group votes (the coefficient suggests they are slightly more left-wing on redistribution relative to other votes, but it is very small). Other than that difference, the two methods provide a very similar assessment of the DLC’s ideology relative to other Democrats: the DLC is in general more conservative, but especially on predistribution.

6 Rising influence of the educated within the Democratic Party

In the previous section we documented that in the 1970s the Democratic Party moved away from the predistribution policies differentially favored by less-educated voters. Our model suggests such a pivot can occur due to : (a) a decline in demand for predistribution among the party’s voters; (b) an increase in within-party bargaining power of the anti-predistribution faction (i.e., the more educated). Section 4 casts doubt on (a), as we have shown the educational gradient in support for predistribution is unchanged and the overall level of support for predistribution also remains steady. In this section, we provide direct evidence in support of (b). We document the rising influence of educated individuals both in the Democratic Party candidate-selection process and as candidates themselves. While we were often unable to compare Democrats directly to Republicans in the previous section (in part because Republicans so rarely controlled Congress during our main period of interest) in this section we show that beginning in the 1970s and 1980s, educated individuals begin to play more important roles in the Democratic than the Republican party.

6.1 Parties’ primary donor base

We begin our examination of the changing influence of the educated in party politics by examining financial donations to *primary* campaigns. If a party’s primaries are increasingly reliant on certain groups for their fund-raising, the candidate the party ultimately supplies to general-election voters will likely reflect the preferences of those groups.

Campaign donations are of particular interest to our story because the 1974 Federal Election Campaign Act transformed campaign finance in both primary and general elections. In particular, it allowed a much larger role for *corporate* political action committees. Before 1974, unions ran the largest political action committees, so this act shifted influence away

from them.¹⁶ The timing of these institutional changes with the rise of the “New Democrat” movement is consistent with an exogenous shock to λ , the bargaining power of the anti-predistribution faction within the party. Note as well that since unions have never had much influence in the Republican party, this reform is likely to mostly affect bargaining power among factions within the Democrats.

6.1.1 Data

Political contributions come from Bonica (2014) which contains more than 170 million political contributions from 1979 to 2018. The data contains both contributions made by individual donors and organizations. For this analysis, we keep only contributions that occurred during the primary period.¹⁷ We focus on House elections, both to remain consistent with the previous section and to facilitate within-state analysis, but show similar patterns for Senate elections in Appendix Figure A.10.¹⁸

Donations data do not report individuals’ education. We instead take an ecological approach and map each donation to the average years of schooling in the census tract of the donor’s address.¹⁹ While an imperfect proxy for individual-level education, the education level of key fund-raising census tracts is of interest in their own right. Primary candidates will likely have to spend more time—in terms of campaigning and holding fund-raisers—in the census tracts that constitute their party’s primary donor base.

6.1.2 Increasing donations from educated census tracts in Democratic Primaries

In Figure 4 we plot, separately for Democrats and Republicans, the years of schooling of the average census tract when census tracts are weighted by individual donations (we show in Appendix Figure A.9 the same patterns hold when we instead weight by dollars).

¹⁶See Cohen *et al.* (2009) and Hilton (2021) on the post-1968 Democratic primary reforms and Ansolabehere *et al.* (2003) on modern history of U.S. campaign finance regulation.

¹⁷After 2004, for a substantive share of contributions the indicator for election type is missing for a large share of the dataset (between half and three quarters of the data). We therefore systematically test the robustness of our results to the exclusion of 2004 onwards.

¹⁸We do not examine presidential races in this analysis given the lack of primary races when presidential incumbents run.

¹⁹To approximate years of schooling in a census tract, we use data from the US Census (Manson, 2020), which bins years of schooling into six categories: less than high school (which we code as seven years), some high school (ten years), high school graduate (twelve years), some college (14 years), and college or more (17 years). Results are very similar if we simply use share college for each census tract. We use data from each decennial census years and interpolate for years in between.

Panel (a) shows that in the earliest election cycle in our data, Republican donations come from more educated census tracts, but that Democrats close the gap in the early 1980s. From the mid-1990s onward, Democrats rely more on educated census tracts for primary donations. Unfortunately, data limitations prevent us from determining whether the late 1970s are an inflection point, but we do observe that the Democratic reliance on educated census tracts for primary donations is steadily increasing relative to Republicans over our entire sample period.

The next two panels separate the data into out-of-district versus within-district donations. For out-of-district donations, census tract education levels for Democratic primary candidates overtakes that of Republicans in the 1980s, whereas the analogous moment for within-district donations is in the early 2000s. Thus, the influence of the educated in the primary process appears to have started first as a *national*, not local, movement and was driven by individuals who, by dint of living outside the district, could not vote for the candidates they were funding. Had the shift in primary donations been driven by actual voters in the local elections, we would have instead worried that it was driven by changing preferences of educated voters.²⁰

6.1.3 Educated donors' support for DLC primary candidates

We now ask whether the anti-predistribution faction of the party, the DLC, are more reliant on these educated areas than are other Democrats. Instead of comparing Democrats to Republicans as we did in the previous analysis, we examine only Democrats and compare the education level in the average census tract for a DLC donation to that for other Democrats.

The first column of Table 3 documents the differential DLC reliance on educated census tracts when we pool both within and out-of-district donations. Here we restrict the sample to general-election winners. We start with this sample because we can most accurately assess DLC status. Recall that in many years, we define a politician as a member of the DLC if she is a member of the New Democrat Caucus, a group of DLC-affiliated Members of Congress, so winning the general election is often a necessary condition for our defining someone as DLC. For this sample, the average DLC primary donation comes from a census tract that is 0.09 years more educated relative to the average non-DLC donation.

Including only general-election winners has the virtue of allowing a more accurate definition of DLC membership as well as comparing similarly successful candidates (i.e., all of

²⁰Interestingly, setting aside the educational gradient, Democratic primary candidates rely more on out-of-district individual donations than do Republicans over our entire sample period (this differential reliance holds in general elections as well).

them won their general election). But it of course introduces sample-selection bias that is hard to sign *a priori*. We thus perform complementary analysis that includes all candidates. As we cannot observe DLC status for those who do not go on to win the general election, we use machine-learning techniques to predict it. As there is limited systematic data collected on primary election candidates, we use variables from the donations: political action committee (PAC) contributions as well as zip code demographics of individual donations (excluding, of course, education, the dependent variable). Appendix E provides greater detail and robustness to alternative prediction methods.

The DLC’s reliance on educated census tracts is still highly significant and in fact slightly larger in this sample, despite the fact that inevitable prediction error in our DLC classification will bias the coefficient toward zero. While both the winners-only and the predicted-DLC-status approaches are imperfect, we find some reassurance that they produce similar results as there is no obvious reason that the biases from each would push in the same direction.

The remaining columns show that the DLC’s reliance on educated census tracts comes entirely from out-of-district donations. Indeed, when only local donations are considered in cols. (3) and (4), any difference in the education level of donation census tracts is small and insignificant. But for out-of-district donations (cols. 5 and 6), the typical DLC census tract is over 0.1 years more educated than that for other Democrats. Thus, DLC candidates are relying on educated donations from *outside* their district, which reduces the concern that changes in local demand from voters in their district are compelling DLC candidates to vote against policies favored by less-educated voters such as predistribution. The educated are increasing their power within the Democratic party as national contributors to out-of-district primary races, not by being a prioritized or pivotal constituency in local district primary races.

6.2 Changing class composition of Democratic politicians

While the previous subsection showed the growing influence of educated donors in choosing Democratic candidates (relative to their influence in Republican primary campaigns), we now focus on the politicians themselves. A key limitation in the previous analysis was that the donations data do not exist before the 1980 election and thus we cannot test whether an inflection point exists in the 1970s. We now look at traits of politicians that we can observe over the entire post-war period.

First, we examine the educational biographies of the candidates themselves, and in par-

ticular whether they attended an Ivy-League school. We show in panels (a) and (b) of Figure 5 that, while Democrats in the House and Senate were significantly less likely to come from Ivy League institutions than were Republicans in the immediate decades after World War II, they begin to reduce this gap in the 1970s. Today they are approximately twenty (ten) percentage points more likely to come from an Ivy League institution than are their Republican counterparts in the House (Senate).

While we cannot perform extensive robustness on the Senate results given the limited number of observations per state-year, we show in panel (c) that the House result is robust to controlling for *state* \times *year* fixed effects and dropping the states that host an Ivy school. This panel also makes clear that the inflection point in the Dem-versus-GOP difference is the early 1970s.

As a final exhibit showing that in the 1970s Democratic politicians begin to court more educated voters, Appendix Figure A.11 documents that, relative to Republicans, Democratic Congresspeople begin to speak in increasingly sophisticated language while giving speeches in Congress. While Democrats and Republicans speak at roughly the same level of sophistication in the immediate post-war decades, in the 1970s Democratic politicians begin to speak in a manner that requires significantly greater education to understand than did their Republican counterparts.

In summary, we show that in the 1970s Democratic politicians, relative to their Republican counterparts, began to come from more elite educational institutions and adopt speech more removed from their traditional base of less-educated voters. While not directly related to their supply of economic policy, we find these results reassuring as further evidence that educated individuals gained greater influence within party leadership.

7 How do voters react to changing supply of economic policies?

The model suggests that unchanging policy demand among voters (Section 4) coupled with a shift in within-Democratic-party bargaining power toward educated voters (Section 6) will result in reduced supply of predistribution (Section 5) as well as educational realignment of voters. This section provides evidence for the latter result. It goes beyond existing evidence on realignment in pinpointing the *timing* as well as showing that educated voters differentially support the economically and socially *conservative* DLC.

7.1 When and why did educational realignment begin?

7.1.1 When did educational realignment begin?

In Figure 6 we estimate equation (1) with an indicator variable for Democratic party identification as the outcome variable (so Republicans, Independents or any other response are coded as zero). As usual we include age-in-five-year-bin fixed effects and survey fixed effects. We perform this regression separately by year, so the coefficients on the age fixed effects are unrestricted across time.

Relative to existing literature, the novel result from this analysis is the clear inflection point in the 1970s. From the 1940s until the mid-1970s, an additional year of education predicted that a respondent was roughly three percentage-points less likely to identify as a Democrat. While largely stable, the magnitude of this negative relationship if anything grew in the 1960s and early 1970s. But shortly thereafter, the slope goes from roughly flat to positive and remains so to this day. As a result, the relationship today is almost exactly the inverse of that after World War II: an additional year of education predicts a respondent is three percentage points *more* likely to be a Democrat. The figure also shows that the general shape is robust to dropping non-whites (second series) or using unadjusted years of schooling (third series).

The large dataset we assembled ($N \approx 2$ million) allows us to credibly estimate an inflection point (see Appendix Figure A.16). We pick the year τ that minimizes the sum of squared residuals of this equation: $Dem_{it} = \beta_0 + \beta_1 edu_{it} + \beta_2 edu_{it} \cdot \mathbb{1}\{t > \tau\} + X_{it} + u_{it}$ (Bai and Perron, 1998). We normalize the survey weights so that each year is given the same weights in the estimation. The data choose 1976 as the inflection point. We also show in the Appendix that finding an inflection in the 1970s is robust to various changes in specification or subgroups (e.g., restricting to whites yields a 1978 inflection point).

Appendix Figure A.12 shows that there is little corresponding movement in overall Democratic party identification, either considering just Democrats and Republicans, or including independents. Educational re-alignment thus did not come with a large change in overall party identification. Through the lens of our model, this is because political competition considerations still constrain the outcomes of party bargaining. While the supply of policies changes the composition of party supporters, it has a negligible effect on overall support. Indeed, our data shows little evidence of any major partisan shift after Civil Rights wins non-white voters to the Democratic Party; in contrast to the sudden change in the educa-

tional gradient in party support, there is only a slow decline in overall Democratic party identification in the last 40 years.

7.1.2 Role of economic policy in realignment

A key goal of this paper is to assess the ability of the parties' economic policies in explaining educational realignment. We now examine how Americans have viewed the economic performance of the two major parties over the past eight decades. Since the 1940s, Gallup has asked respondents the following question (with only small variations over eighty years): "Looking ahead for the next few years, which political party—the Republicans or the Democrats—do you think will do the better job of keeping the country prosperous." Appendix Figure A.19 shows the time series of the share of respondents who say that Democrats are better in this regard—it varies as expected, with the public generally taking a dim view of the economic performance of the party in the White House during recessions.

The first series of Figure 7 panel (a) replicates our main realignment graph in Figure 6, but instead of identifying as a Democrat, choosing Democrats as the party better for the economy is the outcome variable. In general, the pattern is similar—in the immediate post-war decades, an additional year of education predicted a 2-3 percentage point reduction in the likelihood of identifying the Democrats as the better party for a strong economy. Today, it predicts a two percentage point increase in that same likelihood.

One natural concern is that the result is merely a consequence of party identification more generally. Suppose that less-educated individuals leave the Democratic Party entirely over cultural issues and actually still prefer Democratic economic policies. We might worry that, merely to avoid cognitive dissonance, such respondents—now identifying as Republicans or at least no longer Democrats—prefer not to admit that the Democrats in fact remain the better party on economic issues. We do our best to address this concern by including indicator variables for party identification (dummies for Democrat and Republican, leaving independents/others as the omitted group). This exercise asks, compared to other respondents with the same party identification, how does education predict the party you report having the better economic platform? This specification is very demanding, as most individuals do indeed answer that the party to which they belong has the better plan to keep the country prosperous. We also allow, as usual, these control variables to have different effects over time.

The final series shows that, even controlling flexibly for party identification, the same

basic pattern holds. While the coefficients are (as expected) much smaller in magnitude, in the four decades after WWII, education is a negative predictor of identifying the Democrats as the better party to guide the economy. By the 1990s, the gradient is zero or somewhat positive.

A final question is how much less-educated voters’ increasing dissatisfaction with the Democratic party on economic issues explains the overall partisan realignment in Figure 6. In panel (b) of Figure 7, we replicate the overall partisan realignment analysis, but only on the subset of data that also includes the prosperity question (first series). We then add a control for the response to the prosperity question (second series). Roughly half of the overall change is explained by adding this control.

We prefer to take a conservative interpretation of this result. Controlling for “which party is best for prosperity” in a regression explaining partisan identity is a potentially “bad control” because of reverse causality—which party you belong to likely influences your answer on which party is best for the economy. In that sense, we view explaining half of total realignment via the parties’ economic policies as an upper bound. And, of course, even taking the upper bound as the referred estimate, it still leaves half the realignment to be explained by other factors, a point we return to in the conclusion.

7.2 Who supported the “New Democrats?”

So far in this section we have shown that less-educated voters leave the Democratic party at the same time that the “New Democrats” gain power and predistribution fades from the party’s agenda. Below we show that the “New Democrats” were indeed less attractive to less-educated voters than were the old-style New Dealers.

7.2.1 Hypothetical election match-ups

After the unrest at the 1968 Democratic Convention, the party adopted rules that reduced the influence of establishment groups in the party, namely labor unions. In particular, state primaries and caucuses took the place of “smoke-filled back rooms” at the national conventions, where unions and other powerbrokers had once controlled who would appear on the national ticket. The 1972 election was the first under the new system and thus polls such as Gallup would ask respondents their views on potential presidential candidates.

In most of the 1972–1992 Democratic primaries, politicians who clearly identified with one of the two wings of the party (“New Democrats” versus “old-style” New Dealers) either

won or were runners-up to the nomination. In 1972, while Humphrey represented the legacy of LBJ’s Great Society, McGovern had no close ties to labor and in fact unions launched an “ABM” (Anyone but McGovern) campaign (and McGovern was one the only Democrat to not receive the AFL-CIO endorsement in the general election). In 1980, Ted Kennedy runs against the incumbent Carter after a first term that many historians date as the birth of “neoliberalism” in the US.²¹ Carter had so disappointed labor that the United Auto Workers took the unusual step of endorsing Kennedy’s upstart campaign against an incumbent Democrat. In 1984, Mondale beats Gary Hart, one of the most prominent “New Democrats.” Hart had served as McGovern’s campaign manager and once described his New Democrat brethren by saying: “We are not a bunch of little Hubert Humphreys.” In 1988, Mike Dukakis beats Jesse Jackson, one of the DLC’s *bêtes noires* who mocked the group as “Democrats for the Leisure Class.” Finally, in 1992, while Jerry Brown and Bill Clinton were both DLCers, we rank Clinton as “newer” since he was in fact the president of the group.²² The 1976 election is hard to categorize as Carter ran mostly on anti-corruption themes and did not face a formidable opponent, but for consistency we label him as “new” and the runner-up, Mo Udall, as “old”.²³

We make use of surveys that ask all respondents (regardless of party affiliation) for whom they would vote if one of the two Democratic politicians faced the eventual Republican candidate. For example, respondents are asked in 1984 their preferred candidate in a Hart v. Reagan election and, in a separate question, in a Mondale v. Reagan election.

Table 4 shows that years of education predict greater support for the “New Democrat,” relative to the “old” Democrat, when each is paired against the eventual GOP nominee. Col.

²¹? writes of Carter’s single term: “What Carter did not do was advocate policies that might win the support of poor and working-class Americans buffeted by job insecurity and high inflation. Instead, he peered at social programs, old and new, through an austerity-tinted lens. Explaining that he needed ‘to enhance an image of fiscal responsibility,’ the president sought to balance the budget and opposed a national health insurance plan written by Ted Kennedy. He also signed bills to deregulate the airline and trucking industries....In sum, these moves, and the support they drew from other party leaders, signified a momentous retreat. No longer would Democrats maintain that government had an obligation to set strict rules to protect workers and demand that corporations obey them. No longer would union power be viewed as an unambiguous boon to party fortunes as well as prime evidence that Democrats were the natural home of wage earners of any race, religion, or region.”

²²We conclude this analysis in 1992 as there was no 1996 Democratic primary and in 2000 the front-runners (Al Gore and Bill Bradley) were both New Democrats. As noted, even in 1992, both candidates are DLCers. By the 1990s DLC types were firmly in control of the party.

²³While Udall was the (distant) runner-up in the delegate count, he consistently registered low name-recognition in Gallup surveys from that time. The 1976 primary is complicated largely by the role of Hubert Humphrey. He consistently polls as Democrats’ top choice, but he never officially entered the race though at the same time said he might accept the nomination in the case of a brokered convention.

(1) shows that during this period, there is still a significantly negative educational gradient in support for the New Democrat over the Republican (not surprising, as these races are mostly in the 1970s and 1980s when education realignment has only just begun) but it is smaller in magnitude than the negative educational gradient when the “old” Democrat is pitted against the Republican (col. 2). Note the differences in the mean of the dependent variables: New Democrats tend to out-perform their “old”-style counterparts in these hypothetical races.

In many cases, respondents are asked both match-up questions, so we can subtract the “old”-Dem-versus-Republican-nominee response from the “new”-Dem-versus-Republican-nominee response to facilitate more succinct analysis (the dependent variable thus takes the values 1, 0 and -1). Note that the dependent-variable means, reported below each column, tend to be small in this analysis, because most individuals tend to vote for the same party regardless of the identity of the actual candidates.

Col. (3) shows that the main result—more-educated voters prefer the New Democrats—holds once we difference the hypothetical responses. As presidential candidates tend to pick up extra support from their own and neighboring states, we add *state* \times *election* fixed effects in col. (4) and the result remains unchanged. Given the interest in *white* voters in the realignment discourse, col. (5) replicates col. (3) but drops non-whites. The magnitude of the gradient retains two-thirds of its magnitude and remains highly significant.

The final column shows that the result holds in each election in our sample period with the exception of 1976. We are not surprised by the 1976 result given (a) that Carter avoided discussing economic policy in his first presidential campaign and (b) how little attention and name-recognition the runner-up Mo Udall enjoyed.²⁴ Only those following politics very closely would have an opinion, positive or negative, of Udall, and such respondents tend to be more educated.

7.2.2 House election returns

In this section we examine *actual* election results as opposed to hypothetical matchups, so we take an ecological approach as individual-level votes are never observed. King and Palmquist (1997) provide 1980s election results at the *minor-civil-division-group* (MCDG) level. There are roughly sixty MCDG neighborhoods per Congressional District. We match 1980 census data to MCDGs to calculate average years of education among adults and other demographic

²⁴As just one example, in a May 1976 Gallup poll (taken *after* most of the primaries already took place) 28% of respondents did not know who Udall was (compared to two percent for Humphrey and four percent for Carter).

characteristics in each MCDG. This MCDG-level election returns is only available for the 1980s, so we cannot perform a long-run analysis, but we can examine the early years of educational realignment.

To capture how neighborhood education predicts voting patterns in these House elections, we estimate the following equation:

$$Dem_{mt} = \beta DLC_{mt} \times Education_m + \gamma_1 DLC_{mt} + \gamma_2 Education_m + \eta_{d(m)} + e_{mt}, \quad (4)$$

where Dem_{mt} is the share of MCDG m in election year t that votes Democratic; DLC_{mt} is an indicator variable for whether the Democratic candidate in district $d(m)$ in year t is a DLC member; $Education_m$ is the education level of MCDG m (based on the 1980 Census), and $\eta_{d(m)}$ are a vector of Congressional District fixed effects (as each MCDG is mapped to a unique Congressional district).

The results in Table 5 document a strong preference of educated neighborhoods for DLC candidates relative to other Democrats. We begin with a sample of elections where Democrats all win (again, for this sample, we can identify DLC members with greater precision). In col. (1), the main effect of years of education is to substantially depress Democratic vote share, not surprising as in the 1980s less-educated voters were still significantly more likely to be Democrats (even if we identify a turning point in the 1970s, the overall gradient is still markedly negative in the 1980s). But the coefficient of interest—the interaction term—indicates that, as we predict, DLC Democrats out-perform other Democrats in more educated areas. In fact, this effect is so large that, among DLC Democrats, roughly two-thirds of the large Democratic underperformance in educated areas is erased.

One concern is that, perhaps for strategic reasons, DLC candidates run in more educated districts or in districts where the educated voters are more open to voting Democratic. In Col. (2) we include $District \times Year$ fixed effects (i.e., a fixed effect for each House election), so the coefficient of interest is identified by comparing the areas where DLC candidates over-perform in a given election to the areas where non-DLC Democrats over-perform in their elections, so across-election comparisons no longer contribute to the estimate. We find similarly large and significant results.

Cols. (3) and (4) show that this basic pattern—DLCers over-performing in educated areas compared to other Democrats—holds when we include election-losers and predict DLC status via machine learning techniques as in Table 3. Note that, compared to the first two columns, the mean of the dependent variable is much closer to fifty percent (it is still above fifty, as

we would expect, given that during the 1980s the Democrats still held firm control over the House and thus won most elections) as we are no longer restricting the sample to elections where Democrats win.

While we have been focusing on the educational level of MCDGs, it might be the case that the DLC over-performance in these areas is better explained by characteristics merely correlated with education. In col. (5) we include controls for share under age 35 and share white and their interactions with DLC. Both of these main effects have large and significant coefficients in the expected direction—in general Democrats over-perform with young and non-white voters. Interestingly, we see that white areas prefer the DLC to other Democrats (and there is no differential effect by age). But adding these additional controls have no effect on the coefficient of interest. Col. (6) shows robustness to adding $District \times Year$ to the col. (5) specification.

8 Alternative explanations for realignment

In this section, we briefly examine other explanations, especially those that might have particular resonance during the 1970s and 1980s, when realignment begins.

8.1 Demand for and supply of social issues

Are educated voters becoming more left-wing on social issues and thus moving toward the Democratic party (which in recent decades has taken more liberal positions on issues related to race, gender and religion)? Figure A.20 explores this possibility using the same specification as equation (1), again using a set of survey questions that are asked consistently over time. These questions include attitudes towards African-Americans, Jews, abortion, gender equality, church attendance, and affirmative action. Given that some of these questions are about attitudes toward Black Americans, we include only whites in this graph (and repeat the analysis for the full sample in Appendix Figure A.20).

As is clear from Figure A.20, more educated respondents are more socially liberal on almost all of these issues (interestingly, one exception is affirmative action, perhaps because they are more sensitive to university admissions policies). But this tendency is not new nor did it show any marked change in the 1970s when realignment begins.

On the supply side, recall that DLC members are significantly more *conservative* than

other Democrats on social issues (see col. 5 of Table 1).²⁵ Educated voters differentially supporting the DLC *despite* the group’s conservative stances on social issues suggests that their agreement on economic policy was particularly strong. In any case, that educated voters first supported DLC Democrats in the first few decades of realignment suggests that social issues if anything deterred or delayed this process.

8.2 Salience of economic versus non-economic issues

A related possibility is that while the answers that more- and less-educated individuals give on social-issue questions have not diverged over time, the weights that voters put on these non-economic issues have increased over time. Enke *et al.* (2021) proposes a model in which the weight put on social over economic increases with income (so that concern for social issues is a type of “luxury good”). One implication of this model is that over time we might expect that the salience of social issues increases so long as a society is growing richer.

We can directly assess this argument using Gallup’s “Most Important Issue” question, which consistently asks respondents “What do you think is the most important issue facing American?”, with responses harmonized into topics by CAP. Appendix Figure A.21 panel (a) shows that while there is an obvious counter-cyclical to respondents saying the economy is the most important problem, there is no systematic trend, suggesting that economic policy has not become less important—at least based on self-reports—as Americans have gotten richer. We further show in panel (b) of the Appendix figure that the educational gradient in salience of economic issues is relatively constant—and close to zero—since 1960. Interestingly, before 1960, more-educated Americans were more likely to choose foreign-policy issues as the most important problem, but this tendency ends in the 1960s. Since then, there is no evidence that respondents’ education has any strong predictive power over the (self-reported) weight they place on economic issues.

8.3 The role of Civil Rights

The Democrats’ increasingly liberal position on Civil Rights since the 1940s—and especially their role in passing the signature Civil Rights bills in the first half of the 1960s—led to a major partisan realignment whereby a significant share of Southern whites left the party

²⁵One of the most memorable moments of the 1992 presidential campaign was at an event organized by Jesse Jackson, where Bill Clinton chastised rap artist and fellow guest Sister Souljah over her lyrics. This episode is the origin of the phrase “Sister Souljah moment” used to describe a gesture taken by a politician seeking to prove he is not beholden to his party’s base.

(Kuziemko and Washington, 2018). Here, we examine what role if any this large Civil-Rights-driven realignment plays in the educational realignment.

Appendix Figure A.17 shows our main realignment result (Figure 6) separately by the four Census regions, panel (a) for the full sample and (b) for whites only). In the South, white Democrats are becoming *less* educated from the 1940s through the 1960s. Put differently, just as the Democrats are making moves in the liberal direction on Civil Rights, educated white Southerners are *leaving* the party. This pattern is consistent with a model as in Lee and Roemer (2006) where educated, racially conservative white Southerners had supported the Democrats primarily for their defense of Jim Crow and *in spite of* their relatively left-wing economic policies. But as the Democratic party grew more liberal on this key issue, these voters no longer faced a trade-off between their economic interests and their support of segregation, so they left the Democrats. This pattern suggests that partisan shifts due to Civil Rights is unlikely to explain *educational* realignment

A final point to note is that non-whites also experience substantial educational realignment. In Appendix Figure A.13 we reproduce our main realignment figure separately for whites and non-whites (note such an exercise is only possible for non-whites because of the large dataset we constructed). While non-white Americans are much more strongly attached to the Democratic Party, we nonetheless can see significant realignment in this subgroup, especially in the 1980s and 1990s.

8.4 What about the Republicans?

As noted, we cannot always observe the supply of Republican economic policy as easily as we can for Democrats due to the former's infrequent control of Congress over most of our sample period. But in this subsection we briefly address concerns that that Democratic Party was merely reacting to economic policy changes made by Republicans.

First, when we replicate our main realignment Figure 6 but put Republican partisan identification instead of Democratic identification as the outcome, the inflection point is significantly later, in 1992 (see Appendix Figure A.18 for the realignment figure and Appendix Figure A.16 for tests of the inflection point). Roughly speaking, the educational gradient first changes in the mid-1970s along the Democratic-versus-all-other margin, and then only in 1992 along the Republican margin. Voters' reaction is consistent with the Democrats' agenda changing first.

Second, while Ronald Reagan is widely viewed as a transformative political figure, his role

appears rather muted on the more narrow question of educational realignment. In Appendix Figure A.22 we present a version of our main realignment figure but instead of Democratic Party identification we use self-reported Democratic presidential vote as the outcome. If anything, the 1980 election is a bit of a retrenchment, where less-educated voters returned home (briefly, as it would turn out) to the Democratic Party. In neither Reagan election were less-educated voters particularly disloyal to their traditional Democratic home (and this result holds, as we show in the figure, when non-white voters are excluded).

Our paper has focused on positive shifts in λ^D , the bargaining power of the educated within the Democratic Party, but of course Republican factions exist as well. Two new voting factions appear on the Republican scene in the 1970s. One is the rise of Southern Republicans in response to Civil Rights, as discussed above, and the other is the rise of evangelical Christians, previously bi-partisan, as a pro-Republican force. A third faction of economists and business groups, latent throughout the post-war period, gains influence within the party in the 1970s, pivoting the party’s fiscal policies towards supply-side economics over the traditional focus on balanced budgets.²⁶

But none of these schisms line up neatly with *education*. The Southern Republicans attract both high-education and low-education whites, and by the 1970s the South looks no different in its educational gradient in party identification than the other regions. Evangelicals tend to have average or slightly above-average education levels. Finally, while there is an educational gradient on preferences for predistribution vs redistribution, within redistribution there is no educational gradient on tax cuts vs balanced budget. This lack of clear educational cleavages within the Republicans is in contrast to the Democratic party, where the decline of labor union influence explicitly lines up with an educational gradient, given the preponderance of less-educated workers in labor unions in this period (Farber *et al.*, 2021). While further research on the changing policies and composition in the Republican party is certainly welcome, our sense is that the leading changes in party factions seem to not contribute to realignment *by education*.

Of course, this short section is hardly the final word on alternative explanations for

²⁶Traditionally, more strident anti-government types such as Goldwater (presidential nominee in 1964) and Reagan (1980 and 1984) vied for influence with more moderate establishment figures such as George H.W. Bush (who in the 1980 primary memorably called Reagan’s supply-side agenda “voodoo economics”). And even looking beyond the 1970s, while Reagan’s agenda appeared to hold sway in the early and mid 1980s, Republicans saw a return of the moderates with George H.W. Bush (1988 and 1992) and then Bob Dole (1996) as the party’s standard bearer. By the late 1980s, a pro-predistribution, anti-trade faction within the Republican party has also emerged, perhaps personified by Pat Buchanan, although it is not ascendant until Donald Trump.

educational realignment. While we believe that some of the patterns we document here suggest that a social-issues-alone (or a Civil-Rights-alone) model of educational realignment is unlikely to be consistent with the data, more nuanced models of how social and economic issues interact may better accommodate these patterns. Similarly, educational realignment is not the *only* shift in political coalitions over the past several decades. The parties’ positions on social issues likely have predictive power over realignments across other dimensions besides education. For example, the rise of abortion as a major political issue after *Roe v. Wade* (1973) likely contributes to the Republicans’ ability to mobilize evangelical Christians and conservative Catholics.

does not mean that abortion or other social issues haven’t had profound effects on U.S. political economy.

9 Conclusion

We have presented evidence that less-educated voters have long favored more pre-tax-and-transfer interventions (“predistribution”) in the economy and labor market. Beginning in the 1970s, Democrats—once champions of these New-Deal-type policies—backed away from this agenda. Coincident with this change was the rise of the “New Democrats”—who were openly skeptical of predistribution and drew their financial support disproportionately from educated, out-of-state donors. We construct a large dataset of partisan identification dating back to 1942 and identify the 1970s as the key starting point for education realignment. In particular, educated voters disproportionately support the “New Democrats” who come onto the political scene at this time, both in survey questions posing hypothetical election match-ups and in House elections.

In this paper we have not taken up the question of *why* Democrats change their economic agenda in the 1970s. We have noted some potential proximate causes—the change in primary rules after the 1968 election, the loosening of campaign finance rules beginning in the 1970s—but the question merits deeper analysis in future work. It is also important not to over-fit explanations to the U.S. context. Pundits in the 1970s frequently used the term “the crisis of the West” to describe the unprecedented mix of energy crises, inflation, slow economic growth and high unemployment facing industrialized democracies (see Stein, 2010 and Bartel, 2022). These factors among others may have convinced center-left parties to tack right on economic policy, from Francois Mitterrand’s *Parti Socialiste* in the 1980s to Tony Blair’s Labour Party in the 1990s.

While our paper identifies the 1970s as the key moment when predistribution fades from the Democratic policy agenda and less-educated voters leave the party, recent events suggest a potential return of predistribution in U.S. politics. President Biden’s clean-energy agenda has been called a revival of “industrial policy” (and was preceded by Senator Bernie Sanders’ call for a “Green New Deal”). On the other side of the aisle, President Trump increased tariffs on China (which Biden has not reserved). Outside of politics, corporations discuss “re-shoring” and “friend-shoring” jobs in light of the fragility of supply chains exposed by the Covid-19 pandemic and geopolitical tensions. Only future work can assess whether these recent developments herald a more interventionist mode of policy-making with the potential to create new political coalitions.

References

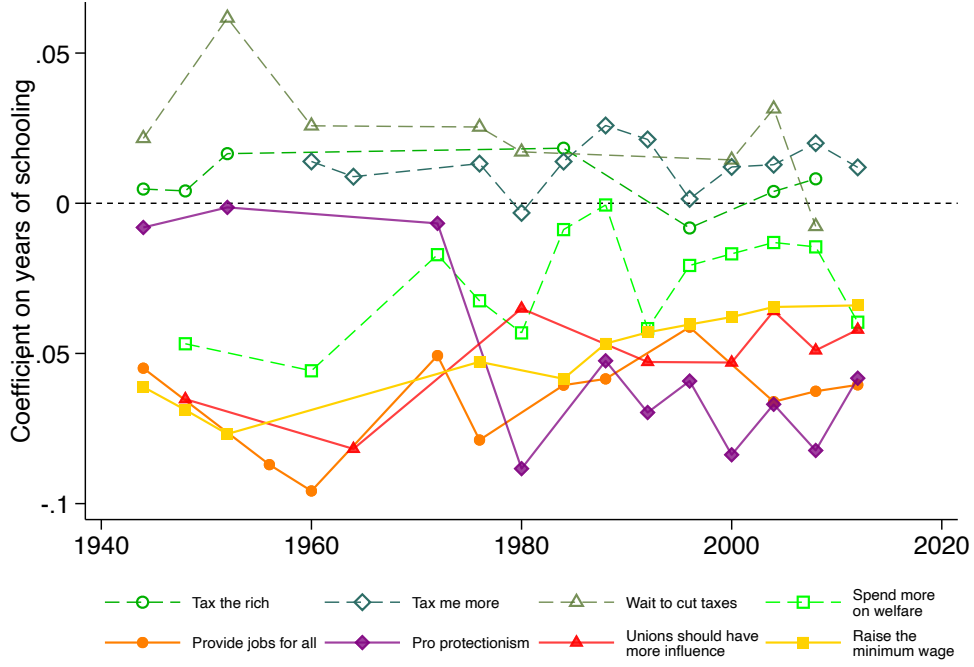
- ACEMOGLU, D., NAIDU, S., RESTREPO, P. and ROBINSON, J. A. (2015). Democracy, redistribution, and inequality. In *Handbook of income distribution*, vol. 2, Elsevier, pp. 1885–1966.
- ALESINA, A., STANTCHEVA, S. and TESO, E. (2018). Intergenerational mobility and preferences for redistribution. *American Economic Review*, **108** (2), 521–54.
- ANSOLABEHERE, S., DE FIGUEIREDO, J. M. and SNYDER JR, J. M. (2003). Why is there so little money in us politics? *Journal of Economic perspectives*, **17** (1), 105–130.
- , RODDEN, J. and SNYDER, J., JAMES M. (2006). Purple america. *Journal of Economic Perspectives*, **20** (2), 97–118.
- ASH, E. and HANSEN, S. (2023). Text algorithms in economics. *Annual Review of Economics*, **15**.
- ASHOK, V., KUZIEMKO, I. and WASHINGTON, E. (2015). Demand for redistribution in an age of rising inequality: Some new stylized facts and tentative explanations. *Brookings Papers for Economic Activity*.
- BAI, J. and PERRON, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, pp. 47–78.
- BARTEL, F. (2022). *The triumph of broken promises: The end of the Cold War and the rise of neoliberalism*. Harvard University Press.
- BATEMAN, D. A., KATZNELSON, I. and LAPINSKI, J. S. (2018). *Southern nation: Congress and white supremacy after reconstruction*, vol. 158. Princeton University Press.
- and LAPINSKI, J. (2016). Ideal points and american political development: Beyond dw-nominate. *Studies in American Political Development*, **30** (2), 147–171.
- BÉNABOU, R. and TIROLE, J. (2016). Mindful economics: The production, consumption, and value of beliefs. *Journal of Economic Perspectives*, **30** (3), 141–64.
- BESLEY, T. (2007). Electoral strategy and economic policy!
- BLANCHET, T., CHANCEL, L. and GETHIN, A. (2020). Why is europe more equal than the united states?
- BONICA, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science*, **58** (2), 367–386.
- (2018). Inferring roll-call scores from campaign contributions using supervised machine learning. *American Journal of Political Science*, **62** (4), 830–848.
- , MCCARTY, N., POOLE, K. T. and ROSENTHAL, H. (2013). Why hasn’t democracy slowed rising inequality? *Journal of Economic Perspectives*, **27** (3), 103–124.

- BOZIO, A., GARBINTI, B., GOUPILLE-LEBRET, J., GUILLOT, M. and PIKETTY, T. (2020). Predistribution vs. redistribution: Evidence from france and the us.
- BREIMAN, L. (2001). Random forests. *Machine learning*, **45**, 5–32.
- CHOI, J., KUZIEMKO, I., WASHINGTON, E. L. and WRIGHT, G. (2021). *Local Economic and Political Effects of Trade Deals: Evidence from NAFTA*. Tech. rep., National Bureau of Economic Research.
- COHEN, M., KAROL, D., NOEL, H. and ZALLER, J. (2009). *The party decides: Presidential nominations before and after reform*. University of Chicago Press.
- COX, G. W. and MCCUBBINS, M. D. (2005). *Setting the agenda: Responsible party government in the US House of Representatives*. Cambridge University Press.
- CRUCES, G., PEREZ-TRUGLIA, R. and TETAZ, M. (2013). Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics*, **98**, 100–112.
- DECHEZLEPRÊTRE, A., FABRE, A., KRUSE, T., PLANTEROSÉ, B., SANCHEZ CHICO, A. and STANTCHEVA, S. (2022). *Fighting Climate Change: International Attitudes Toward Climate Policies*. Working Paper 30265, National Bureau of Economic Research.
- DELLAVIGNA, S. and KAPLAN, E. (2007). The fox news effect: Media bias and voting. *The Quarterly Journal of Economics*, **122** (3), 1187–1234.
- DEVLIN, J., CHANG, M.-W., LEE, K. and TOUTANOVA, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- DIAMOND, P. A. and MIRRLEES, J. A. (1971). Optimal taxation and public production i: Production efficiency. *The American economic review*, **61** (1), 8–27.
- DORN, D., HANSON, G., MAJLESI, K. *et al.* (2020). Importing political polarization? the electoral consequences of rising trade exposure. *American Economic Review*, **110** (10), 3139–83.
- ENKE, B., POLBORN, M. and WU, A. A. (2021). Morals as luxury goods and political polarization.
- FARBER, H. S., HERBST, D., KUZIEMKO, I. and NAIDU, S. (2021). Unions and inequality over the twentieth century: New evidence from survey data. *The Quarterly Journal of Economics*, **136** (3), 1325–1385.
- FRIEDMAN, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pp. 1189–1232.
- GANS, J. S. and SMART, M. (1996). Majority voting with single-crossing preferences. *Journal of public Economics*, **59** (2), 219–237.

- GEISMER, L. (2022). *Left Behind*. Public Affairs.
- GENNAIOLI, N. and TABELLINI, G. (2019). Identity, beliefs, and political conflict.
- GENTZKOW, M., SHAPIRO, J. M. and TADDY, M. (2019). Measuring group differences in high-dimensional choices: method and application to congressional speech. *Econometrica*, **87** (4), 1307–1340.
- GETHIN, A., MARTÍNEZ-TOLEDANO, C. and PIKETTY, T. (2021). Brahmin Left Versus Merchant Right: Changing Political Cleavages in 21 Western Democracies, 1948–2020*. *The Quarterly Journal of Economics*, qjab036.
- HACKER, J. (2011). The institutional foundations of middle-class democracy. *Policy Network*, **6** (2011), 33–37.
- HILTON, A. (2021). *True blues: The contentious transformation of the Democratic Party*. University of Pennsylvania Press.
- HOBOLT, S. B. (2016). The brexit vote: a divided nation, a divided continent. *Journal of European Public Policy*, **23** (9), 1259–1277.
- HOPKINS, D. J., COFFEY, D. J., GALVIN, D. J., GAMM, G., HENDERSON, J., PADDOCK, J. W. and SCHICKLER, E. (2022). Select American State Party Platforms, 1846–2017.
- KAPLOW, L. and SHAVELL, S. (1994). Why the legal system is less efficient than the income tax in redistributing income. *The Journal of Legal Studies*, **23** (2), 667–681.
- KILLIAN, M., SCHOEN, R. and DUSSO, A. (2008). Keeping up with the joneses: The interplay of personal and collective evaluations in voter turnout. *Political Behavior*, **30** (3), 323–340.
- KING, G. and PALMQUIST, B. (1997). The record of american democracy, 1984–1990. *PS: Political Science & Politics*, **30** (4), 746–747.
- KITSCHOLT, H. P. and REHM, P. (2019). Secular partisan realignment in the united states: The socioeconomic reconfiguration of white partisan support since the new deal era. *Politics & Society*, **47** (3), 425–479.
- KUZIEMKO, I., NORTON, M. I., SAEZ, E. and STANTCHEVA, S. (2015). How elastic are preferences for redistribution? evidence from randomized survey experiments. *American Economic Review*, **105** (4), 1478–1508.
- and WASHINGTON, E. (2018). Why did the democrats lose the south? bringing new data to an old debate. *American Economic Review*, **108** (10), 2830–67.
- LEE, W. and ROEMER, J. E. (2006). Racism and redistribution in the United States: A solution to the problem of American exceptionalism. *Journal of Public Economics*, **90** (6), 1027–1052.

- MANSON, S. M. (2020). Ipums national historical geographic information system: version 15.0.
- MARTIN, G. J. and YURUKOGLU, A. (2017). Bias in cable news: Persuasion and polarization. *American Economic Review*, **107** (9), 2565–99.
- MAYDA, A. M., PERI, G. and STEINGRESS, W. (2022). The political impact of immigration: Evidence from the united states. *American Economic Journal: Applied Economics*, **14** (1), 358–89.
- MCCARTY, N., POOLE, K. T. and ROSENTHAL, H. (2006). *Polarized America: The dance of ideology and unequal riches*, vol. 5.
- MUDGE, S. L. (2018). *Leftism reinvented: Western parties from socialism to neoliberalism*. Harvard University Press.
- POOLE, K. T. and ROSENTHAL, H. (2000). *Congress: A political-economic history of roll call voting*. Oxford University Press, USA.
- ROEMER, J. E. (1998). Why the poor do not expropriate the rich: An old argument in new garb. *Journal of Public Economics*, **70** (3), 399–424.
- ROTHENBERG, R. (1984). *The neoliberals: Creating the new American politics*. Simon & Schuster.
- SANCES, M. W. (2019). How unusual was 2016? flipping counties, flipping voters, and the education–party correlation since 1952. *Perspectives on Politics*, **17** (3), 666–678.
- SCHUMAN, H. and PRESSER, S. (1996). *Questions and answers in attitude surveys: Experiments on question form, wording, and context*. Sage.
- STANTCHEVA, S. (2022). *Understanding of Trade*. Working Paper 30040, National Bureau of Economic Research.
- STEIN, J. (2010). *Pivotal decade: How the United States traded factories for finance in the seventies*. Yale University Press.

Figure 1: Preferences for pre- and re-distribution by education



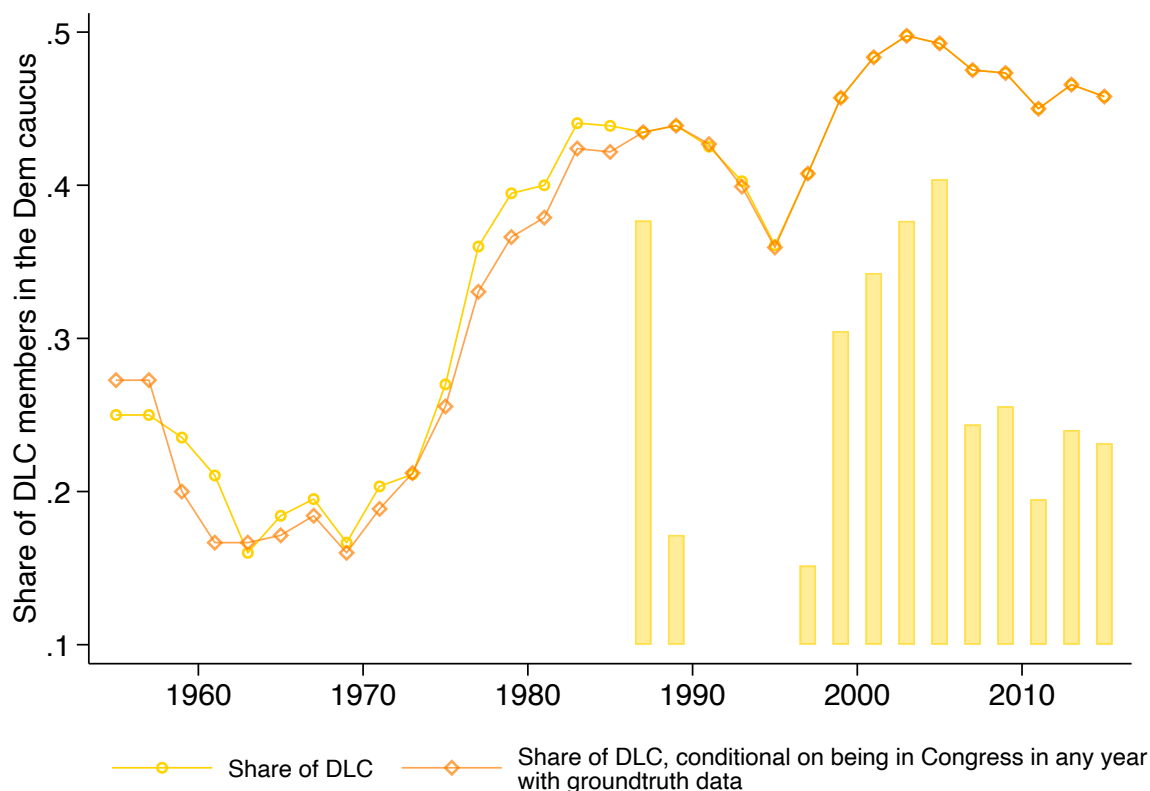
Source: ANES, Gallup, and GSS along with various other survey organizations hosted on iPoll. See Appendix Table G.7 for more detail. We use survey weights if provided (we divide by the mean of the weights by survey to ensure they all average to one).

Notes: This figure plots the coefficients β^p from the following regression estimated separately for each five-year period p :

$$y_{is} = \beta^p \text{Adj. years of school}_i + \mu_s + \mathbf{Age}_i + e_{is},$$

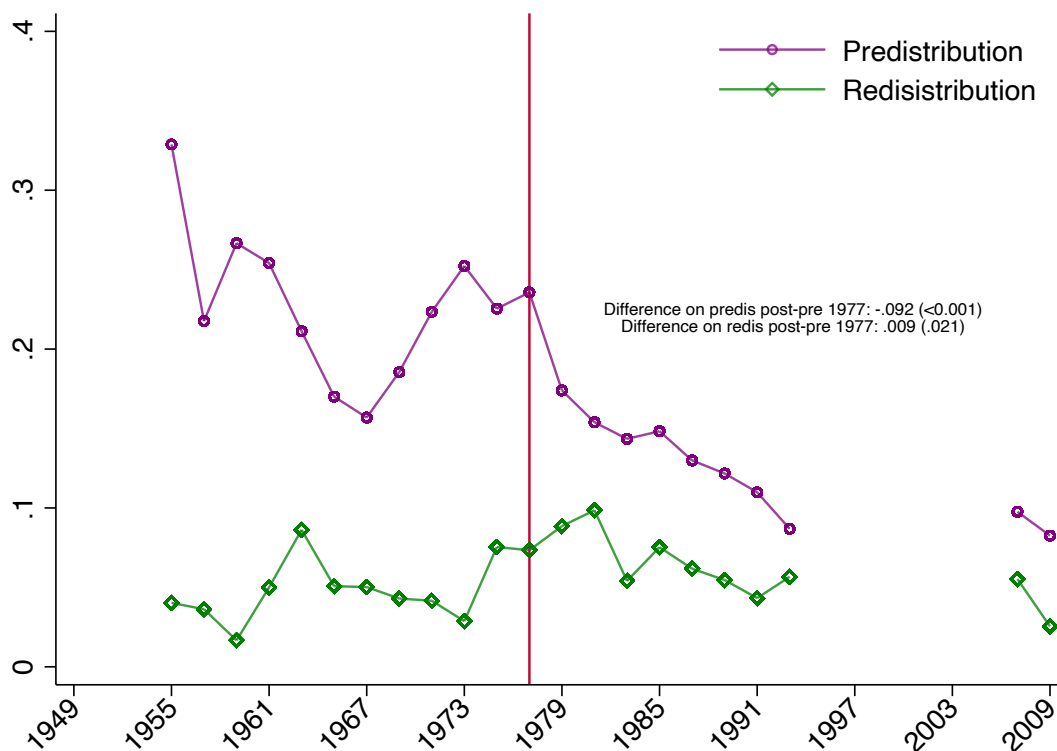
where y_{is} is the outcome variable for person i in survey s ; *Adjusted years of school* is our predicted years of school based on the self-reported educational category provided by the respondent along with demographics and years (see Section 3), μ_s are survey (which subsume date) fixed effects, as we often have several surveys per period p and \mathbf{Age}_i is a vector of age-in-five-year-bin dummies. The outcome variables are standardized measures (mean zero and variance 1) of support for the given policy. Detailed survey questions for each policy are displayed in Appendix B. We suppress confidence intervals to reduce clutter, but we plot each policy question separately along with 95-percent confidence intervals in Appendix Figure A.1.

Figure 2: Evolution of Democratic Leadership Council (DLC) membership in Congress



Notes: The figure shows the evolution of the share of House Democrats belonging to the DLC. We combine several sources (e.g., original membership lists as well as membership in DLC-aligned House caucuses) to determine DLC membership, detailed in Appendix E. The solid bars show the years for which we actually have records of membership. The series with hollow circles (yellow) shows the raw evolution of (interpolated) DLC membership. We assume that any politician on any DLC list is a member for life, so the DLC share is non-zero even in years where we are missing actual membership information. The series with hollow diamonds (orange) shows the evolution of the share of DLC members among politicians who were serving in the House during at least one year for which we have actual DLC membership lists (so this series captures the DLC share among representatives who could possibly appear on one of our DLC lists). Given data collection, DLC is likely underestimated from 1987 – 1996 as we will likely miss any DLC-er who served *only* during those years.

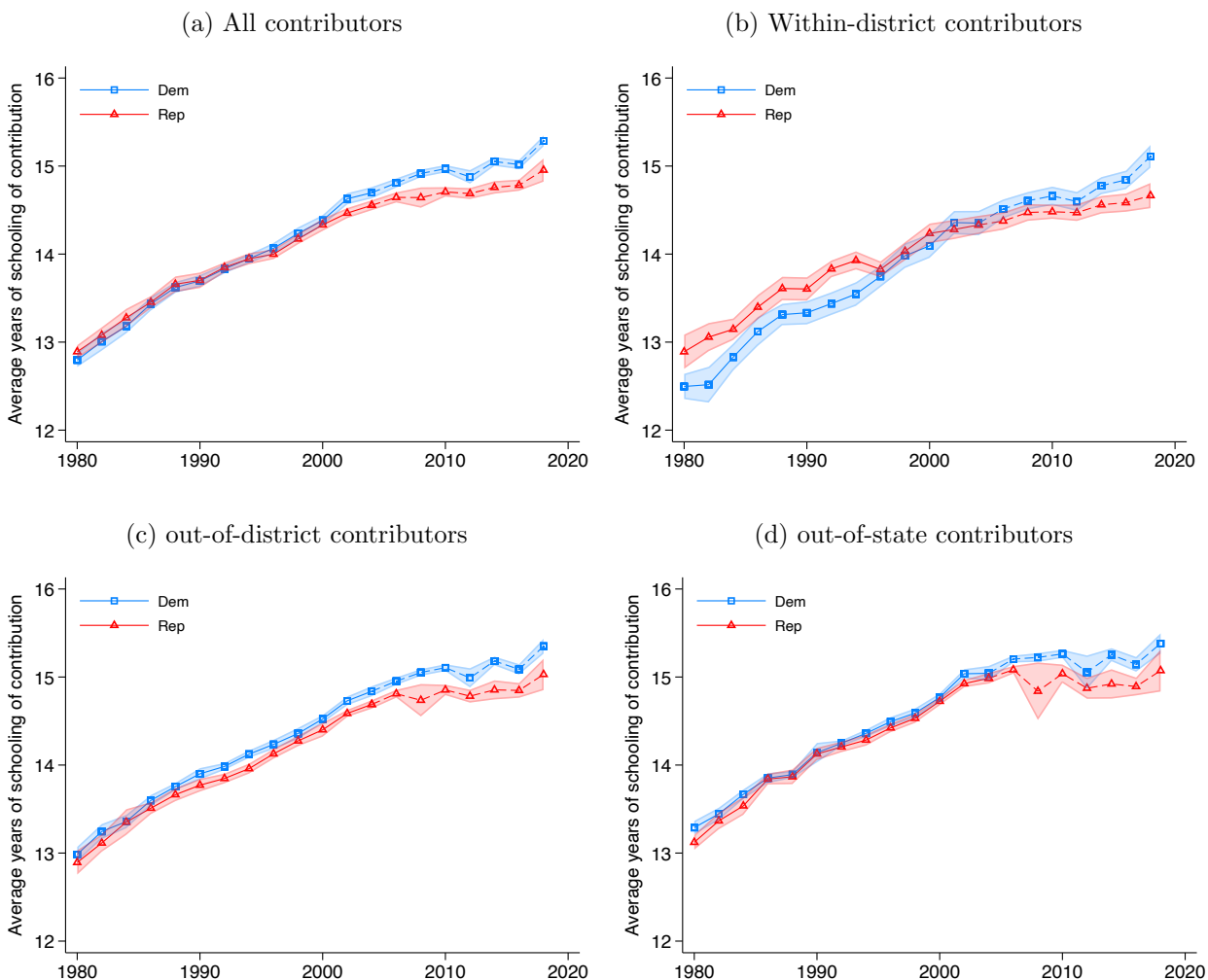
Figure 3: The pre- and re-distribution share of House roll-call votes under Democratic leadership



Source: Comparative Agenda Project (CAP), described in Section 5.1.1.

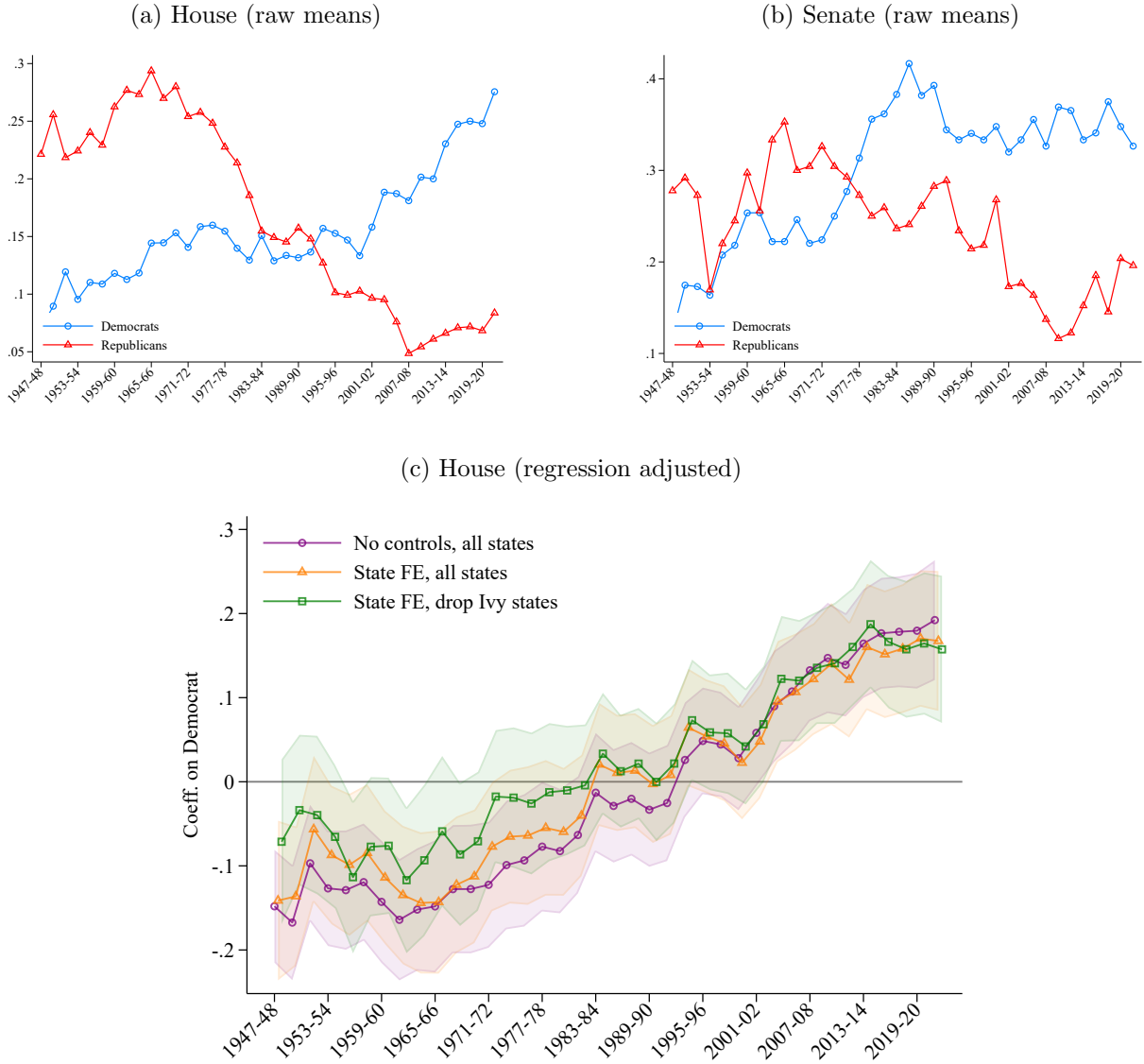
Notes: This figure shows the share of pre- and re-distribution roll-call votes every year that the Democrats are in control of the House from 1947 until 2009. Breaks in the series are when Republicans control the House. Robustness to using alternative classification from Bateman *et al.* (2018) is shown in Appendix Figure D.23. See those sections for how we define pre- and re-distribution but generally predistribution involves labor and employment topics, industrial policy, and public works while redistribution involves taxes, transfer programs and the budget process. We test whether the share of rollcalls devoted to predistribution (redistribution) is different before or after 1977, the first year of the Carter administration. We report the p -value of the test in parenthesis.

Figure 4: Average level of schooling of primary election contributors for House elections.



Source: Contributions data from Bonica (2014) as described in Section 6.1. *Notes:* We plot, in each year, the average years of schooling (based on the 1980 census) at the census tract level for all primary-election donations, separately by party. We weight this average by total contributors (see Appendix Figure A.9 for analogous results, but weighted by total dollars). Panel (a) shows results for *all* primary contributors; (b) for only within-congressional- district contributors (i.e., they are donating to candidates for whom they are eligible to vote); (c) for only out-of-district contributors (i.e., they live outside the district in which their recipient is running); (d) for only out-of-state contributors (i.e., they live outside the state containing the district in which their recipient is running). Appendix Figure A.10 shows the same patterns for the Senate.

Figure 5: Ivy-League backgrounds of members of Congress



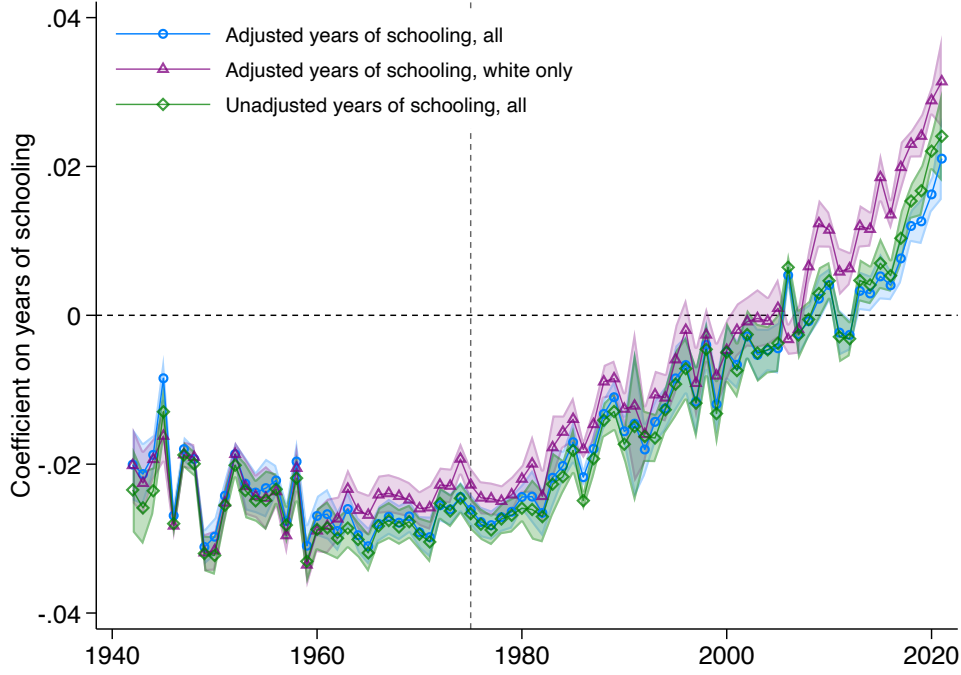
Source: Biographical Directory of the United States Congress.

Notes: Panels (a) and (b) plot the share of House and Senate members, respectively, who attended an Ivy League institution. Panel (c) shows the coefficients β^p from variants of the following regression estimated separately by two-year Congressional term p :

$$Ivy_{is} = \beta^p Democrat_{is} + \mu_s + e_{is},$$

where Ivy_{is} is an indicator variable for whether politician i from state s attended an Ivy League institution; $Democrat_{is}$ is an indicator for whether politician i from state s is a Democrat; and μ_s are state fixed effects. The first series, shown in purple, of panel (c) omits the fixed effects, so just plots the Dem-versus-all-other difference. The second series, shown in orange, includes state fixed effects (which are allowed to vary in an unrestricted manner across time, as we estimate the coefficient separately by two-year term). The third, shown in green, is identical to the second but we drop all states with an Ivy institution.

Figure 6: Democratic Party identification as a function of education



Source: ANES, CCES, Gallup, and GSS as described in Table G.7. We use survey weights (transferred to average to one within each survey) if provided.

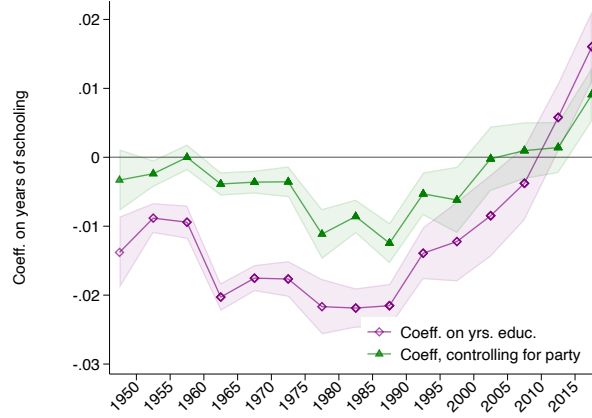
Notes: The first series plots the estimated β_t from the following regression estimated separately for each year t :

$$Democrat_i = \beta_t Adj. \text{ years of school}_i + \mathbf{Age}_i + \mu_{s(i)} + e_i,$$

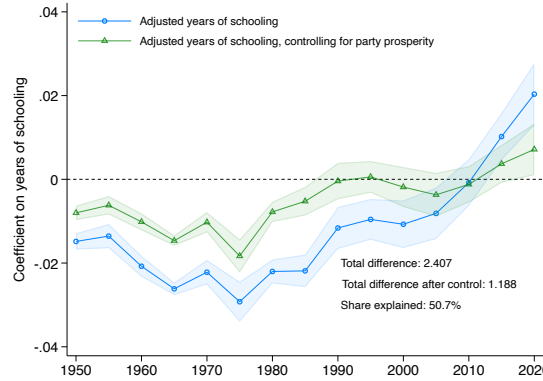
where $Democrat_{is}$ is an indicator for whether person i identifies as a Democrat (as opposed to a Republican, Independent, other or nothing, all coded as zero); $Adjusted \text{ years of school}_i$ is our predicted years of school based on the self-reported educational category provided by respondent i along with demographics and years (see Section 3); \mathbf{Age}_i is a vector of age-in-five-year-bin dummies; $\mu_{s(i)}$ are survey (which subsume date) fixed effects, as we often have several surveys per year (for surveys like the GSS that span years, we have a separate fixed effect for each year). The second series replicates the first but includes only white respondents. The third series replicates the first but instead of using our $Adj. \text{ years of school}$ assigns those with “high school or less” as ten, “some college” as 14 and “college or more” as 16 years of schooling. A Chow test identifies 1976 as the break point in trend.

Figure 7: Opinion of parties' economic policies, by respondent education

(a) Democrats better to keep the country prosperous



(b) Views on economics policy explains half of re-alignment



Source: Gallup as described in Table G.7

Notes: Panel (a) plots the coefficients ω_p from the following regression, estimated separately by five-year period:

$$DemocratsBetter_{it} = \omega_t Adj. \text{ years school}_{it} + \mu_s + \gamma X_{it} + e_{it},$$

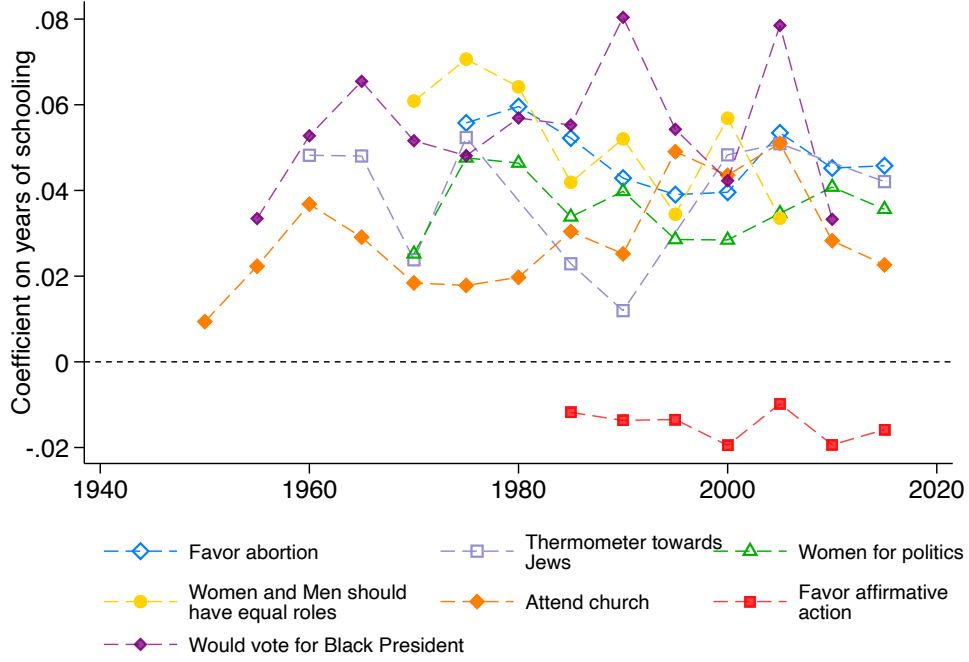
where $DemocratsBetter_{it}$ is an indicator for respondent i answering that the Democratic Party is the best to keep the country prosperous. Panel (b) plots the coefficients of a regression of party identification on years of schooling similar to figure 6:

$$Democrat_{is} = \beta_t Adj. \text{ years of school}_{it} (+ DemocratsBetter_{it}) + \mu_s + \mathbf{Age}_{it} + e_{is},$$

we estimate both the unconditional regression and controlling for the views of respondents on Democratic Party's economic policy. We report the share of the total difference in partisan identification by education that is explained by changes in opinion of parties' economic policies. This share explained can be written as one minus the ratio between the difference in the partisan alignment conditional on economic policy opinion and the difference in the unconditional alignment: $1 - \frac{\widetilde{\beta}_{last} - \widetilde{\beta}_{first}}{\beta_{last} - \beta_{first}}$ with $\widetilde{\beta}_\tau$ the conditional estimator and β_τ the unconditional estimator. The *last* years are 2001-2020 and the *first* years are 1948-1967.

Figure A.19 in Appendix shows the share of respondent answering that the Democratic or the Republican Party is the best to keep the country prosperous.

Figure 8: Social Issue Preferences by education (white only)



Source: ANES, Gallup, and GSS as described in Table G.7

Notes: As in Figure 1, this figure plots the coefficients β^p from the following regression estimated separately for each five-year period p :

$$y_{is} = \beta^p \text{Adj. years of school}_i + \mu_s + \mathbf{Age}_i + e_{is},$$

where y_{is} is the outcome variable for person i in survey s ; *Adjusted years of school* is our predicted years of school based on the self-reported educational category provided by the respondent along with demographics and years (see Section 3), μ_s are survey (which subsume date) fixed effects, as we often have several surveys per period p and \mathbf{Age}_i is a vector of age-in-five-year-bin dummies. The outcome variables are standardized measures (mean zero and variance 1) of support for the given policy. Detailed survey questions for each policy are displayed in Appendix B. We suppress confidence intervals to reduce clutter, but we plot each policy question separately along with 95-percent confidence intervals in the Appendix. We restrict the sample to white respondents only.

10 Tables

Table 1: Democratic House members position depending on DLC status

	Dept Var: Ideal Points (incr. in conservativeness)					
	(1)	(2)	(3)	(4)	(5)	(6)
DLC	0.091*** (0.005)	0.359*** (0.020)	0.264*** (0.031)	0.211*** (0.031)	0.268*** (0.046)	0.106*** (0.005)
Predistribution x DLC						0.267*** (0.016)
Redistribution x DLC						0.191*** (0.015)
Ideal Point Type	DW-N 1	Predis	Predis	Redis	Social	DW-N 1, Predis, Redis
Congress FE	X	X				
State x Congress FE			X	X	X	X
Average difference between parties	.772	1.805	1.805	1.901	1.755	1.486
Predistribution x DLC -						0.076***
Redistribution x DLC						(0.022)
Observations	3,404	3,348	3,226	3,237	3,230	10,110

Notes: The table shows the position of House members as defined by DW-Nominate or their topic-specific ideal points. We use the rollcall topic-classification described in Appendix section D and the topic-specific ideal point estimation, described in Appendix section F. Column 1 regresses the DW-Nominate first dimension on a DLC dummy, controlling for Congress fixed effects, Column 2 regresses the predistribution ideal point, controlling for congress FE, while column 3 onwards control for state by congress FE. Column 4 looks at the redistribution ideal point and column 5 at the ideal point on social issues. Column 6 pools together the DW Nominate first dimension, the predistribution and the redistribution ideal point and looks at the relative effect of DLC on economic policies, with respect to DW-Nominate first dimension. We report the average difference between the parties in the table notes and we test the significance of the difference between predistribution and redistribution in column 6. Standard errors clustered by Congress. The sample is restricted to Democratic House members. The same table using the Bateman et al. classification is displayed in Appendix Table D.4.

* p<0.1, ** p<0.05, *** p<0.001. Standard errors clustered by congress in parentheses.

Table 2: Democratic House members' alignment with Republican voting

	Dept Var: Voted Yea			
	(1)	(2)	(3)	(4)
DLC \times MeanRepub	0.148*** (0.002)	0.141*** (0.002)	0.142*** (0.002)	0.145*** (0.002)
DLC \times Predis		-0.033*** (0.004)	-0.032*** (0.004)	-0.032*** (0.004)
DLC \times Predis \times MeanRepub		0.051*** (0.006)	0.050*** (0.006)	0.049*** (0.006)
DLC \times Redis			0.010 (0.007)	0.009 (0.007)
DLC \times Redis \times MeanRepub			-0.010 (0.009)	-0.010 (0.009)
Rollcall FE	X	X	X	X
State \times Congress FE				X
Linear combination of estimates: Predis - Redis			0.060*** (0.010)	0.059*** (0.010)
Observations	4,031,972	4,031,972	4,031,972	4,031,972

Notes: The table shows the extent to which Democrats are voting with Republicans. The dependent variable is a dummy equal to one if the House member votes Yea. The independent variable *MeanRepub* is the share of Republicans House members who voted Yea. *DLC* is a dummy variable equal to one if the House member is part of the DLC caucus. *Predis* (resp. *Redis*) are dummy variables equal to one if the vote is about predistribution (resp. redistribution), as classified by CAP. All Columns control for rollcall fixed effects and Column 5 control for state and year fixed effects. We show the difference between the coefficient on DLC \times Predis \times MeanRepub and DLC \times Redis \times MeanRepub in the Table statistics. The equivalent table with the Bateman et al. classification is displayed in Appendix Table D.5. ** p<0.1, *** p<0.05, **** p<0.001. Standard errors clustered by Rollcall in parentheses.

Table 3: Primary donors' census tract average education by candidate DLC status

Panel A: Weighted by distinct contributors

	All contributions		Within district		Out-of-district	
	(1)	(2)	(3)	(4)	(5)	(6)
DLC	0.089*** (0.020)	0.057*** (0.018)	0.002 (0.049)	0.008 (0.037)	0.110*** (0.017)	0.069*** (0.019)
Year x State FE	X	X	X	X	X	X
Sample	Winners	All	Winners	All	Winners	All
ML type		RF		RF		RF
Number of DLC candidates	1816	2889	1652	2520	1799	2766
Mean of dependent variable	14.693	14.681	14.372	14.353	14.823	14.816
Observations	585,847	887,383	101,330	161,253	484,499	726,114

Panel B: Weighted by amount

	All contributions		Within district		Out-of-district	
	(1)	(2)	(3)	(4)	(5)	(6)
DLC	0.046** (0.020)	0.038** (0.016)	-0.026 (0.050)	0.019 (0.038)	0.066*** (0.016)	0.047*** (0.013)
Year x State FE	X	X	X	X	X	X
Sample	Winners	All	Winners	All	Winners	All
ML type		RF		RF		RF
Number of DLC candidates	1816	2889	1652	2520	1799	2766
Mean of dependent variable	14.757	14.735	14.382	14.36	14.895	14.881
Observations	584,977	884,778	101,247	160,851	483,712	723,911

Notes: The Table shows the difference in the average education level of Primary contributors giving to DLC candidates compared to non-DLC Democratic candidates. We regress the average years of schooling of each contributor for the Democrats at the Primaries on whether the candidates are affiliated with the DLC. Panel A weight the results by distinct contributors while panel B weight them by amount. Columns 1, 3, and 5 of both panels use only winners of the general elections, for whom we have the actual DLC status. Columns 2, 4, and 6 of both panel use our 2-step Random Forest prediction algorithm as described in Appendix section E. Education of contributors is measured by their census tract average education level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by candidate by year.

Table 4: Votes for “new” versus “old” Democrats in hypothetical general-election match-ups

	Nem Dem	Old Dem	New Dem minus Old style Dem				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Years educ.	-0.00881*** (0.000896)	-0.0119*** (0.00106)	0.00541*** (0.00109)	0.00534*** (0.00109)	0.00366*** (0.00109)	0.00353*** (0.00116)	
Female					0.00142 (0.00625)		
White					0.100*** (0.0104)		
Years educ. x 1972							0.00846*** (0.00204)
Years educ. x 1976							-0.00758*** (0.00286)
Years educ. x 1980							0.0128*** (0.00253)
Years educ. x 1984							0.00405* (0.00231)
Years educ. x 1988							0.00870*** (0.00306)
Years educ. x 1992							0.00436* (0.00234)
Dept. var. mean	0.438	0.352	0.0752	0.0752	0.0752	0.0873	0.0752
Sample	All	All	All	All	All	Whites	All
State FE	X	X	X			X	X
State x Election FE				X	X		
Observations	41365	28134	28205	28202	28202	24742	28205

Notes: Column 1 shows the probability to vote for a “New Democrat” rather than a Republican as a function of years of education. Column 2 shows the probability to vote for an “Old-style Democrat” and Columns 3 to 6 show the difference in the probability to vote for the Democratic candidate versus the Republican candidate if the Democratic nominee is a “New Democrat” versus an “Old-style Democrat”. In 1972, the dependent variable is equal to $(vote_{McGovern} - vote_{Nixon}) - (vote_{Kennedy} - vote_{Nixon})$. In 1976, the dependent variable is equal to $(vote_{Carter} - vote_{Ford}) - (vote_{Humphrey} - vote_{Ford})$. In 1980, the dependent variable is equal to $(vote_{Carter} - vote_{Reagan}) - (vote_{Kennedy} - vote_{Reagan})$. In 1984, the dependent variable is equal to $(vote_{Mondale} - vote_{Reagan}) - (vote_{Hart} - vote_{Reagan})$. In 1988, the dependent variable is equal to $(vote_{Dukakis} - vote_{Bush}) - (vote_{Jackson} - vote_{Bush})$. In 1992, the dependent variable is equal to $(vote_{Clinton} - vote_{Bush}) - (vote_{Brown} - vote_{Bush})$. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

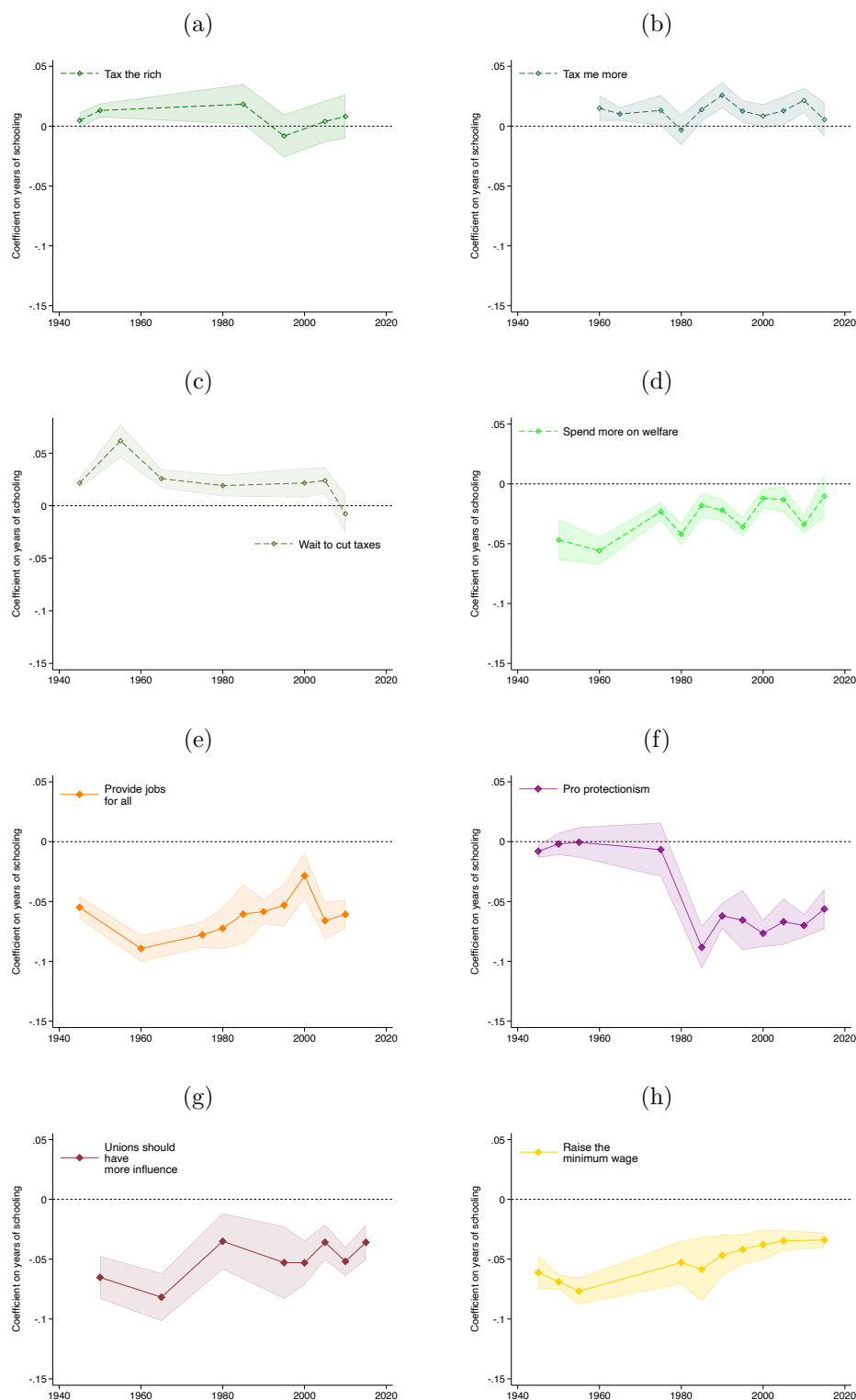
Table 5: House Democratic vote share by education of neighborhood

	Dept Var: Share Dem					
	(1)	(2)	(3)	(4)	(5)	(6)
Educ yrs	-0.014*** (0.005)	-0.011** (0.005)	-0.010*** (0.002)	-0.009*** (0.001)	-0.009*** (0.001)	-0.008*** (0.001)
Years educ. \times DLC	0.009*** (0.003)	0.010*** (0.003)	0.006*** (0.002)	0.005** (0.002)	0.005*** (0.002)	0.004** (0.002)
DLC	-0.182*** (0.040)		-0.003 (0.008)		-0.017* (0.010)	
Share white					-0.298*** (0.017)	-0.301*** (0.017)
Share white \times DLC					0.057*** (0.015)	0.066** (0.027)
Share below 35					0.036*** (0.008)	0.036*** (0.007)
Share below 35 \times DLC					-0.008 (0.022)	0.003 (0.021)
Year FE	X		X		X	
District FE	X		X		X	
District \times Year FE		X		X		X
Sample	Winners	Winners	All	All	All	All
ML type			RF	RF	RF	RF
Controls	Basic	Basic	Basic	Basic	Extended	Extended
Number of DLC candidates	211	211	306	306	306	306
Mean of dependent variable	.669	.669	.529	.529	.529	.529
Observations	22,210	22,243	53,262	53,208	53,262	53,208

Notes: The table shows the Democratic vote share obtained by House Democratic Candidates in MCD-groups for the period 1984–1990. We regress the share of votes obtained by the Democratic candidate for the House in each MCD-group on the average years of education (*Years educ.*) of that MCD-group and the interaction of this term with *DLC*, a dummy equals to one if the Democratic candidate is part of the DLC. Column 1, 3, and 5 include year and district fixed effects while column 2, 4, and 6 include district by year fixed effects, meaning that we only at variation within congressional district for each year, for a given candidate. Any non-varying candidate attribute (such as *DLC*) is therefore subsumed by the fixed effects. Columns 1 and 2 keep only elected Democratic candidates since we do not have any information on caucus membership on those who lost at the general election. Columns 3 to 6 include predictions from out machine learning algorithm, as described in Appendix section E. Standard errors are clustered by district. We exclude the bottom and top 1% of MCD groups in terms of poulation. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix A. Supplementary Figures and Tables Noted in the Text

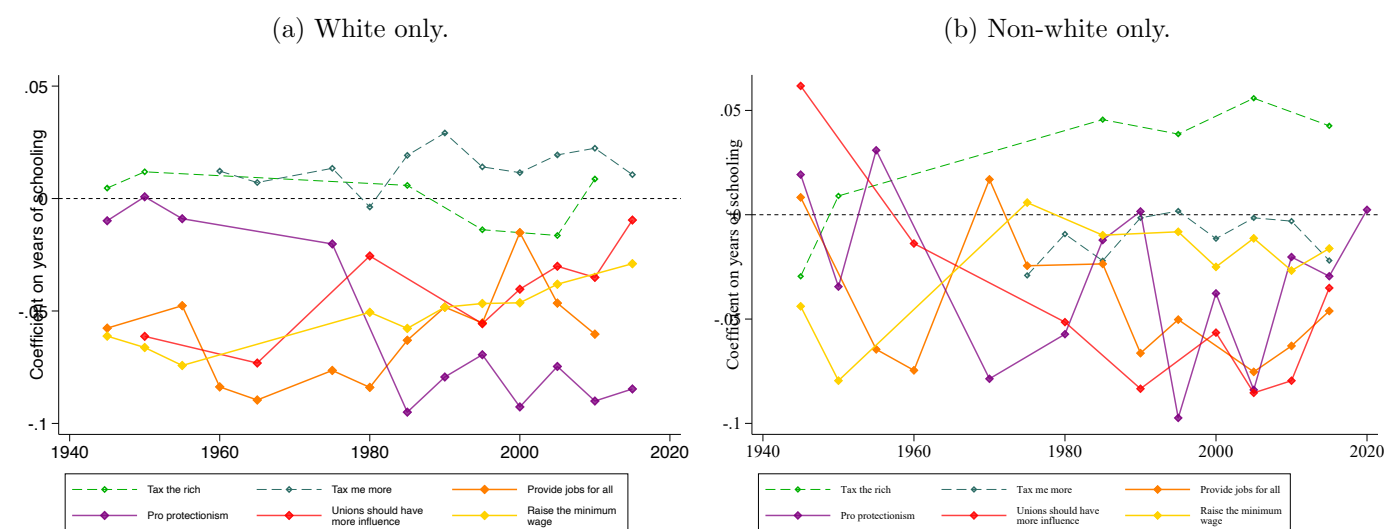
Appendix Figure A.1: Preferences for pre- and re-distribution by education (question by question and including confidence intervals)



Source: ANES, Gallup, and GSS as described in Table G.7

Notes: As in Figure 1, this Figure shows the coefficients β^p from a regression of for each five-year period from specification 1 for each standardized survey question separately.

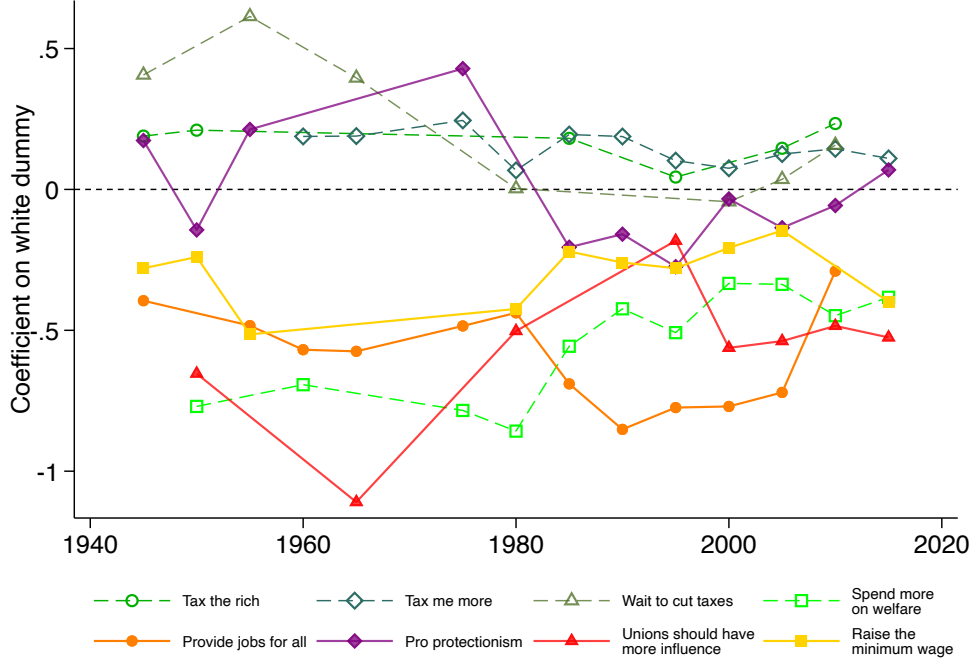
Appendix Figure A.2: Preferences for pre- and re-distribution by education (separately for white and non white)



Source: ANES, Gallup, and GSS as described in Table G.7

Notes: This figure replicates Figure 1 but includes only white respondents (left panel) and only non-white respondents (right panel).

Appendix Figure A.3: Preferences for pre- and re-distribution by race

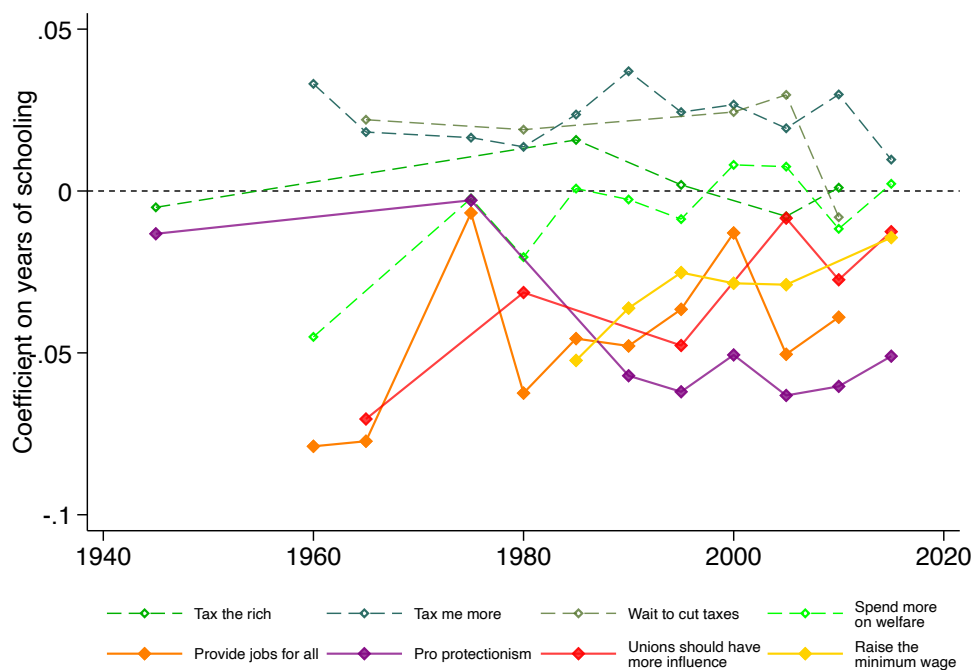


Notes: This figure replicates Figure 1 but instead of years of schooling as the main explanatory variable, a dummy variable for identifying as white is the main explanatory variables. That is, the figure plots the estimated β^p from the following regression estimated separately for each five-year period p :

$$y_{is} = \beta^p White_i + \mu_s + Age_i + e_{is},$$

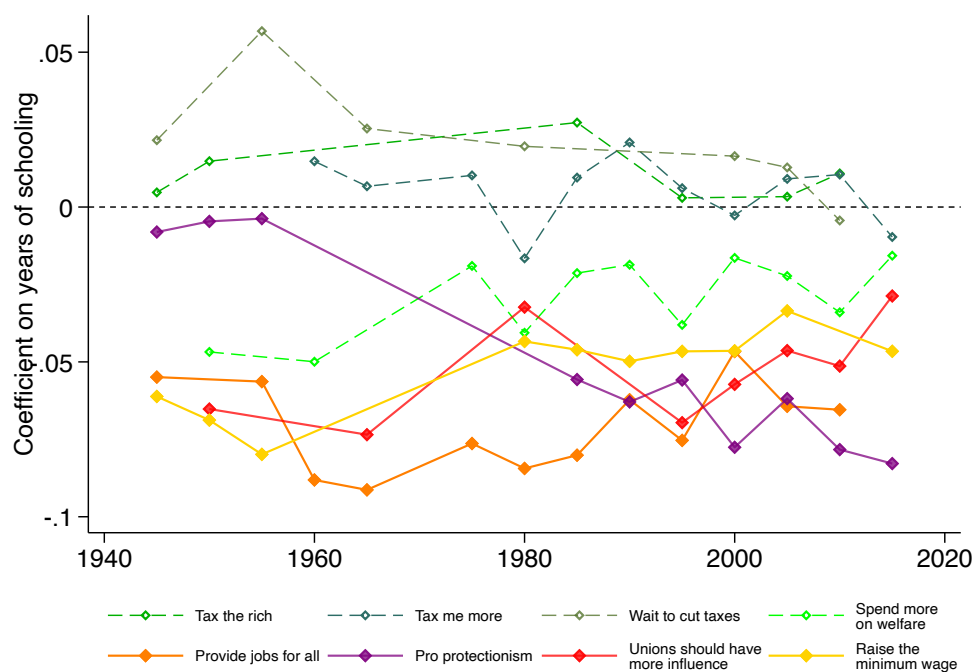
where y_{is} is the outcome variable for person i in survey s ; *Adjusted years of school* is our predicted years of school based on the self-reported educational category provided by the respondent along with demographics and years (see Section 3), μ_s are survey (which subsume date) fixed effects, as we often have several surveys per period p and Age_i is a vector of age-in-five-year-bin dummies. The outcome variables are standardized measures (mean zero and variance 1) of support for the given policy. Detailed survey questions for each policy are displayed in Appendix B.

Appendix Figure A.4: Preferences for pre- and re-distribution, controlling for income, gender and age.



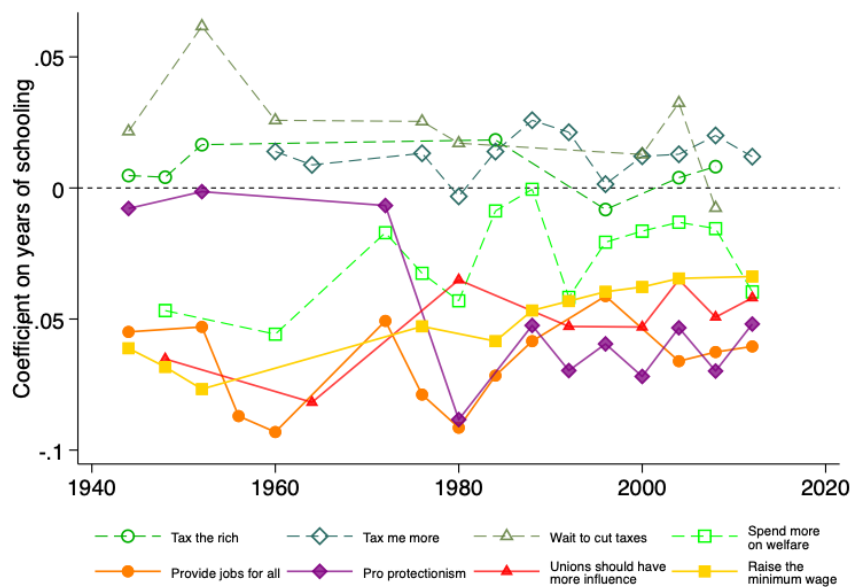
Notes: This figure replicates Figure 1 but flexibly controls for income, gender and age by five-year period. Detailed survey questions for each policy are displayed in Appendix B.

Appendix Figure A.5: Preferences for pre- and re-distribution, holding the composition of the population as fixed.



Notes: This figure replicates Figure 1 but holds the composition of the population as fixed in terms of education and race. Detailed survey questions for each policy are displayed in Appendix B.

Appendix Figure A.6: Preferences for pre- and re-distribution, controlling for text features.



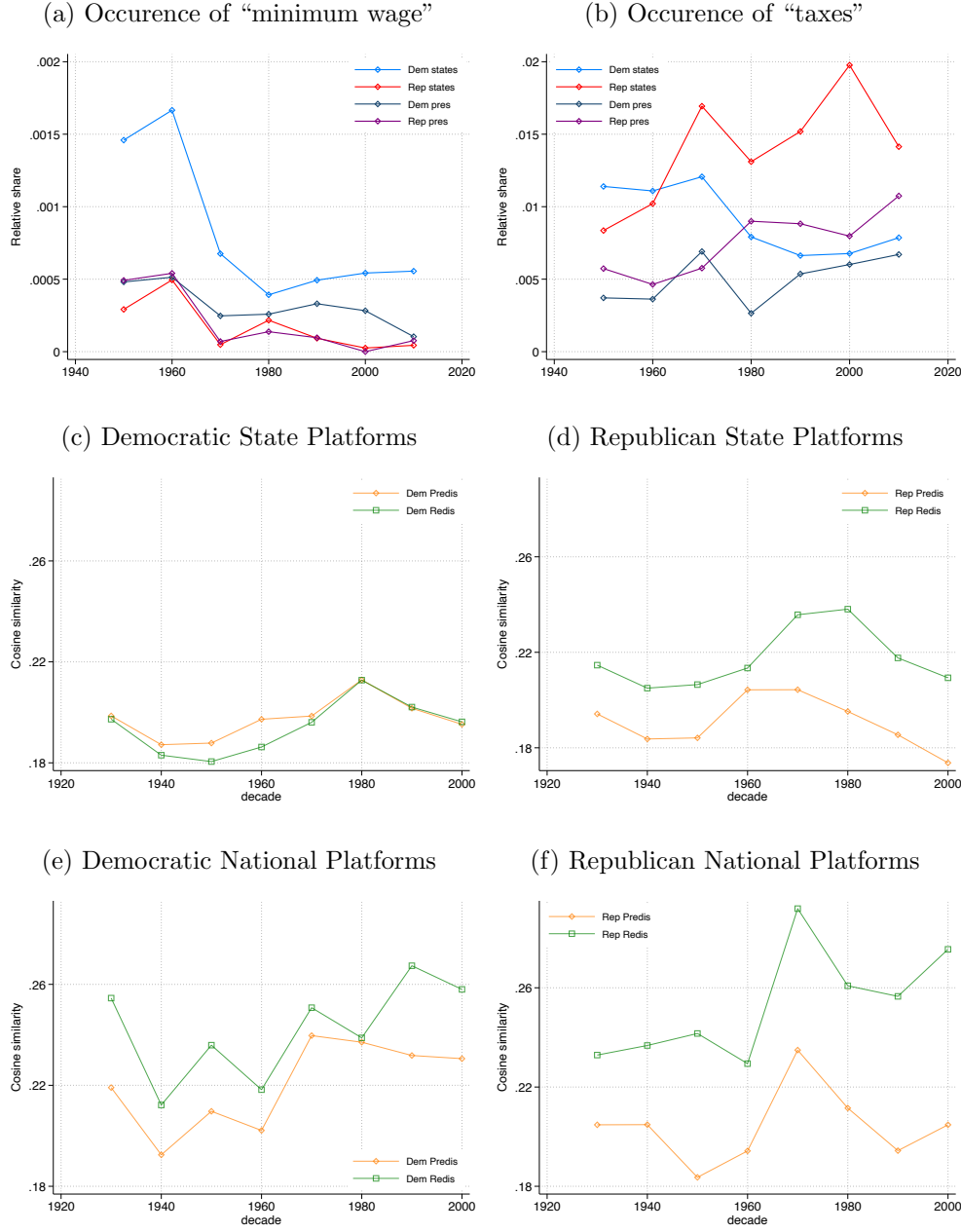
Notes: This figure replicates Figure 1 but flexibly controls for text features: question sentiment and redability as well as principal components of questions as predicted by SBERT (see Appendix G for more details).

Appendix Figure A.7: The DLC share of Democratic Senators



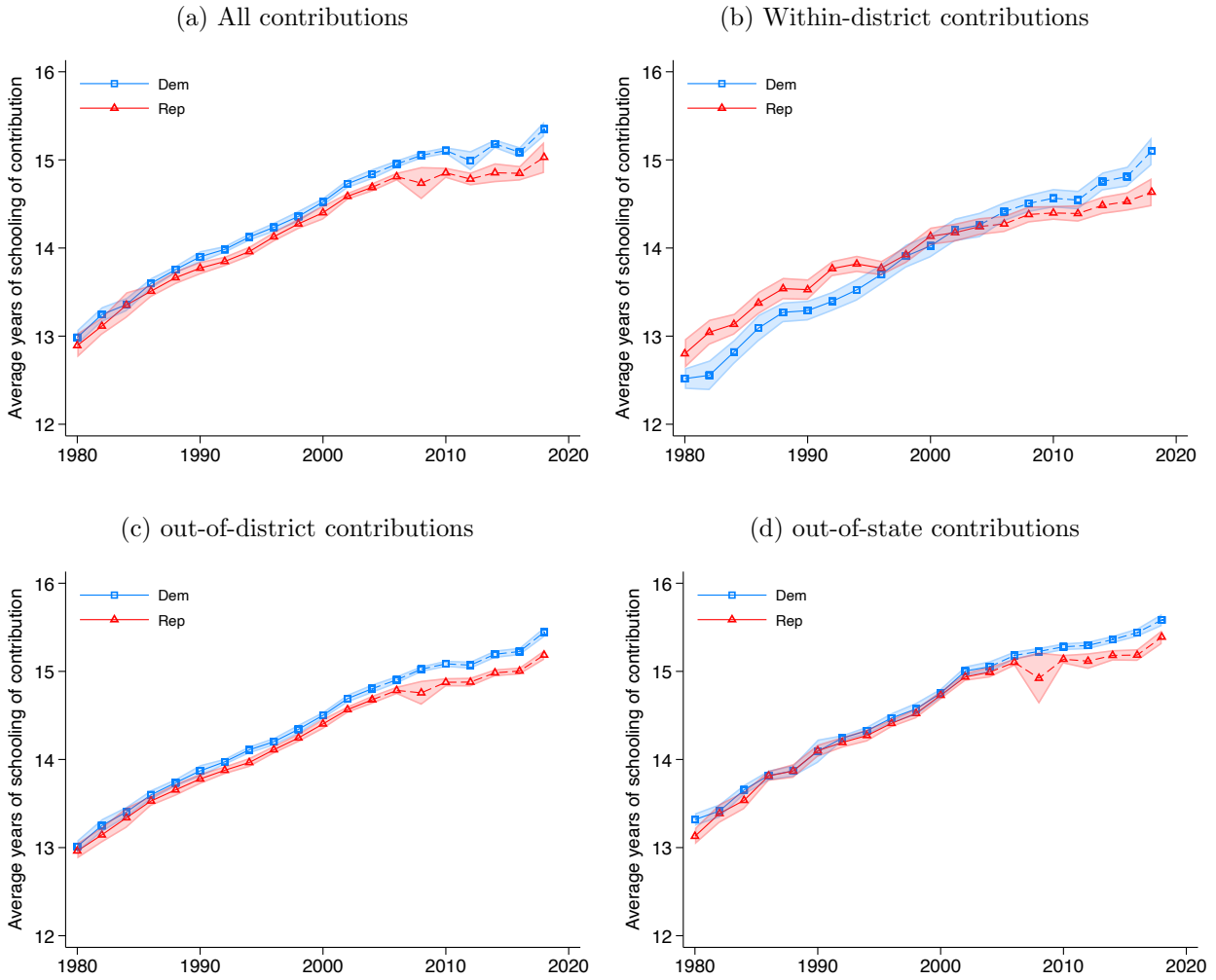
Notes: As in Figure 2, this figure shows the evolution of the share of Democrats in the Senate who are members of the New Democratic Caucus. We combine several sources for caucus membership, detailed in Appendix Section E. The solid bars show the years for which we actually have records of membership. The yellow line shows the raw evolution of (interpolated) DLC membership while the orange line shows the evolution of the share of DLC members for politicians who were seated in the 100th Congress, elected in 1986 to control for compositional effect. A politician is defined as DLC if they are ever listed as a member or are in the New Dem Caucus. Given data collection, DLC is likely underestimated from 1987 – 1996.

Appendix Figure A.8: Changing Language of Party Platforms



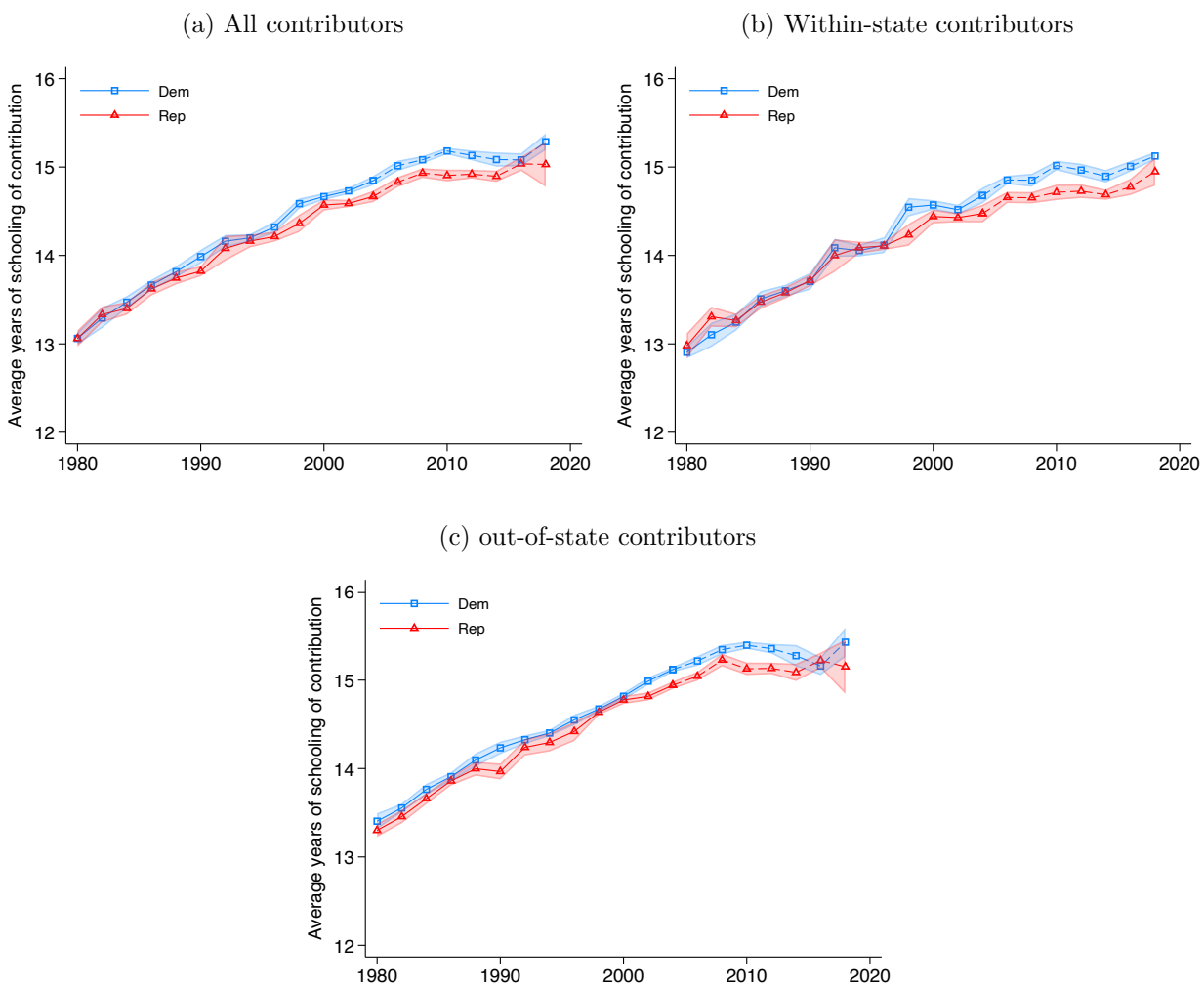
Notes: Data from Hopkins *et al.* (2022). Panel (a) and (b) show the relative occurrences of the words “minimum wage” and “tax” in a balanced panel of state platform and the Presidential platforms for each party. Panel (c), (d), (e), and (f) plots the average cosine similarity between the SBERT embedding vector of survey question text q and party p 's platform in year y . Panel (c) and (d) look only at state platforms while panels (e) and (f) look at national party platforms. It shows that the national Democratic platform has diverged more from the language of the redistribution questions in our surveys relative to both Republicans and the redistribution questions.

Appendix Figure A.9: Average level of schooling of Primary dollars for House elections.

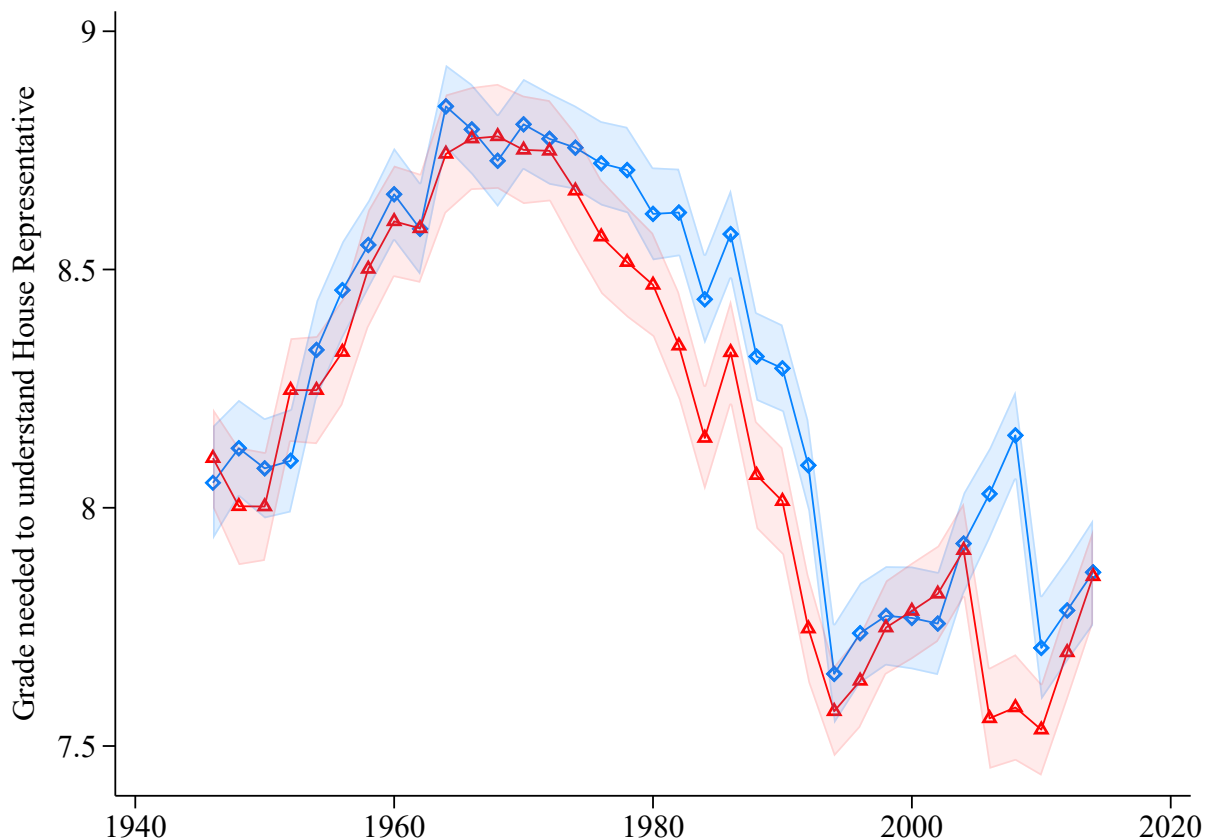


Notes: As in Figure 4, This figure shows the average level of education of each Primary contribution, weighted by amount, panel (a) shows all the contributors, panel (b) shows the within congressional district donors, panel (c) shows the out-of-district donors, and panel (d) shows the out-of-state donors. We use the census tract level average level of education. Contributions data from Bonica (2014) as described in section 6.1.

Appendix Figure A.10: Average level of schooling of Primary contributors for Senate elections.



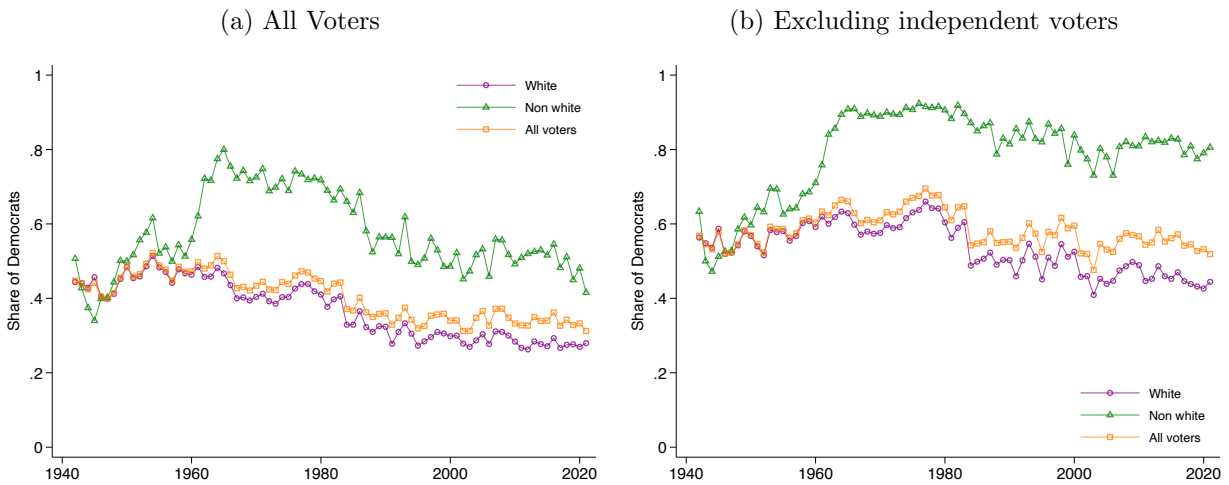
Notes: As in Figure 4, This figure shows the average level of education of each Primary contributor, panel (a) shows all the contributors, panel (b) shows the within state donors, panel (c) shows the out-of-state donors. We use the census tract level average level of education. Contributions data from Bonica (2014) as described in section 6.1.



Appendix Figure A.11: Average school grade needed to understand U.S. House Representative speeches.

Notes: The Congressional Speeches data comes from Gentzkow *et al.* (2019). For each speech, we compute the average school grade needed to understand the speech given by Members of Congress. We show the average across Members of the weighted average grade of their speeches. We exclude speeches of less than ten words. We use the average grade between the Flesch-Kincaid Level, the Fog Scale, the SMOG Index, the Automated Readability Index, the Coleman-Liau Index, the Linsear Write Index, and the Dale-Chall Score. The first year for which the difference between Democratic and Republican is statistically significant at the 5% level is 1976.

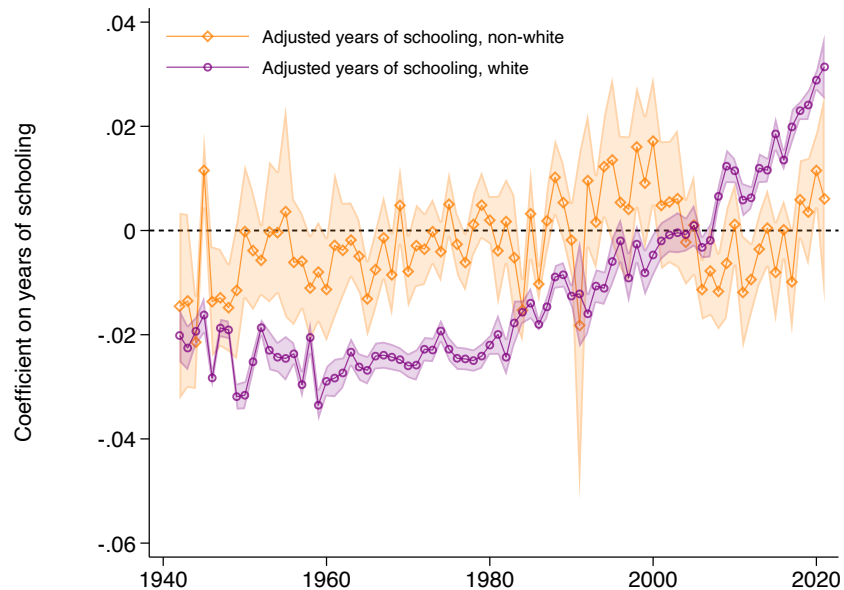
Appendix Figure A.12: Democratic Party identification over time.



Source: ANES, CCES, Gallup, and GSS as described in Table G.7

Notes: This figure shows the average share of respondents who identify as Democrats by race, panel (a) shows the raw shares among all respondents while panel (b) excludes independent respondents.

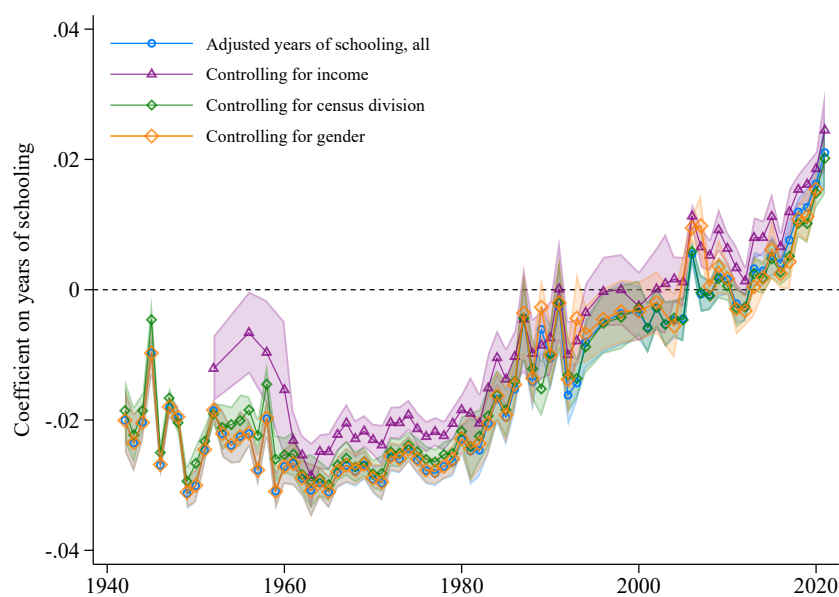
Appendix Figure A.13: Democratic Party identification by education, whites and nonwhites



Source: ANES, CCES, Gallup, and GSS as described in Table G.7

Notes: As in Figure 6, this Figure shows the estimated β^t from a regression of an indicator for Democratic affiliation on our *Adjusted years of schooling* measure. The first series, shown in orange, restricts the sample to respondents identifying as non-white. The second series shows the same coefficients for respondents identifying as white.

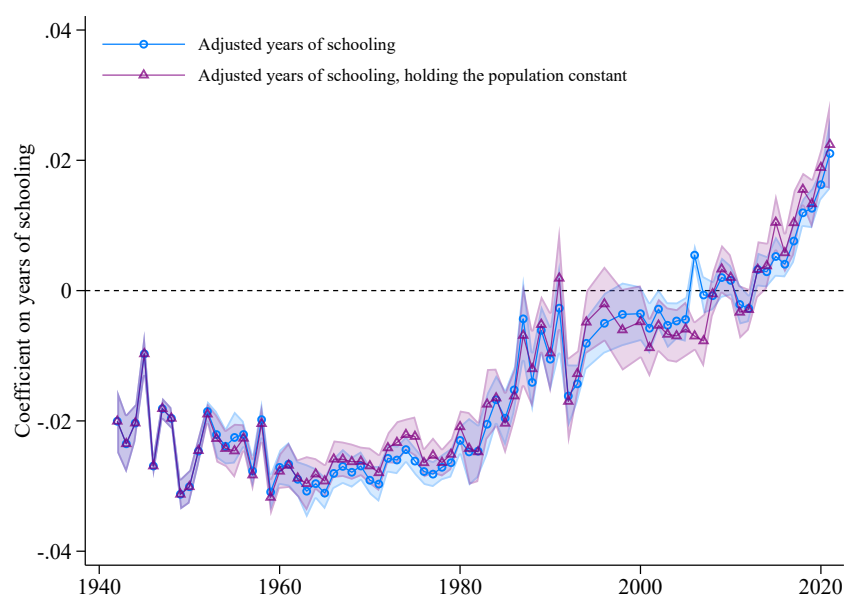
Appendix Figure A.14: Democratic Party identification controlling for basic covariates



Source: ANES, CCES, Gallup, and GSS as described in Table G.7

Notes: As in Figure 6, this Figure shows the estimated β^t from a regression of an indicator for Democratic affiliation on our *Adjusted years of schooling* measure. The first series, shown in blue, shows the unconditional effect of education, controlling only for age. The second series, shown in purple, shows the same coefficients flexibly controlling for income. The third series, shown in green, shows the same coefficients, flexibly controlling for census divisions. The fourth series, shown in orange, shows the same coefficients, flexibly controlling for gender.

Appendix Figure A.15: Democratic Party identification holding fixed population composition

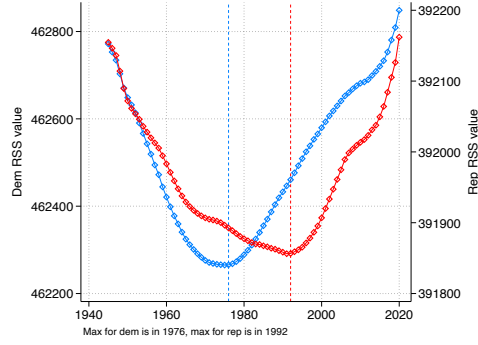


Source: ANES, CCES, Gallup, and GSS as described in Table G.7

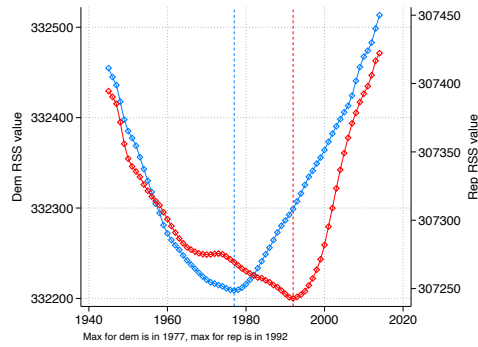
Notes: As in Figure 6, this Figure shows the estimated β^t from a regression of an indicator for Democratic affiliation on our *Adjusted years of schooling* measure. The first series, shown in blue, shows the effect of education. The second series, shown in purple, shows the effect of education, holds the educational and racial composition of the population constant.

Appendix Figure A.16: Estimating inflection points in partisan realignment

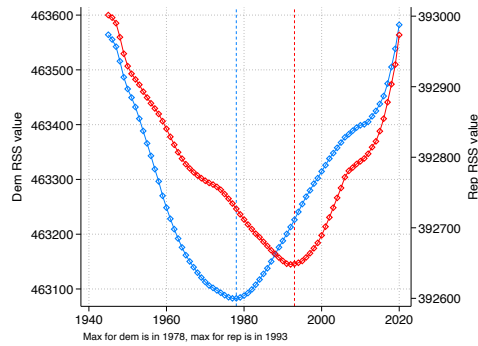
(a) All respondents



(b) Whites only



(c) All respondents, unadjusted years of schooling

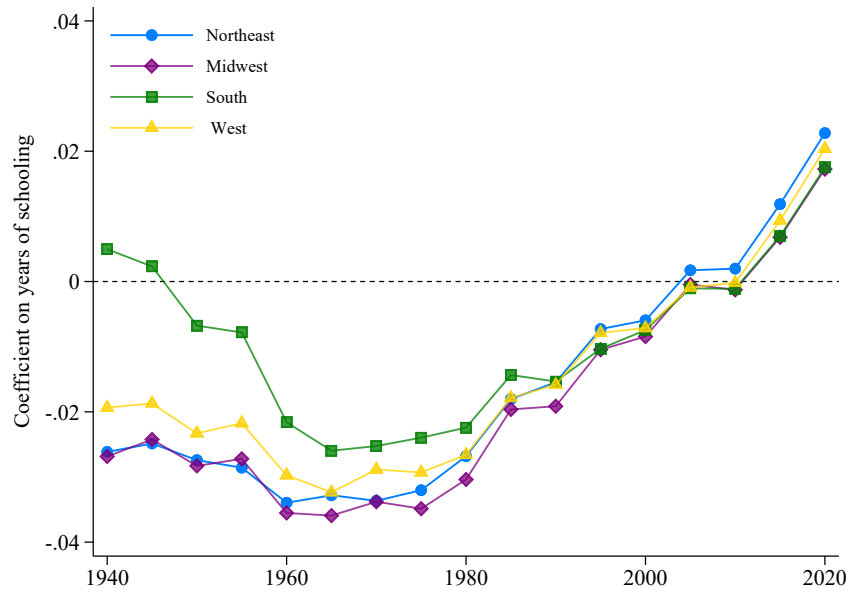


Source: ANES, CCES, Gallup, and GSS as described in Table G.7

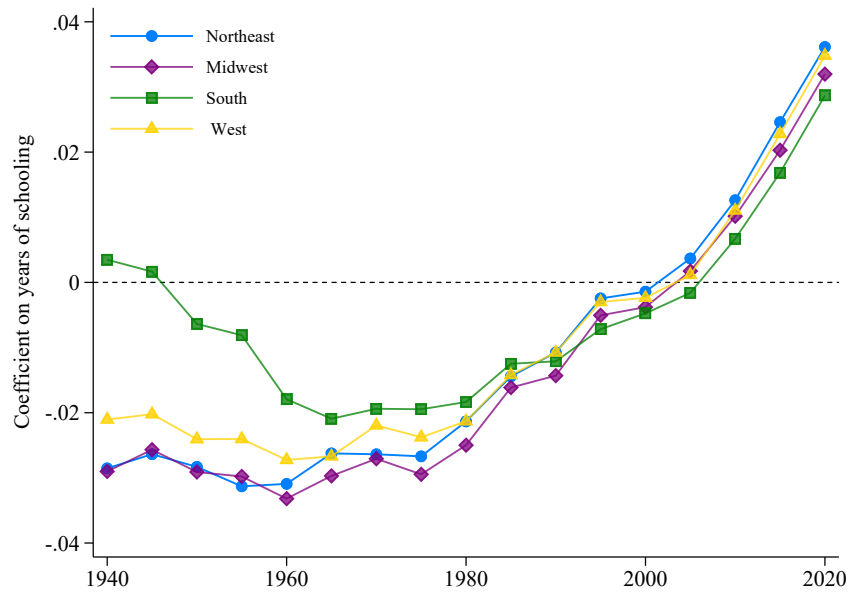
Notes: The graph shows the residual sum of squares from regressions with a linear trend cutoff at each year. Survey-year fixed effects, flexibly controlling for age by year. We normalize the survey weights so that each year has the same weights in the regressions.

Appendix Figure A.17: Democratic Party identification by education, by region

(a) All respondents



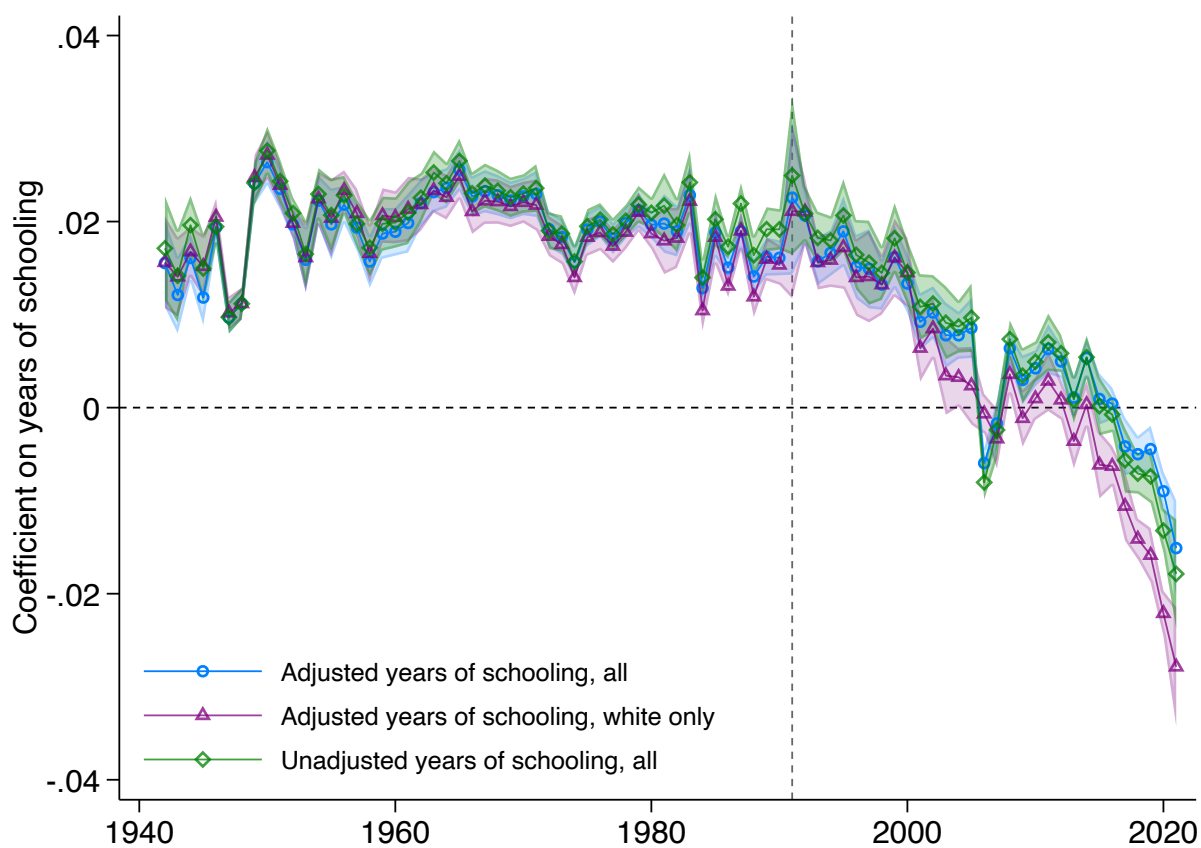
(b) Whites only



Source: ANES, CCES, Gallup, and GSS as described in Table G.7

Notes: As in Figure 6, this Figure shows the estimated β^t from a regression of an indicator for Democratic affiliation on our *Adjusted years of schooling* measure. The four series show the coefficients β^p for each census region separately. Panel (a) shows the results for all the population while panel (b) displays the results, restricting the sample to respondents identifying as white.

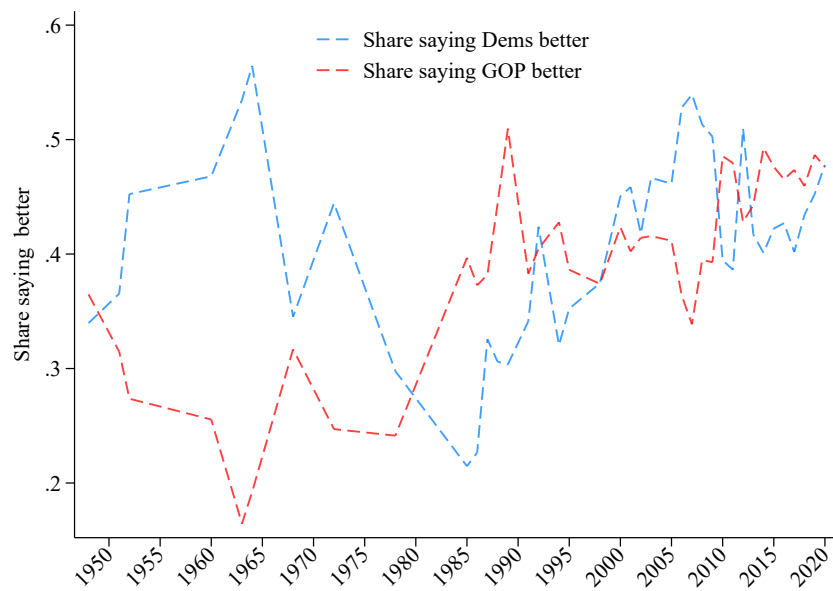
Appendix Figure A.18: Republican Party identification by education



Source: ANES, CCES, Gallup, and GSS as described in Table G.7

Notes: As in Figure 6, this Figure shows the estimated β^p from a regression of an indicator for Republican affiliation on our *Adjusted years of schooling* measure.

Appendix Figure A.19: Respondents' views of the parties' economic policies

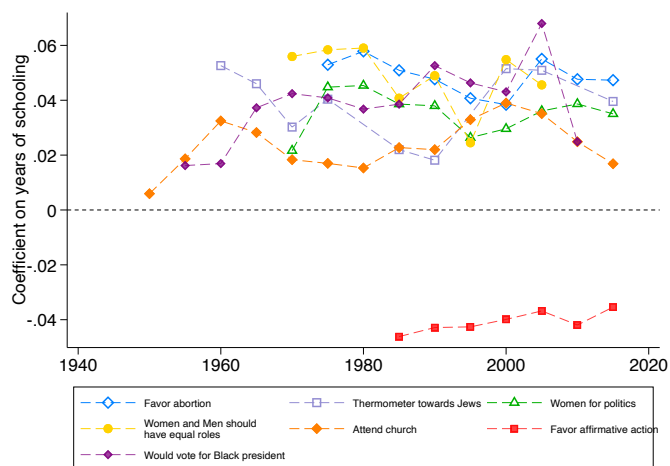


Notes: This figure plots the coefficients β^p from the following regression, estimated separately by five-year period:

$$Democrats_{it} = \beta^t Adj. \text{ years school}_i + \gamma X_i + e_{it},$$

where $Democrats_{it}$ is an indicator for respondent i answering that the Democratic Party is the best to keep the country prosperous.

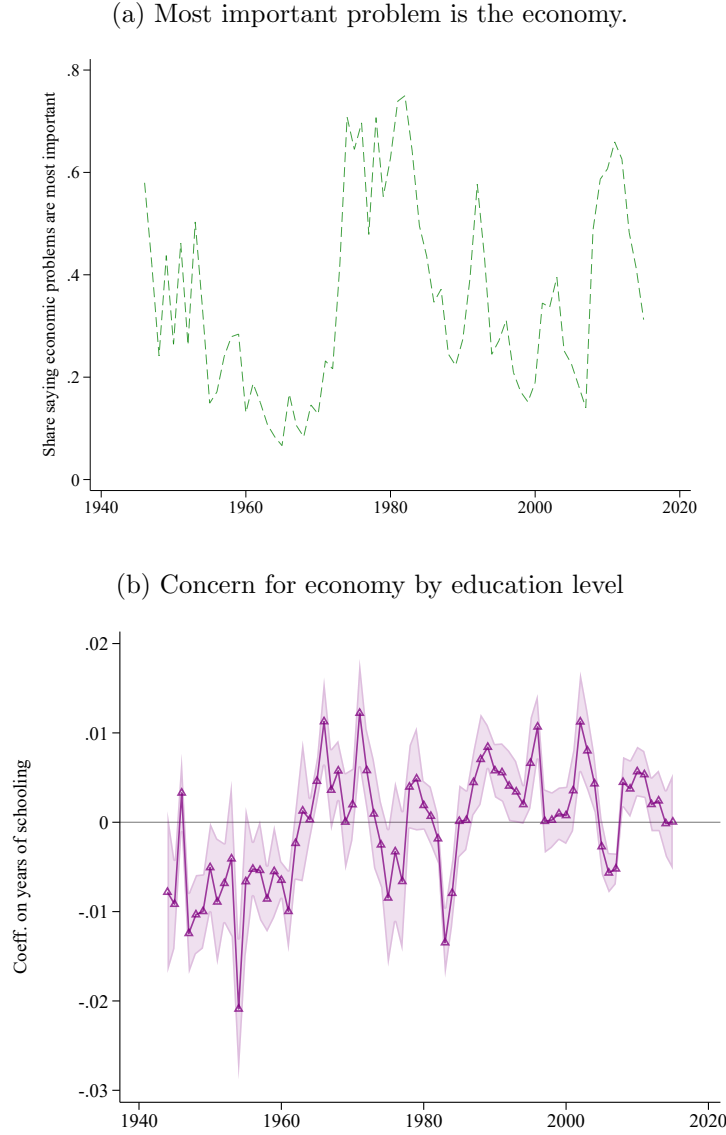
Appendix Figure A.20: Social issue preferences by education (all respondents, including non-whites)



Source: ANES, CCES, Gallup, and GSS as described in Table G.7

Notes: As in Figure 8, this figure shows coefficients β^p for each five-year period from specification 1 for each standardized survey question. Detailed survey questions are displayed in Appendix B.

Appendix Figure A.21: Economic-issue salience over time and by education



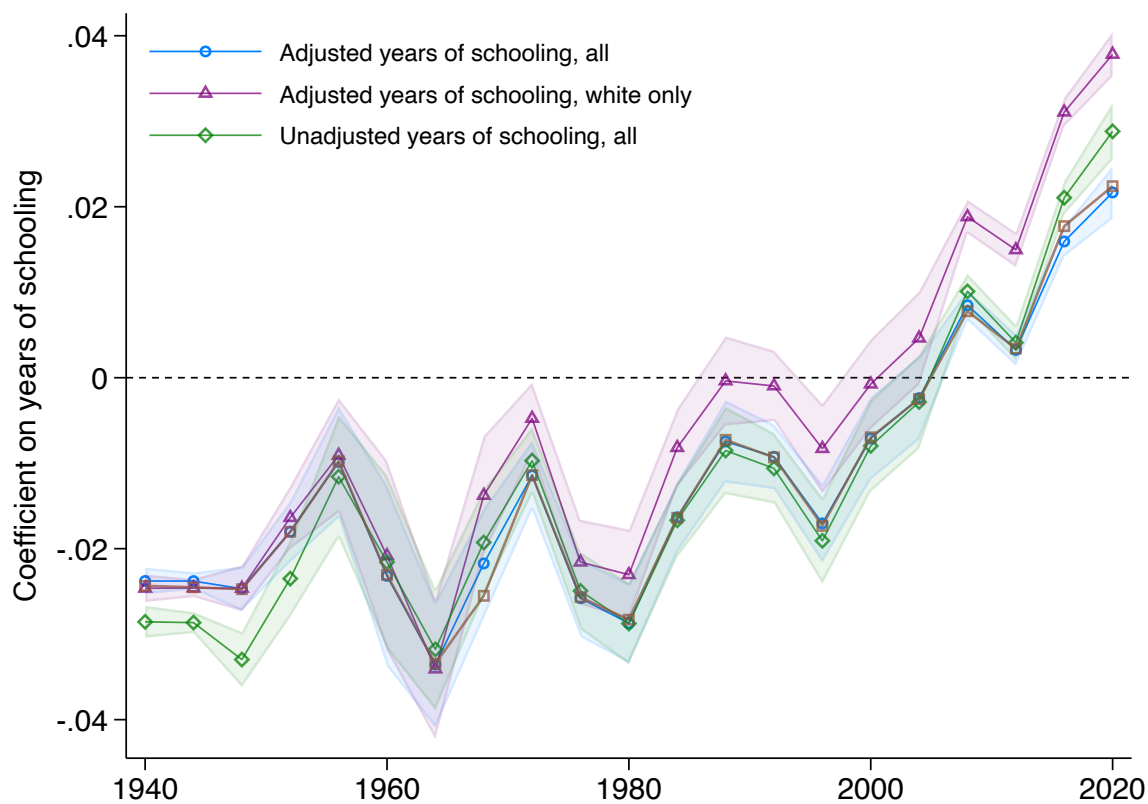
Source: Gallup as described in Table G.7

Notes: Panel (a) plots the share of all respondents who choose an economic problem as the most important faced by the country. Panel (b) plots the coefficient β^t from estimating the following regression separately by year t :

$$Econ\ most\ important_{is} = \beta Adj.\ years\ of\ school_i + \mathbf{Age}_i + \mu_s + e_{is},$$

where $Econ\ most\ important_{is}$ is an indicator for respondent i in survey s saying that an economic problem is the most important; $Adj.\ years\ of\ school_i$ is respondent i 's predicted years of schooling (as defined in Section 3), μ_s are survey (which subsume date) fixed effects; \mathbf{Age}_i are age-in-five-year-bins fixed effects.

Appendix Figure A.22: Votes for Democratic President, as a function of education



Notes: This graph is analogous to the main Figure 6 but substitutes voting for a Democratic president (coded as one, all other answers coded as zero, with non-voters dropped) for Democratic party ID. We also group answers by election instead of by year of observation.

Appendix Table A.1: Trends In Relative Support For Predistribution

	(1)	(2)	(3)	(4)	(5)	(6)	
predist					0.203*** (0.0393)	0.176*** (0.0432)	
Predist X Year	0.00181 (0.00110)	0.000599 (0.00303)	0.0000316 (0.00199)	0.00116 (0.00117)	0.00327*** (0.00113)	0.00277** (0.00137)	
Question FE	Yes	Yes	Detailed	Yes	No	No	<i>Notes:</i>
Survey FE	No	Yes	No	No	No	No	
Text Features	No	No	No	Yes	Yes	Yes	
Age FE	Yes	Yes	Yes	Yes	Yes	Yes	
Sample	All	All	All	All	All	All	
Addl. Controls	Basic	Basic	Basic	Basic	Basic	Extended	
Observations	264868	264868	264868	264868	264868	237438	

This table shows estimates from variants of a regression at the question by individual level of the form $Support_{qiy} = \beta Predist_q \times Trend_t + \delta_{age(i)} + \delta_y + \delta_q + \epsilon_i$, where q denotes survey question type (four predistribution and four redistribution, as in Figure A.1), y denotes year, and i denotes individual. Age in 5-year bins. All specifications include year fixed effects. Column 1 additionally includes only the question type fixed effects and age fixed effects. Column 2 adds survey \times year fixed effects, restricting the identification of the trend to occur between predistribution and redistribution question asked in the same survey. Column 3 includes fixed effects for each specific question wording (95 different wordings). Column 4 replaces these with 5 principal components and 8 hierarchical cluster indicators constructed from the embedding vectors. Column 5 and column 6 do not include question fixed effects which allows us to recover the overall differential level of support for predistribution vs. redistribution. Column 6 includes extended controls: region, race, gender, and education by year fixed effects, in addition to our regular age group by year fixed effects. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$. Standard errors clustered by survey in parentheses.

Appendix Table A.2: Robustness of Educational Gradient

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Predist X Yrs Ed.	-0.0524*** (0.00538)	-0.0529*** (0.00547)	-0.0490*** (0.00483)	-0.0487*** (0.00487)	-0.0620*** (0.00576)	-0.0416*** (0.00600)	-0.0635** (0.0241)	-0.0561*** (0.00465)	-0.0487*** (0.00722)	-0.0517*** (0.00689)	-0.0483*** (0.00505)
Predist X Yrs Ed. X Trend				0.0257 (0.0276)	0.150*** (0.0323)	0.0424 (0.0366)	0.305** (0.129)	0.0101 (0.0269)	0.0497 (0.0396)	0.0651* (0.0385)	0.0305 (0.0297)
Predist X Non-White											0.279*** (0.0380)
Predist X Female											0.140*** (0.0164)
Predist X South											-0.0772*** (0.0125)
Sample	All	All	All	All	Indep.	Dem.	No Coll.	Whites	Women	South	All
Controls	Basic	Basic	Ext.	Ext.	Ext.	Ext.	Ext.	Ext.	Ext.	Ext.	Basic
Text Features	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	257021	257017	236438	236438	65708	73404	93573	198966	121671	72392	257017

Notes: This table shows estimates from variants of a regression at the question by individual level of the form $Support_{qiy} = \beta Predist_q \times AdjEduc_i + \delta_{age(i) \times q \times y} + \delta_{AdjEduc \times y} + \epsilon_i$, where q denotes survey question type (four predistribution and four redistribution, as in Figure A.1), y denotes year, and i denotes individual. Age is in 5-year bins and $AdjEduc$ is years of schooling as described in the text. All specifications include age by question by year fixed effects. Column 1 additionally includes only the education by year fixed effects. Column 2 adds age and year-specific effects of the 5 principal components and 8 k-means cluster indicators constructed from the embedding vectors. Column 3 further adds nonwhite, female, and South all interacted with question fixed effects interacted with year fixed effects. Column 4 adds an interaction of $Predist_q \times AdjEduc_i \times Trend_y$ to the specification in (3), and Column 5-9 restrict attention to Democrats, independents, nonwhite, female, and South subsamples, respectively. Finally, Column 10 removes the question by demographic controls and examines the differential support for predistribution by demographic subgroups. ** $p < 0.1$, * $p < 0.05$, *** $p < 0.001$. Standard errors clustered by survey in parentheses.

Appendix B. Data Appendix

B.1. ANES repeated cross-sectional data and cumulative file

We use both the individual files for each year and the cumulative file that ANES creates for convenience. The individual files have questions that are not included in the cumulative file. We use every year of data from 1948 to 2018 that includes one of the questions of interest. We use the partisan affiliation variable to define Democratic affiliation. We code as democrats any individual who describe themselves as either “strongly Democrat” or “not strongly Democrat”.

B.2. Gallup and other historical opinion polls data

As in Farber *et al.* (2021), we use historical opinion polls from survey corporations, for the most part housed by iPoll at the Roper Center at Cornell. The majority of these data come from Gallup, which beginning in 1942 asked respondents in most of their surveys both their educational attainment and their self-reported partisan identification (Gallup surveys begin in 1935 and since then have always asked age, race and state of residence).

B.3. General Social Survey

The GSS surveys a sample of around 2,000 nationally representative Americans yearly since 1972 (GSS has been implemented once every other year only since 1994). We use the partisan affiliation variable (*partyid*) to define Democratic affiliation. We code as democrats any individual who describe themselves as either “strongly Democrat” or “not strongly Democrat”.

B.4. Cooperative Election Study

The CCES is a survey administered by YouGov to a very large sample of Americans (typically over 50,000 people) since 2006. We use the partisan identity question *pid3* to measure Democratic affiliation. The question asks whether the respondent think of themselves as a (Democrat/Republican/Independent/Other/Not Sure).

B.5. Comparative Agendas Project

CAP has harmonized legislation, executive orders and other political material across dozens of countries. We use the roll-call data, which they have grouped into categories and sub-categories based on subject matter.

Appendix Table B.3: Surveys used in the main analysis, grouped by decade

Survey	Decade	Observations	Share college graduates	Share women	Share white
ANES	1950	6,549	0.08	0.54	0.89
ANES	1960	11,568	0.11	0.56	0.87
ANES	1970	9,909	0.15	0.57	0.87
ANES	1980	9,094	0.19	0.56	0.80
ANES	1990	8,932	0.22	0.54	0.76
ANES	2000	6,656	0.27	0.54	0.73
ANES	2010	9,880	0.32	0.52	0.71
CCES	2000	90,284	0.27	0.52	0.74
CCES	2010	366,988	0.28	0.52	0.73
CCES	2020	58,653	0.32	0.52	0.70
GSS	1970	10,187	0.25	0.53	0.88
GSS	1980	13,782	0.32	0.56	0.82
GSS	1990	12,880	0.43	0.55	0.82
GSS	2000	14,502	0.48	0.54	0.77
GSS	2010	11,476	0.51	0.55	0.74
Gallup	1940	48,848	0.09	0.50	0.93
Gallup	1950	18,183	0.08	0.50	0.92
Gallup	1960	1,562	0.11	0.52	0.90
Gallup	1970	10,055	0.18	0.52	0.87
Gallup	1980	6,845	0.20	0.53	0.87
Gallup	1990	5,522	0.25	0.52	0.80
Gallup	2000	25,196	0.30	0.52	0.79
Gallup	2010	58,757	0.32	0.51	0.68
Gallup	2020	2,861	0.37	0.51	0.68

Notes: The table shows the number of observations for each data sources used through the analysis, grouped by decades. The third column shows the share of college graduates in each dataset, the fourth column shows the share of women and the last column shows the share of white respondents.

Appendix C. Theory Appendix

In this appendix we present a model of electoral competition between two parties with intra-party bargaining, building on work by Roemer (1998) and Besley (2007). The timing of the game is that parties commit to a level of utility to offer the pivotal voter. Elections happen, and then factions within the party bargain over how to deliver the utility promised. Voters have education level h that corresponds to their pre-tax earnings under laissez-faire, and care about both their post-tax income and their pre-tax income. The predistribution policy is given by θ , while the redistribution policy is given by τ . We focus on the last stage of this game, where two party factions, indexed by high and low levels of education, bargain over the policies to be implemented.

Voters possess utility functions, denoted as $V(\tau, \theta|h)$, that are continuously differentiable. Here, τ represents redistribution via a tax-and-transfer system, while θ denotes predistribution (e.g. minimum wages or job guarantees), and h signifies educational attainment or pre-tax earnings. We suppose that $V_\theta > 0$, indicating a universal preference for predistribution.²⁷ In the next subsection, we alternatively posit quadratic (single-peaked) preferences.

We invoke a Spence-Mirrlees condition (Gans and Smart, 1996), stipulating that

$$\frac{d\left(\frac{\frac{dV(\tau, \theta|h)}{d\tau}}{\frac{dV(\tau, \theta|h)}{d\theta}}\right)}{dh} > 0$$

implying that more educated voters (higher h) adjust their valuations of τ more significantly than θ . Importantly, we do not impose any presuppositions regarding $V_{\theta\tau}$, allowing predistribution and redistribution to be either complements or substitutes. Given the intricate interactions within extensive policy domains associated with predistribution and redistribution, it is may be important to avoid imposing *a priori* constraints on these cross-partials in the indirect utility function.

Thus there is a well-defined median or swing voter given by h^{swing} , although we do not require this for the existence of equilibrium, because parties compete on utility levels, not policies. We assume party competition is strict, so that both parties offer policies so that $V(\theta^R, \tau^R|h^{swing}) = V(\theta^D, \tau^D|h^{swing}) = V^s$. While parties cannot commit to particular policies, they can commit to providing a level of utility.

After the election, we assume intra-party bargaining over the policy mix between $J = 2$ factions, indexed by the human capital of their representative member h^k , $k \in H, L$ where the high education group in party j has bargaining power λ^j . We assume that $h^H > h^{swing} > h^L$ but also that $h^H - h^L < \epsilon$ where ϵ is a small positive number. This last assumption is required to use a first-order approximation in the proof without making assumptions about the concavity of V in h . λ could also be the share of the highly-educated group who are party activists or the probability of each group being pivotal in a primary election. The constraint on this problem is the swing voter's utility, given by $V(\tau, \theta|h^{swing}) = V^s$. Focusing on the D party, the post-election bargaining problem thus solves:

²⁷An alternative assumption could posit a universal preference for redistribution, represented as $V_\tau > 0$, leaving V_θ unrestricted. However, we adopt the former assumption in alignment with standard models depicting agents with single-peaked preferences over redistribution.

$$\max_{\theta, \tau} W^D(\tau, \theta) = (1 - \lambda)V(\tau, \theta|h^{low}) + \lambda V(\tau, \theta|h^{high}) \quad (5)$$

subject to:

$$V(\tau, \theta|h^{swing}) = V^s \quad (6)$$

Denote the Lagrangian as $\widetilde{W^D}(\theta, \tau)$ and the optimal solution as the policy $x^*(\lambda) = (\theta^*(\lambda), \tau^*(\lambda))$. We can show two comparative statics in response to an increase in the bargaining power of the relatively educated faction, λ :

- First, $\frac{d^2 V(\tau^*(\lambda), \theta^*(\lambda)|h)}{dh d\lambda} > 0$. An increase in the bargaining power of the relatively educated faction leads to an increase in the educational gradient in voting for the D platform.
- Second, $\frac{d\theta^*(\lambda)}{d\lambda} < 0 < \frac{d\tau^*(\lambda)}{d\lambda}$, so an increase in the bargaining power of the relatively educated leads to a larger decrease in predistribution and an increase in redistribution.

Proof: 1) Denote the Lagrangian as $\widetilde{W^D}(\theta, \tau)$ and note that at the optimum x^* we have:

$$\nabla_{\lambda} x^*(\lambda) = -\nabla_{xx} \widetilde{W^D}(x^*)^{-1} \nabla_{\lambda} W^D(x^*) \approx \nabla_{xx} \widetilde{W^D}(x^*)^{-1} \nabla_x V_h(x^*(\lambda))(h^{high} - h^{low}) \quad (7)$$

Where the approximation follows from $h^{high} - h^{low}$ being small. Therefore the first claim follows from expanding:

$$\frac{d^2 V(\tau^*(\lambda), \theta^*(\lambda)|h)}{dh d\lambda} = \nabla_{\lambda} V_h(x^*(\lambda)) \cdot \nabla_{\lambda} x^*(\lambda) = -\nabla_x V_h(x^*(\lambda)) \nabla_{xx} \widetilde{W^D}(x^*)^{-1} \nabla_x V_h(x^*(\lambda))(h^{high} - h^{low}) > 0 \quad (8)$$

where the inequality follows from $\nabla_{xx} \widetilde{W^D}$ and its inverse being negative definite (since V is always concave, the Lagrangian is concave, and so the Hessian is negative definite).

2) The swing voter constraint defines a function $\theta(\tau, V^s)$ with $\frac{d\theta}{d\tau} = -\frac{V_{\tau}(\theta, \tau|h^{swing})}{V_{\theta}(\theta, \tau|h^{swing})} < 0$ that we can substitute out for θ in W^D and so the optimal tax rate chosen by D will have comparative static $\frac{d\tau}{d\lambda}$ signed the same direction as:

$$W_{\lambda\tau} = (V_{\theta}(\theta(\tau), \tau|h^{high}) - V_{\theta}(\theta(\tau), \tau|h^{low})) \frac{d\theta}{d\tau} + (V_{\tau}(\theta(\tau), \tau|h^{high}) - V_{\tau}(\theta(\tau), \tau|h^{low})) \quad (9)$$

Which is positive iff:

$$(V_{\theta}(\theta(\tau), \tau|h^{high}) - V_{\theta}(\theta(\tau), \tau|h^{low})) \frac{V_{\tau}(\theta(\tau), \tau|h^{swing})}{V_{\theta}(\theta(\tau), \tau|h^{swing})} < (V_{\tau}(\theta(\tau), \tau|h^{high}) - V_{\tau}(\theta(\tau), \tau|h^{low})) \quad (10)$$

Which, for h^{high} and h^{low} are sufficiently close to h^{swing} then a first-order approximation of V_{τ} and V_{θ} around h^{swing} will be sufficiently close to yield:

$$\frac{V_{\theta h}(\tau, \theta(\tau)|h^{swing})}{V_{\theta}(\theta(\tau), \tau|h^{swing})} < \frac{V_{\tau h}(\tau, \theta(\tau)|h^{swing})}{V_{\tau}(\theta(\tau), \tau|h^{swing})} \quad (11)$$

Which is implied by single crossing, since:

$$\frac{dV_\tau}{dh} > 0 \iff \frac{V_{\tau h}V_\theta - V_{\theta h}V_\tau}{V_\theta^2} > 0 \iff \frac{V_{\theta h}}{V_\theta} < \frac{V_{\tau h}}{V_\tau} \quad (12)$$

So we have that $\frac{d\tau}{d\lambda} > 0$. The constraint $V(\theta, \tau|h^{swing}) = V^s$ implies that $\frac{d\theta}{d\tau} < 0$ and 2) follows.

A natural question is what effect does an increase in λ within the D party have on the policies of the R party. In our model there is no effect because all of the effects of political competition are summarized in the swing voter utility being held constant. There is no interaction between the parties that is not captured by the swing voter utility constraint facing the ex-post bargaining problem. Sharp political competition between the two parties means that any within-party change in bargaining power must keep the swing voter indifferent, otherwise the party will lose.

Suppose instead of a within-party increase in bargaining power of the relatively educated, the swing-voter's education increases. Then *both* parties policies would change to favor redistribution over predistribution.

C.1. Quadratic Preferences

While single-crossing is an attractive assumption on preferences, as it imposes no restriction on the relationship between preferences for redistribution and prediction (i.e. the cross-partial $V_{\theta\tau}$ is unrestricted), it does require utility over predistribution to be monotonic, rather than single-peaked. Further, for tractability, the previous model only allowed two factions and two policy areas. In this subsection we focus on quadratic preferences, which imposes additive separability across policy domains, but has no monotonicity requirement (only single-peakedness), and obtain similar results with an arbitrary number of policy domains and an arbitrary number of factions.

We order J factions by their level of human capital or education. There are I policy domains. Each faction has ideal point for policy i given by y_i^j , while the swing voter has ideal point y_i^m . We assume that the highest education faction has higher education than the median or swing voter.

The right wing party offers the pivotal voter utility normalized to V^m . The problem of the left party is to solve a weighted welfare function of its groups, subject to giving the swing voter the same utility V^m :

$$\max_{x_i} W^D(x_i) = \sum_j^J \lambda_j V^j = - \sum_i \sum_j \lambda_j (x_i - y_i^j)^2 \quad (13)$$

subject to:

$$- \sum_i (x_i - y_i^m)^2 = V^m \quad (14)$$

the resulting optimal choice (with multiplier μ) is given by:

$$x_i^* = \frac{\sum_j \lambda_j y_i^j + \mu y_i^m}{\sum_j \lambda_j + \mu} \quad (15)$$

The effect of an increase in the within party bargaining power of group j is given by:

$$\frac{dx_i^*}{d\lambda_j} = \frac{y_i^j - x_i^*}{\sum_j \lambda_j + \mu} \quad (16)$$

If y_i^j is greater than x_i^* then this is positive, otherwise it is negative. If we further assume there is a policy domain i_R for redistribution, where $j > j' \implies y_{i_R}^j > y_{i_R}^{j'}$ and similarly i_P for predistribution where $j > j' \implies y_{i_P}^j < y_{i_P}^{j'}$, we can easily see that $\frac{dx_{i_R}}{d\lambda_j} > 0$ and $\frac{dx_{i_P}}{d\lambda_j} < 0$.

Looking at effects on voting, the utility of voter group j is then:

$$V^j(x_i^*) = - \sum_i (x_i^* - y_i^j)^2 \quad (17)$$

And so an increase in the within party bargaining power of group j is given by:

$$\frac{V^j(x_i^*)}{d\lambda_j} = C \sum_i (x_i^* - y_i^j)^2 > 0 \quad (18)$$

where $C \equiv \frac{2}{(\sum_j \lambda_j + \mu)^3} > 0$. Unsurprisingly, an increase in the bargaining weight of a group inside the party results in a platform more preferred by that group in the electorate. If J is the highest human capital faction, then their bargaining power λ_J increases, and so the vote share of group J for D will also increase, while keeping the median voter at V^m .

Appendix D. Classifying votes as predistribution or redistribution

D.1. Comparative Agendas Project

In order to classify policies between *Predistribution*, *Redistribution*, and *Social Issues*, we make heavy use of the Comparative Agendas Project (CAP) dataset. CAP groups all Congressional votes since 1947 into policy-related categories and sub-categories. We take these categories as given and then place them into pre- and re-distribution groups.

The table below shows the CAP topics and subtopics and our classification between the various categories.

General topic	Subtopic	Classification
1. Macroeconomics	100: General 101: Interest Rates 103: Unemployment Rate 104: Monetary Policy 105: National Budget 107: Tax Code 108: Industrial Policy 110: Price Control 199: Other	Predis Predis Predis Predis Redis Redis Predis Predis Both
2. Civil Rights	200: General 201: Minority Discrimination 202: Gender Discrimination 204: Age Discrimination 205: Handicap Discrimination 206: Voting Rights 207: Freedom of Speech 208: Right to Privacy 209: Anti-Government 299: Other	
3. Health	300: General 301: Health Care Reform 302: Insurance 321: Drug Industry 322: Medical Facilities 323: Insurance Providers 324: Medical Liability 325: Manpower 331: Disease Prevention 332: Infants and Children 333: Mental Health 334: Long-term Care 335: Drug Coverage and Cost 341: Tobacco Abuse 342: Drug and Alcohol Abuse 398: R&D 399: Other	Social Issues Social Issues
4. Agriculture	400: General 401: Trade 402: Subsidies to Farmers 403: Food Inspection & Safety 404: Food Marketing & Promotion 405: Animal and Crop Disease 408: Fisheries & Fishing 498: R&D 499: Other	Predis Predis
5. Labor	500: General 501: Worker Safety 502: Employment Training 503: Employee Benefits 504: Labor Unions 505: Fair Labor Standards 506: Youth Employment 529: Migrant and Seasonal 599: Other	Predis Predis Predis Predis Predis Predis Predis Predis Predis

General topic	Subtopic	Classification
6. Education	600: General 601: Higher 602: Elementary & Secondary 603: Underprivileged 604: Vocational 606: Special 607: Excellence 698: R&D 699: Other	Predis
7. Environment	700: General 701: Drinking Water 703: Waste Disposal 704: Hazardous Waste 705: Air Pollution 707: Recycling 708: Indoor Hazards 709: Species & Forest 711: Land and Water Conservation 798: R&D 799: Other	
8. Energy	800: General 801: Nuclear 802: Electricity 803: Natural Gas & Oil 805: Coal 806: Alternative & Renewable 807: Conservation 898: R&D 899: Other	Predis Predis Predis Predis
9. Immigration	900: Immigration	
10. Transportation	1000: General 1001: Mass 1002: Highways 1003: Air Travel 1005: Railroad Travel 1007: Maritime 1010: Infrastructure 1098: R&D 1099: Other	Predis
12. Law and Crime	1200: General 1201: Agencies 1202: White Collar Crime 1203: Illegal Drugs 1204: Court Administration 1205: Prisons 1206: Juvenile Crime 1207: Child Abuse 1208: Family Issues 1210: Criminal & Civil Code 1211: Crime Control 1227: Police 1299: Other	Social Issues Social Issues Social Issues Social Issues Social Issues Social Issues Social Issues Social Issues Social Issues Social Issues Social Issues Social Issues

General topic	Subtopic	Classification
13. Social Welfare	1300: General 1302: Low-Income Assistance 1303: Elderly Assistance 1304: Disabled Assistance 1305: Volunteer Associations 1308: Child Care 1399: Other	Redis
14. Housing	1400: General 1401: Community Development 1403: Urban Development 1404: Rural Housing 1405: Rural Development 1406: Low-Income Assistance 1407: Veterans 1408: Elderly 1409: Homeless 1498: R&D 1499: Other	Redis Redis
15. Domestic Commerce	1500: General 1501: Banking 1502: Securities & Commodities 1504: Consumer Finance 1505: Insurance Regulation 1507: Bankruptcy 1520: Corporate Management 1521: Small Businesses 1522: Copyrights and Patents 1523: Disaster Relief 1524: Tourism 1525: Consumer Safety 1526: Sports Regulation 1598: R&D 1599: Other	
16. Defense	1600: General 1602: Alliances 1603: Intelligence 1604: Readiness 1605: Nuclear Arms 1606: Military Aid 1608: Personnel Issues 1610: Procurement 1611: Installations & Land 1612: Reserve Forces 1614: Hazardous Waste 1615: Civil 1616: Civilian Personnel 1617: Contractors 1619: Foreign Operations 1620: Claims against Military 1698: R&D 1699: Other	Predis Predis Predis Predis

General topic	Subtopic	Classification
17. Technology	1700: General 1701: Space 1704: Commercial Use of Space 1705: Science Transfer 1706: Telecommunications 1707: Broadcast 1708: Weather Forecasting 1709: Computers 1798: R&D 1799: Other	
18. Foreign Trade	1800: General 1802: Trade Agreements 1803: Exports 1804: Private Investments 1806: Competitiveness 1807: Tariff & Imports 1808: Exchange Rates 1899: Other	Predis Predis Predis Predis Predis Predis Predis Predis
19. International Affairs	1900: General 1901: Foreign Aid 1902: Resources Exploitation 1905: Developing Countries 1906: International Finance 1910: Western Europe 1921: Specific Country 1925: Human Rights 1926: Organizations 1927: Terrorism 1929: Diplomats 1999: Other	
20. Government Operations	2000: General 2001: Intergovernmental Relations 2002: Bureaucracy 2003: Postal Service 2004: Employees 2005: Appointments 2006: Currency 2007: Procurement & Contractors 2008: Property Management 2009: Tax Administration 2010: Scandals 2011: Branch Relations 2012: Political Campaigns 2013: Census & Statistics 2014: Capital City 2015: Claims against the government 2030: National Holidays 2099: Other	Predis Predis Predis Redis Social Issues Social Issues Social Issues
21. Public Lands	2100: General 2101: National Parks 2102: Indigenous Affairs 2103: Public Lands 2104: Water Resources 2105: Dependencies & Territories 2199: Other	
23. Culture	2300: General	Social Issues

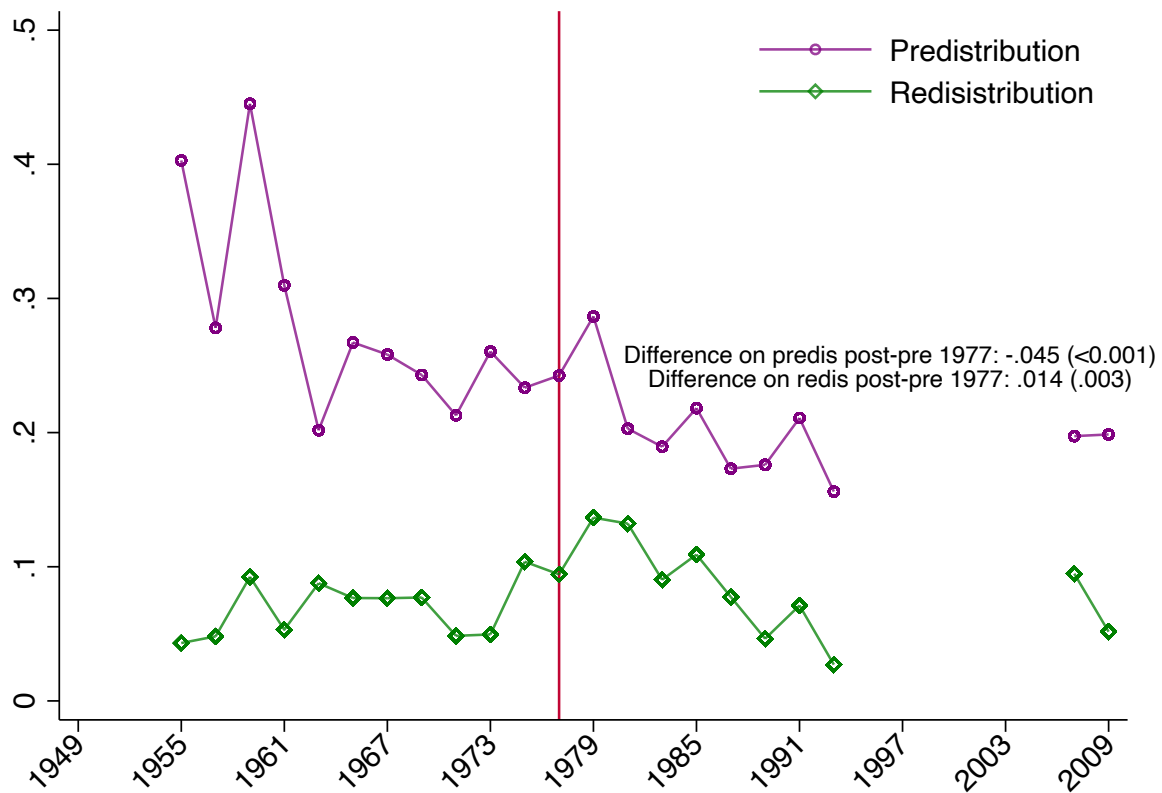
D.2. Alternative classification

To test the robustness of our results to alternative rollcall classification, we use a second topic classification from (Bateman *et al.*, 2018). Table ?? below shows which topics have been classified as predistribution, redistribution, and social issues. We also reproduce below all the Tables and Figures of the main paper that are making use of the CAP classification with the (Bateman *et al.*, 2018).

General topic (Tier 1)	Tier 2	Tier 3	Classification
Sovereignty	Liberty	Loyalty and Expression	Social Issues
		Religion	Social Issues
		Privacy	Social Issues
	Membership and Nation	Commemorations and	Social Issues
		National Culture	
		Immigration and Naturalization	
	Civil Rights	African American	Social Issues
		Native Americans	Social Issues
		Other Minority Groups	Social Issues
		Women	Social Issues
Organization and Scope	Boundaries	Voting Rights	Social Issues
		Frontier Settlement	Social Issues
		Indian Removal and	Social Issues
		Compensation	
		State Admission	Social Issues
		Union Composition	
		Territories and Colonies	Social Issues
	Gov Organization	Congressional Organizational	
		Executive Org	
		Impeachment	
International Relations	Representation	Judicial	
		Census	Social Issues
		Elections	
		Groups and Interests	
	Constitutional Amendments	Federalism and terms of office	
		Political participation and rights	
		Other	
	Defense	Air Force	
		Army Organization	
		Conscription	
		Militias	
		Naval	
		Organization	
		Military Installations	
	Geopolitics	Civil Homeland Defense	
		Diplomacy	
		Foreign Aid	
		International Org	
	International Political Economy	Maritime	
		Trade Tariffs	Predis
	Economic International Organizations		

General topic (Tier 1)	Tier 2	Tier 3	Classification
Domestic Affairs	Agriculture and Food	Agricultural Technology	
		Farmers / Farming support	Predis
	Planning and Resources	Fishing and Livestock	
		Corporatism	Predis
		Environment	
		Infrastructure and Public Works	Predis
		National Resources	
		Social Knowledge	Social Issues
		Post office	Predis
		Transportation	Predis
		Wage and price control	Predis
		Interstate compacts	
	Political Economy	Urban rural and regional dev	Predis
		Appropriation	
		Multi agency	
		Business / Capital Markets	Predis
		Fiscal taxation	Redis
		Labor Markets unions	Predis
	Social Policy	Monetary	Predis
		Regulation economic	Predis
		Children / Youth	
		Crime	Social Issues
		Disaster	
		Education	
		Handicapped	
		Health	
		Housing	
		Military pensions	
		Public works and volunteer employment	Predis
		Regulation, social	Social Issues
		Social insurance	

Appendix Figure D.23: The pre- and re-distribution share of House roll-call votes under Democratic leadership (Bateman et al. data)



Notes: As in Figure 3, this figure shows the share of pre- and re-distribution roll-call votes every year that the Democrats are in control of the House from 1947 until 2009. Breaks in the series are when Republicans control the House. The Figure uses data from Bateman *et al.* (2018)

Appendix Table D.4: Democratic House members position depending on DLC status

	Dept Var: Ideal Points (incr. in conservativeness)					
	(1)	(2)	(3)	(4)	(5)	(6)
DLC	0.092*** (0.005)	0.315*** (0.022)	0.239*** (0.041)	0.211*** (0.037)	0.208*** (0.039)	0.105*** (0.005)
Predistribution x DLC						0.222*** (0.019)
Redistribution x DLC						0.192*** (0.023)
Ideal Point Type	DW-N 1	Predis	Predis	Redis	Social	DW-N 1, Predis, Redis
Congress FE	X	X				
State x Congress FE			X	X	X	X
Average difference between parties	.772	1.805	1.805	1.898	1.77	1.486
Predistribution x DLC -						0.030
Redistribution x DLC						(0.025)
Observations	3,404	3,350	3,228	3,231	3,267	10,106

Notes: As in table 1, this table shows the position of House members as defined by DW-Nominate or their topic-specific ideal points. We use the rollcall topic-classification from Bateman *et al.* (2018) and the topic-specific ideal point estimation, described in Appendix section F. Column 1 regresses the DW-Nominate first dimension on a DLC dummy, controlling for Congress fixed effects, Column 2 regresses the predistribution ideal point, controlling for congress FE, while column 3 onwards control for state by congress FE. Column 4 looks at the redistribution ideal point and column 5 at the ideal point on social issues. Column 6 pools together the DW Nominate first dimension, the predistribution and the redistribution ideal point and looks at the relative effect of DLC on economic policies, with respect to DW-Nominate first dimension. We report the average difference between the parties in the table notes and we test the significance of the difference between predistribution and redistribution in column 6. Standard errors clustered by Congress. The sample is restricted to Democratic House members. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$. Standard errors clustered by congress in parentheses.

Appendix Table D.5: Democratic House members' alignment with Republican voting

	Dept Var: Voted Yea			
	(1)	(2)	(3)	(4)
DLC \times MeanRepub	0.148*** (0.002)	0.141*** (0.002)	0.142*** (0.002)	0.144*** (0.002)
DLC \times Predis		-0.008** (0.004)	-0.008** (0.004)	-0.008* (0.004)
DLC \times Predis \times MeanRepub		0.015** (0.006)	0.015** (0.006)	0.013** (0.006)
DLC \times Redis			0.008 (0.006)	0.007 (0.006)
DLC \times Redis \times MeanRepub			-0.006 (0.008)	-0.008 (0.008)
Rollcall FE	X	X	X	X
State \times Congress FE				X
Linear combination of estimates: Predis - Redis			0.021** (0.010)	0.021** (0.010)
Observations	4,031,972	3,743,700	3,743,700	3,743,700

Notes: As in Table 2, this table shows the extent to which Democrats are voting with Republicans using the Bateman *et al.* (2018) classification instead of the CAP. * p<0.1, ** p<0.05, *** p<0.001. Standard errors clustered by Rollcall in parentheses.

Appendix E. Predicting DLC Membership

We have relied on several sources to obtain the list of House and Senate members that belong to the Democratic Leadership Council (DLC). It is important to note that we are not necessarily looking for a strict definition of who belonged to this caucus at what exact time, but rather which Democrats were part of the ideological movement corresponding to the DLC. We combine five main sources of DLC membership:

1. A list of neoliberal Democrats obtained from Rothenberg (1984)
2. A list of founding members of the DLC obtained from the Appendix of the first volume of the *Mainstream Democrat* published in September 1989
3. The list of participants to the third annual DLC’s conference in Philadelphia in March 1989, obtained through the University of Florida
4. Lists of members obtained through archives of the DLC website and the New Democrats Online website
5. List of members of the DLC caucus in Congress obtained through the archives of the House and Senate websites.

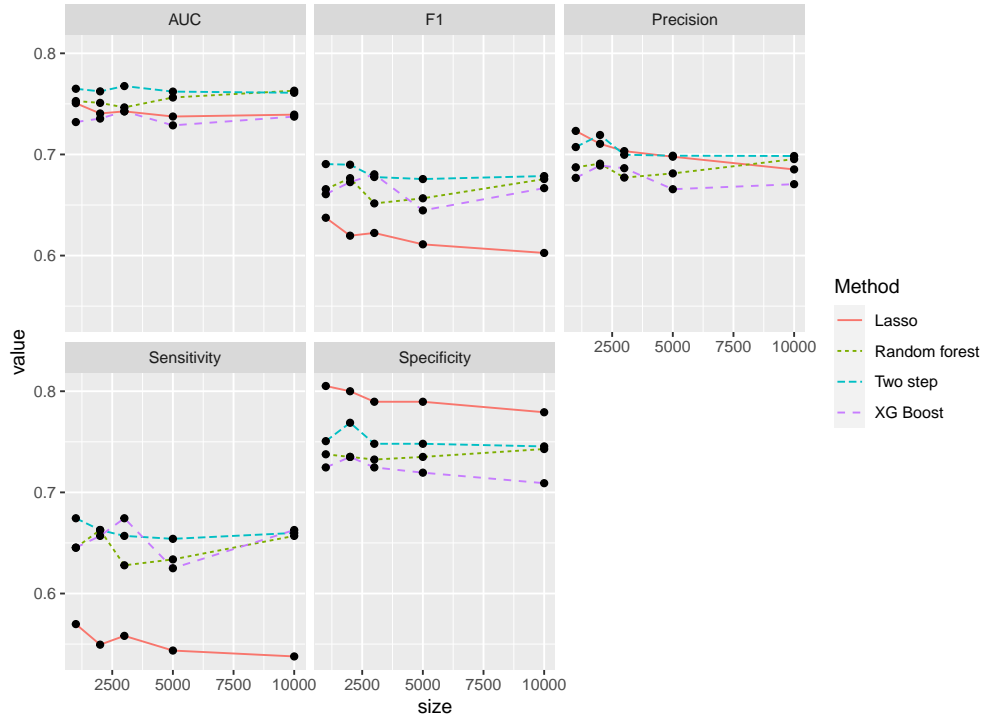
Since the historical lists of DLC members that we have had access to only include elected members, we only observe DLC members who ran for office and got elected. Most of the results we present in the paper use only this set of "groundtruth" winner DLC members.

To assess the robustness of our results to the inclusion of potential "losers", we predicted DLC membership based on campaign contributions. We use data from Bonica (2014) to obtain information on all the campaign contributions received by Congress candidates since 1980. These contributions have been used by several papers to estimate the ideology of Congress members, here we want to use them to estimate their membership to a specific faction of the Democratic party.

We train several Machine Learning models to predict DLC membership using information on contributions and the identity of contributors. We train a Random Forest algorithm (Breiman, 2001), an extreme gradient boosting algorithm (Friedman, 2001), a Lasso and a two-step process where we first select the list of relevant variables with a Lasso and apply a Random Forest on the set of selected variables.

Similarly to Bonica (2018), since the number of distinct contributors is too large to include them all as explanatory variables, we keep the 5,000 largest contributors and include the amount received from each of them, interacted with the campaign period. For all other contributors, we compute summary statistics of their zip code demographic variables. For each candidate, we include the share of their contributions coming from deciles of the following variables: candidate, share of money coming from zipcodes in each decile of the following: population, median home value, median household income, share of African American, share of white, share of Hispanic, share of homeowners, share of the population below 25, share of the population above 65, share of the population working in the agricultural sector, share of the population working in the secondary sector, share of the population working in the tertiary sector, share of the population working in the public sector, share of the population

Appendix Figure E.24: Performance of DLC status prediction



Notes: We report five statistics about the performance of our four prediction models. The x-axis shows the number of distinct contributors included as explanatory variables. The remaining contributors are included as summary statistics. We report the Area Under the Receiver Operating Characteristic Curve (AUC), the F-score, the Precision, the Sensitivity, and the Specificity.

working in health, share of the population working in finance. Some specific PACs such as the *New Democratic Network* (NDN) are strong predictors of DLC membership.

Figure E.24, we report the statistics regarding our predictions. The two-step algorithm seems to outperform the other models.

Appendix F. Estimating Topic-specific Ideal Points

Ideal-point estimation is widely used in political science to summarize dimensions of ideology based on political behavior (roll-call votes, contributions, speech). The standard two-dimensional “DW-Nominate” score from Poole and Rosenthal (1997)²⁸ estimates 2-dimensions of ideology, and allows politician ideal points to linearly drift with time. The first dimension of DW-Nominate explains over 70% of variation, widely used as summary measure of left-right ideology, and is interpreted by its creators as largely capturing divisions over economic policy in the post-war period (Poole and Rosenthal, 2000).

We are interested in disaggregating economic policy into predistribution and redistribution, and recovering separate ideal points for each congressperson for each of these dimensions. Following the notation of the model, each politician i has ideal point for policy dimension j (predistribution or redistribution) given by y_{ij} . We suppose politicians have quadratic utility function over voting Y and N over roll-call r in congress c given by: $U(Y(r)) = -(y_{icj(r)} - x_r^Y)^2 + \nu_{ri}$ and $U(N(r)) = -(y_{icj(r)} - x_r^N)^2 + \nu'_{ri}$, where $j(r)$ designates whether the issue is predistribution or redistribution. If ν and ν' are jointly normal so that $E[\nu] - E[\nu'] = 0$ and $Var(\nu - \nu') = \sigma^2$, then we have that

$$Pr(Vote_{ir} = Y) = Pr(\nu'_{ir} - \nu_{ir} < (x_r^Y - x_r^N)y_{icj(r)} + (x_r^Y)^2 - (x_r^N)^2) = \Phi(\beta_r y_{icj(r)} - \alpha_r) \quad (19)$$

Where $\beta_r = (x_r^Y - x_r^N)$ and $\alpha_r = (x_r^Y)^2 - (x_r^N)^2$. Estimating this separately for redistribution and predistribution roll-calls yields estimates of topic-specific roll-call votes in congress c $y_{ic,predist}$ and $y_{ic,redist}$. We use the `pscl` package in R, which provides functionality for estimating ideal points from rollcall data, and normalize the ideal points on each topic to be mean 0 and unit standard deviation within each congress to ensure identification, restricting attention to bills (coded as redistribution and predistribution by CAP) with less than 90% support.

In order to make ideal points comparable over time, Bateman and Lapinski (2016) use a procedure that regression-adjusts the congress-specific ideal points over time to make them comparable. Within a congress, they regress the congress-specific individual ideal points on the mean for each congressperson over all congresses and then subtracting the intercept from each congress-specific ideal point and divide by the coefficient. We follow this procedure to get predistribution and redistribution ideal points that are comparable across congresses.

²⁸Which we downloaded from voteview.com.

Appendix G. Using text-as-data to assess question wording effects

As discussed in the text, we use text embeddings as continuous representations of words to ensure that variants in question wording are not driving any of our results. We calculate BERT (Bidirectional Encoder Representations from Transformers) embeddings (Devlin *et al.*, 2018) for each of the raw survey question texts. BERT returns embeddings for each token in a piece of text. These embedding vectors are optimized to predict nearby words that are omitted (masked), and capture the semantic meanings of words and phrases (a widely-used example is that the embedding of “king” minus the embedding for “queen” yields the embedding for “man”). Embeddings are widely used in deep learning models of language, taking a sentence, breaking it up into tokens (words, subwords and punctuation), and uses both the position of the token as well as a 1024-dimensional vector to predict the omitted tokens in the text (BERT uses both the tokens before and after the missing token for prediction, hence “bidirectional”). The 1024-dimension vector for each token is then tuned to optimally predict the omitted tokens in a 24-layer neural network. BERT (specifically the version we use, BERT-large) is a 340 million parameter model trained on over three billion words of text, and the embeddings perform very well on standard NLP tasks and are widely used in text analysis.²⁹

We begin by showing that our embeddings are good at picking up the substance of a question. We have 95 question wordings, with 60 predistribution questions and 35 redistribution questions. In this sample of 95 questions, the first principle component of the embedding vector explains over 20% of the R^2 and the 5 top principal components explain over 80% of the variation in whether the question is predistribution or redistribution, while explaining 40% of the variation in the embeddings themselves. Appendix Figure G.25 shows that a LASSO prediction of predistribution using all the embeddings perfectly separates predistribution and redistribution. 95

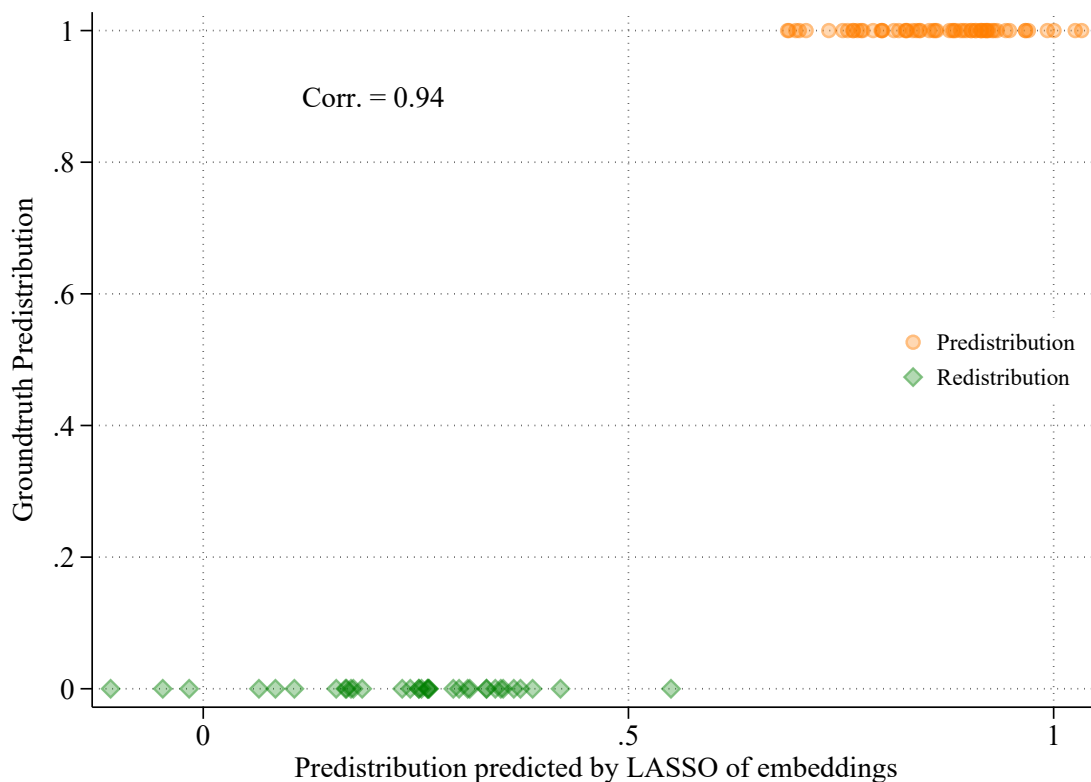
We also calculate eight clusters (as we have eight policy domains) using the Ward hierarchical clustering algorithm applied to the 1024-dimensional vector of embeddings. The Ward clustering algorithm defines clusters to maximize the between-cluster sum of squares (of the embedding vector in our case).

Appendix Table G.6 shows that these purely data-driven clusters nonetheless correspond closely to our policy categories, with most policies falling in no more than 2 clusters, and a few being perfectly collinear with the clustered embeddings. Further, the predistribution questions tend to fall in clusters 1-4, while the redistribution questions tend to fall in clusters 5-8. Together Figure G.25 and Table G.6 show that there is a great deal of semantic information in the embeddings, as they are able to recover our high-level policy classifications fairly reliably. In linguistic terms, embeddings show that questions on the same issue are semantically similar despite being syntactically different.

We next turn to examining whether support is systematically associated with linguistic features of the question. In addition to the embeddings, we calculate dictionary-based measures of sentiment, such as polarity (how positive or negative the sentiment of a text is) and subjectivity (how subjective is a text, e.g. reporting a particular state of mind via “I feel”).

²⁹For an overview of the use of embeddings in economics see (Ash and Hansen, 2023).

Appendix Figure G.25: Embeddings Perfectly Separate Predistribution from Redistribution.



We also calculate Fleish-Kincaid scores for each survey question, measuring the reading level of the text. In order to maximize the predictive power of the embeddings, we predict mean support for each of the 95 questions with the 1024-embeddings using LASSO. The in-sample correlation between the LASSO prediction and the mean level of support is only 0.7, while it was over .95 for predistribution. While LASSO selected 24 embedding dimensions to predict predistribution, it only selected 9 to predict support.

Consistent with the suggestion from the LASSO regression, Appendix Table G.7 shows that support for a policy is not highly influenced by question wording in our context. We examine how support for policies is affected by our dictionary-based scores, as well as our embedding based measures. Regardless of what text features we include, our ability to predict support with either embedding-based or dictionary-based text features is extremely limited. Only when we directly include the support as predicted by the embeddings do we get a significant predictor. Table G.7 shows that wording differences, while able to predict the specific dimensions of policies, are not able to predict support for policy, either across our 8 broad categories or within. Together, these regressions suggest that in our context, differences in wording, at least as measured by embeddings, are not a major determinant of support for policies. Nonetheless, we control for the five principal components and eight clusters of embeddings (along with the other text features) in Appendix Tables A.1 and A.2, and show that neither the level of support nor the educational gradient is affected by inclusion of these text features.

Appendix Table G.6: Embedding Clusters Capture Policy Areas

Question type	Cluster number							
	1	2	3	4	5	6	7	8
Job guarantee	7	4	0	0	0	0	0	0
Minimum wage	0	0	0	0	0	17	5	9
Pro unions	4	1	0	0	0	0	0	0
Protectionsim	1	2	10	0	0	0	0	0
Transfer payments	2	3	0	0	0	0	0	0
Tax cuts not priority	0	9	0	7	7	0	0	0
Tax me more	0	2	0	0	0	0	0	0
Tax rich more	0	5	0	0	0	0	0	0

G.1. Analyzing party platforms

The text of party platforms comes from Hopkins *et al.* (2022) who digitized state and national party platforms, going back to the mid-19th century. In Panel A and B of Appendix Figure A.8 we first transparently show a few select word-counts in national platforms. In Panel A we show the occurrences of the phrases “minimum wage” in national and state platforms from 1948 to 2010. While Republican platforms never use these words very often, we see Democratic platforms reduce the use of these terms in the 1970s and 1980s. In contrast, in Panel B we see that the the occurrence of “taxes” is constant over the same period.

In Panels C-F we examine this more systematically. We use SBERT language embeddings to analyze changing patterns in the text of party platforms³⁰. We calculate embeddings for all the national and state platforms, and then measure the cosine similarity to each survey question embedding vector (so each platform gets 95 distances). We then examine how the linguistic similarity between platforms change over time, focusing on the differential change in the similarity to predistribution questions relative to redistribution questions by Democratic platforms. For example, Panel C of Appendix Figure A.8 shows that after many years of Democratic state platforms being closer to predistribution than redistribution, redistribution catches up after the 1970s, without any corresponding convergence in Republican state platforms as seen in Panel D. Panel E shows national Democratic platform distances to predistribution and redistribution moving together, but diverging after the 1980s, where Democratic national platforms become more similar to redistribution survey questions without any increase in similarity to predistribution questions. In contrast, predistribution and redistribution question text distances to Republican national platforms often move together. While there is a pivot towards redistribution and away from predistribution in the Republicans in the 1970s, it does not appear large given the large difference in levels throughout the sample period.

³⁰SBERT is a variant of BERT that produces sentence-level embeddings, which work better in long documents like party platforms.

Appendix Table G.7: Text Features and Policy Support

	(1)	(2)	(3)	(4)
predist				
predistXyear				
Polarity	-0.0490 (0.165)	-0.0335 (0.0465)	-0.120 (0.114)	-0.120 (0.114)
Subjectivity	-0.384** (0.158)	0.0879 (0.0721)	0.0362 (0.0848)	0.0362 (0.0848)
Reading Grade	0.00596 (0.00446)	-0.000713 (0.00152)	0.00745*** (0.00260)	0.00745*** (0.00260)
Scores for component 1			0.00707* (0.00388)	0.00707* (0.00388)
Scores for component 2			-0.00173 (0.00428)	-0.00173 (0.00428)
Scores for component 3			0.00965*** (0.00310)	0.00965*** (0.00310)
Scores for component 4			0.00146 (0.00343)	0.00146 (0.00343)
Scores for component 5			-0.00217 (0.00326)	-0.00217 (0.00326)
R2	0.0292	0.168	0.144	0.144
Sample	All	All	All	All
Policy FE	No	Yes	No	No
Text Features	No	No	Yes	Yes
Observations	266638	266638	266638	266638

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

Notes: