VALUES AS LUXURY GOODS AND POLITICAL POLARIZATION

Benjamin Enke
Mattias Polborn
Alex Wu

Values as Luxury Goods and Political Polarization
Benjamin Enke, Mattias Polborn, and Alex Wu
NBER Working Paper No. 30001
April 2022, Revised November 2022
JEL No. D03,D72

ABSTRACT

This paper develops a theory of political behavior in which values are a luxury good: the relative weight that voters place on values rather than material considerations increases in income. This idea both generates new testable implications and ties together a broad set of empirical regularities about political polarization in the U.S. The model predicts (i) voters who are sufficiently rich to afford voting left;(ii) that more rich than poor people vote against their material interests; (iii) that Democrats are internally more fragmented than Republicans; and (iv) widely-discussed realignments: rich moral liberals who swing Democrat, and poor moral conservatives who swing Republican. Assuming that parties set policies by aggregating their supporters' preferences, the model also predicts increasing social party polarization over time, such that poor moral conservatives swing Republican even though their relative incomes decreased. We relate these predictions to known stylized facts, and test our new predictions empirically.

Benjamin Enke
Department of Economics
Harvard University
Littauer M008
Cambridge, MA 02138
and CESifo
and also NBER
enke@fas.harvard.edu

Mattias Polborn
Department of Economics
Vanderbilt University
Nashville, TN 37235
polborn@gmail.com

Alex Wu
Harvard University
Baker Library 420
Soldiers Field
Boston, MA 02163
alex_wu@fas.harvard.edu

"Erst kommt das Fressen, dann kommt die Moral." – *"A hungry man has no conscience."*

Bertolt Brecht, The Threepenny Opera

1 Introduction

A growing body of empirical work documents that people's policy views and voting behavior depend both on the economic incentives that are traditionally emphasized by political economists, and on people's values. For instance, moral values along the universalism vs. particularism cleavage (Tabellini, 2008) are strongly predictive of policy views and voting.[1] The empirical insight that both material and non-material considerations matter for political behavior raises important theoretical questions, such as how people trade off these concerns when they are in conflict with each other, and how this affects political polarization. For example, in a widely-discussed realignment, rich people with liberal values (such as the "educational elite") swung Democratic, while poor people with conservative values (such as the "rural working class") swung Republican. Both of these trends are often interpreted as "people voting against their material interests" (e.g., Frank, 2007; Gethin et al., 2021), but it is not well-understood what brought about these changes.

To shed light on how the tradeoff between material incentives and values shapes voting behavior and polarization, we propose a theory of political behavior in which *values are luxury goods*: the relative weight that people place on non-material versus material issues increases in their absolute income. This assumption does not mean that the poor have "weaker" values, or that their values are more or less extreme. For example, rich and poor voters can both strongly believe that loyalty is an important value. Instead, the assumption of luxury goods only captures the idea that – relative to the poor – the rich are more likely to vote based on their values rather than their pocketbook. Our main objectives are (i) to present motivating evidence for the luxury goods assumption; (ii) to show that it explains a large variety of empirical regularities about U.S. political polarization, both in the cross-section and over time; and (iii) to derive new predictions that follow from the luxury goods assumption, and to empirically test them on a correlational basis.

---

[1]See Luttmer (2001); Graham et al. (2009); Haidt (2012); Enke (2020); Andre et al. (2021); Enke et al. (2022); Cappelen et al. (2022); Fehr et al. (2022); Landier and Thesmar (2022).

Motivating evidence. Our approach incorporates into formal political economy analyses the main idea behind the influential "postmaterialism" literature initiated by Inglehart (1997, 2020). Using the World Values Survey, this body of work documented that, time and again, richer people report placing a greater emphasis on their values relative to their material concerns, both in the cross section and over time.

The insight that the rich place greater emphasis on their values arguably also has intuitive appeal in the political domain, where the rich can "afford" to prioritize their views on moral and cultural issues because they need to worry less about how to feed their family. To provide further motivating evidence that this is true empirically, we show in two different datasets that the link between values and voting is considerably stronger among the rich than among the poor, such that a voter's location on the universalism-particularism continuum appears more decisive for the vote choice of the rich. This pattern is in line with the results of a recent large-scale focus group study in the U.S. that likewise identified the financially well-off as those who are the most pre-occupied with moral and cultural issues, on both the left and the right (Hawkins et al., 2019).

Model setup. Motivated by this broad body of evidence, we present a model in which voters' primitive types consist of their income and their values. We refer to people with conservative / liberal values as "moral conservatives / liberals," with the implicit understanding that these values could also be cultural in nature. In the model, values determine the voter's ideal social policy. A voter's ideal economic policy such as the desired level of redistributive taxation is determined by their relative income position in society and, potentially, also by their values. We allow values to affect both the voter's ideal social and the ideal economic policy in ways that are in line with the empirical evidence, though almost none of our results depend on this assumption.

Our main assumption is that while all voters potentially place non-zero utility weight on their values, this weight strictly increases in absolute income. As a result, voters express their values more strongly at the voting booth as they get richer.

Implications with fixed party positions. To cleanly isolate the implications of values as luxury goods for individual voter behavior, we first assume that party positions

2

are exogenously fixed (which, for individual voters, they are). To see the most basic implication of the luxury goods assumption, consider the canonical comparative static effect of an increase in a voter's relative income on his vote choice. In traditional models such as Meltzer and Richard (1981), richer voters have more conservative ideal economic policies and are, hence, more likely to vote conservatively. In our model, however, income-dependent utility weights generate an additional indirect effect of higher income. For moral conservatives, whose ideal social policy is more conservative than their materially-optimal economic policy, a higher weight on values moves the voter in a conservative direction. For moral liberals, however, this indirect effect points in the opposite direction as they can now "afford" to prioritize their liberal values over their material concerns.

As a result, our model predicts that the voting behavior of moral conservatives is more responsive to income than that of moral liberals. Indeed, the model makes the stark prediction that the effect of income on voting can even be zero or negative for voters who are sufficiently morally liberal. This produces a version of the "gauche caviar:" rich voters who support left-wing policies. Contrast this with two-dimensional voting models with fixed utility weights. In those models, every voter – no matter how morally liberal – has a threshold income level above which s/he would vote for the conservative party. In our model, this prediction – which we view as counterintuitive – does not hold.

We empirically study these predictions in two datasets, the American National Election Studies (ANES) and a representative Moral Foundations Questionnaire (MFQ) dataset. In line with our model, we find that the positive effect of income on voting Republican is substantially more pronounced among moral conservatives than among moral liberals. Moreover, again consistent with the model, the effect of income on voting Republican within the set of very morally liberal voters is even weakly *negative*.

In our model, this has direct implications for who is more likely to "vote against their economic interests." Consider the two "off-diagonal" groups of voters that have attracted considerable attention in the recent popular debate: voters who are either rich-and-morally-liberal or poor-and-morally-conservative, such that they align with each party on only one dimension. An influential body of narratives argues that the Republican party has convinced poor moral conservatives to accept low taxes on the rich, in exchange for the conservative social policies that they supposedly desire (Frank, 2007; Hacker and Pierson,

2020). Yet, while much of the public discussion tends to focus on the "puzzle" that poor moral conservatives vote Republican, it could be considered equally puzzling that rich moral liberals (such as the educational elite) often vote Democrat. Indeed, our model predicts that rich moral liberals are *more* likely to vote against their material interests than poor moral conservatives. The simple intuition is that both groups need to trade off their economic incentives and their values, and rich moral liberals place a higher weight on their values than poor moral conservatives. We again test this new prediction empirically. In both the ANES and the MFQ, voters who are rich and morally liberal are more than 20 percentage points more likely to vote Democratic than poor and morally conservative people are to vote Republican.

The asymmetry in terms of whether the rich and poor vote based on material or non-material considerations also has implications for understanding within-party supporter heterogeneity. In our model, people vote Democratic for two fundamentally different reasons: either because they are poor (in this case, more-or-less independent of their values, because the weight on these is relatively small), or because they are rich and have very liberal values. For Republican voters, on the other hand, the mirror image of this pattern is less likely to happen: the very poor will less often vote Republican even if they have conservative values because – being poor – they place a lower weight on their values. As a result, our model predicts that the set of Democratic voters is internally more heterogeneous than the set of Republican ones.

This prediction jives with popular media accounts that highlight the internal tensions between various factions within the Democratic party. At the same time, we are not aware of formal models that have made such a prediction. We again take this prediction to the data in the ANES and MFQ. Consistent with the model, Democratic voters exhibit a significantly greater standard deviation in both income and values than Republican ones.

Secular change. We then explore the implications of the luxury goods assumption for changes in voting patterns and party positions over time. We focus on the effects of long-run economic growth. For the sake of analytical simplicity, we first consider a setup in which all voters' incomes increase by the same multiplicative factor. Thus, all voters' relative economic positions remain constant, but the weight that everyone in society places

on values increases.

The implications of such a model for changes in voting behavior over time are somewhat involved because two effects occur at the same time: (i) For fixed party positions, changes in income induce voters to place a higher weight on values, which affects their vote choice. (ii) Party positions may change in response to changes in voter demand for different positions, which itself leads to changes in voting behavior.

While we ultimately analyze these mechanisms jointly through simulations in a general equilibrium framework, it is instructive to additionally analyze them separately in partial equilibrium. Regarding (i), rising absolute incomes mean that voters place a higher weight on values in the later period. As a result, "off-diagonal" voters become more likely to vote in line with their values rather than their material concerns, such that poor moral conservatives (such as the white rural working class) swing towards the Republican party, while rich moral liberals (the canonical "educational elite") swing towards the Democrats.

Regarding (ii), we analytically show that exogenous social party polarization likewise generates the realignment discussed in the previous paragraph: when the parties diverge on social issues, voters' values become more decisive for their vote choice because their utility difference from the two parties' social policies increases, even for fixed utility weights. As a result, if social party positions polarize over time, poor moral conservatives (liberals) again become more likely to vote Republican (Democratic).

In Section 6, we put these pieces together in a framework in which party positions are endogenous and determined in equilibrium, along with agents' vote choices. This requires us to explicitly model how party positions are determined. We do so by assuming that party positions aggregate the preferences of party supporters. We discuss the various potential interpretations of this assumption in detail in Section 6.

Because this two-dimensional model cannot be solved in closed form, we proceed numerically. The objective of this analysis is to work out the comparative statics implications of our model, rather than to offer a calibrated model of the U.S. economy. A robust result of our simulations is that, as incomes increase over time, parties polarize on social issues. Intuitively, as voters get richer, they care more about their values, inducing some rich moral liberals to switch from Republicans to Democrats and some poor moral conservatives to swing Republican. Because parties aggregate their supporters' preferences, this

5

realignment induces the Democratic party to move to the left and the Republicans to the right on social issues. This, in turn, induces further realignment among voters with relatively liberal or conservative values, which triggers new changes in party positions, until this process equilibrates at a higher level of social polarization. Thus, our "values as luxury goods" approach offers a new lens through which the increased relevance of cultural and moral rather than class-based dimensions of political conflict can be understood.

By a similar logic, our simulations robustly generate the well-known empirical regularity that the distributions of values of Republican and Democratic voters become less overlapping and more extreme over time, which is often referred to as "moral / cultural polarization in the electorate". Yet, in our model, this is not because the distribution of values in the population has changed, but purely a result of voter sorting.

In our baseline analysis, we assume that all voters' incomes increase by the same multiplicative factor over time. This is analytically convenient but counterfactual. Extending our approach to increasing income inequality, we use a combination of analytical derivations and model simulations to shed some light on the oft-discussed puzzle that rural working-class voters swung Republican over time *even though* their relative incomes decreased over time. In our model, this pattern can easily occur as an equilibrium outcome even if income increases are concentrated at the top. The logic is that when the rich get disproportionately richer, they place a higher weight on their values, which induces some rich moral liberals to swing Democratic. By our assumed party objective function, this induces the parties to polarize on social issues because the Democrats now partly desire to make the rich moral liberals happy by moving left, while the Republicans move right because they are no longer constrained by the rich moral liberals. Faced with such socially increasingly polarized parties, a poor morally conservative voter may well become more likely to vote Republican, even when income inequality increased. Importantly, in our model – and in line with the facts – only those poor voters swing Republican who are morally sufficiently conservative.

Thus, in the inequality extension of our model, the values-based conflict between different subsets of the rich spills over to the poor because the parties strategically reposition themselves to accommodate the changing priorities of the rich. Despite increased inequality, a poor moral conservative can become more likely to vote his / her values over

6

time as an equilibrium outcome. We view these model predictions as consistent with a large body of sociological work that argues that poor, rural whites feel that the Democratic party has "abandoned" them in favor of "extremist" social policy positions that cater to urban moral liberals (e.g., Cramer, 2016; Sherman, 2009).

Contribution and related literature. To sum up, a simple and evidence-backed idea – that the relative utility weights of material and non-material considerations depend on income – sheds light on three of the most widely-discussed empirical regularities regarding U.S. political polarization: increasing party polarization on social policy issues, realignment of rich moral liberals and poor moral conservatives, and increasing moral polarization of the electorate. Moreover, our approach generates various new testable predictions that find correlational support in the data, including (i) steeper voting-income gradients among moral conservatives than among moral liberals; (ii) that the income effect can even be zero or negative for morally sufficiently liberal voters; (iii) that more rich moral liberals than poor moral conservatives vote against their material interests and (iv) greater within-party heterogeneity on the left.

Our framework with two-dimensional types differs from much of the formal literature on electoral competition, which adopts a one-dimensional framework.[2] Our work builds on contributions that model competition in multi-dimensional settings but feature fixed weights (e.g., Schofield, 2007; Krasa and Polborn, 2014b; Dragu and Fan, 2016; Nunnari and Zápal, 2017; Xefteris, 2017). Pástor and Veronesi (2021) highlight how diminishing marginal utility of money can make a preference for equality a luxury good, but they do not study implications for polarization or realignment.

Our paper also relates to recent theoretical work on values or ethics in political economy (e.g., Tabellini, 2008; Strulovici, 2020), as well as to models that link polarization to psychological factors (e.g., Besley and Persson, 2019; Rodrik et al., 2021; Gennaioli and Tabellini, 2021; Callander and Carbajal, 2022). Our main innovation to this literature is the idea of income-dependent utility weights and the many new implications it generates.

---

[2]A framework with one-dimensional types is unappealing for our purposes because (i) the assumption that economic incentives overwhelmingly determine how people vote is empirically counterfactual, especially in recent elections; and (ii) a one-dimensional type space precludes analyses of the realignment patterns observed in practice, such as that poor-and-morally-conservative voters swung Republican, while rich-and-morally-liberal voters swung Democratic.

The paper proceeds as follows. Section 2 presents motivating evidence for the luxury goods assumption. Section 3 exposits the model. Sections 4 and 5 analyze voter behavior, taking party positions as exogenous. Section 6 sets up a simple model of endogenous party positions and simulates the general equilibrium effects of long-run economic growth in this framework. Section 7 concludes.

2   Motivating Evidence

The broad idea that the importance placed on non-material considerations increases in income has appeared in different terminology across the social sciences. In psychology, this idea is typically referred to in the context of "Maslow's hierarchy of needs" (Maslow, 1943). The most widely known empirical evidence for this idea was developed as part of the "postmaterialism" agenda of Ronald Inglehart. He and his collaborators devised the now-famous World Values Survey to document that richer people indeed place a higher weight on non-material versus material considerations, both in the cross-section and over time (Inglehart, 1997, 2020).

Evidence from the MFQ and ANES.   A question is whether the logic of "postmaterialism" also applies in the political domain. To study this, we use the same two datasets that we also use to test our new model predictions: the Moral Foundations Questionnaire (MFQ) and the American National Election Studies (ANES). The MFQ is an influential psychological questionnaire that measures people's moral values (Haidt, 2012). Based on the MFQ, Enke (2020) develops a summary statistic of the relative importance of communal (particularist) versus universalist values for a nationally representative sample of 4,011 respondents. For simplicity, we refer to this index as "moral conservatism."

Unlike the MFQ, the ANES has the advantages that it allows for (i) time series analyses and (ii) the construction of an economic conservatism index in addition to a moral conservatism index. As described in Appendix E.1, we follow the methodology of Ansolabehere et al. (2006) in constructing an index of economic conservatism from survey questions about, e.g., social security and federal spending. Our ANES moral conservatism index follows the same methodology, except that we restrict attention to questions that
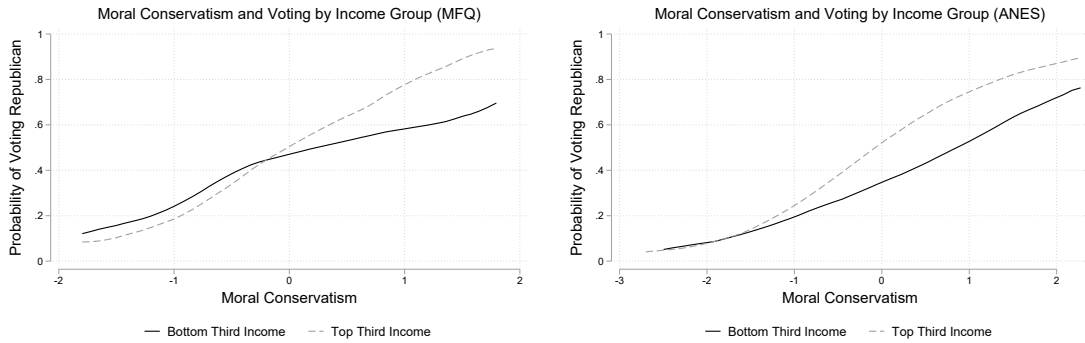
Figure 1: The figure shows local polynomial plots of the link between voting Republican and moral conservatism, separately for respondents in the top / bottom third of the income distribution. The left panel shows the link with respondents' moral values along the particularism vs. universalism cleavage in the MFQ for 2016 (Enke, 2020), while the right panel shows the link with a moral conservatism scale in the ANES (data cover 1988–2020). Both indices are normalized into z-scores. See Appendix E for details on the construction of both indices.

are related to the universalism vs. particularism cleavage, such as questions about family values, LGBTQ rights and tolerance. We do this for the sake of comparability with the MFQ results, but our results are almost identical when we include all survey questions.

The left panel of Figure 1 shows the relationship between voting Republican in 2016 and the moral conservatism index in the MFQ, separately for respondents in the top and bottom third in the income distribution. The right panel shows the same pattern for the years 1988–2020, except that the x-axis represents the social conservatism index in the ANES.[3] In both panels, we see that the link between values and voting is considerably stronger among the rich. For example, in the MFQ data, in the top third income group, the probability of voting Republican increases by 60 percentage points going from moral conservatism of less than (-1) to conservatism greater than 1, while the corresponding increase in the bottom third income group is only 46 percentage points. Appendix Table 2 formally confirms the statistical significance of this pattern. The table also shows that this pattern is considerably more pronounced for income than for education.[4]

These empirical patterns are at odds with popular qualitative accounts that assert that the voting behavior of the poor is to a much larger extent driven by values than

---

[3]We cannot implement a median split because the middle income group in the ANES data comprises respondents between the 34th and 67th percentile of the income distribution.

[4]A potential concern is that the results in Figure 1 reflect higher measurement error and, hence, greater attenuation among the poor. However, as we show in Appendix Figure 8, we don't see a similar pattern for the link between economic policy views and voting. This suggests that the main mechanism is not more noise for poorer people.

that of the rich (e.g., Frank, 2007; Hacker and Pierson, 2020). Rather, we interpret the data as saying that values are a luxury good: whichever values people hold, they matter more for people's vote choice the richer they are. Also note that our account differs from the idea that income affects *which* values people hold. Indeed, the correlations between income and moral conservatism are small and inconsistent in sign ($r = 0.06$ in the MFQ and $r = -0.05$ in the ANES).

A second, perhaps surprising, pattern in Figure 1 is that, within the set of morally relatively liberal people, the rich are no more likely to vote Republican than the poor are. Indeed, in a two-dimensional voting model with fixed utility weights for material concerns and non-material (moral) issues, the probability of voting Republican would always strictly increase in income. Our model with income-dependent utility weights, on the other hand, will be consistent with the zero correlation within the set of morally liberals voters (see Section 4.2).

Related Evidence. There is also other, independent evidence to suggest that values are particularly important for the political views and voting of the rich (also see the discussion in Gelman et al. (2007)). Based on a large focus group study of 8,000 Americans, the recent "Hidden tribes" report (Hawkins et al., 2019) partitions voters into different groups according to their primary political concerns. The two groups that are most preoccupied with moral and cultural issues are labeled "progressive activists" on the left and "devoted conservatives" on the right. The report describes these two voter groups by highlighting that their material positions are unusually secure:

> *"Progressive activists: Their own circumstances are secure, which perhaps frees them to devote more attention to larger issues of justice in society... [such as] issues of fairness and equity."*
> *"Devoted conservatives: They are one of the highest income-earning groups, and feel more secure than most other Americans. They are highly engaged in social issues...[and] feel that traditional values are under assault."*

Arguably, this implicity or explicitly alludes to the idea that materially secure voters can "afford" to prioritize their values. On the flipside, the report highlights that those voter groups that are less concerned with moral or cultural issues are materially less secure.

Similarly, U.S. Senator Sherrod Brown recently asserted that:[5] "*When you're on the edge, worried about the next layoff or health setback and struggling to make ends meet, the latest Twitter feud or cable news controversy [about culture wars] is just background noise.*"

3 Model

3.1 Preferences

We write all relationships that depend on income as functions of the natural logarithm of income, denoted y. We assume that the deterministic part of an agent's overall utility is given by a weighted average of two components, which we call material (M) and non-material (or values-based) utility, P:

$$U = [1 - w(y)]M + w(y)P. \tag{1}$$

Here, the weight function ranges from 0 to 1, and is increasing and concave. For example, in the general equilibrium section and for some illustrations, we use $w(y) = \frac{c + \rho(y-k)}{1 + \rho(y-k)}$, where $c \in (0, 1)$ is a lower bound on the weight placed on moral utility, and $k \in [0, y)$ and $\rho > 0$ are parameters. We can approach a standard framework (i.e., one where the weight placed on moral utility is constant in income) by letting $\rho \to 0$, so that the utility function collapses into one in which voters place constant weights c and $1 - c$ on non-material and material utility, respectively. We do not assume that the values-based utility weight is zero (or negligible) for relatively poor voters. It may well be that the poorest voter in the U.S. places a non-trivial weight on moral and cultural issues – we only assume that the weight increases in income.

Discussion. In standard household theory, some physical goods are inferior and some are normal. In much the same way, it would be very surprising if the rate at which voters trade off material wealth and their values is unaffected by their wealth. While diminishing marginal utility is a plausible reason for why higher incomes might lead to

---

[5]See https://www.sherrodbrown.com/news/2022/the-american-prospect-op-ed-becoming-the-workers-party-again/.

a higher effective weight on values, this is not mathematically true for all concave utility functions. The reason is that there are two countervailing effects. First, as people get richer, their marginal utility of money decreases, which makes values more important at the margin. Second, however, as people get richer, economic policy also becomes more important to them in absolute terms as their tax base increases. Nevertheless, it is clearly conceivable that the effective relative weight that a voter puts on values changes with income due to diminishing marginal utility. We are agnostic about whether the weight function $w(y)$ reflects the effects of diminishing marginal utility or some other effect.

Material utility. Material utility is a function of individual income and the implemented economic policy x_e. Rather than setting up an explicit model of material utility, we employ a reduced-form one, which is compatible with several interpretations:

$$M(x_e, y, \bar{y}) = y - \frac{1}{2} \left[x_e - \underbrace{(r(y - \bar{y}))}_{\substack{\text{Materially-optimal} \\ \text{economic policy}}} \right]^2. \tag{2}$$

Here, y denotes individual (log) income, and $\bar{y}$ average (log) income in society. We assume that $r(\cdot)$ is an increasing and differentiable function of the individual's relative economic position in society. Many natural economic models have the property that materially-optimal policies are monotone in relative income. For example, in a model of redistribution, richer agents prefer lower taxes. Similarly, if the level of taxation determines the amount of public goods that can be provided to citizens, the ideal level of taxation will also depend on relative income.

Note that (2) implies that, if all citizens' incomes increase by the same factor (i.e., log incomes increase by the same constant for all voters), the respective materially-optimal policy is unchanged for all citizens. Also, the equivalent variation of the utility loss from any non-optimal economic policy, relative to the individual's optimal policy, is a constant fraction of income that is unchanged if all citizens' incomes increase by the same factor. Thus, economic interests as captured by (2) scale proportionately.

Values-based (non-material) utility. Non-material utility P depends on the deviation of social policy, x_s, from one's values, captured by μ. Here, we think of higher

12

μ and x_s as reflecting more conservative values and social policies, respectively. This is the first component of eq. (3). We emphasize that μ does not capture whether a voter has "better" or "worse" values but, instead, whether a voter has more conservative (e.g., particularist) or liberal (e.g., universalist) values. Our model is one of horizontal rather than vertical differences in values. Henceforth, we refer to voters with high (low) μ as "moral conservatives (liberals)," with the implicit understanding that some of the relevant values need not be moral in nature (e.g., cultural values).

We also allow values to potentially affect how the voter thinks about economic policy. For example, the voter may have a bliss point that determines which tax rate or level of redistribution he finds morally fair. In line with the evidence, we assume that voters who have more particularist (less universalist) values desire more conservative economic policies. This assumption implies that voters' economic and social ideal points are correlated. There is a considerable body of evidence to support this assumption. In the economics literature, both Enke et al. (2022) and Cappelen et al. (2022) document that a person's degree of universalism is strongly predictive of both left-wing social policy views (e.g., immigration and affirmative action) and left-wing economic policy views (e.g., redistribution and universal health care). Similarly, in psychology, the meta-study by Kivikangas et al. (2021) documents that universalism (as measured using the MFQ) is predictive of both social and economic left-wing orientation. While we wish to allow for these empirical regularities, almost all of our results are driven by the assumption of values as luxury goods (i.e., that $w(y)$ is strictly increasing), rather than by the relation between ideal social and economic policies. To make this transparent, we formulate overall non-material utility as

$$P(x_e, x_s, \mu) = -\frac{1}{2}(x_s - \mu)^2 - \frac{\lambda}{2}\left[x_e - \underbrace{h(\mu)}_{\substack{\text{Morally-optimal} \\ \text{economic policy}}}\right]^2, \qquad (3)$$

where $h(\cdot)$ is an increasing differentiable function capturing how the values-optimal economic policy depends on μ. The parameter $\lambda \geq 0$ determines how much ideal economic policies depend on values. If $\lambda = 0$, the voter's ideal economic policy only depends on income and is uncorrelated with his ideal social policy. Below, all propositions will clearly indicate if they require $\lambda > 0$ (only Proposition 1 does).

13

Utility function. Substituting (2) and (3) into (1) yields

$$U = [1 - w(y)]\left\{ y - \frac{1}{2} \left[x_e - r \left(y - \bar{y} \right) \right]^2 \right\} - w(y) \left[\frac{\lambda}{2} (x_e - h(\mu))^2 + \frac{1}{2} (x_s - \mu)^2 \right] \qquad (4)$$

The key characteristic of this utility function is that the utility derived from social policy is a strongly superior good (luxury good). To see this, consider the equivalent variation of an individual's utility loss due to social policy differing from the individual's ideal position μ. Suppose the individual's income increases. For fixed weight w, the equivalent variation would increase proportionately with income. However, since w is increasing in y, the equivalent variation increases more than proportionately with income.

3.2 Ideal Policies

Turning to a voter's ideal economic and social policies that maximize (4), we have $x_s^* = \mu$. Maximizing with respect to x_e gives

$$\begin{aligned} x_e^* &= \frac{1 - w(y)}{1 - (1 - \lambda)w(y)} r \left(y - \bar{y} \right) + \frac{w(y)\lambda}{1 - (1 - \lambda)w(y)} h(\mu) \\ &= [1 - W(y, \lambda)] r \left(y - \bar{y} \right) + W(y, \lambda)h(\mu), \end{aligned} \qquad (5)$$

where $W(y, \lambda) \equiv \frac{w(y)\lambda}{1-(1-\lambda)w(y)}$. Thus, x_e^* is a weighted average of the materially-optimal economic policy and the values-optimal economic policy, where the weight on the values-optimal policy increases in income if and only if $\lambda > 0$. The relationship between income and the ideal economic policy is more involved than in standard models because of the indirect effect that is due to $W(y, \lambda)$.

3.3 Voting

Let x_e^i and x_s^i be the economic and social positions of party i. We assume that there are exactly two parties, called Democrats and Republicans.[6] In Sections 4 and 5, party positions are exogenous, and we assume that, on both dimensions, Republican positions are more conservative than Democratic ones (i.e., $x_e^D < x_e^R$ and $x_s^D < x_s^R$). In Section 6,

---

[6]In the United States, plurality rule ensures that only two parties are viable (for a normative justification of the optimality of two parties, see Lizzeri and Persico (2005)).

we endogenize the party positions.

In addition to their policy utility, voters also receive an idiosyncratic utility shock ε, which could reflect local candidates' competence or the voter's evaluation of differences between the parties beyond the economic and social policy dimensions that are at the center of our model. The utility shock enters additively to utility from party positions, and is measured in a way that positive realizations indicate a non-policy preference for the Democratic party. That is, voter j supports the Republican party if and only if

$$U_j^R - U_j^D > \varepsilon_j,$$

where the left-hand side is voter j's net policy preference for the Republican positions. Let the density of the net utility shock ε be denoted by ψ, and assume that it is always positive, with associated cumulative distribution Ψ. Then, the probability that a voter votes for the Republican candidate is given by $\Psi(U^R - U^D)$, where, of course, the policy utility difference depends on the voter's type (y, μ).

It is useful and without loss of generality to normalize policies such that they are symmetric around zero (i.e., such that $x_e^D = -x_e^R$ and $x_s^D = -x_s^R$). Let $\Delta x_e \equiv x_e^R - x_e^D$ and $\Delta x_s \equiv x_s^R - x_s^D$ denote the degree of economic and more polarization between Republicans and Democrats. Using this, we can write the utility difference ΔU as

$$\Delta U \equiv U^R - U^D = [1 - w(y)]\Delta x_e r\,(y - \bar{y}) + w(y)\left[\lambda \Delta x_e h(\mu) + \Delta x_s \mu\right]. \qquad (6)$$

4 Policy Views and Voting in the Cross-Section

Our model analysis proceeds in three steps. In Section 4, we present predictions about voters' ideal policies and voting behavior that follow purely from the preferences described in the previous section, holding party positions fixed. We test these model predictions using cross-sectional data. Second, in Section 5, we present results on changes in voting behavior over time that result from assumed, exogenous changes in economic conditions and / or party positions. Then, in a third step, Section 6 analyzes endogenous party positions, how they change over time, and corresponding implications for voting behavior.

Throughout the paper, all empirical analyses are correlational in nature. Our objective is to test whether our model predictions are consistent with reality, rather than to supply causally-identified tests of each of our predictions.

4.1 Economic Policy Views in the Cross-Section

To begin, it is instructive to consider how support for a higher (more conservative) economic policy changes as a function of income:

$$\frac{\partial x_e^*}{\partial y} = \underbrace{[1 - W(y, \lambda)]r'(y - \bar{y})}_{\text{direct income effect, } +} + \underbrace{\frac{\partial W(y, \lambda)}{\partial y}[h(\mu) - r(y - \bar{y})]}_{\text{indirect values-based effect, ?}} \tag{7}$$

The first term (direct effect) is positive, as in standard political economy models: from the perspective of material incentives, higher income pushes the voter to prefer a more conservative economic policy, as in Meltzer and Richard (1981). However, the sign of the second term (indirect effect) depends on whether an individual's values-optimal economic policy, $h(\mu)$, is to the left or to the right of his materially-optimal economic policy, $r(y - \bar{y})$. For voters who are sufficiently morally conservative (high μ), both the direct and the indirect income effect are positive. For morally liberal people, on the other hand, the two effects point in opposite directions because the higher weight on values leads them to favor left-wing economic policies. Thus, the desired economic policies of rich and poor moral liberals are more similar to each other than those of rich and poor moral conservatives. This gives rise to the concept of the "gauche caviar": people who are sufficiently rich that their values make them economically left-wing.

Proposition 1. *Suppose that $\lambda > 0$. Then, the relationship between economic policy views and income is stronger among moral conservatives. Equivalently, the relationship between economic and social policy views is stronger among the rich. Formally,*

$$\frac{\partial^2 x_e^*}{\partial y \partial x_s^*} = h'(\mu)\frac{\partial W(y, \lambda)}{\partial y} > 0.$$

The proof is obvious from (7) by noting that $x_s^* = \mu$. This proposition is the only one in the paper that relies on the assumption that values affect not only voters' social policy

views but also their economic policy views ($\lambda > 0$).

Empirical evidence. This is a new (and distinctive) prediction of our model that crucially depends on the idea that the non-material weight increases in income. For example, two-dimensional models with constant weights do not make such a prediction. We therefore proceed by investigating the empirical patterns in the ANES. We use the economic conservatism index as proxy for x_e^* and the moral conservatism index as proxy for $x_s^* = \mu$. Figure 2 summarizes the results.[7] We plot standardized beta coefficients that report by which fraction of a standard deviation economic conservatism changes in response to a one standard deviation increase in log income. We show the results of two regression specifications, one without controls and one that conditions on state and year FE, age, age squared, gender and education. Throughout the paper, we reference regression evidence in figure notes and relegate the tables to appendices. Consistent with Proposition 1, we find that the coefficient in a regression of economic conservatism on income is more than twice as large among above- than among below-median moral conservatism respondents.

4.2   Voting in the Cross-Section

Recall that the probability of voting R is given by $\Psi(\Delta U) = \Psi(U^R - U^D)$. Differentiating (6) with respect to μ shows that

$$\frac{dPr(R|y,\mu)}{d\mu} = \psi(\Delta U)\frac{d\Delta U}{d\mu} = \psi(\Delta U)w(y)\left(\lambda h'(\mu)\Delta x_e + \Delta x_s\right) > 0, \qquad (8)$$

which implies that morally-conservative types are more likely to vote Republican. In contrast, the effect of an increase in income y on voting is more complicated:

$$\frac{dPr(R|y,\mu)}{dy} = \psi(\Delta U)\left\{\underbrace{[1 - w(y)]\Delta x_e r'(y - \bar{y})}_{\text{direct income effect, +}} + \underbrace{w'(y)[\lambda \Delta x_e(h(\mu) - r(y - \bar{y})) + \Delta x_s \mu]}_{\text{indirect values-based effect, ?}}\right\} \qquad (9)$$

As in the analysis of economic policy views, there is a direct income effect and an indirect values-based effect. The sign of the direct income effect is positive because, from a purely

---

[7]We use coefficient plots based on sample splits to visualize our results, but none of the results depend on the arbitrary median splits, see e.g. Appendix Table 3.
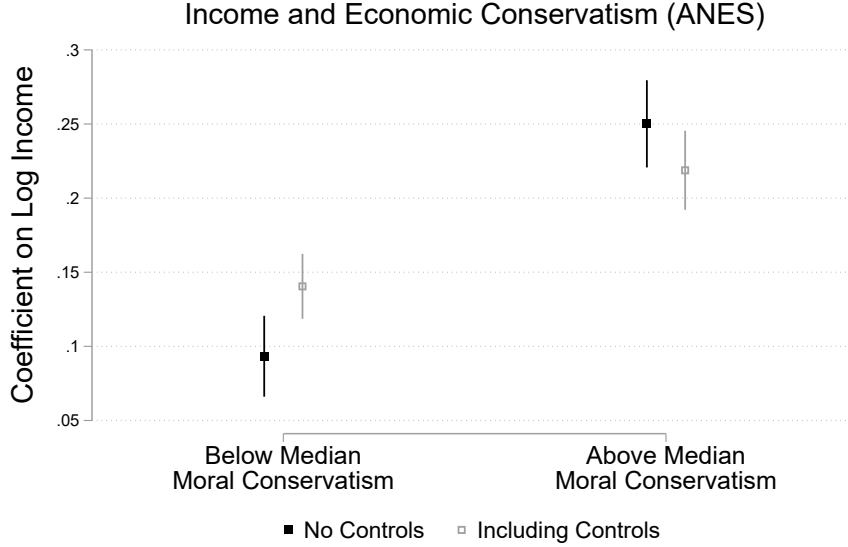
Figure 2: Coefficients from OLS regressions of the economic conservatism index on log income, split by moral conservatism. 95% confidence intervals reported. In a Wald test of the equality of coefficients, $p < 0.001$. Data from ANES, total $N = 18,193$. Income and the conservatism indices all have mean zero and standard deviation one. Controls include state and year FE, gender, age, age squared and education. Regression results are reported in Appendix Table 3.

material perspective, richer voters prefer the Republicans' more conservative economic policy (i.e., $r(\cdot)$ is increasing). The sign of the indirect effect again depends on a voter's values. It is positive for morally conservative voters, which implies that the probability of voting Republican is unambiguously increasing in income for these voters. However, the indirect effect is negative for morally sufficiently liberal voters. Intuitively, if an increase in income makes a voter rich enough to "afford" voting based on values, they can become more likely to vote Democrat if they are sufficiently morally liberal. We now state results based on this intuition formally in Proposition 2.

Proposition 2. *1. For any y, there exists a critical level $\mu^*(y)$ such that $\frac{dPr(R)}{dy} \gtreqqless 0$ if and only if $\mu \gtreqqless \mu^*(y)$.*

2. We have $\frac{d^2 Pr(R)}{d\mu dy}$, as given by (10) below, positive if $|\psi'(\Delta U(y, \mu))|$ is sufficiently small. In particular, this is the case if ψ is uniform.

Proof. Part 1 follows directly from arguments given above. For Part 2, differentiating (9)

with respect to μ yields

$$
\begin{aligned}
\frac{d^2 Pr(R)}{d\mu dy} &= \psi(\Delta U) w'(y)[\lambda h'(\mu) \Delta x_e + \Delta x_s] + \\
&\psi'(\Delta U) \frac{d\Delta U}{d\mu} \left\{ [1 - w(y)] \Delta x_e r'(y - \bar{y}) + w'(y)[\Delta x_e(\lambda h(\mu) - r(y - \bar{y})) + \Delta x_s \mu] \right\}.
\end{aligned}
\tag{10}
$$

Positivity of (10), if $|\psi'|$ is sufficiently small, follows because then the second term is close to zero, and the term in the first line is positive: the density, the derivative of the weight, and both terms in the parenthesis are positive. □

The first item in Proposition 2 shows that, for each income level y, there is a critical level of moral conservatism, $\mu^*(y)$, such that the marginal income effect is negative for more liberal voters, and positive for more conservative ones. This claim follows immediately because the terms in curly brackets (9) are strictly increasing in μ in an unbounded way. This first part of the proposition has two closely related implications. First, among morally very liberal voters, the income effect on voting Republican can be zero or even negative. Second, unlike in a two-dimensional model with fixed utility weights, in our model it is not true that every voter – no matter how morally liberal – could state a cutoff income level above which s/he would vote Republican: in our model, not every moral liberal "has a price" because even a very high income need not push in a conservative direction if it increases the weight placed on values sufficiently strongly.

The second item in Proposition 2 contains a somewhat stronger claim, namely that the marginal effect of income is increasing in the voter's moral conservatism, μ. This is not true in general because there are two effects. The first one, corresponding to the intuition discussed above, has a definitive sign. The second one is related to the fact that an increase in μ changes ΔU and thus $\psi(\Delta U)$, the value of the density of ε at ΔU. If the distribution of ε is sufficiently close to a uniform distribution, then the size of the second effect is small, and the first effect dominates. If $\psi' \approx 0$, so that we can neglect the effect of changes in the density, Proposition 2 states that voting is more responsive to variation in income among moral conservatives than among moral liberals.[8] As the proof highlights, this prediction is again driven by the assumption of income-dependent utility weights, $w'(y) > 0$.

_____

[8]We discuss the case where ψ is not approximately constant in Online Appendix A.
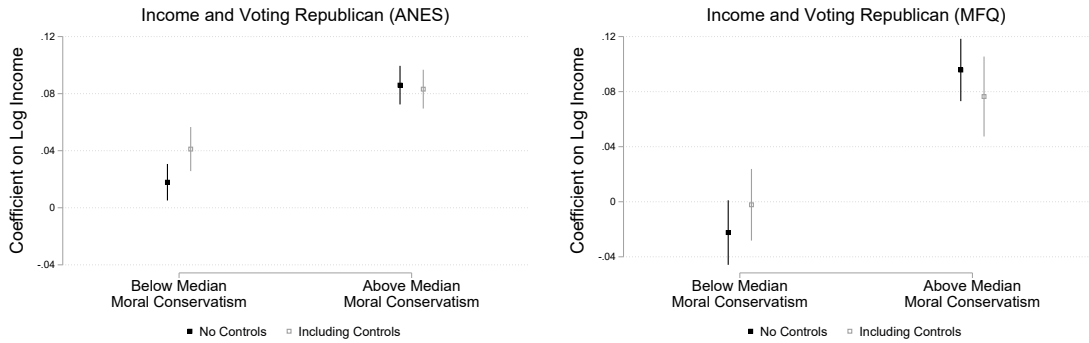
Figure 3: Log income coefficient in a voting regression, by level of moral conservatism. The left panel uses data from the ANES ($N = 16,164$) and the right panel from the MFQ ($N = 3,172$). In a Wald test of the equality of coefficients, $p < 0.0001$ in both datasets. Log income is standardized to have mean zero and standard deviation one. Controls include state and year FE, gender, age, age squared and education. Regression results are reported in Appendix Tables 4 and 5.

Empirical evidence. The prediction implied by the first part of Proposition 2 clarifies that our model is consistent with one of the main stylized facts that emerged from the discussion of the motivating evidence in Section 2: that rich moral liberals and poor moral liberals may exhibit the same voting patterns. In other words, the empirical evidence summarized in Figure 1 in Section 2 indeed strongly suggests that, within the set of morally liberal voters, higher incomes are not associated with a higher probability of voting Republican. Indeed, in the MFQ data, the correlation between income and voting Republican is even significantly *negative* when we restrict attention to the 25% most morally liberal voters in the data ($p < 0.05$).

Regarding the second part of Proposition 2, Figure 1 already provided suggestive evidence in favor of this prediction.[9] To test it more rigorously, we regress a binary voting indicator on household income, separately for above- and below-median moral conservatism. As shown in Figure 3, in both the ANES and the MFQ, the effect of a one standard deviation increase in log income on the probability of voting Republican is about 8-10 percentage points larger among moral conservatives. Appendix Tables 4 and 5 document that this result does not hinge on the median sample split but also holds with the continuous moral scale.

---

[9]The idea that the voting-income gradient is steeper among moral conservatives has loosely been discussed in the literature. For example, McCarty et al. (2016) and Gelman et al. (2007) both report that the income gradient is particularly steep among religious people.

4.3 Voting Against One's Material Interests

The insight that values are more relevant for the voting behavior of the rich affords interesting insights into the behavior of "off-diagonal" voters: those (y, μ) types who are either rich-and-morally-liberal or poor-and-morally-conservative. These voters align with the Democrats and the Republicans on one dimension each, such that they are potential swing voters in elections. Moreover, these groups have recently received a large amount of attention in the popular discussion. For instance, in the bestseller "What's the matter with Kansas?," Thomas Frank (2007) argues that the swing of relatively poor white working-class voters to the Republican party reflects that people vote "against their interests," which Frank identifies exclusively with their economic interests. Interestingly, while Frank accuses the rich moral liberals in the Democratic party of being responsible for Democrats losing the heartland working class, he does not discuss that they themselves often vote against their (economic) interests, too. In fact, the existing literature has rarely compared how much poor moral conservatives actually vote Republican *relative to* how much rich moral liberals vote Democratic.

Proposition 3 compares the voting behavior of a rich moral liberal (RML) and a poor moral conservative (PMC) who are, in a sense, mirror images of each other: they have moral ideal points that are symmetric around zero and their materially-optimal economic policies are also symmetric around zero. Thus, RML is better off economically with the Republican platform, but prefers the Democrats values-wise, and vice versa for PMC.

Proposition 3 shows that RML is actually more likely to vote against their material interests than PMC. The simple intuition is that, because of the income effect on the utility weights, it is more important for RML to align with the Democrats values-wise than it is for PMC to align with the Republicans. Thus, the following proposition is again a reflection of the assumption of income-dependent utility weights.

Proposition 3. *Suppose that h is a linear function (i.e., $h(\mu) = \ell\mu$ for some $\ell > 0$), and that the function ψ is symmetric around 0 (i.e., no party has an expected valence advantage). Consider two voters, RML and PMC, such that $-\mu_{PMC} = \mu_{RML}$ and $r(y_{RML} - \bar{y}) = -r(y_{PMC} - \bar{y})$. Then, $Pr(D|RML) > Pr(R|PMC)$.*

The proof is in Section 8. Since we cannot observe individual *probabilities* of voting
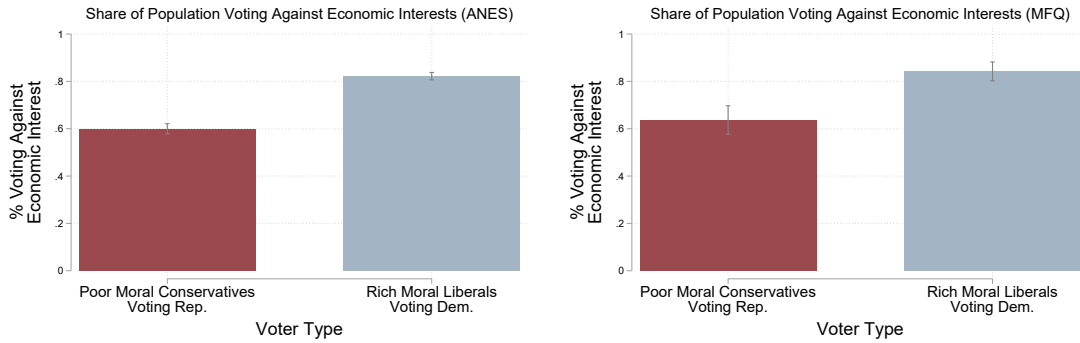
Figure 4: Sample mean and 95% confidence interval of the share of poor moral conservatives who vote Republican, and rich moral liberals who vote Democrat. Here, we define "rich" and "poor" as the top/bottom third of household income, and "morally conservative / liberal" as top/bottom third of the moral conservatism indices. The left panel uses ANES data ($N = 4,274$) and the right panel MFQ data ($N = 608$). In t-tests, $p < 0.001$ in both datasets.

Republican, an empirical test of Proposition 3 requires that we aggregate across voters. If, in such an aggregation, every voter in the set of "rich moral liberals" is perfectly matched with a mirror "poor moral conservative" type as defined in Proposition 3, then the percentage of Democratic votes in the former set should be larger than the percentage of Republican votes in the latter set. Moreover, by continuity, the result holds in such an aggregation even if the two empirical groups are slightly unbalanced (i.e., do not consist of voters that are *exact* mirror images of each other.

Empirical evidence. Given that this new prediction is in stark contrast with many popular (qualitative) narratives, we again test it using ANES and MFQ data. Figure 4 plots the fraction of rich moral liberals and poor moral conservatives who vote "against their material interests." Here, we define "rich" and "poor" as the top/bottom third of household income, and "morally conservative" and "morally liberal" as top/bottom third of the moral conservatism indices. Figure 4 shows that rich moral liberals are about 35% more likely to vote Democrat than poor moral conservatives are to vote Republican. These comparisons are statistically significant in both datasets.

4.4 Within-Party Supporter Heterogeneity

The higher emphasis on values among the rich also has direct implications for within-party constituent heterogeneity in income and values. Our main insight is that the set of

Democratic voters is more fragmented than the set of Republican ones. To show this, we first introduce the concept of an iso-probability curve.

Definition 1. *Iso-probability curves are given by the set of (y, μ) for which the voting probability is constant, i.e., for which ΔU is constant.*

A special case of an iso-probability curve is the level set for which the voting probabilities are given by $1/2$. All voters who are located to the right of this curve are more likely to vote R rather than D, and vice versa.

Proposition 4 below analyzes properties of the iso-probability curves. Iso-probability curves are usually downward-sloping. We show that this holds also in our model for voters who are not too morally liberal, and also applies for all voters if the weight function is of the logistic form (i.e., $w(y) = \frac{c+\rho(y-k)}{1+\rho(y-k)}$, with ρ sufficiently small).[10]

More importantly, if the weight function is either linear or logistic, then iso-probability curves are always strictly convex. Convexity establishes a sense in which the set of D voters is more heterogeneous than the set of R voters. Intuitively, the reason for this asymmetry is that rich moral liberals are more willing to vote against their material interests than poor moral conservatives. Thus, in our model, the D party is a "big tent" that attracts both the very poor and some of the very rich.

This prediction on differential within-party heterogeneity is quite particular to our model in that it crucially depends on $w'(y) > 0$. If, in contrast, the utility weight of non-material issues is constant ($\rho = 0$), then the iso-probability curves are linear and downward-sloping. This clarifies that what generates our prediction about differential supporter heterogeneity is indeed values as luxury goods, rather than a two-dimensional setup as such.

Proposition 4.   *1. The slope of the isoprobability curve is given by*

$$\frac{dy}{d\mu} = -\frac{w(y)\left[\lambda h'(\mu)\Delta x_e + \Delta x_s\right]}{[1-w(y)]\Delta x_e r'(y-\bar{y}) + w'(y)[\lambda \Delta x_e(h(\mu) - r(y-\bar{y})) + \Delta x_s \mu]}$$
$$= -\frac{\frac{w(y)}{1-w(y)}\left[\lambda h'(\mu) + \frac{\Delta x_s}{\Delta x_e}\right]}{r'(y-\bar{y}) + \frac{d}{dy}\left(\frac{w(y)}{1-w(y)}\right)[1-w(y)][\lambda h(\mu) - r(y-\bar{y}) + \frac{\Delta x_s}{\Delta x_e}\mu]} \tag{11}$$

---

[10]In principle, the indirect income effect that operates through $w(\cdot)$ could be so strong for moral liberals that it makes some iso-probability curves upward-sloping. As the indirect income effect is proportional to ρ, assuming that ρ is not too large effectively limits the size of this effect.
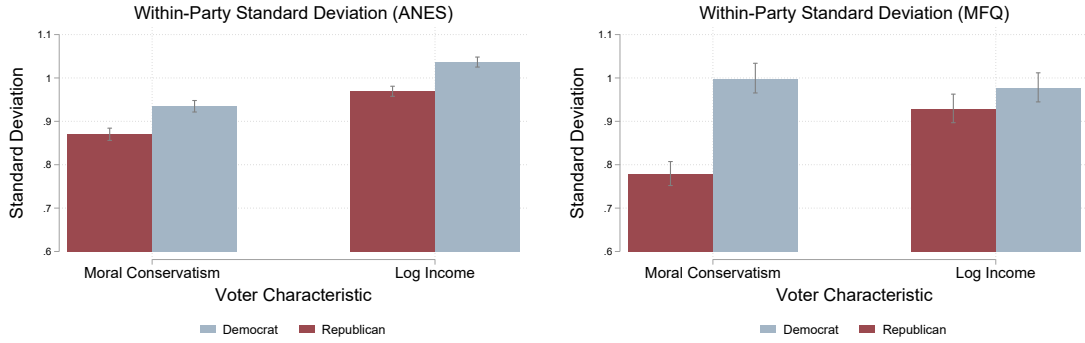
Figure 5: Sample standard deviation and 95% confidence interval of moral conservatism and log income, split by whether a respondent voted Republican or Democrat. The top panel uses ANES data ($N = 13,175$) and the bottom panel MFQ data ($N = 3,471$). In F-tests of the equality of standard deviations across groups, the p-value is $p < 0.01$ for moral conservatism and log income in the ANES, and for moral conservatism in the MFQ. For log income in the MFQ it is $p = 0.08$.

A sufficient conditions for (11) *to be negative at* (μ, y) *for any weight function is that* $\lambda h(\mu) - r(y - \bar{y}) + \frac{x_s^R - x_s^D}{x_e^R - x_e^D} \mu \geq 0$ *(i.e., isoprobability curves are downward-sloping for all individuals who are not too liberal).*

An alternative condition for (11) *to be negative for all voters is that* $w(y) = \frac{c + \rho(y-k)}{1 + \rho(y-k)}$, *with* ρ *sufficiently small.*

2. *Suppose that* $r(\cdot)$ *and* $h(\cdot)$ *are linear functions. If the weight function satisfies* $\frac{d}{dy}\left(\frac{w'(y)}{w(y)}\right) \leq 0$ *and* $w''(y)(1 - w(y)) + 2[w'(y)]^2 \geq 0$ *for all* y, *then all downward-sloping iso-probability curves are convex.*

In particular, a linear weight function, and $w(y) = \frac{c + \rho(y-k)}{1 + \rho(y-k)}$ *with* $\rho > 0$ *satisfy these conditions.*

The proof is in Section 8. Proposition 4 suggests that it is interesting to empirically compare intraparty standard deviations of income and values for Democrats and Republicans. For example, when the type distributions of y and μ are uniform, then the proposition says that the standard deviation of values and income should be higher in the set of Democratic voters.

Empirical evidence. To test the novel prediction that the set of Democratic supporters is more heterogeneous, we compute the standard deviation of income and moral conservatism within the sets of ANES and MFQ respondents that voted Republican and Democrat, respectively. Figure 5 summarizes the results. All four comparisons point in

the hypothesized direction, as Democratic voters are more heterogeneous in terms of their values and income across datasets. Three of these comparisons are statistically highly significant. This evidence is also consistent with the data summarized in Rodden (2019), who argues that the Democratic party represents a more heterogeneous set of districts than the Republicans. Moreover, Bock (2020) finds that Democrats exhibit greater within-party variation in views on immigration than Republicans.

5   Changes over Time: Partial Equilibrium

In studying changes over time, we focus on secular change and the consequences of long-term economic growth. We view this medium- or long-run analysis as a primary application of our theory because we believe that the relevant utility weights might be relatively slow to change. For such an analysis, it is analytically convenient to analyze a framework in which voters' relative incomes remain constant. Thus, we assume that the income of all voters is scaled multiplicatively by $\beta > 1$ between an "early" period and a "late" period. Since the income variable in the model, y, is the logarithm of income, this means that y grows by $\ln(\beta)$.

In principle, an increase in aggregate income could have two effects that are relevant for voting behavior. First, for fixed party positions, voters' preferences could change in a way that affects their voting behavior. Second, party positions could change in response to changes in voter preferences. In Section 6, we will simultaneously study these effects in a general equilibrium analysis through simulations.

However, as a prequel, it is instructive to first study each of these two cases separately analytically. To do this, we first ask how voting behavior changes in response to an increase in incomes, holding party positions fixed. Then, in a second step, we ask how exogenous party polarization on social issues affects voting behavior, holding incomes fixed. This latter thought experiment is relevant because in Section 6 we will see that increasing incomes indeed induce social party polarization.

The bottom line of this analysis is that these two effects point in the same direction: a realignment of rich-and-morally-liberal voters (such as the "educational elite") from Republicans to Democrats, and a realignment of poor-and-morally-conservative voters

(such as the rural working class) from Democrats to Republicans. These are arguably two of the most widely discussed voter realignments that have taken place over the last 50 years (Frank, 2007; Hacker and Pierson, 2020; Gethin et al., 2021).

5.1   Realignment Through Increasing Incomes

The following Proposition shows that, as incomes increase over time, poor moral conservatives become more likely to vote Republican, and rich moral liberals become more likely to vote for Democrats.

Proposition 5. *Assume that each voter's income increases by a factor of $\beta - 1$ (i.e., a voter's y increases by $\ln(\beta)$), while party positions remain fixed. Then a voter's probability of voting Republican increases if and only if*

$$\lambda h(\mu) + \frac{(x_s^R - x_s^D)}{(x_e^R - x_e^D)}\mu - [r\,(y - \bar{y})] > 0. \tag{12}$$

Proof. Consider how $U^R - U^D$ changes with β. Since a voter's relative income remains unchanged by β, we have

$$\frac{\partial(U^R - U^D)}{\partial\beta} = \frac{w'(y + \ln(\beta))}{\beta}\Delta x_e \left\{ \lambda h(\mu) + \frac{\Delta x_s}{\Delta x_e}\mu - [r\,(y - \bar{y})] \right\} \tag{13}$$

As the first two terms are positive, the sign of this expression is the same as the sign of the term in curly brackets, as claimed. $\qquad\square$

To interpret this proposition, define a moral conservative (liberal) as a voter for whom $\lambda h(\mu) + \frac{x_s^R - x_s^D}{x_e^R - x_e^D}\mu > (<)0$. Likewise, define a voter as rich (poor) if $r\,(y - \bar{y}) > (<)0$.

For morally conservative and poor voters, the first term in (12) is positive, and the second (deducted) term is negative, so that the left-hand side of (12) is positive. Thus, poor moral conservatives become more likely to vote Republican. Analogously, for rich moral liberals, both effects go in the opposite direction, so they become more likely to vote for Democrats. These effects are intuitive: because voters weight their values more highly after everyone's income increases, a realignment based on values takes place.

5.2 Realignment Through Party Polarization

It is useful to build intuition for the model mechanics by considering how voting behavior would change *if* one knew that party positions changed in particular ways.[11] To analyze the effect of party polarization on the different types of voters, we consider a symmetric change in party positions that leaves the midpoint between the two parties' positions unchanged; that is, when Democrats move to the left (on economic and/ or social issues), then Republicans move to the right by the same amount.[12] As before, and without loss of generality, we normalize the mid-point of party positions to zero.

The first part of Proposition 6 deals with a balanced polarization on both economic and social issues and shows that this intensifies the probability that any given type votes for his policy-preferred party ("sorting"). In contrast, the second part deals with polarization on social issues, which leads to a systematic realignment of which voter types prefer which party, on average. Any mixed change in polarization can be decomposed into the two cases covered by the proposition.

Proposition 6.   *1. If both economic and social party platforms polarize symmetrically over time, then voter sorting increases. Formally, if $\frac{x_{s,1}^R - x_{s,1}^D}{x_{s,0}^R - x_{s,0}^D} = \frac{x_{e,1}^R - x_{e,1}^D}{x_{e,0}^R - x_{e,0}^D} = k > 1$, then $\Psi(U_1^R - U_1^D) \gtreqqless \Psi(U_0^R - U_0^D)$ if and only if $\Psi(U_0^R - U_0^D) \gtreqqless 1/2$.*

   2. If social party platforms polarize symmetrically over time and economic platforms are unchanged, a realignment based on values takes place, meaning that moral conservatives (liberals) become more (less) likely to vote R.
   Formally, if $\frac{x_{s,1}^R - x_{s,1}^D}{x_{s,0}^R - x_{s,0}^D} > 1 \quad \wedge \quad x_{e,1}^R = x_{e,0}^R \quad \wedge \quad x_{e,1}^D = x_{e,0}^D$, then $\Psi(U_1^R - U_1^D) \gtreqqless \Psi(U_0^R - U_0^D)$ if and only if $\mu \gtreqqless 0$.

Proof.   1. Let $\Delta x_{p,t} = x_{p,t}^R - x_{p,t}^D$ be the policy polarization between Republicans and Democrats in policy $p \in \{e, s\}$ at time $t \in \{0, 1\}$. Inspection of (6) shows that, if $\frac{\Delta x_{s,1}}{\Delta x_{s,0}} = \frac{\Delta x_{e,1}}{\Delta x_{e,0}} = k > 1$, then $U_1^R - U_1^D = k(U_0^R - U_0^D)$. Thus, if $U_0^R - U_0^D > 0$ (so that $\Psi(U_0^R - U_0^D) > 1/2$), then $U_1^R - U_1^D > U_0^R - U_0^D$, and hence $\Psi(U_1^R - U_1^D) > \Psi(U_0^R - U_0^D)$. The case that $U_0^R - U_0^D < 0$ is analogous.

---

[11]Our exposition here partly builds on Krasa and Polborn (2014a) who show how changes in voter behavior over time can be used to recover information about the changes in party positions.

[12]The reason for this assumption is that asymmetric movements will clearly benefit one side, but the reason is less "polarization," but rather that one party takes up less competitive positions.

2. If $\frac{\Delta x_{s,1}}{\Delta x_{s,0}} > 1$ while economic positions are unchanged, then inspection of (6) shows that $U^R - U^D$ increases if and only if $\mu > 0$. This immediately implies the claim. $\qquad \square$

Part 1 of the Proposition clarifies how party polarization induces more sorting of voters into parties based on their ideological utility. This result is helpful because it shows that the increasing amount of voter sorting that is often discussed in political science (Levendusky, 2009; Mason, 2018) can rationally occur even without any changes in voters' preferences, purely as a result of party polarization.

More important for our purposes, Part 2 of Proposition 6 implies a form of realignment based on values. When only social polarization increases, then the utility difference changes by $w(y)\mu\Delta x_s$, which is positive if and only if $\mu > 0$. As a result, values become "more important" for voting: any fixed voter type becomes more likely to vote R if and only if he is morally conservative. Note that this occurs without any voter becoming more extreme (in terms of their values), or putting more weight on non-material issues (in terms of the weight w). This effect is of particular relevance for those voters who are close to being indifferent, such as "off-diagonal" voters that are either poor-and-morally-conservative or rich-and-morally-liberal.

Part 2 of Proposition 6 arguably informs the discussion about the oft-discussed puzzle that many of socially conservative working-class voters have swung Republican *even though* their relative economic position deteriorated over time, which should push them towards the Democrats. In our model, a decrease in relative income implies a more left-wing optimal economic policy. However, if the decrease in relative income coincides with increased social party polarization, these voters experience the offsetting effect identified in Proposition 6.2 and thus may still end up being more likely to vote for the Republican party. This is stated formally in the following Corollary.

Corollary 1. *Consider a voter of a fixed type μ_j, y_j, and suppose that there is symmetric social polarization, as well as growth in average income from $\bar{y}_{old}$ to $\bar{y}_{new}$ (voter j's income remains fixed). Then, the change in voter j's policy utility difference is given by*

$$\Delta(U^R - U^D) = [1 - w(y_j)](x_e^R - x_e^D)\left[r\left(y_j - \bar{y}_{new}\right) - r\left(y_j - \bar{y}_{old}\right)\right]$$
$$+ w(y_j)\mu\left[x_{s,new}^R - x_{s,old}^R + x_{s,old}^D - x_{s,new}^D\right]$$

This illustrates a logic that will also be at play in our general equilibrium analysis: if social party polarization increases, then morally conservative voters can become more likely to vote Republican *even when* their own relative income declines and their materially-optimal economic policy becomes more left-wing.

6 Endogenous Party Positions

6.1 Overview

We now endogenize party positions and allow them to respond to changes in voter income over time. Unlike all previous analyses, this requires us to model how party positions are determined. Most models of candidate competition in the literature focus on single-candidate elections and assume that candidates can choose, before the election, a position and commit to implementing it if elected to a majority. It is well-known that such models, with a multidimensional policy space, generically do not have a pure strategy equilibrium, and if they do have an equilibrium, it would involve complete convergence of party positions (McKelvey, 1976; Krasa and Polborn, 2012). Observing party positioning in real life, this appears counterfactual, and there is also a large body of empirical evidence in political science that documents that candidate positions do not converge (e.g., Fiorina, 1999; Ansolabehere et al., 2001; Burden, 2004).

For this reason, we follow the spirit of Polborn and Snyder (2017) who conceive of parties not as unitary vote-maximizing agents, but rather argue that parties are better thought of as an aggregation of different agents, each with their own objectives. Specifically, we assume that a party's position aggregates their supporters' political preferences, i.e., the party's position is the one that maximizes the aggregate utility of the party's voters. This reduced-form assumption lends itself to three interpretations.

1. In the spirit of the model in Polborn and Snyder (2017), we can view the agents in our model as being the "representative" voters in each of multiple Congressional districts. Congressional elections can be viewed as a contest in which voters in different districts choose between local candidates. Once elected, these local candidates all have an incentive to influence their party's position in the direction preferred by

their local constituents (e.g., to foster their chances in the next general election). Under this interpretation, each legislator would like their party to set policies that maximize the utility of their own district's representative (or median) voter. Then, actual party positions reflect a convex combination of the interests of each elected legislator of a party: party positions maximize the average policy utility of the voters in those districts that a party actually represents. This interpretation acknowledges that parties are not cohesive organizations under the full control of a single leader, but rather that they are influenced by the demands of elected representatives who face heterogeneous reelection incentives themselves.

2. A slightly different interpretation is that the set of agents consists of all voters. Then, loosely, our assumption says that each party sets policies to make those people happy who affiliate with that party. For example, party leaders could have incentives to make their own affiliates happy because they vote in the party's primaries.

3. A variation of the previous interpretation is that the set of agents in our model is given by rich donors. Then, our assumption says that each party sets social and economic policy positions to make those people happy who donate large amounts.

In summary, the interpretation of our model is flexible in terms of what constitutes the set of agents that influence the parties' positions. Furthermore, while we assume that party positions are only determined by their supporters' preferences, we could assume more generally that parties maximize some convex combination of the average utility of their supporters, and the utility of the average voter in the general election (in order to capture, in addition, the parties' incentives to win the general election). This would lead to quantitatively less polarization between the parties than in our setup, but the qualitative comparative static results (i.e., the effect of increased income on social and economic polarization) would be unchanged.

6.2 Formal Equilibrium Definition

An equilibrium is a partition of voters and a set of party positions such that (i) voters maximize utility by choosing the optimal party given expected party positions, (ii) parties

choose policy positions to maximize the weighted average utility of the voters which the party represents, and (iii) voters' expectations about party positions are correct.

Definition 2. *Let $\phi_t(\mu, y)$ denote the density of the voter type distribution at time t. Let S_i denote the set of voters supporting party i. A political equilibrium is defined by a tuple $(S_D, S_R, x_e^D, x_e^R, x_s^D, x_s^R)$ such that*

1. *The density of S_D is given by $\xi_D(\mu, y) = \frac{\phi_t(\mu,y)\Psi(\Delta(\mu,y))}{\int_{(\mu,y)} \Psi(\Delta(\mu,y))d\Phi_t(\mu,y)}$, where*

$$\Delta(\mu, y) = \frac{(1 - w(y))}{2} \left(\left[x_e^R - r(y - \bar{y}) \right]^2 - \left[x_e^D - r(y - \bar{y}) - \eta \right]^2 \right) +$$
$$\frac{w(y)}{2} \left(\lambda(x_e^R - h(\mu))^2 - (x_e^D - h(\mu))^2 + \alpha(x_s^R - \mu)^2 - \alpha(x_s^D - \mu)^2 \right)$$

   is the utility difference of a type (μ, y) between the Democratic and Republican policy positions. Let Ξ_D denote the corresponding cdf.

2. *Similarly, the density of S_R is given by $\xi_R(\mu, y) = \frac{\phi_t(\mu,y)[1-\Psi(\Delta(\mu,y))]}{\int_{(\mu,y)}[1-\Psi(\Delta(\mu,y))]d\Phi_t(\mu,y)}$*

3. *The parties' policies maximize the average utility of their respective supporters. For instance, for the Democrats, we have*

$$(x_e^D, x_s^D) \in \arg\max \int_{(\mu,y)} U(x_e, x_s; \mu, y)d\Xi_D(\mu, y)$$

Like in Polborn and Snyder (2017), there is always a trivial equilibrium with policy convergence, in which both parties have positions at the average (μ, y) so that no voter perceives a difference between the parties; in this case, only the random shocks determine whether a voter ends up as Democrat or Republican, and thus the two parties' voter type distributions are identical. However, by the same reasoning as in Polborn and Snyder (2017), this equilibrium is unstable for sufficiently small shocks, i.e., once party positions are slightly differentiated, adjustment processes lead to increased divergence between parties. We therefore focus on the equilibrium in which the party positions are differentiated.

6.3 Analytical Approach

Our objective in this section is to understand how an increase in incomes over time affects party positions and voting behavior. Because voter iso-probability curves in (y, μ) space

31

are nonlinear (see Section 4.4), it is impossible to derive the general equilibrium of our model in closed form. Therefore, we proceed numerically, through simulations. However, we do this for a wide range of parameter values so as to analyze which results are robust and which depend on specific parameters. It turns out that the model essentially always delivers the same directional predictions.

We do not view our approach as a full structural calibration of our model, but rather as a proof-of-concept that our model robustly generates three widely discussed comparative statics phenomena: an increase in social party polarization, increased moral polarization in the electorate, and the familiar realignment patterns emphasized in Section 5.2.

Parameterization of simulations. We model log income as $y \sim N(11 + \beta, 0.75)$, which is roughly calibrated to the U.S. household income distribution in 1988 (which is when the ANES data that we use start). To assess the effect of increases in aggregate income, we simulate the model separately for an "early period" with $\beta = 0$ and a "later period" with $\beta = 0.23$. This income increase corresponds to the approximate rise in median log household income in the U.S. between 1988 and 2020. Mirroring the patterns in the data, we assume that the cross-sectional distribution of values follows a normal distribution, and is independent of y. We normalize mean and standard deviation such that $\mu \sim N(0, 1)$.

The main degrees of freedom in the model are the parameters of the weight function (c, ρ and k) as well as the preference parameters that determine voters' materially- and values-optimal economic policies. We assume that the latter two functions are linear, such that $h(\mu) = h_0 \cdot \mu$ and $r(y - \bar{y}) = r_0 \cdot (y - \bar{y})$. We set $k = 8.99$, which implies that the poorest voters in the early period place the minimum weight of c (which may be zero) on non-material utility.

We then simulate the model for a large range of economically reasonable combinations of parameters c, ρ, r_0, h_0 and λ. When we say "economically reasonable," we refer to the fact that some parameters are unbounded, in principle, such as $r_0, h_0 \in [0, \infty)$. However, as r_0 or h_0 go to infinity, the model collapses to one dimension. Similarly, $c = 1$ implies that even the poorest voters only care about values, which eliminates the role of relative income for voting. Since our objective is to analyze how tradeoffs between economic

incentives and values affect political outcomes, these parameter values are not particularly useful. We, hence, simulate over the grid: $\rho \in [0.1, 0.2, \ldots, 0.9] \times h_0 \in [1, 2, \ldots, 9] \times r_0 \in [1, 2, \ldots, 9] \times c \in (0, 0.1, \ldots, 0.4) \times \lambda \in \{0, 0.1, \ldots, 0.9, 1\}$ for a total of 40,095 simulations. The net idiosyncratic (valence) shock is modeled as $\varepsilon \sim N(0, 0.5)$.

These combinations of parameters imply a broad range of relative weights and distributions of ideal policy points. To illustrate this, consider the following:

1. The parameters imply a large range of implied weight functions, see Appendix Figure 11. For instance, for $c = 0$ and $\rho = 0.1$, the non-material weight $w(y)$ ranges between 0% for the poorest and 30% for the richest voters, while for $c = 0.4$ and $\rho = 0.9$ it ranges between 40% and 90%.

2. The parameters allow for large variation in how much values, μ, and relative income, $y - \bar{y}$, affect a voter's ideal economic policy, respectively, as Appendix Figure 12 illustrates. For instance, for $r_0 = 9$ and $h_0 = 1$ (relative income matters much more for economic conservatism than values), the distribution of ideal economic policies (x_e^*) of top third income voters does not at all overlap with that of bottom third income voters (fixing $\lambda = 1$). For $r_0 = 1$ and $h_0 = 9$ (values matter much more), on the other hand, the distributions of top and bottom third income voters have a large common support and are visually very similar.

3. Consider a voter with median income y and median values μ and fix $\lambda = 1$. Now increase the voter's moral conservatism by one standard deviation – how much would we have to decrease the voter's income to keep his ideal economic policy unaffected? For $r_0 = h_0 = 1$, the required compensation is a 1.6 SD reduction in income. For $r_0 = 9$ and $h_0 = 1$, a 0.06 SD decrease in income is required. For $r_0 = 1$, $h_0 = 9$, a 2.1 SD decrease in income is required. Thus, again, our parameters allow for a wide range of scenarios.

The simulations are conducted with $N = 90,000$ voters, drawn from the joint distribution $\Phi_{\mu,y}$. We set the initial party positions to be $x_s^D = x_e^D = -0.5$ and $x_s^R = x_e^R = 0.5$, and calculate the voters' probability of voting for each party given the initial positions.[13]

---

[13]In Appendix D, we show that the simulations are robust to a wide range of initial values, subject to the constraint that the parties do not offer the same policies.

Then, we calculate updated Republican positions as those which maximize the weighted average utility of voters, where the weights are determined by the probability of voting Republican (and similarly for the Democrats). These new party positions induce new voting probabilities, which induce new positions, and so forth. We iterate this process until the party positions converge to a fixed point.[14]

6.4 Party Polarization on Social Issues Over Time

We again intersperse simulation "predictions" with a discussion of the available evidence.

Simulation prediction 1. *Social party polarization increases over time (as β increases).*

Discussion of simulation results. Table 1 summarizes the results of our 40,095 simulations regarding party polarization. We are interested in whether the distance between equilibrium party positions increases from the early period ($\beta = 0$) to the late period ($\beta = 0.23$). We see that social party polarization weakly increases in all simulations. To illustrate the intuition behind this social polarization result, consider a scenario in which the parties are initially somewhat differentiated in both dimensions. As all voters get richer, they place a higher weight on non-material utility, which makes their indifference curves in (y, μ) space steeper. Thus, some rich-and-morally-very-liberal voters will switch to the Democrats and some poor-and-morally-very-conservative ones to the Republicans. As parties aggregate the utilities of their supporters, and both parties' supporter sets have become more extreme in terms of their values, this induces both Democrats and Republicans to move their social policy platform to the left and right, respectively. This party movement then means that values now become even more important for the vote choice, which triggers a further round of realignment. Ultimately, this process equilibrates at a higher level of social polarization. Our simulations show that this intuition holds true for a wide range of parameter configurations.

On economic issues, the picture is considerably more mixed: in only 36% of our simulations does economic party polarization increase. Appendix Figure 13 dissects how

---

[14]We iterate each simulation 50 times, and define convergence as cases in which the average difference in party positions between subsequent iterations is less than 1%. The simulation converges in 99.8% of cases. The reported results below include the remaining 0.2% of cases, but the results are virtually identical if we exclude them.

Table 1: Simulation results regarding effects of increase in β

	Fraction of simulations in which variable in col. (1)		
	Increases	Constant	Decreases
Party polarization on social issues	91.6	8.4	0.0
Party polarization on economic issues	36.0	3.7	60.3
Moral polarization in electorate	91.8	8.2	0.0
Income polarization in electorate	0.0	7.1	92.9
Rich moral liberals voting Democratic	65.1	34.9	0.0
Poor moral conservatives voting Republican	81.9	5.4	12.7

Notes. Fraction of simulations in which the variable in col. 1 increases, remains constant, or decreases over time (as β increases). Social (economic) party polarization increases if the distance between party positions on social (economic) issues increases by more than 0.1% of a standard deviation of social (economic) preferences. Moral (income) polarization in the electorate increases if the difference in average moral conservatism μ (income y) between voters who vote Republican and those who vote Democrat increases by more than 0.1% of a standard deviation of moral conservatism (income). To compute this, we calculate weighted averages of moral conservatism and income in the population, where the weights are given by the probability of voting Republican and Democrat, and take the difference between the two. Rich moral liberals are defined as top 1/3 income and bottom 2/5 moral conservatism. Poor moral conservatives are defined analogously.

increases or decreases in economic polarization depend on parameter values.[15]

Empirical evidence.   The simulation prediction of increased social party polarization is supported by a wealth of data on elite polarization in domains such as immigration, affirmative action and minorities, abortion, religion and morality. For instance, Moskowitz et al. (2018) document using candidate surveys that increasing elite polarization in the U.S. Congress is largely driven by social and cultural issues. Gentzkow et al. (2019) and Enke (2020) provide evidence for increasing elite polarization on social issues using text analyses of Congressional speeches.[16]

---

[15]Economic polarization generally increases over time when the quantity $r_0 - h_0$ is not too large, especially when it is negative. To see why, note that in our model, economic party polarization increases in the variance of the distribution of voters' ideal economic policy positions. This variance, in turn, is jointly determined by voters' materially-optimal economic policies (as partly determined by r_0) and their values-optimal economic policies (as partly determined by h_0). Thus, party polarization will tend to increase (decrease) as β increases if the increased utility weight on non-material issues implies that a higher weight is placed on that utility component that has higher variance in the population, which is determined by whether $r_0 - h_0$ is large.

[16]For example, Figure 6 in Gentzkow et al. (2019) shows that polarization increased more strongly in domains such as immigration, religion, or minorities than in economic domains such as money, education, health or labor.

While the available evidence strongly suggests that increasing elite polarization is largely driven by social and moral issues, the evidence is a bit mixed regarding polarization on economic issues. For instance, while Moskowitz et al.'s data on candidate surveys show no evidence of increased economic polarization (e.g., taxes, welfare, health insurance, and employment), Gentzkow et al.'s text analyses reveal some evidence for increased elite polarization also on economic issues. Overall, it is probably fair to summarize the empirical literature as saying that economic party polarization is non-decreasing, and possibly increasing.

Finally, there is also recent empirical evidence that supports our model mechanism of non-material utility weights that increase in income. In a decomposition exercise, Danieli et al. (2022) find that, over the past 30 years, voters place an increasing weight on social and cultural issues relative to canonical economic ones.

6.5 Sorting and Polarization Based on Values

Simulation prediction 2. *Over time (as β increases):*

1. *The strength of the relationship between values and voting increases.*

2. *The supporters of both parties exhibit "moral polarization": the distributions of values of Republican and Democratic voters diverge.*

Discussion of simulation results. Regarding the first part of the prediction, we find that, in every single one of our simulations, the OLS coefficient in a regression of voting R on values μ strictly increases as β increases. Moreover, we also find that the raw correlation between these quantities always increases. Thus, the slope of the relationship between these variables becomes larger and the fit becomes tighter. Appendix Figure 14 illustrates by showing the distribution of the changes in regression / correlation coefficients across simulation parameterizations.

Intuitively, as anticipated in the analytical derivations in Section 5, two interrelated model mechanisms that point in the same direction produce this result. First, as people get richer, they place a higher utility weight on values. This directly implies that – holding party positions fixed – values become more strongly related to voting (Proposition 5).

Second, because social polarization between parties increases, it becomes more important for the voters' decision in the later period, which again increases the slope between voting and values (Proposition 6).

It is instructive to compare this pattern with the strength of the relationship between income and voting. As illustrated in Appendix Figure 15, the OLS coefficient in a regression of voting on log income y strictly decreases as β increases in the vast majority of simulations (95%). The intuition is twofold. First, social party polarization generally increases. Second, voters now place a higher utility weight on their values relative to their material interests.[17]

Table 1 summarizes the second part of the simulation prediction: the idea that the average Republican and the average Democratic voter diverge over time. We see that in all of our simulations, moral polarization (defined as the difference in average values between groups) increases. Because the underlying distribution of values is constant across time, this only reflects increased values-based sorting of voters into the parties. In contrast, income-based polarization in the electorate decreases in all of our simulations. Intuitively, because voting is now based on values to a larger degree, the income distributions of Republican and Democratic voters become more similar to each other.

Empirical evidence. A number of contributions have pointed to empirical regularities that are in line with these predictions. First, it is widely known that in voting regressions the coefficient of social or moral views has increased over time (e.g., Ansolabehere et al., 2006). Meanwhile, in line with our model's predictions, Appendix Figure 10 uses the ANES data to show that the income coefficient has decreased over time.

Second, regarding diverging distributions of voter types, a large body of work in political science has documented that increased polarization in the general public indeed reflects increased sorting, rather a change of the distribution of issue positions (e.g., Fiorina et al., 2006; Fiorina and Abrams, 2008; Levendusky, 2009). We view our simulation results as squarely in line with this well-documented empirical regularity.

---

[17] The small minority of cases in which income becomes more strongly predictive of voting as β increases are simulations in which r_0 is very large relative to h_0, generally when $r_0 - h_0 \geq 8$. See Appendix Figure 16. This is (i) because economic party polarization may strongly increase in those cases, as discussed above and (ii) because high r_0 means that when the parties are strongly polarized economically the voter will vote his income rather than his values.

6.6   Realignment of "Off-Diagonal" Voters

Simulation prediction 3. *Over time (as β increases), rich moral liberals become more likely to vote Democrat and poor moral conservatives more likely to vote Republican.*

Discussion of simulation results.   This prediction is a corollary of the patterns described in the previous section. If the strength of the relationship between voting and values (income) increases (decreases), then a values-based realignment will take place. In particular, "off-diagonal" voters who agree with each party on only one dimension, and who voted in line with their material incentives in the early period, become more likely to vote in line with their values in the later period. To visually illustrate these results using a specific set of parameter values ($c = 0$, $\rho = 0.2$, $h_0 = 1$, $r_0 = 1$, and $\lambda = 1$), the left panel of Figure 6 shows the change in voting behavior as a function of fundamentals (y, μ) from the earlier to the later period. The figure's main takeaway is that morally conservative (liberal) voters swing Republican (Democrat), in particular the "off-diagonal" voters in the bottom right and top left corners.

To document the generality with which our model simulations produce the realignment of rich moral liberals and poor social conservatives, Table 1 reports the share of simulations in which rich liberals (poor conservatives) become more likely to vote Democrat (Republican). It shows that rich moral liberals swing weakly left in all of the simulations, and poor moral conservatives swing weakly right in 87% of the simulations. The 13% of cases in which poor moral conservatives swing left are largely those cases discussed in footnote 17 in which $r - h$ is very large or λ equals zero, such that almost exclusively relative income matters for a voter's ideal economic policy. See Appendix Figure 17.

Empirical evidence.   Krasa and Polborn (2014a) use the ANES to analyze which types of voters "switched" between the 1970s and the 2000s. They find that voters who are economically conservative and socially-liberal tended to switch from voting Republican to Democrat. Analogously, voters who are economically liberal and socially conservative tended to switch from voting Democrat to voting Republican. Relatedly, Gingrich and Häusermann (2015) and Gethin et al. (2021) provide evidence that the educational elite (which researchers often think of as relatively rich and morally liberal) mostly voted
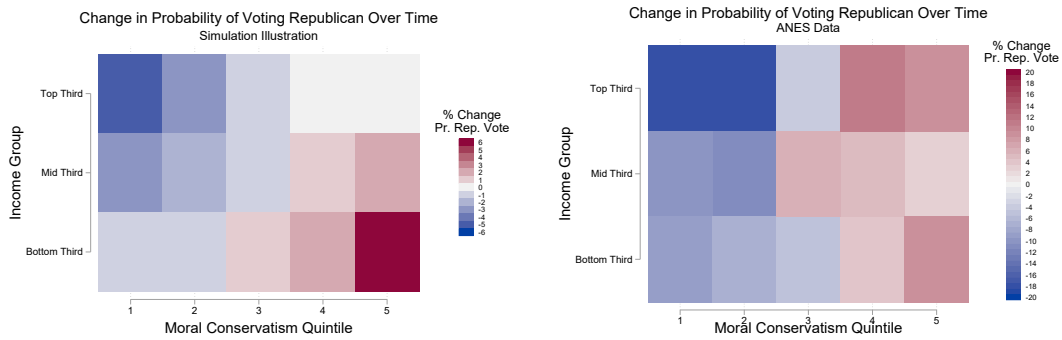
Figure 6: Percentage point change in Republican vote share over time as a function of fundamentals (y, μ). The left panel shows the result of a simulation (with $c = 0$, $\rho = 0.2$, $h_0 = 1$, $r_0 = 1$, and $\lambda = 1$). The right panel shows the change in Republican vote share in the ANES, from time period 1988-2000 to 2004-2020.

Republican in the 1950s-60s, yet increasingly voted Democrat over time, such that today high educational attainment is positively correlated with voting Democrat. The right panel of Figure 6 illustrates this well-known empirical pattern using ANES data. In the data, values becomes substantially more predictive of voting over time, such that poor moral conservatives swing Republican and rich moral liberals Democrat.

The figure also serves as an important reminder that – in contrast to some popular narratives – it is not generically true that poor voters have trended Republican. Instead, only those poor voters who are socially conservative have done so – which is a pattern that our simulations also generate, see the left panel of Figure 6.

6.7 Income Inequality

In all simulations reported thus far, we increase incomes uniformly for all voters by adding a constant (in log space), which leaves inequality unchanged over time. However, an important feature of the American political landscape has been growing income inequality, or, more precisely, increasing incomes at the top and largely stagnating incomes at the bottom. Such inequality partly counteracts the mechanics in our simulations above because higher inequality means that people's materially-optimal economic policies become more dispersed. As a result, relative income should matter more for voting, which partly counteracts the mechanics highlighted in the previous subsections.

At the same time, as illustrated by Corollary 1 in Section 5.2, even with increasing inequality our model can generate increasing social party polarization and values-based

realignment. There are two countervailing effects. First, because many voters get richer, they place a higher weight on their values, which induces social party polarization. This, in turn, leads to the realignment of poor moral conservatives and rich moral liberals. Second, because of the increase in inequality, the voters' economic conflicts increase, which, ceteris paribus, decreases social polarization and therefore reduces values-based realignment.

To examine this in the simulations, we scale the log income of all voters multiplicatively by $\kappa = 1.015$; that is, $y_{new} = \kappa y_{old}$. In contrast to our additive scaling above, this multiplicative scaling increases inequality because it implies that the logarithm of income increases more for the rich than the poor, so that the ratio of (absolute) income of the rich to income of the poor increases. We choose κ such that our setup roughly matches the change in income among the 25th percentile, median, 75th percentile, 90th percentile, and 95th percentile of households going from 1988 to 2020.

Appendix D discusses the results. We find that in all simulations – conducted with the same parameters as in the simulations above – social party polarization weakly increases. Moreover, we again find that the simulations generally produce strong values-based re-alignment. Thus, even as inequality increases, our model can generate that poor moral conservatives swing Republican. The logic is that when the rich get disproportionately richer, they place a higher weight on values, which induces some rich moral liberals to swing Democratic. This induces the parties to polarize on social issues, which, in turn – as an equilibrium effect – "pushes" some poor moral conservatives to vote Republican.

This logic highlights that our mechanism of values as luxury goods can account for empirically observed polarization and realignment patterns, even if income increases are concentrated at the top. Indeed, the emphasis our model places on the rich inducing social party polarization that then propagates to the poor is consistent with work in sociology that argues that rural working class voters feel that the Democratic party "abandoned" them to cater to the values of coastal elites (e.g., Sherman, 2009; Cramer, 2016).

7   Conclusion

This paper has shown that the simple idea of income-dependent utility weights – which is bolstered by a considerable body of evidence – generates a host of new testable predictions

and sheds light on various stylized facts about the nature of political polarization and realignment patterns in the U.S. In particular, our approach offers a new lens through which the increased salience of moral and cultural dimensions of political conflict can be understood.

Relationship to informal narratives of realignment. While income in our model only affects the relative weight of values, but not values as such, some narratives assert that income and/ or higher education directly affect a voter's values. According to this argument, modernization made the elite's values more cosmopolitan and universalist, which may have alienated working class voters and led to cultural backlash (Norris and Inglehart, 2019; Fukuyama, 2018). Relatedly, authors such as Frank (2007) and Hacker and Pierson (2020) argue that many working-class voters are convinced by Republican party elites to vote against their economic interests, in exchange for the socially conservative policies that the poor supposedly desire. In a nutshell, these are all stories of conflict between rich and poor. Our account is complementary in that we instead highlight the cultural or moral conflict between different subsets of the elite. This conflict among elites induces party polarization, which then propagates into changes in voting behavior among the poor. While this perspective may have received less attention in popular writings due its slightly higher complexity, it is backed by various pieces of evidence, such as our insight that the effect of values on voting and policy views is larger among the rich. Indeed, political scientists have long documented that the recent increase in the importance of social issues for voting is concentrated in the elite (Bartels, 2006), in contrast to what narratives such as "What's the matter with Kansas" suggest.

Affective polarization. One aspect of polarization that we did not formally touch upon is affective polarization: the stylized fact that people's dislike of supporters of the other party has strongly increased over time (Iyengar et al., 2012, 2019). Through the lens of our model, a casual interpretation of this pattern is that it reflects that the distribution of values of Republican and Democrat voters have diverged over time as a result of sorting processes that are triggered by our account of values as luxury goods. However, while much psychological research suggests that people get angry if others don't share their

basic moral convictions (Haidt, 2012), more research is needed to establish a direct link between increased voter sorting based on values and affective polarization.

Western Europe. An important open question is to which extent our insights are applicable to the Western European context. Some of the empirical regularities that motivate this paper are also visible in Europe, while others are not. On the one hand, the realignment of the educated elite towards the left (e.g., Gingrich and Häusermann, 2015; Gethin et al., 2021), and the contemporaneous importance of values (Enke et al., 2022; Cappelen et al., 2022) are both recurring phenomena throughout the Western world. On the other hand, measures of affective polarization have not increased nearly as much in Europe as in the U.S. Moreover, there is anecdotal evidence that some European countries such as Germany have seen a *convergence* of the main political parties. One potential reason for this complicated pattern is differences in institutional environments.

One reason why we believe that our approach could plausibly shed light also on the European context is our model prediction that the voters of left-wing parties will be more internally heterogeneous than those of right-wing parties. This prediction jives with a casual observation of the party landscape in some European countries. For example, in Germany, until the 1980s, politics was largely organized around a traditional economic left-right divide between the Social Democrats and the Christian Democrats. Then, the internal divisions within the Social Democratic Party became so large that the Green party (essentially the culturally very liberal wing of the Social Democrats) emerged. Relatedly, in both France and Italy, a widely-discussed stylized fact is that the party landscape on the left is more fragmented than on the right, which puts left-wing parties at a "structural disadvantage" in terms of winning elections.

8   Proofs not Given in Main Text

Proof of Proposition 3. The probability of voting against one's economic interest for RML is $Pr(D|RML) = 1 - \Psi(U^R - U^D) = \Psi(U^D - U^R)$, and for PMC is $Pr(R|PMC) = \Psi(U^R - U^D)$. Given that Ψ is symmetric around zero, the claim holds if and only if

$(U^D - U^R)_{RML} > (U^R - U^D)_{PMC}$. By (6), we have for the RML and PMC, respectively

$$(U^D - U^R)_{RML} = -[1 - w(y_{RML})]\Delta x_e \left[r\left(y_{RML} - \bar{y}\right)\right] - w(y_{RML})\left[\Delta x_e \lambda h(\mu_{RML}) + \Delta x_s \mu_{RML}\right],$$

$$(U^R - U^D)_{PMC} = [1 - w(y_{PMC})]\Delta x_e \left[r\left(y_{PMC} - \bar{y}\right)\right] + w(y_{PMC})\left[\Delta x_e \lambda h(\mu_{PMC}) + \Delta x_e \mu_{PMC}\right].$$

Differencing these two expressions, and substituting $r\left(y_{PMC} - \bar{y}\right) = -r\left(y_{RML} - \bar{y}\right)$ and $\mu_{RML} = -\mu_{PMC}$ yields $\quad (U^D - U^R)_{RML} - (U^R - U^D)_{PMC} =$

$$[w(y_{RML}) - w(y_{PMC})]\left\{\Delta x_e \lambda h(\mu_{PMC}) + \Delta x_s \mu_{PMC} + \Delta x_e r\left(y_{RML} - \bar{y}\right)\right\} > 0, \quad (14)$$

because $w(y_{RML}) > w(y_{PMC})$ and all terms in the curly brackets are positive. Thus, it follows that the probability that RML votes against his economic interest is larger than that PMC votes against his economic interest. $\qquad \square$

Proof of Proposition 4. 1. The first line in (11) follows directly from applying the implicit function theorem to (6). The second line follows from dividing both numerator and denominator by $[1 - w(y)]\Delta x_e$, and observing that $\frac{d}{dy}\left(\frac{w(y)}{1-w(y)}\right) = \frac{w'(y)}{(1-w(y))^2}$. Since all terms in the numerator are positive, and both $r' > 0$ and $\frac{d}{dy}\left(\frac{w(y)}{1-w(y)}\right) > 0$, a sufficient condition for an iso-probability curve to be downward sloping is that $[\lambda h(\mu) - r(y - \bar{y}) + \frac{\Delta x_s}{\Delta x_e}\mu] \geq 0$, though, clearly, the same result would still hold if it is slightly negative.

Similarly, if $w(y) = \frac{c+\rho(y-k)}{1+\rho(y-k)}$, then (11) simplifies to

$$\begin{aligned}\frac{dy}{d\mu} &= -\frac{\frac{c+\rho(y-k)}{1-c}\left[\lambda h'(\mu) + \frac{\Delta x_s}{\Delta x_e}\right]}{r'(y - \bar{y}) + \frac{\rho}{1-c}\frac{1-c}{1+\rho(y-k)}[\lambda h(\mu) - r(y - \bar{y}) + \frac{\Delta x_s}{\Delta x_e}\mu]}\\ &= -\frac{\lambda h'(\mu) + \frac{\Delta x_s}{\Delta x_e}}{r'(\cdot)\frac{1-c}{c+\rho(y-k)} + \frac{\rho(1-c)}{[c+\rho(y-k)][1+\rho(y-k)]}[\lambda h(\mu) - r(y - \bar{y}) + \frac{\Delta x_s}{\Delta x_e}\mu]}.\end{aligned} \quad (15)$$

The numerator of the term in the second line in (15) is independent of y and μ (because h is linear). If $\rho \to 0$, then (15) goes to $-\frac{c}{(1-c)r'}\left[\lambda h'(\mu) + \frac{\Delta x_s}{\Delta x_e}\right] < 0$.

2. Using $r(y) = r_0(y - \bar{y})$ and $h(\mu) = h_0\mu$ with $r_0, h_0 > 0$, we can rewrite (11) as

$$\frac{dy}{d\mu} = -\frac{\lambda h_0 + \frac{\Delta x_s}{\Delta x_e}}{\frac{1-w}{w}r_0 + \frac{w'}{w}\left(\lambda h_0\mu - r_0(y - \bar{y}) + \mu\frac{\Delta x_s}{\Delta x_e}\right)} \quad (16)$$

43

For downward-sloping isoprobability curves, the denominator is positive. Thus,

$$\lambda h_0 \mu - r_0(y - \bar{y}) + \mu \frac{\Delta x_s}{\Delta x_e} \geq -\frac{1-w}{w'} r_0. \tag{17}$$

To show the claim, we have to show that the value of the denominator in (16) increases in μ (because the numerator is a constant in μ and y). Differentiating the denominator with respect to μ, taking into account that, along an isoprobability curve, y is a function of μ, yields

$$\left[-\frac{w'}{w^2} r_0 + \frac{w''w - (w')^2}{w^2} \left(\lambda h_0 \mu - r_0(y - \bar{y}) + \mu \frac{\Delta x_s}{\Delta x_e} \right) - \frac{w'}{w} r_0 \right] \frac{dy}{d\mu} + \frac{w'}{w} \left(\lambda h_0 + \frac{\Delta x_s}{\Delta x_e} \right). \tag{18}$$

Clearly, the second summand is positive, so, given that $dy/d\mu < 0$, to show that (18) is positive, it is sufficient to show that the term in square brackets is negative. By assumption, $\frac{w''w - (w')^2}{w^2} = \frac{d}{dy} \left(\frac{w'(y)}{w(y)} \right) \leq 0$. Together with (17), this implies that the term in square brackets is smaller than

$$-\frac{w'}{w^2} r_0 + \frac{w''w - (w')^2}{w^2} \left(-\frac{1-w}{w'} r_0 \right) - \frac{w'}{w} r_0. \tag{19}$$

Putting all terms in (19) in a common fraction, we have

$$-\frac{w'^2 + w''w(1-w) - (1-w)(w')^2 + (w')^2 w}{w^2 w'} r_0 = -\frac{w''(1-w) + 2w'^2}{ww'} r_0 \leq 0. \tag{20}$$

Observe that, since $r_0 > 0$, (20) holds if and only if

$$w''(1-w) + 2w'^2 \geq 0 \quad \forall y \tag{21}$$

This is satisfied whenever w is linear or convex,[18] and when it is not too concave. If $w(y) = \frac{c + \rho(y-k)}{1 + \rho(y-k)}$, then $w'(y) = \frac{\rho(1-c)}{(1+\rho(y-k))^2}$ and $w''(y) = -\frac{2\rho^2(1-c)}{(1+\rho(y-k))^3}$. Substituting, the left hand side of (21) is identically zero for all y. (Note that this does not mean that $w(y) = \frac{c + \rho(y-k)}{1 + \rho(y-k)}$ is a knife-edge case. After all, the second summand in (18) is strictly positive.) $\qquad \square$

---

[18]Though it also can't be too convex since we assumed that $\frac{d}{dy} \left(\frac{w'(y)}{w(y)} \right) \leq 0$.

References

Andre, Peter, Teodora Boneva, Felix Chopra, and Armin Falk, "Fighting climate change: The role of norms, preferences, and moral values," 2021.

Ansolabehere, Stephen, James M Snyder, and Charles Stewart III, "Candidate positioning in US House elections," *American Journal of Political Science*, 2001, pp. 136–159.

— , **Jonathan Rodden, and James M Snyder**, "Purple america," *Journal of Economic Perspectives*, 2006, *20* (2), 97–118.

Bartels, Larry M, "What's the Matter with What's the Matter with Kansas?," *Quarterly Journal of Political Science*, 2006, *1* (2), 201–226.

Besley, Timothy and Torsten Persson, "The rise of identity politics," *LSE Documents*, 2019.

Bock, Sean, "Divided Dems, United Reps: Party strategy and within-party variation in immigration attitudes," 2020.

Burden, Barry C, "Candidate positioning in US congressional elections," *British Journal of Political Science*, 2004, pp. 211–227.

Callander, Steven and Juan Carlos Carbajal, "Cause and Effect in Political Polarization: A Dynamic Analysis," *Journal of Political Economy*, 2022, *130* (4), 825–880.

Cappelen, Alexander W, Benjamin Enke, and Bertil Tungodden, "Universalism: Global Evidence," Technical Report, National Bureau of Economic Research 2022.

Cramer, Katherine J, *The politics of resentment: Rural consciousness in Wisconsin and the rise of Scott Walker*, University of Chicago Press, 2016.

Danieli, Oren, Noam Gidron, Shinnosuke Kikuchi, and Ro'ee Levy, "Decomposing the Rise of the Populist Radical Right," *Working Paper*, 2022.

Dragu, Tiberiu and Xiaochen Fan, "An agenda-setting theory of electoral competition," *Journal of Politics*, 2016, *78* (4), 1170–1183.

Enke, Benjamin, "Moral Values and Voting," *Journal of Political Economy*, 2020, pp. 3679–3729.

_ , **Ricardo Rodríguez-Padilla, and Florian Zimmermann**, "Moral Universalism and the Structure of Ideology," *Review of Economic Studies*, 2022.

Fehr, Dietmar, Johanna Mollerstrom, and Ricardo Perez-Truglia, "Your Place in the World: Relative Income and Global Inequality," *American Economic Journal: Economic Policy*, 2022.

Fiorina, Morris P, "Whatever happened to the median voter?," in "MIT Conference on Parties and Congress, Cambridge, MA," Vol. 2 1999, p. 1999.

Fiorina, Morris P. and Samuel J. Abrams, "Political polarization in the American public," *Annual Review of Political Science*, 2008, *11* (1), 563–588.

_ , _ , **and Jeremy C. Pope**, *Culture war? The myth of a polarized America*, Pearson Longman, 2006.

Frank, Thomas, *What's the matter with Kansas?: How conservatives won the heart of America*, Picador, 2007.

Fukuyama, Francis, *Identity: The demand for dignity and the politics of resentment*, Farrar, Straus and Giroux, 2018.

Gelman, Andrew, Boris Shor, Joseph Bafumi, and David Park, "Rich State, Poor State, Red State, Blue State: What's the Matter with Connecticut?," *Quarterly Journal of Political Science*, 2007, *2*, 345–367.

Gennaioli, Nicola and Guido Tabellini, "Identity, Beliefs, and Political Conflict," *Quarterly Journal of Economics*, 2021.

Gentzkow, Matthew, Jesse M Shapiro, and Matt Taddy, "Measuring group differences in high-dimensional choices: method and application to congressional speech," *Econometrica*, 2019, *87* (4), 1307–1340.

Gethin, Amory, Clara Martínez-Toledano, and Thomas Piketty, "Brahmin Left versus Merchant Right: Changing Political Cleavages in 21 Western Democracies, 1948-2020," *WID. world Working Paper*, 2021.

Gingrich, Jane and Silja Häusermann, "The decline of the working-class vote, the reconfiguration of the welfare support coalition and consequences for the welfare state," *Journal of European Social Policy*, 2015, *25* (1), 50–75.

Graham, Jesse, Jonathan Haidt, and Brian A. Nosek, "Liberals and Conservatives Rely on Different Sets of Moral Foundations," *Journal of Personality and Social Psychology*, 2009, *96* (5), 1029.

Hacker, Jacob S and Paul Pierson, *Let them Eat Tweets: How the Right Rules in an Age of Extreme Inequality*, Liveright Publishing, 2020.

Haidt, Jonathan, *The Righteous Mind: Why Good People are Divided by Politics and Religion*, Vintage, 2012.

Hawkins, Stephen, Daniel Yudkin, Miriam Juan-Torres, and Tim Dixon, "Hidden tribes: A study of America's polarized landscape," 2019.

Inglehart, Ronald, *Modernization and postmodernization: Cultural, economic, and political change in 43 societies*, Princeton university press, 1997.

— , *Modernization and postmodernization: Cultural, economic, and political change in 43 societies*, Princeton university press, 2020.

Iyengar, Shanto, Gaurav Sood, and Yphtach Lelkes, "Affect, not ideologya social identity perspective on polarization," *Public Opinion Quarterly*, 2012, *76* (3), 405–431.

— , **Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J Westwood**, "The origins and consequences of affective polarization in the United States," *Annual Review of Political Science*, 2019, *22*, 129–146.

Kivikangas, J Matias, Belén Fernández-Castilla, Simo Järvelä, Niklas Ravaja, and Jan-Erik Lönnqvist, "Moral foundations and political orientation: Systematic review and meta-analysis.," *Psychological Bulletin*, 2021, *147* (1), 55.

Krasa, Stefan and Mattias K Polborn, "Political competition between differentiated candidates," *Games and Economic Behavior*, 2012, *76* (1), 249–271.

— **and** — , "Policy divergence and voter polarization in a structural model of elections," *Journal of Law and Economics*, 2014, *57* (1), 31–76.

— **and** — , "Social ideology and taxes in a differentiated candidates framework," *American Economic Review*, 2014, *104* (1), 308–22.

Landier, Augustin and David Thesmar, "Who is Neoliberal? Durkheimian Individualism and Support for Market Mechanisms," Technical Report, National Bureau of Economic Research 2022.

Levendusky, Matthew, *The partisan sort: How liberals became Democrats and conservatives became Republicans*, University of Chicago Press, 2009.

Lizzeri, Alessandro and Nicola Persico, "A drawback of electoral competition," *Journal of the European Economic Association*, 2005, *3* (6), 1318–1348.

Luttmer, Erzo FP, "Group loyalty and the taste for redistribution," *Journal of Political Economy*, 2001, *109* (3), 500–528.

Maslow, Abraham Harold, "A theory of human motivation.," *Psychological review*, 1943, *50* (4), 370.

Mason, Lilliana, *Uncivil agreement: How politics became our identity*, University of Chicago Press, 2018.

McCarty, Nolan, Keith T Poole, and Howard Rosenthal, *Polarized America: The dance of ideology and unequal riches*, mit Press, 2016.

McKelvey, Richard D, "Intransitivities in multidimensional voting models and some implications for agenda control," *Journal of Economic Theory*, 1976, *12* (3), 472–482.

Meltzer, Allan H and Scott F Richard, "A rational theory of the size of government," *Journal of Political Economy*, 1981, *89* (5), 914–927.

Moskowitz, Daniel J, Jon C Rogowski, and James M Snyder, "Parsing Party Polarization in Congress," 2018.

Norris, Pippa and Ronald Inglehart, *Cultural backlash: Trump, Brexit, and authoritarian populism*, Cambridge University Press, 2019.

Nunnari, Salvatore and Jan Zápal, "Dynamic elections and ideological polarization," *Political Analysis*, 2017, *25* (4), 505–534.

Pástor, Lúboš and Pietro Veronesi, "Inequality aversion, populism, and the backlash against globalization," *Journal of Finance*, 2021, *76* (6), 2857–2906.

Polborn, Mattias K and James M Snyder, "Party polarization in legislatures with office-motivated candidates," *The Quarterly Journal of Economics*, 2017, *132* (3), 1509–1550.

Rodden, Jonathan A, *Why cities lose: The deep roots of the urban-rural political divide*, Hachette UK, 2019.

Rodrik, Dani, Elliot Ash, and Sharun W Mukand, "Economic Interests, Worldviews, and Identities: Theory and Evidence on Ideational Politics," *HKS Faculty Research Working Paper Series*, 2021.

Schofield, Norman, "The mean voter theorem: necessary and sufficient conditions for convergent equilibrium," *Review of Economic Studies*, 2007, *74* (3), 965–980.

Sherman, Jennifer, *Those who work, those who don't: Poverty, morality, and family in rural America*, U of Minnesota Press, 2009.

Strulovici, Bruno, "Can society function without ethical agents? an informational perspective," *arXiv preprint arXiv:2003.05441*, 2020.

Tabellini, Guido, "The Scope of Cooperation: Values and Incentives," *Quarterly Journal of Economics*, 2008, *123* (3), 905–950.

Xefteris, Dimitrios, "Multidimensional electoral competition between differentiated candidates," *Games and Economic Behavior*, 2017, *105*, 112–121.

ONLINE APPENDIX

A   Discussion of uniformity of ψ

What happens in in Proposition 2 if ψ is not constant, but rather single-peaked (for example, if it is normal)?

Intuitively, under normal circumstances, increases in both μ and y will increase a voter's likelihood of voting Republican.[19]

Moreover, as μ increases, the effect of an income increase *on the utility difference* ΔU increases. Thus, μ and y are complements in terms of shifting up the cutoff value of the critical ε shock (remember that voters vote Republican if their ε is smaller than the critical value).

The only case in which this complementarity in terms of shifting the cutoff does not translate into complementarity in terms of the probability of voting Republican is if the density ψ decreases rapidly.

To see this more formally, observe first that $d\Delta U/d\mu$ is positive. As we show in the proof of Proposition 4, the term in curly brackets in the second line of (10) is positive whenever the slope of the isoprobability curve is negative (see discussion in footnote 19 why this is arguably the standard case).

In this case, $\psi' \geq 0$ is sufficient (but not necessary) for $\frac{d^2 Pr(R)}{d\mu dy} > 0$ because the density effect further strengthens the shift of the cutoff. In contrast, if $\psi' << 0$, it is possible that μ and y are not complements in terms of the probability of voting Republican.

---

[19]As shown, this is definitely the case for an increase in μ, and while an opposite effect is possible for y for very socially-liberal voters, the standard case is probably that richer people are more likely to prefer the Republicans.

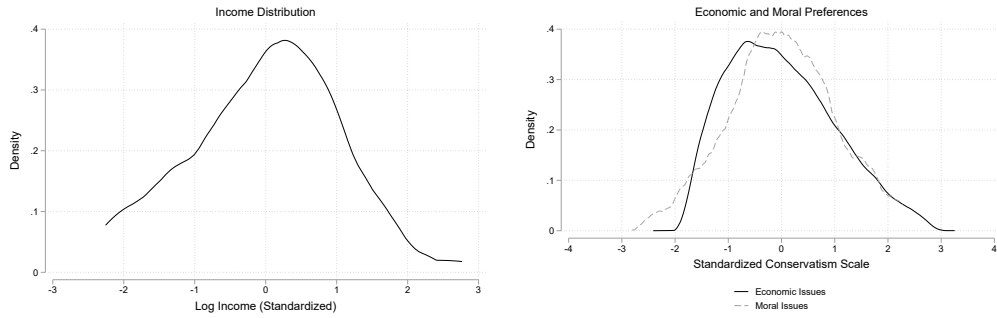# B Additional Figures for Empirical Analyses (ANES and MFQ)

Figure 7: This figure shows a kernel denisty plot of standardized log income on the left, and standardized economic and moral conservatism on the right. All variables are measured from the ANES (data cover 1988–2020). See Appendix E for details on the construction of both indices.
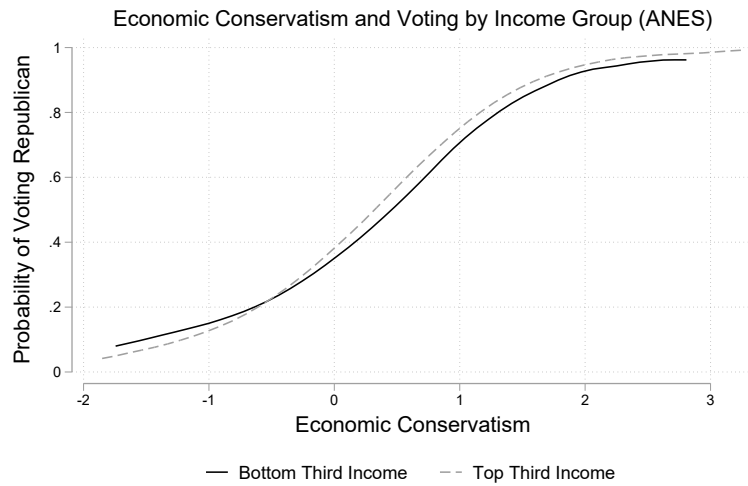

Figure 8: This figure shows a local polynomial plot of the link between voting Republican and economic conservatism, separately for respondents in the top / bottom third of the income distribution. Economic conservatism is measured from the ANES (data cover 1988–2020), and the index is normalized into a z-score. See Appendix E for details on the construction of both indices.

Demeaned Share of Population Voting Against Economic Interests (ANES)

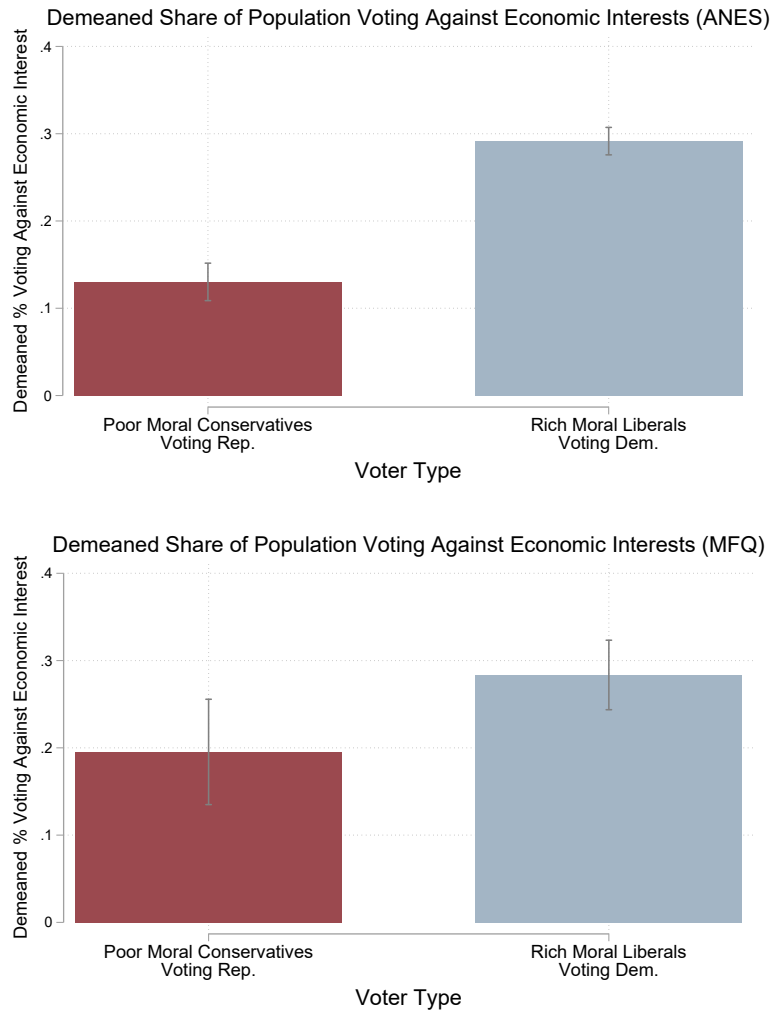**Demeaned Share of Population Voting Against Economic Interests (MFQ)**

Figure 9: This figure reports the demeaned share and 95% confidence interval of the share of poor social conservatives who vote Republican, and rich social liberals who vote Democrat. The demeaning is done by subtracting the overall Republican vote share from the poor conservative Republican share, and the overall Democratic vote share from the rich liberals Democratic share. Here, we define "rich" and "poor" as the top/bottom third of household income, and "socially conservative / liberal" as top/bottom third of the social / moral conservatism indices. The panel on the top uses ANES data ($N = 4,274$) and the panel on the bottom uses MFQ data ($N = 608$). In t-tests of the equality of means across groups, the p-value is $p < 0.01$ in both the ANES and the MFQ.
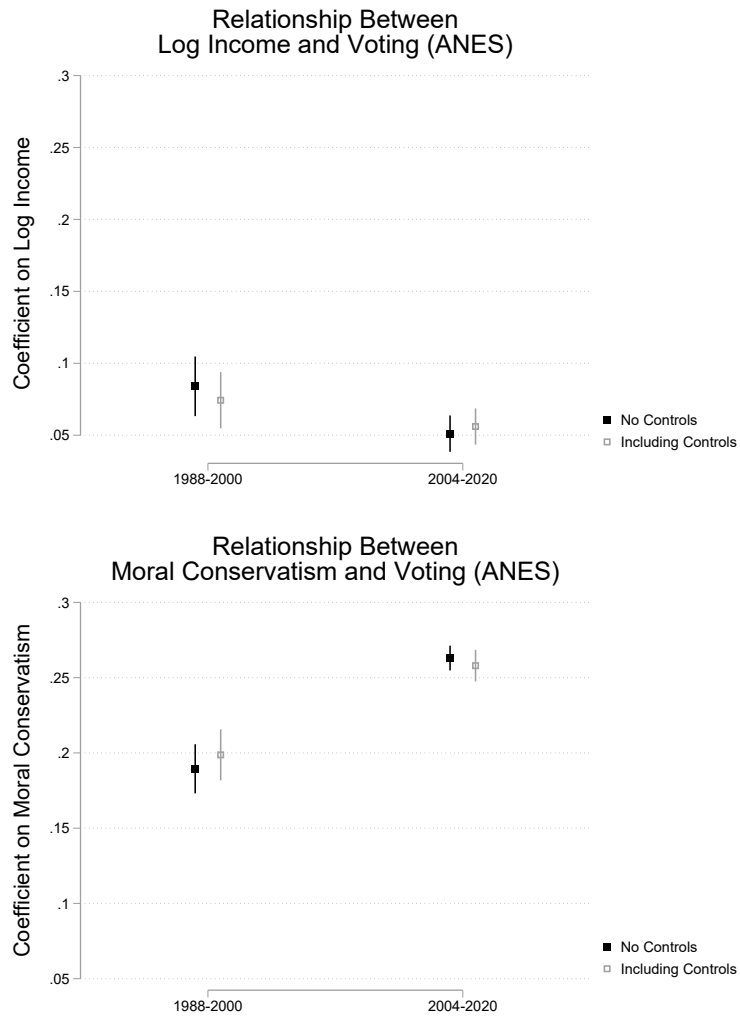
Figure 10: This figure reports coefficients from OLS regressions of an indicator for voting Republican on log income (top panel) and social conservatism (bottom panel) split by whether the survey year is pre- or post-2000. 95% confidence intervals reported. In a Wald test of the equality of coefficients between groups, the p-value is $p = 0.005$ for the top panel, and $p < 0.0001$ for the bottom panel. Data are from the ANES. Total $N = 15,505$. The conservatism index has mean zero and standard deviation one. The corresponding regression results are reported in Table 6 in Appendix C

C   Additional Tables for Empirical Analyses (ANES and MFQ)

Table 2: Voting and Moral Conservatism By Income and Education (MFQ)

	Dependent variable:			
	Indicator for Voting Republican (MFQ)			
	Bottom 1/3 Income	Top 1/3 Income		
	(1)	(2)	(3)	(4)
Moral Conservatism	0.23***	0.30***	-0.12	-0.12
	(0.02)	(0.01)	(0.10)	(0.11)
Log [Household income]			0.0072	0.0071
			(0.01)	(0.01)
Moral Conservatism x Log Income			0.036***	0.037***
			(0.01)	(0.01)
College				0.00066
				(0.01)
Moral Conservatism x College				-0.00038
				(0.01)
State FE	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes
Observations	914	791	2714	2714
R^2	0.21	0.41	0.30	0.30

Notes. OLS estimates, clustered standard errors in parentheses. The dependent variable is an indicator for voting Republican. Moral conservatism is standardized into a z-score. Controls include age, age squared, and gender fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Economic Conservatism and Income (ANES)

	Dependent variable:			
	Economic conservatism (ANES)			
	(1)	(2)	(3)	(4)
Log [HH income]	0.15***	0.19***	0.20***	0.18***
	(0.02)	(0.01)	(0.01)	(0.01)
Log [HH income] × Moral Conservatism		0.068***	0.072***	0.071***
		(0.01)	(0.01)	(0.01)
Moral Conservatism		0.38***	0.39***	0.39***
		(0.01)	(0.01)	(0.01)
State FE	No	No	Yes	Yes
Survey year FE	No	No	Yes	Yes
Controls	No	No	No	Yes
Observations	22972	18193	18193	18193
R^2	0.03	0.22	0.30	0.31

Notes. OLS estimates, clustered standard errors in parentheses. The dependent variable is the economic conservatism index, standardized into a z-score. Log income is standardized into a z-score. Controls include age, age squared, gender fixed effects, and educational attainment fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Income Gradient in Voting (ANES)

	Dependent variable:			
	Indicator for Voting Republican (ANES)			
	(1)	(2)	(3)	(4)
Log [HH income]	0.037***	0.062***	0.062***	0.061***
	(0.01)	(0.00)	(0.00)	(0.00)
Log [HH income] $\times$ Moral Conservatism		0.027***	0.027***	0.025***
		(0.00)	(0.00)	(0.00)
Moral Conservatism		0.22***	0.23***	0.23***
		(0.00)	(0.00)	(0.00)
State FE	No	No	Yes	Yes
Survey year FE	No	No	Yes	Yes
Baseline controls	No	No	No	Yes
Observations	28554	16164	16164	16164
R^2	0.01	0.25	0.28	0.28

Notes. OLS estimates, clustered standard errors in parentheses. The dependent variable is an indicator which equals 100 if a respondent votes Republican and 0 if they vote Democrat. Controls include age, age squared, gender fixed effects, and educational attainment fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Income Gradient in Voting (MFQ)

| | Dependent variable: | | | |
| | Indicator for Voting Republican (MFQ) | | | |
	(1)	(2)	(3)	(4)
Log Income [std.]	0.040***	0.027***	0.035***	0.027***
	(0.01)	(0.01)	(0.01)	(0.01)
Log Income [std.] × Moral conservatism		0.035***	0.032***	0.032***
		(0.01)	(0.01)	(0.01)
Moral conservatism		0.22***	0.21***	0.21***
		(0.01)	(0.01)	(0.01)
State FE	No	No	Yes	Yes
Baseline controls	No	No	No	Yes
Observations	3172	3172	3172	3172
R^2	0.01	0.23	0.25	0.27

Notes. OLS estimates, clustered standard errors in parentheses. The dependent variable is an indicator which equals 1 if a respondent votes Republican and 0 if they vote Democrat. Controls include age, age squared, gender fixed effects, and educational attainment fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Importance of Income and Social Conservatism in Voting Over Time (ANES)

| | Dependent variable: | | | |
| | Indicator for Voting Republican (ANES) | | | |
	(1)	(2)	(3)	(4)
Log [HH income]	0.059***	0.085***	0.091***	0.088***
	(0.01)	(0.01)	(0.01)	(0.01)
Moral Conservatism	0.24***	0.12***	0.11***	0.12***
	(0.00)	(0.01)	(0.01)	(0.01)
Survey year × Log [HH Income]		-0.0010**	-0.0011**	-0.0010**
		(0.00)	(0.00)	(0.00)
Survey year × Moral Conservatism		0.0057***	0.0055***	0.0055***
		(0.00)	(0.00)	(0.00)
State FE	No	No	Yes	Yes
Baseline controls	No	No	No	Yes
Observations	16164	16164	16164	16164
R^2	0.25	0.27	0.28	0.28
Prob > F		2.4e-15	7.9e-16	1.3e-15

Notes. OLS estimates, clustered standard errors in parentheses. The dependent variable is an indicator for voting Republican. "Survey year" is year of survey minus 1988. Controls include age, age squared, gender fixed effects, and educational attainment fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Details and Analyses for Model Simulations

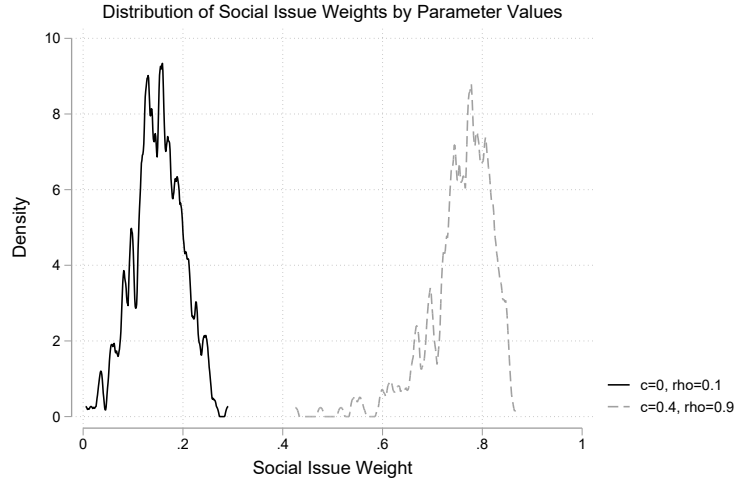## D.1 Figures for Baseline Simulations

Figure 11: This figure plots the distribution of weights that voters place on social issues for two parameter combinations, $c = 0, \rho = 0.1, k = 9$ and $c = 0.4, \rho = 0.9, k = 9$.
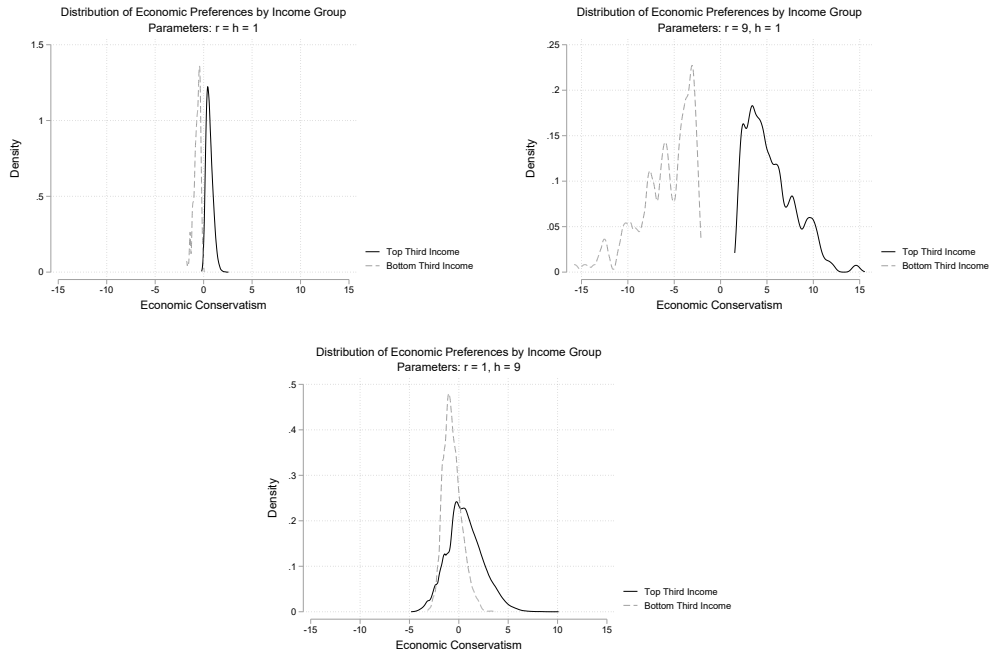

Figure 12: This figure plots the distribution of voters' economic views, split by whether voters are in the top or bottom third of the income distribution. The parameters are $r_0 = h_0 = 1$ in the top panel (income and values are equally important for economic views), $r_0 = 9$ and $h_0 = 1$ (income is significantly more important than values) in the middle panel, and $r_0 = 1$ and $h_0 = 9$ (values are significantly more important than income)in the bottom panel. Across all three panels, $c = 0$, $\rho = 0.1$, and $k = 9$.
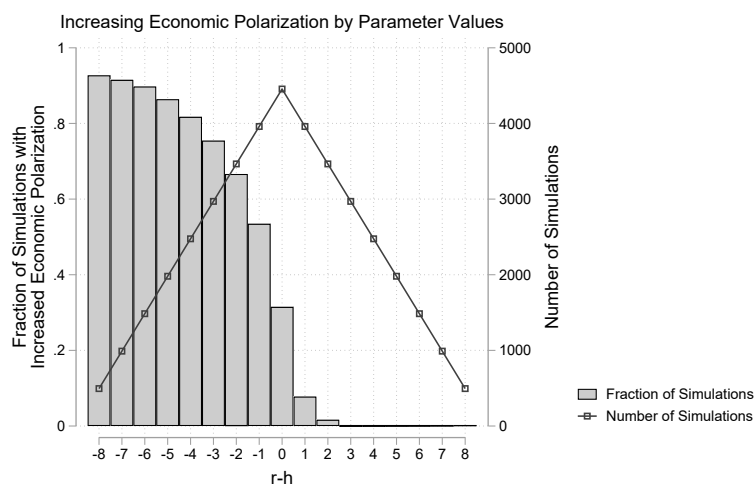
Figure 13: This bar graph plots the share of simulations in which economic polarization increases over time, separately for different values of $r_0 - h_0$. The line graph with the y-axis on the right plots the number of simulations for each $r_0 - h_0$ value.
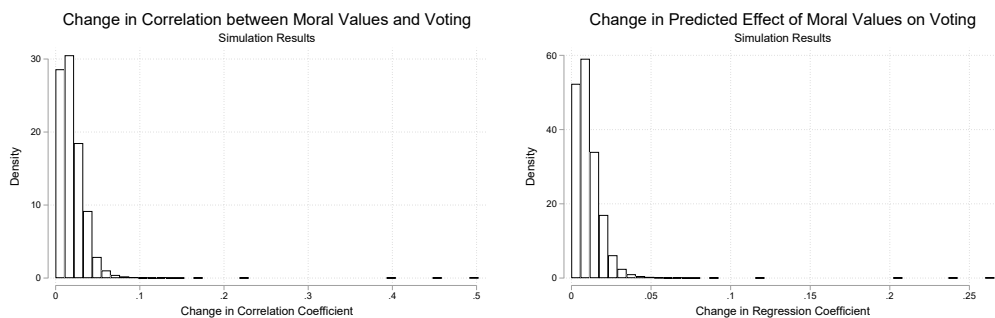

Figure 14: To analyze how the relationship between values and voting changes over time in the simulations, we calculate two measures of the relationship in the "early" ($\beta = 0$) and "late" ($\beta = 0.23$) periods of the simulation. First, we compute the correlation between moral conservatism and an indicator for voting Republican (observations are individual voters with varying levels of income and moral conservatism). Then, we calculate the change in the correlation from the early period to the late period, where positive values mean the correlation between values and voting increases over time. The top panel plots the distribution of the change in correlation between moral conservatism and voting Republican, where variation in changes comes from different parameters used in the simulation. The second measure we calculate is the coefficient on moral conservatism in a linear regression of an indicator for voting Republican on moral conservatism. We calculate the change in this regression coefficient from the early period to the late period, and the bottom panel plots the distribution of changes across simulation parameters.
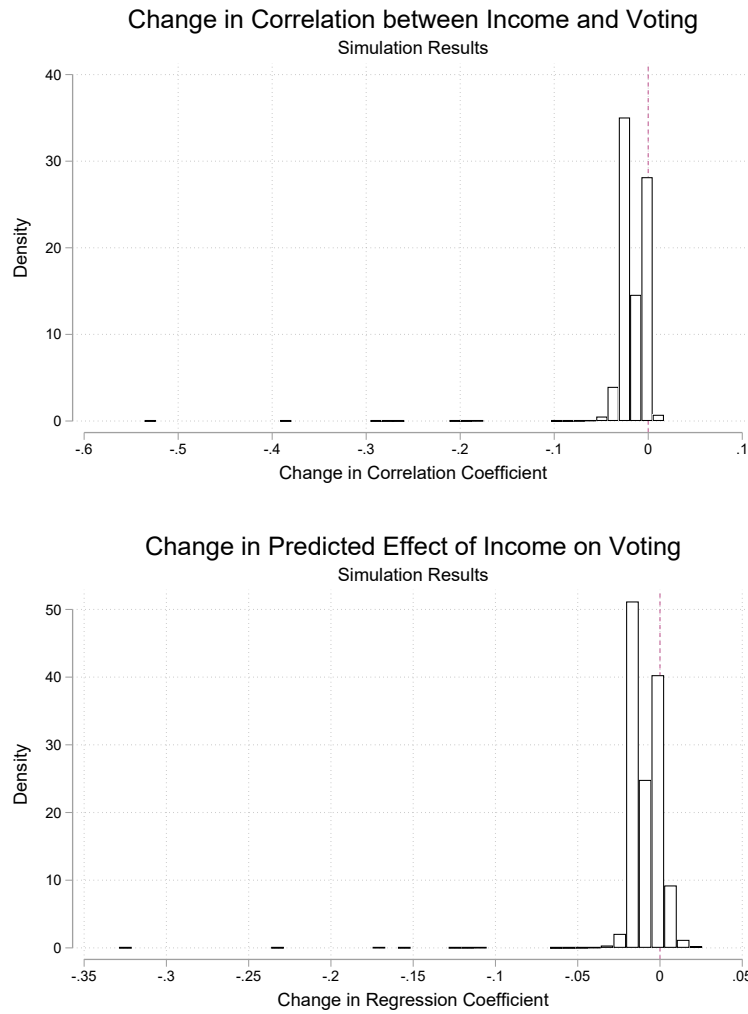
Figure 15: To analyze how the relationship between (log) income and voting changes over time in the simulations, we calculate two measures of the relationship in the "early" ($\beta = 0$) and "late" ($\beta = 0.23$) periods of the simulation. First, we compute the correlation between log income and an indicator for voting Republican (observations are individual voters with varying levels of log income and moral conservatism). Then, we calculate the change in the correlation from the early period to the late period, where negative values mean the correlation between log income and voting decreases over time. The top panel plots the distribution of the change in correlation between log income and voting Republican, where variation in changes comes from different parameters used in the simulation. The second measure we calculate is the coefficient on log income in a linear regression of an indicator for voting Republican on log income. We calculate the change in this regression coefficient from the early period to the late period, and the bottom panel plots the distribution of changes across simulation parameters.
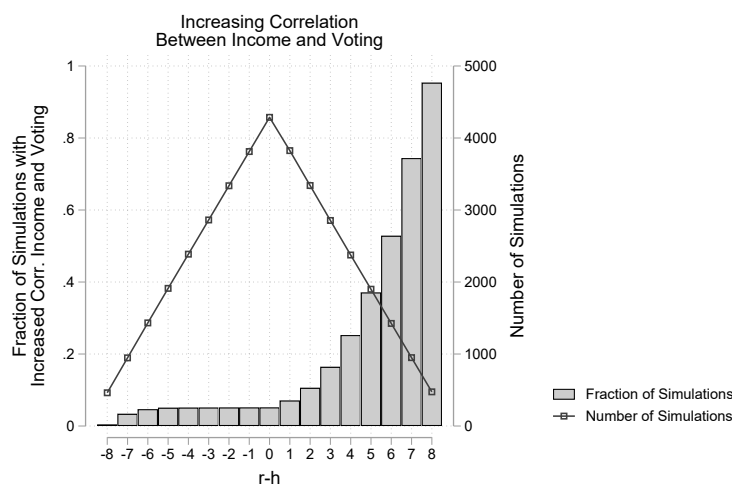
Figure 16: This figure plots the fraction of simulations in which the correlation between log income and an indicator for voting Republican increaes over time, separately for each value of $r_0 - h_0$. The line graph plots the number of simulations with a given value of $r_0 - h_0$.
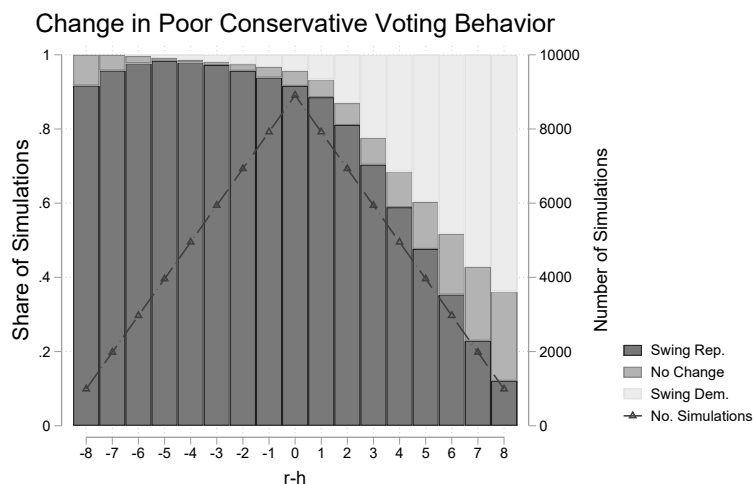

Figure 17: The bar graph plots the fraction of cases in which poor conservatives (bottom $1/3$ income, top $2/5$ moral conservatism) are more likely, equally as likely, or less likely to vote Republican over time as income increases. This is done separately for each value of $r_0 - h_0$. The line graph plots the number of simulations within each $r_0 - h_0$ value.

D.2 Details for Income Inequality Extension Simulations

D.2.1 Procedure

As in the main analyses, we model log income in the "early" period as $y \sim N(11, 0.75)$, which is roughly calibrated to the U.S. household income distribution in 1988. To assess the effect of increases in income inequality (largely driven by the growing incomes of

the rich), we scale log income multiplicately by a scaling factor κ. Thus, log income is distributed $\kappa \cdot y \sim N(\kappa \cdot 11, \kappa^2 \cdot 0.75)$ in the "late period". Since the scaling is multiplicative, the absolute log income of rich voters increases more than the absolute log incomes of the poor; thus, the difference between their log incomes increases, i.e., the ratio of the absolute income of the rich relative to the income of the poor grows. This is in contrast to the additive scaling in the analyses in the main text, which preserves relative income in the population.

We choose $\kappa = 1.015$, which corresponds to the approximate change in absolute income for the 25th percentile, median, 75th percentile, 90th percentile, and 95th percentile of U.S. households from 1988 to 2020. Table 7 illustrates this by presenting data on the U.S. household income distribution, along with the corresponding incomes in our simulated dataset.

Table 7: Income Distribution in U.S. Data and Inequality Simulation

	1988		2020	
Income Percentile	U.S. Data	Simulated Data	U.S. Data	Simulated Data
5th Pctile	$7,000	$14,500	$9,400	$16,800
10th Pctile	$17,600	$20,000	$16,000	$23,000
25th Pctile	$32,000	$33,000	$34,000	$39,000
Median	$59,500	$59,500	$68,000	$70,500
75th Pctile	$90,000	$107,000	$123,000	$128,000
90th Pctile	$176,000	$180,000	$203,000	$208,000
95th Pctile	$206,000	$246,000	$273,000	$276,000

This table presents data from the Census Bureau on the U.S. household income distribution in 1988 and 2020, along with the simulated income distribution implied by model parameters. In the "early" period of the simulation (1988), log income is distributed $N(11, 0.75)$. In the "late" period of the simulation (2020), log income is distributed $N(\kappa \cdot 11, \kappa^2 \cdot 0.75)$ where $\kappa = 1.015$. This table reports the corresponding levels of income associated with various percentiles of both log income distributions.

The increase in income inequality has meaningful effects on voters' policy preferences. Consider a voter in the "early" period with log income equal to the 30th percentile of the early income distribution. For this voter to be indifferent between the parties under the

standard set of parameters used to illustrate our simulation in the main text and fixed party positions,[20] they would need to have values equal to the 75th percentile of the moral conservatism distribution. Next, consider that same voter, who now lives in a society with the "late" income distribution. Although this voter has the same level of absolute income as before, he is relatively poorer because the rich have become richer. He now has a greater economic incentive to vote Democrat, and for him to be indifferent between parties, he needs to have values equal to the 90th percentile of the moral conservatism distribution.

D.2.2   Results

Table 8: Simulation results regarding effects of increase in κ

	Fraction of simulations in which variable in col. (1)		
	Increases	Constant	Decreases
Party polarization on social issues	90.7	9.3	0.0
Party polarization on economic issues	70.9	3.5	25.6
Moral polarization in electorate	91.2	8.8	0.0
Income polarization in electorate	42.8	2.6	54.6
Rich moral liberals voting Democratic	63.0	36.3	0.7
Poor moral conservatives voting Republican	79.2	5.6	15.2

Notes. This table presents the fraction of inequality simulations in which the variable in column 1 increases, remains constant, or decreases over time (as κ increases). Social (economic) party polarization increases if the distance between party positions on social (economic) issues increases by more than 0.1% of a standard deviation of social (economic) preferences. Moral (income) polarization in the electorate increases if the difference in average moral conservatism μ (income y) between voters who vote Republican and those who vote Democrat increases by more than 0.1% of a standard deviation of moral conservatism (income). To compute this, we calculate weighted averages of moral conservatism and income in the population, where the weights are given by the probability of voting Republican and Democrat, and take the difference between the two. Rich moral liberals are defined as top 1/3 income and bottom 2/5 moral conservatism. Poor moral conservatives are defined analogously.

---

[20]The parameters are $c = 0$, $\rho = 0.2$, $k = 9$, $h_0 = 1$, and $r_0 = 1$. The Democrat positions are -0.5 on both social and economic issues, and the Republican positions are 0.5 on both issues.
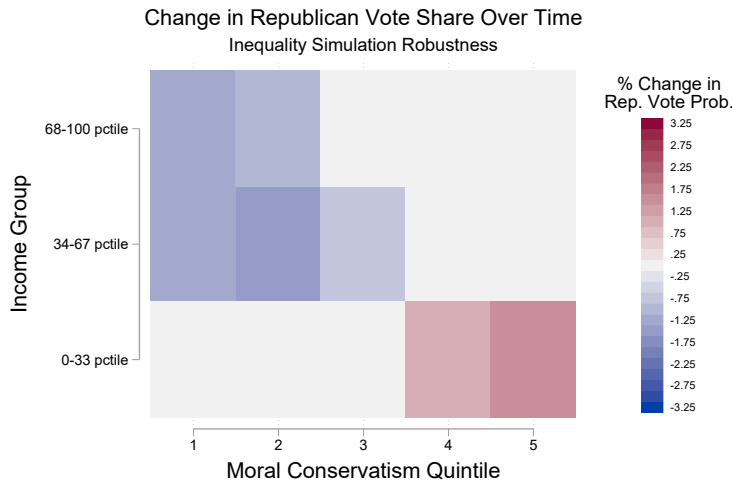
Figure 18: Percentage point change in simulated Republican vote share as a function of fundamentals (y, μ) between the early period and late period in which incomes are scaled multiplicatively by $\beta = 1.015$. This figure shows average results across all simulations.

Table 8 shows that social polarization increases in essentially all of the simulations and economic polarization increases in 71% of simulations. We find more cases of economic polarization in the inequality simulations than in the baseline simulations - this is because growing income inequality increases the likelihood that voters sort along income rather than values, and parties respond by setting more extreme economic policy to cater to their supporters.

Table 8 shows that in all of our simulations, the average Republican and the average Democratic voter diverge in terms of their values, and moral polarization in the electorate increases. Income-based polarization in the electorate decreases in 55% of the inequality simulations - this is less than the 93% of cases in the baseline simulation, and it is driven by cases where the growing income inequality outweighs increasing social polarization for a subset of voters.

On average across all the simulations, rich liberals become more likely to vote Democrat and poor conservatives become more likely to vote Republican. However, the extent of sorting of poor conservatives into the Republican party is less extreme relative to the baseline simulations. This is because income inequality makes poor conservatives more economically liberal, which pushes against their moral conservatism.

Figure 19 shows that values always become more important for voting over time as incomes rise and income inequality increases. Figure 20 shows that income almost always

becomes less important for voting, except for a handful of cases. As in the baseline simulations, these cases have large values of $r_0 - h_0$. Intuitively, when voters care more about income than values, rising income inequality leads poor voters (and in particular poor conservatives) to vote Democrat, which produces an stronger correlation between income and voting Republican.
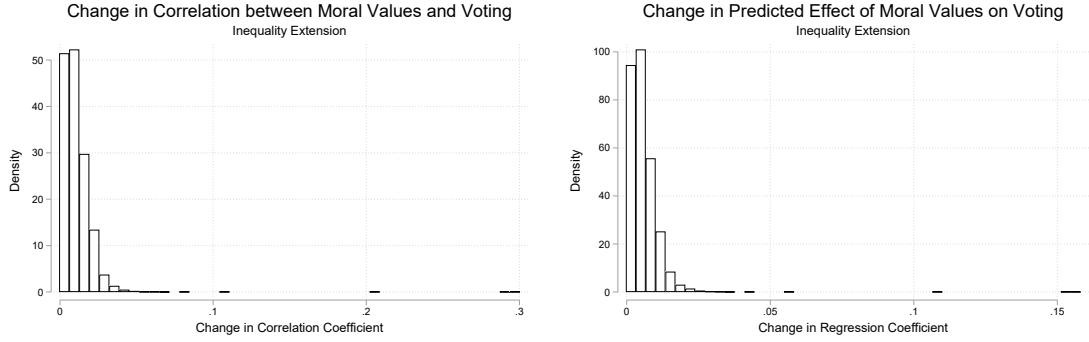

Figure 19: To analyze how the relationship between values and voting changes over time in the inequality simulations, we calculate two measures of the relationship in the "early" ($\beta = 1$) and "late" ($\beta = 1.015$) periods of the simulation. First, we compute the correlation between moral conservatism and an indicator for voting Republican (observations are individual voters with varying levels of income and moral conservatism). Then, we calculate the change in the correlation from the early period to the late period, where positive values mean the correlation between values and voting increases over time. The top panel plots the distribution of the change in correlation between moral conservatism and voting Republican, where variation in changes comes from different parameters used in the simulation. The second measure we calculate is the coefficient on moral conservatism in a linear regression of an indicator for voting Republican on moral conservatism. We calculate the change in this regression coefficient from the early period to the late period, and the bottom panel plots the distribution of changes across simulation parameters.
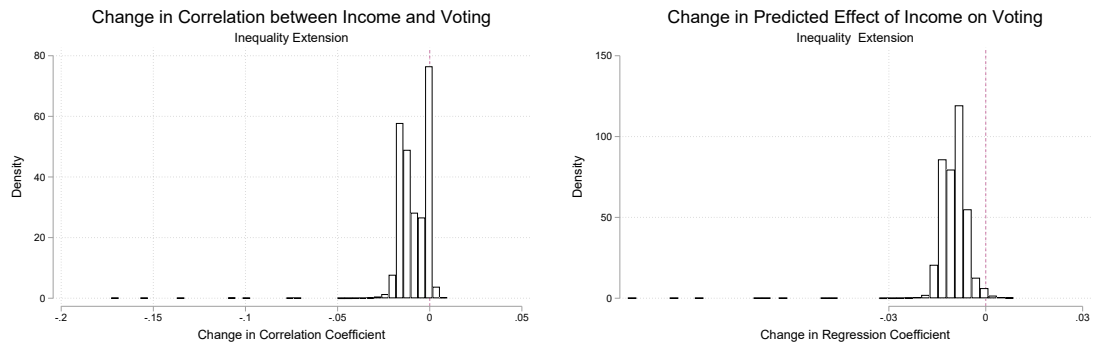
Figure 20: To analyze how the relationship between log income and voting changes over time in the inequality simulations, we calculate two measures of the relationship in the "early" and "late" ($\kappa = 1.015$) periods of the simulation. First, we compute the correlation between log income and an indicator for voting Republican (observations are individual voters with varying levels of log income and log income). Then, we calculate the change in the correlation from the early period to the late period, where positive values mean the correlation between values and voting increases over time. The top panel plots the distribution of the change in correlation between log income and voting Republican, where variation in changes comes from different parameters used in the simulation. The second measure we calculate is the coefficient on log income in a linear regression of an indicator for voting Republican on log income. We calculate the change in this regression coefficient from the early period to the late period, and the bottom panel plots the distribution of changes across simulation parameters.
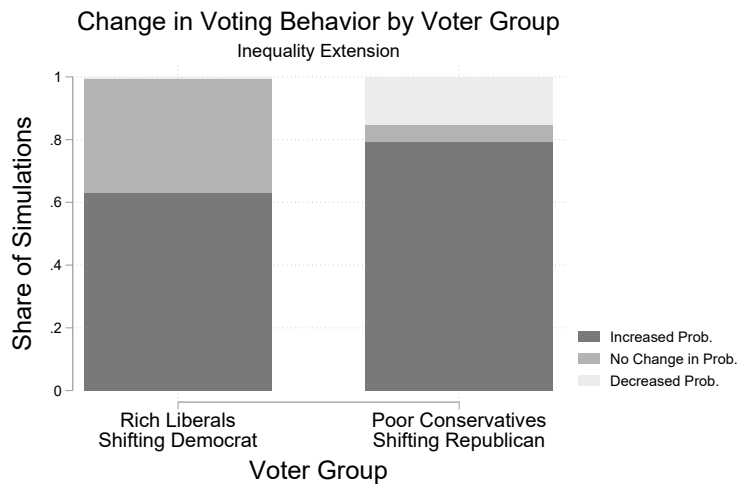

Figure 21: This figure plots the share of inequality robustness simulations in which rich liberals (top 1/3 income and bottom 2/5 moral conservatism) become more likely to vote Democrat and poor conservatives (bottom 1/3 income and top 2/5 moral conservatism) become more likely to vote Republican.
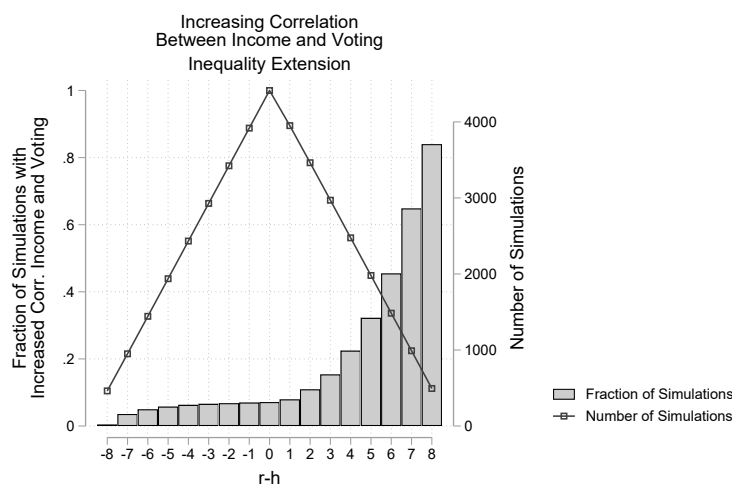
Figure 22: This figure plots the fraction of inequality simulations in which the correlation between log income and an indicator for voting Republican increaes over time, separately for each value of $r_0 - h_0$. The line graph plots the number of simulations with a given value of $r_0 - h_0$.
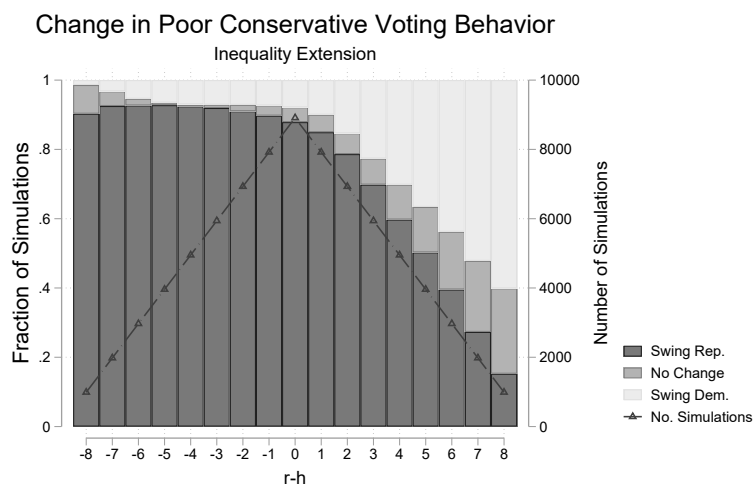

Figure 23: This figure plots the distribution of $r_0 - h_0$, split by whether poor conservatives (bottom 1/3 income, top 2/5 moral conservatism) are more likely, equally as likely, or less likely to vote Republican over time as income inequality increases.

D.3 Robustness to Initial Party Positions

In this main text, we simulate equilibrium party positions starting from the initial positions $(x_s^D, x_e^D) = (-0.5, -0.5)$ and $(x_s^R, x_e^R) = (0.5, 0.5)$. In this section, we show that our simulation results are robust to a range of initial party positions - that is, the simulation delivers the same equilibrium party positions irrespective of the starting points.

To do so, we simulate over a grid of initial positions: $x_s^D \in [-5, -1, -0.5, -0.1, 0, 0.1, 0.5, 1, 5] \times$ $x_e^D \in [-5, -1, -0.5, -0.1, 0, 0.1, 0.5, 1, 5] \times x_s^R \in [-5, -1, -0.5, -0.1, 0, 0.1, 0.5, 1, 5] \times x_e^R \in$ $[-5, -1, -0.5, -0.1, 0, 0.1, 0.5, 1, 5]$ for a total of 6561 simulations. These positions cover a wide range of initial cases for $(x_s^D, x_e^D, x_s^R, x_e^R)$:

1. Parties converge at various points. For example, $(0, 0, 0, 0)$ or $(5, 5, 5, 5)$.

2. Parties polarize symmetrically to varying degrees. For example, $(-0.1, -0.1, 0.1, 0.1)$ or $(-5, -5, 5, 5)$.

3. One party is more extreme than the other on one or both issues. For example, $(-0.5, -0.5, 5, 0)$ implies that the Republicans are extreme on social issues and moderate on economic issues, while the Democrats are somewhat liberal on both.

4. Parties cater to "off-diagonal" voters who are rich and liberal or poor and conservative. For example, $(-1, 1, 1, -1)$. Importantly, these cases show that our main simulation results which deliver one liberal party and one conservative party (on both issues) do not depend on starting from such points.

Note that in cases where the parameter configurations are symmetric (e.g. $(-1, -1, 1, 1)$ and $(1, 1, -1, -1)$), one of the cases can be ignored without loss of generality, as we simply rename the party with larger (i.e. more conservative) positions as the Republican party. More generally, this logic applies to cases where the equilibrium positions chosen by the Democratic party are more conservative than those chosen by the Republican party. We label the party with the more conservative policies as the Republican party, since the relevant result from the simulation is that there exists a party which optimally chooses the positions, not the identity of the party.

To implement this robustness simulation, we require a set of values for the remaining model parameters, $[c, \rho, k, r_0, h_0]$. If we use the full grid from the simulation in the main text, we quickly face a curse of dimensionality since 6,561 initial positions $\times$ 3,645 parameter values generates nearly 24 million simulations. Instead, we select a handful of parameter values which cover an economically meaningful range of voter preferences and run the initial party position robustness with those parameters.

Table 9 summarizes the results. For each set of parameters, we report the mean,

standard deviation, maximum, and minimum equilibrium party position across all initial party positions. Regardless of the initial position, the parties always converge very closely to the equilibrium position. This table excludes the 81 initial positions where both parties offer the same policies - in all of these cases (and only in these cases), both parties offer the exact same equilibrium positions as one another.

Table 9: Simulation Robustness to Initial Party Positions

	Parameter Set 1				Parameter Set 2				Parameter Set 3			
Variables	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Dem. Social Policy	-0.508	0.000132	-0.508	-0.504	-0.320	0.000373	-0.321	-0.319	-0.625	0.000303	-0.627	-0.622
Dem. Econ Policy	-0.534	0.000120	-0.534	-0.531	-1.455	0.000283	-1.456	-1.455	-3.223	0.000504	-3.225	-3.221
Rep. Social Policy	0.477	0.000186	0.471	0.477	0.176	0.000178	0.175	0.176	0.845	0.000260	0.845	0.846
Rep. Econ Policy	0.447	0.000150	0.440	0.447	0.782	0.000187	0.782	0.783	4.153	0.000627	4.152	4.155

This table presents summary statistics for the equilibrium party positions reached from baseline simulations which are initialized at 6,480 unique starting values. "Parameter Set 1" refers to the model parameters $\beta = 0$, $c = 0.1$, $\rho = 0.2$, $h_0 = 2$, and $r_0 = 1$. "Parameter Set 2" refers to the model parameters $\beta = 0$, $c = 0.1$, $\rho = 0.3$, $h_0 = 1$, and $r_0 = 3$. "Parameter Set 3" refers to the model parameters $\beta = 0$, $c = 0.2$, $\rho = 0.5$, $h = 8$, and $r = 3$.

E   Data Description

E.1   American National Election Studies

Economic and social conservatism indices.   Since 1988, the ANES has consistently included questions on economic issues such as social security and federal spending on public schools, and on social issues such as gay rights and the importance of family values. We adopt the methodology of Ansolabehere et al. (2006) in aggregating survey questions to construct indices of economic and social conservatism. We classify issues as either economic or moral, then perform factor analyses on each set of questions. The factor analysis allows us to construct a weighted average of the questions, where the weight placed on each question is determined its factor loading. We standardize the factor scores to have mean zero and standard deviation one, and arrange them such that higher values correspond to more conservative issue positions. Across years, our sample spans up to 18,432 respondents. The economic conservatism scale is constructed from survey responses to the following ten items:

1. "Some people think the government should provide fewer services, even in areas such as health and education, in order to reduce spending. (2004: Suppose these people are at one end of a scale, at point 1.) Other people feel that it is important for the government to provide many more services even if it means an increase in spending. (2004: Suppose these people are at the other end, at point 7. And of course, some other people have opinions somewhere in between, at points 2,3,4,5, or 6.) Where would you place yourself on this scale, or haven't you thought much about this?"

2. "There is much concern about the rapid rise in medical and hospital costs. Some (1988,1994-LATER: people) feel there should be a government insurance plan which would cover all medical and hospital expenses (1984 AND LATER: for everyone). (1996,2004: Suppose these people are at one end of a scale, at point 1). Others feel that (1988,1994-1996: all) medical expenses should be paid by individuals, and through private insurance (1984 AND LATER: plans) like Blue Cross (1984-1994: or [1996:some] other company paid plans). (1996,2004: Suppose these people are at the other end, at point 7. And of course, some people have opinions somewhere in between at points 2,3,4,5 or 6.) Where would you place yourself on this scale, or haven't you thought much about this?"

3. "Some people feel that the government in Washington should see to it that every person has a job and a good standard of living. (1972-1978,1996-LATER: Suppose these people are at one end of a scale, at point 1). Others think the government should just let each person get ahead on his/their own. (1972-1978,1996: Suppose these people are at the other end, at point 7. And, of course, some other people have opinions somewhere in between, at pints 2,3,4,5 or 6.). Where would you place yourself on this scale, or haven't you thought much about this?"

4. "Should federal spending on improving and protecting the environment (2000,2002: environmental protection; 2008,2012,2016: protecting the environment) be increased, decreased, or stay the same?"

5. "If you had a say in making up the federal budget this year, for which (1986 AND

LATER: of the following) programs would you like to see spending increased and for which would you like to see spending decreased:"

(a) Child care

(b) Welfare programs

(c) Social security

(d) Public schools

6. "We'd also like to get your feelings about some groups in American society. When I read the name of a group, we'd like you to rate it with what we call a feeling thermometer. Ratings between 50 degrees-100 degrees mean that you feel favorably and warm toward the group; ratings between 0 and 50 degrees mean that you don't feel favorably towards the group and that you don't care too much for that group. If you don't feel particularly warm or cold toward a group you would rate them at 50 degrees. If we come to a group you don't know much about, just tell me and we'll move on to the next one."

(a) Labor unions

The social conservatism scale is constructed from survey responses to the following nine items:

1. "Do you favor or oppose laws to protect [homosexuals/gays and lesbians] against job discrimination? Do you favor/oppose such laws strongly or not strongly?"

2. "Now, I am going to read several statements. After each I would like you to tell me whether you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat or disagree strongly with this statement."

(a) The newer lifestyles are contributing to the breakdown of our society.

(b) The world is always changing and we should adjust our view of moral behavior to those changes.

(c) This country would have many fewer problems if there were more emphasis on traditional family ties.

71

(d) We should be more tolerant of people who choose to live according to their own moral standards, even if they are very different from our own.

3. "Do you think gay or lesbian couples, in other words, homosexual couples, should be legally permitted to adopt children?"

4. "There has been some discussion about abortion during recent years. (RESPON-DENT BOOKLET) Which one of the opinions on this page best agrees with your view? You can just tell me the number of the opinion you choose."

   (a) By law, abortion should never be permitted.

   (b) The law should permit abortion only in case of rape, incest, or when the woman's life is in danger.

   (c) The law should permit abortion for reasons other than rape, incest, or danger to the woman's life, but only after the need for the abortion has been clearly established.

   (d) By law, a woman should always be able to obtain an abortion as a matter of personal choice.

   (e) DK; other

5. "Which of these statements comes closest to describing your feelings about the Bible?"

   (a) The Bible is the actual Word of God and is to be taken literally, word for word

   (b) The Bible is the Word of God but not everything in it should be taken literally, word for word

   (c) The Bible is a book written by men and is not the Word of God

   (d) Other; DK

6. "Do you consider religion to be an important part of your life, or not?"

In our baseline specification, we allow the set of questions used to calculate the indices to vary across years, as some questions are added and removed from the survey over time. Two advantages of this approach are that it (i) incorporates the broadest amount of

information possible and (ii) accounts for the fact that some issues which were important in 2020 were not included in the survey in 1988. A concern with this approach is that it skews comparisons across time because liberals and conservatives in 2020 are defined over a different set of questions than their counterparts in 1988. To address this, we construct indices from a factor analysis on a restricted set of issues which are available in all years from 1988 to 2020. The correlation between the baseline and restricted indices is $\rho = 0.95 - 0.97$, and we verify that all of the empirical results are consistent across specifications.

Income. The income variable in the ANES consists of five buckets that partition the population into income percentiles. These buckets are the 0-16th percentile, 17-33rd percentile, 34-67th percentile, 68-95th percentile, and 96-100th percentile. The ANES also provides the lower and upper bounds of the lower four buckets. To impute (log) income from the buckets, we take the average of the upper and lower bounds for the four lowest buckets. To estimate income in the top bucket, we use data from the Tax Policy Center on average US incomes in the 96th to 99th percentile and the top 1% for each year. We take a weighted average of these incomes to get average income within the 96th-100th percentile group.

E.2   Moral Foundations Questionnaire

See Enke (2020) for a detailed description of the dataset and how the moral conservatism index (relative importance of communal vs. universalist moral values) is computed.