

# Why Some Care More About Free Riding Than Others and Why It Matters

## *Abstract*

People support policies that increase their own expected income. They also support policies that move the status quo closer to what is prescribed by agreed-upon norms of fairness. How do these two motives combine? In most circumstances, I argue, people reason as moral agents trying to do the “fair” thing. Only when status quo changing policies have large and certain material consequences will they deviate from saying what is fair and choose to express a self-serving position instead. I apply this simple framework to a form of fairness reasoning that ties social policy preferences to beliefs about the prevalence of free riders among net beneficiaries of social spending. I show that income level and the institutional context affect the extent to which pocketbook concerns overrun free riding ones. I flesh out the implications for the politics of social policy reform in mature welfare states.

**Word Count:** 11,980

# Introduction

People derive utility—in the form of self-respect and social approval—from expressing what is “morally fair.” In studies of people’s attitudes toward redistributive social policies, one type of fairness concern routinely surfaces, namely dissatisfaction with a status quo perceived as unfairly benefiting “undeserving” recipients and unfairly excluding “deserving” ones (Cramer 2016; Lamont and Molnar 2002).<sup>1</sup> Research shows that this form of moral reasoning is important for understanding why individual policy preferences do not align with what one might expect from self-serving economic reasoning alone (Gilens 1999; Fong 2001). Should we conclude that this form of moral reasoning is powerful enough to systematically crowd out pocketbook concerns? Under what conditions might people deviate from what their free riding beliefs prescribe and adjust their policy preferences in line with their material self-interest?

This paper seeks to answer these questions. I start with a simple observation: because people have a hard time processing the pocketbook implications of a given set of policy changes, expressing what is fair is often easier than expressing what is in one’s own material self-interest. As a result, people’s policy preferences are often best predicted by their beliefs about the fairness of the status quo, i.e., beliefs about the prevalence of free riding among net beneficiaries of a social program. Yet, this does not imply that material self-interest plays no role. Indeed, while pocketbook reasoning is difficult for most people most of the time, there are important differences both across individuals and across policy contexts.

First, some policies are more transparent about costs and benefits than others. Second, assuming high enough stakes, people will invest the necessary time to understand how a policy works and how it can be improved to best serve their own economic interests (Roth, Settele and Wohlfart 2020). When a policy’s consequences are large and easy to understand, individuals will depart from supporting the fair policy and support what is economically self-serving instead. Applied

---

<sup>1</sup> For evidence regarding other types of fairness concerns, including the extent to which income differences are tied to differences in merit, see Cavaille (Forthcoming).

to attitudes toward redistributive social policies, this simple argument suggests that the policy preferences people ultimately settle on are a function of 1) the type of free riding beliefs they start from, and 2) the extent to which a policy's implications for one's disposable income will lead them to deviate from the baseline set by these beliefs.

By combining moral concerns over free riding and material self-interest in the same framework, this paper speaks to important debates in the study of social policy preferences. A lot has been said about the “irrational poor” who vote against their interests (Frank 2007). The framework and evidence presented in this paper demonstrate that this statement does not easily extend to social policy preferences. On the one hand, low-income respondents are more likely to find the status quo unfair, i.e., believe that free riding is high prevalence. On the other hand, they are less likely to translate their free riding beliefs into opposition to generous and inclusive social transfers, expressing self-serving support for such policies instead. This means that a weak correlation between income and attitudes toward policies that benefit the worse off is due to higher-than-expected support among high-income respondents who believe it benefits deserving recipients, not lower-than-expected opposition among low-income respondents who believe it benefits undeserving recipients. Relatedly, changes in free riding beliefs that affect all income groups have a larger impact on the social policy preferences of high-income respondents, with only limited impact on that of low-income respondents. This framework, I show, helps explain puzzling variations in attitudinal differences between the rich and the poor. It also sheds a new light on the coalitions underpinning social policy reform in countries with mature welfare states.

This paper's argument is most closely related to work by Rueda and Stegmueller (2019), who argue that, because of declining marginal return to consumption, only the rich have the luxury to reason altruistically. Instead of altruism, I emphasize the role of fairness reasoning, which might or might not result in altruistic preferences. Indeed, for a subset of the population, retrenchment—and not expansion—is the “fair” policy. Furthermore, I examine how institutional factors shape the mix of motives (fairness versus pocketbook reasoning) people rely on when forming social policy preferences. This emphasis on institutional factors builds on previous studies by Beramendi

and Rehm (2016) and Holland (2018), which examine the extent to which the institutional context favors pocketbook reasoning. I extend these contributions by theorizing what happens when the institutional context does *not* favor such reasoning.

## **Fairness Reasoning Comes First, Material Self-Interest Second**

Material self-interest, i.e., the assumption that individuals behave in ways that maximize their income, is best understood “in context.” To understand what I mean by this, it is helpful to contrast the institutional environment of redistributive politics in representative democracies with the “highly structured and competitive environment” of “open markets” (Ostrom and Walker 2003: 25). Actors participating in open markets face a high-stakes environment where information is provided at a manageable cost, feedback is continuous and learning is strongly incentivized. Such institutional set up rewards self-interested economic reasoning, decreases its cognitive costs, and penalizes competing other-regarding and non-economic modes of reasoning. As a result, in the majority of cases, the behavior of individuals interacting through market institutions can be usefully approximated using the homo economicus theoretical tool box.<sup>2</sup>

Yet the *same* set of individuals, when engaged in “producing” redistributive social policies through the institutions of representative democracies, are unlikely to behave as self-interested, materially inclined and well-informed actors. This is partly due to the nature of redistributive politics in countries with mature welfare states. Indeed, self-interested politicians seeking to compromise between conflicting interests and to minimize the risk of a backlash will often propose and implement policies with diffuse redistributive implications. They will also remain equivocal and ambiguous about the redistributive implications of their policy proposals (Downs 1957).

---

<sup>2</sup> There are, of course, differences across types of market actors and market behaviors: stakes are higher for consumers choosing which car to buy than for consumer choosing which snack to indulging in. Similarly, the disciplining effect of market institutions are especially large for entrepreneurs: those who fail to behave as self-interested profit maximizers are quickly pushed out.

Alternative information-aggregating actors such as unions have often stepped in to close the information gap, but, as recent studies have shown, their contribution to public discourse has decreased dramatically in the past few decades (Wallerstein and Western 2000; Farber et al. 2018). Strategic avoidance of policies with large redistributive implications is especially likely in countries with large welfare states and high levels of taxation (Pierson 2001). In addition, in these countries, low stakes often go hand-in-hand with high uncertainty. Indeed, states' capacity to generate cheap sovereign debt not only helps diffuse the costs of redistributive reforms by passing them on to future generations, but also increases uncertainty by blurring the relationship between social spending and taxation. Faced with a complex combination of taxes and transfers and only limited guidance from the political context, voters might still form self-interested preferences through experience or "learning-by-doing," by which I mean the process of acting on one's preferences and then facing the consequences of these actions. Yet, as highlighted by Schumpeter (1950) more than half a century ago, in representative democracies, this specific form of learning is limited as voters rarely get to connect a political outcome to their own individual action, let alone policy preference. In other words, for most people, redistributive politics, at least as they play out in representative democracies with mature welfare states, represent an issue whose personal implications are either low-stakes or unknown.

In such a context, choosing the policy that maximizes one's own income is often difficult. For many, it is not worth the effort. As I will argue later, only when high-stakes policy reforms have credible implications for a voter's pocketbook can one reasonably expect such egocentric and rational behavior. If self-interested economic reasoning is the exception, not the rule, then what is the rule?

## **Fairness Reasoning and the Monitoring of Free Riding**

As always with human behavior, there are many rules, and parsimony requires choosing one among those that matter the most. Here, I make the case for a particular form of fairness-maximizing be-

havior, which manifests itself as the impulse, in a cooperative resource pooling situation, to reward the “deserving cooperator” and punish the “undeserving free rider.” In this case, behaving fairly requires being able to answer the following question: “is the modal beneficiary free riding and undeserving?” Assuming people hold beliefs about the prevalence of free riding, then answering this question is comparatively easier than answering the self-interested one (“will my income be affected, and if so, how?”). This explains why fairness reasoning tends to be the default mode of reasoning.

The existence of free riding beliefs is well documented by studies examining perceptions of welfare recipients as more or less “deserving” (Gilens 1999; Petersen et al. 2011). I prefer the term “free riding beliefs” to that, more common in the literature, of “deservingness perception” for at least three reasons. First, this expression better conveys the specific type of deservingness judgments under scrutiny, namely the extent to which net beneficiaries of social spending are perceived as abusing the kindness of net contributors. Second, it provides an umbrella expression to capture the different types of beliefs that contribute to people’s evaluation of the status quo as fair or not. These include people’s beliefs about recipients’ personal flaws (e.g. “laziness), which are central to the deservingness literature, as well as their beliefs about the behavioral consequences of social transfers, what economists call moral hazard. Both types of beliefs are highly correlated and contribute to the perceived prevalence of free riding (see Table 1). Third, as I explain next, this emphasis on free riding helps connect survey evidence from the “deservingness” literature on the one hand, to research on the role moral reasoning plays in the provision of mutual help on the other.

Faced with markets’ failure to offer affordable insurance against unemployment, disability and old age (Akerlof 1970), governments have stepped in, compelling - to various extents - their citizens to participate in public risk-pooling programs that provide income to those unable to work (Korpi 2006; Moene and Wallerstein 2001; Iversen and Soskice 2001). For similar reasons<sup>3</sup>, states

---

<sup>3</sup> See, for example Barr (1998)

have concomitantly become the main providers of basic education and health care. As a result, citizens are stakeholders in risk and resource-pooling programs of a historically unprecedented scope. Such cooperative and resource pooling endeavor constitutes what researchers call a social dilemma. Specifically, all members have an incentive to free ride, i.e., take advantage of the shared resource without contributing sufficiently to its maintenance. Researchers have documented the existence of a deeply rooted set of norms and behavioral impulses that, in such a context, turn people into conditional cooperators, which helps mitigate free riding and sustain cooperation. (Petersen 2012; Bowles and Gintis 2011; Graham, Haidt and Nosek 2009; Charness and Rabin 2002; Fischbacher, Gächter and Fehr 2001; Ostrom 1998). Specifically, evidence shows that people willingly contribute to a collective endeavor if they feel others are not free riding (positive reciprocity). They punish free riders by either ceasing to cooperate or by excluding them from accessing the goods generated by cooperation (negative reciprocity). Cooperative behavior is thus inherently two-faceted and can be presented in one of two lights. The more positive light casts it as a form of conditional altruism: people's default position is to help others unless others are "antisocial" (Henrich et al. 2001; Fong, Bowles and Gintis 2006). Viewed in a negative light, it is a form of conditional punishment: people's default position is to deny help to others unless they are prosocial.

In the lab, cooperation between two individuals is conditional on past and new information on each individual's situation and perceived intentions.<sup>4</sup> Such context-specific information becomes too complicated to process in larger groups: what stabilizes cooperation is the shared beliefs that free riding is limited (e.g., Fischbacher and Gächter (2010)). In other words, forming expectations about the prevalence of free riding in a cooperative context is something people do by default: a positive or negative opinion about a given cooperative endeavor goes alongside an optimistic or pessimistic assessment of the prevalence of free riding.

---

<sup>4</sup> Research has documented the range of emotions underpinning resource sharing and cooperation, most importantly outrage at being taken advantage of (contributing when others are not), shame for failing to "carry one's weight" (not contributing when others are), or outrage at seeing someone "jump the queue" (receiving the same access despite "slacking" behavior).

In survey data, this form of reasoning manifests itself in the form of a strong correlation between 1) a subset of social policy attitudes, e.g., attitudes toward targeted social benefits, and 2) beliefs about the prevalence of free riding. Table 1 and Figure 1 illustrate this point using the example of attitudes toward tax and spending cuts. Table 1 lists items, available in the 2008 and 2016 waves of the European Social Survey, that tap into perceptions of free riding. The columns to the right of these items reproduce results from a factor analysis examining whether these items all load on the same latent dimension. The large Eigenvalues suggest a latent construct ranking individuals according to how prevalent they believe free riding to be whether due to individual moral failures or moral hazard. Figure 1 uses data from the 2008 wave to plot the country and individual-level correlations between scores on this latent construct and attitudes toward tax and social spending. Echoing individual-level evidence from the deservingness literature (Petersen et al. 2012; Fong 2001; Gilens 1999; Skitka and Tetlock 1993), Figure 1 illustrates fairness reasoning “in action:” people who support tax and spending cuts over increases tend to hold pessimistic free riding beliefs, while people who support increases over cuts tend to hold optimistic beliefs.<sup>5</sup>

Explaining why some people form more optimistic beliefs than others is far from straightforward. As I show in Appendix A.2, free riding beliefs do not correlate with individuals’ economic conditions in expected ways: income, job insecurity and skill levels do not predict free riding beliefs. Furthermore, as shown by Cavaille and Trump (2015), traditional left-wing ideology does not correlate with free riding beliefs: knowing whether someone supports redistribution, believes income inequality is too high, or perceives an inherent conflict between workers and capital owners says little about this person’s perceptions that the unemployed are made lazy by benefits and concerns over the prevalence of benefit fraud. One reason for the limited role of economic condi-

---

<sup>5</sup> Note that this form of fairness reasoning cannot be subsumed under a concept such as altruism (e.g. Rueda and Stegmueller (2019)). Altruism is unconditionally aimed at making the other better-off. In contrast, the type of other-regarding reasoning described above is highly contingent on beliefs about what others are doing, meaning that the same person can appear altruistic in one context and selfish in another, something the concept of altruism cannot capture. Luckily, students of social policy preferences have not waited for breakthroughs in research on cooperation and fairness reasoning to document this form of conditional altruism. However, their findings, mostly inductive, have failed to coalesce into a much needed revision of behavioral assumptions in political economy. This paper is a first step in that direction.

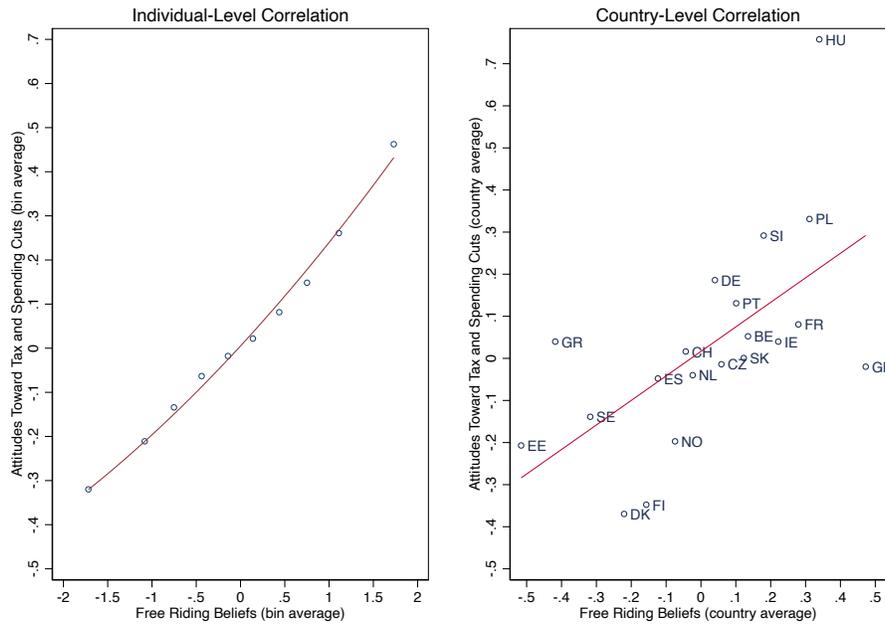
**Table 1: Free Riding Beliefs: Factor Loadings**

Item wording	ESS 2008	ESS 2016
<b>Beliefs about the prevalence of shirking</b>		
Most unemployed people do not really try to find a job	0.54	0.72
Many manage to obtain benefits/services not entitled to	0.42	0.54
Employees often pretend they are sick to stay at home	0.50	NA
<b>Beliefs about the disincentive effects of social benefits</b>		
Social benefits/services make people lazy	0.75	0.84
Social benefits/services make people less willing to care for one another	0.81	0.73
Social benefits/services make people less willing look after themselves/family	0.78	NA
Eigenvalue	2.56	2.04

Exploratory factor analysis on the pooled data using a polychoric correlation matrix adapted to ordinal variables. Extracted method: iterated principal factor method, robust to using other extraction methods. Analysis includes all countries available in both waves namely: GB, FR, IE, BE, PT, DE, NL, CH, NO, AT, ES, FI, SE, GR, as well as EE, HU, PL, SI, CZ.

Source: ESS waves 4 and 8, weighted

**Figure 1: Social Policy Preferences and Free Riding Beliefs**



*X-axis:* Average free riding beliefs are measured by combining answer to the items listed in Table 1 using weights recovered from an exploratory factor analysis ran on the pooled data. The higher a respondent’s score, the more likely she is to be concerned about free riding.  
*Y-axis:* “If the government had to choose between increasing taxes and spending more on social benefits and services, or decreasing taxes and spending less on social benefits and services, which should they do?” Respondents answer using a 0 (Government should decrease taxes a lot and spend much less on social benefits and services) to 10 (Should increase a lot and spend much more) response scale. **The scale is reversed** so that higher values indicate more support for tax and spending cuts. *Left panel:* scores are standardized using **country-specific** mean and SD. *Right panel:* scores are standardized using the mean and SD from the **pooled** data.  
 Source: ESS wave 4, weighted

tions and traditional left-right economic ideology is that free riding beliefs correlate with liberal-authoritarian “second-dimension” values (Häusermann and Kriesi 2011), which in turn, are orthog-

onal to income and traditional left-right economic ideology.<sup>6</sup> As Figure 1 shows, free riding beliefs also diverge significantly across borders, suggesting amplifying and coordinating mechanisms that incentivize large populations to settle on similar beliefs. These mechanisms might include social stratification and socialization; a political landscape that politicizes divergent free riding beliefs or complementarities between welfare state design and the modal free riding belief (Larsen 2008; Mau 2004). It is beyond the scope of this paper to unpack these social processes. For now, I focus on documenting how this form of fairness reasoning conditionally impacts policy preferences.

Note that not all social policy attitudes are shaped by fairness reasoning as defined above. That is because the monitoring of free riding assumes a situation of interdependence and cooperation between stakeholders. Interdependence and cooperation are defining features of the welfare state and “government-provided transfers” or “social spending” broadly defined. In contrast, policies that seek to regulate markets to make them “fair” do not raise free riding concerns because markets are not understood as cooperative resource pooling endeavors. Relatedly, debates over the fairness of income redistribution, to the extent that they are debates over the fairness of differences in *market* income inequality, rarely trigger free riding concerns (Cavaille Forthcoming). Appendix A.3 illustrates this heterogeneity: while free riding beliefs correlate with attitudes toward government-provided income support, they do not correlated with attitudes toward the redistribution of market income.

Having described the existence of a specific type of fairness reasoning – in the form of free riding beliefs and their correlation with social policy preferences –, I now turn to the conditions under which people disregard what fairness reasoning prescribes and settle on a self-serving policy position instead.

---

<sup>6</sup> See Appendix A.2 and B.1 for more evidence.

## Combining Fairness Reasoning and Material Self-Interest

While reasoning according to one's pocketbook is difficult for most people most of the time, there are important differences both across individuals and across different types of redistributive policies. For a given individual, whether a policy proposal is low stakes, high stakes or an unknown quantity is determined by her own individual characteristics, the amount of information available regarding how these characteristics interact with existing policies and proposed policy changes, and the extent to which one faces incentives to acquire new information to turn uncertainty into a known risk. Take, for example, a 60-year-old American man, with adult children and aging high-income parents. Some policies will be high-stakes (e.g., a decrease in the inheritance tax to be paid when his parents die), others low-stakes (e.g., an increase in spending on Medicaid, which he will only need if he falls into poverty) and some will be an unknown quantity (e.g., Obamacare, whose implications for already insured individuals are mediated by long-term adjustments in the healthcare market).

To account for this variation in uncertainty and stakes, I conceptualize preference formation as a two-step process. First, individuals rely on fairness reasoning to anchor their preferences following the “if fair then support”, “if unfair then oppose” heuristic. Second, depending on information and material stakes, individuals will adjust their preferences in line with their material self-interest, something I call *self-serving adjustment* for short. The size of this adjustment will vary with the stakes — the higher the material stakes, the more extensive self-serving adjustment will be.

This two-step process is a stylized representation of reality. Nevertheless, existing studies suggest it is a decent approximation of people's thought process. Take, for example, the finding by Margalit (2013) that job loss triggers an increase in support for unemployment insurance, though only temporarily and mostly among Republicans. In “normal times,” Republicans tend to oppose unemployment insurance because it rewards free riding. Job loss triggers self-serving adjustment among those affected. Over time, as their situation improves, and unemployment insurance is no longer experienced as a high stakes issue, individuals revert to their original position, as defined

by their beliefs about the fairness of the status quo. Using priming experiments, Chong, Citrin and Conley (2001a) document the role contextual and situational factors play in shaping the relative importance of material self-interest over other modes of reasoning. Specifically, they demonstrate that “people are more likely to recognize their own self-interest, and to act upon it, when their stakes in the policy are clear or when they have been primed to think about the personal costs and benefits of the policy.” In contrast, “people with a smaller stake in an issue” are more likely to rely on their “values and symbolic predispositions.”

The assumption that people follow such two-step process generates testable predictions regarding the structure of mass attitudes,<sup>7</sup> i.e., who will be particularly likely to self-servingly adjust, and how this varies across policy issues, institutional contexts and time. One set of predictions pertains to cross-national differences in the extent to which the rich and the poor hold different opinions on social policy. The other pertains to over-time changes in these differences. I examine each in turn.

## **Explaining Cross-National Differences in The Income Gradient**

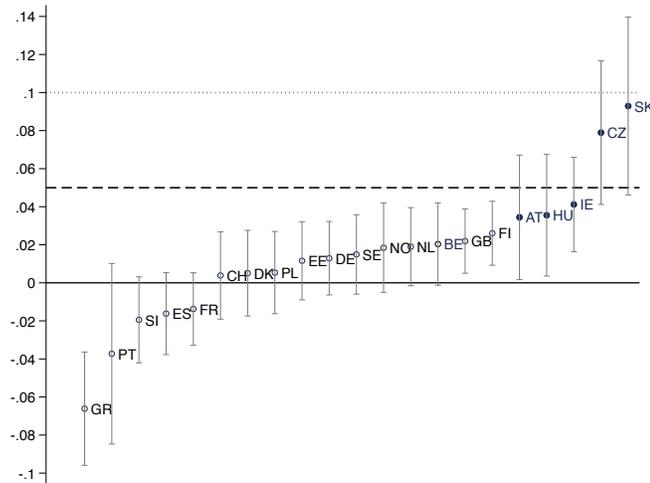
Given that social spending tends to be redistributive, one might reasonably expect attitudes toward taxes and social spending to correlate with income. Specifically, given existing levels of taxation and social spending, support for tax and spending increases (over cuts) should be higher among the poor, and support for tax and spending cuts (over increases) should be higher among the rich. Yet, anyone familiar with European survey data knows that, in most countries, the relationship between income on the one hand and attitudes toward taxation and social spending on the other, is surprisingly flat. Figure 2 illustrates this fact. In most countries, tax and spending attitudes do not substantively vary with income. In Greece, support for cuts in taxes and social spending (over increase) is even *higher* among the poor than it is among the rich. The main exceptions are Austria,

---

<sup>7</sup> In this paper, I focus on people’s attitudes toward social spending broadly defined. To understand how people reason about taxation policies taken in isolation, and policies that address market income inequalities more specifically, a different form of fairness reasoning applies, see Cavaille (Forthcoming) for more detail.

Hungary, Ireland, the Czech Republic and Slovakia, where I cannot reject the null that attitudinal difference between the two groups is at least equal to 0.5 SD.

**Figure 2:** Income Differences in Tax/Spend Attitudes



Coefficients and standard errors are obtained by regressing the tax-spend item described in Figure 1 on income interacted with country fixed effects. The tax-spend item is standardized using the **country-specific** mean and SD. The income variable is a categorical variable designed to capture people's position on the country's income distribution (from the 1st to the 10th income decile). Interpretation: In IE, the predicted difference on the tax-spend response scale between a respondent in the first and a respondent in the 10th income decile is around 0.4 SD.  
*Source:* ESS wave 4, weighted.

The absence of an income gradient illustrates the ascendancy of fairness reasoning: attitudes toward the welfare state are shaped by concerns over free riding (see Figure 1), concerns that are themselves orthogonal to the likelihood of relying on social transfers (see Appendix A.2). This, in turn, explains the small income gradient in Figure 2. Mechanically, a small income gradient has two possible causes. One is higher-than-expected support for cuts (over increases) among cross-pressured low-income individuals saying the fair thing (“cut spending that goes to free riders”) instead of the self-interested thing (“increase spending that benefits me”). Another is higher-than-expected opposition to cuts (over increases) among cross-pressured high-income individuals saying the fair thing (“do not cut spending that goes to deserving recipients”) instead of the self-interested thing (“cut spending, so you can cut my taxes”). If my argument is correct, high-income individuals should contribute more than low-income ones to the flattening of the income gradient.

To understand why this is the case, we can rely on the two-step model described above. When forming an opinion on status quo changing policies, individuals first ask themselves what is the fair

thing to do.<sup>8</sup> People who believe that social spending benefits and encourages free riding will start from a baseline of high support for cuts relative to increases. People who believe beneficiaries are deserving cooperators who need to be treated in kind will start from a baseline of high opposition. In a second step, people ask themselves how much a given policy change will affect them. On average low-income individuals are net beneficiaries of the tax/spend social policy nexus, while high-income individuals are net contributors. For net beneficiaries, tax and spending cuts imply less social benefits. For net contributors, tax and spending cuts imply less taxes. There is a difference between both groups in terms of transparency and stakes. While spending cuts have clear first order effects on the income of the poor, tax cuts' implications, in contrast, are obscured by fiscal policy and the sheer complexity of the welfare state's funding structure. Furthermore, assuming individuals weight credible losses more than uncertain gains (Tversky and Kahneman 1991), then the relative importance of lower benefits is greater for the poor than the relative importance of lower taxes is for the rich (Rueda and Stegmueller 2015: 3). In other words, the redistributive design of the welfare state, the unequal distribution of economic risk and wealth, and the difference between the first order consequence of spending cuts and the second order consequence of tax cuts introduce an asymmetry between the rich and the poor: self-serving adjustment is larger for the former than for the latter.

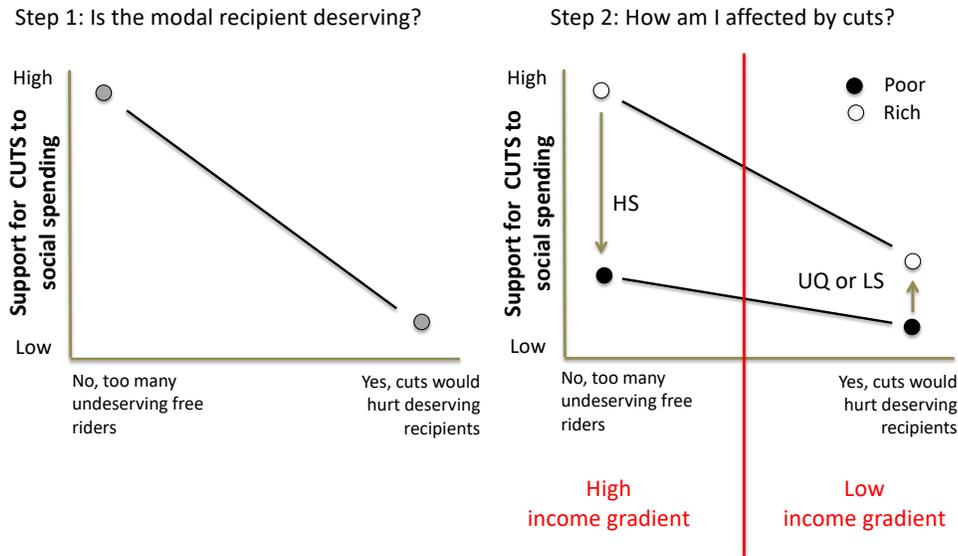
As a result, while the poor's attitudes will reflect their self-interested opposition to spending cuts, the rich's attitudes will reflect what they believe to be fair instead: depending on their free riding beliefs, some will support cuts that punish free riders over increases, and others will oppose the unfair punishment of deserving recipients by supporting tax and spending increases. Opposition to cuts within this group thus follows from the high proportion of respondents who hold optimistic free riding beliefs (e.g., the prevalence of free riding is low). Empirically, this means that, among

---

<sup>8</sup> When asked about taxation and spending, people assess two possible changes to the status quo: cuts in taxes and spending on the one hand, increases on the other. Throughout, I follow the literature, which considers these attitudes on a continuum ranging from strong support for increases to strong support for cuts. When discussing support for cuts, I consequently mean support for cuts over support for increases. As a result, I disregard possible income differences in terms of status quo bias as well as a possible asymmetry between supporting (opposing) cuts and opposing (supporting) increases, though see footnote 16 for a brief discussion of this latter point.

people with optimistic free riding beliefs, there are limited attitudinal differences by income level. Given self-serving adjustment among low-income individuals, this also implies that any evidence of an attitudinal income gap should be limited to the subset of voters with pessimistic free riding beliefs (i.e., the prevalence of free riding is high). This reasoning is sketched in Figure 3.

**Figure 3:** Tax/Spend Attitudes, Free Riding Beliefs and Self-serving Adjustment



HS: high stakes, LS: low stakes, UQ: unknown quantity

The existence of a large proportion of high-income respondents who trust that social spending is done fairly helps explain why the income gradient on the tax-spend item is small “on average.” It does not explain why it varies across countries. Explaining why income is a better predictor in some countries than others amounts to explaining why the income gradient among people with pessimistic free riding beliefs is lower in some countries than others. Such variation might be due to low-income individuals failing to self-servingly adjust and support cuts instead. This would suggest that, in some countries, low-income individuals do not perceive that they benefit from the welfare state (e.g., Holland (2018) for this argument applied to Latin America). Alternatively, high-income individuals might, despite their pessimistic free riding beliefs, oppose cuts and support increases for self-interested reasons. This suggests, more plausibly, that in some countries, social spending benefits all income groups: even high-income individuals fear the consequences

of spending cuts more than the (uncertain) benefits of tax cuts.

If this reasoning is correct, then cross-national differences in the size of the income gradient should be, at least partially, explained by cross-national differences in the distribution of social transfers. Where the probability of relying on publicly-funded transfers is disproportionately higher for the poor, self-serving opposition to cuts among those concerned about free riding will be limited to low-income individuals. Where the probability of becoming a recipient is more evenly distributed among income groups, self-serving opposition to cuts will extend to those higher up the income ladder, explaining why support for cuts (over increases) differs less by income.

What, in turn, explains cross-national differences in how social benefits are distributed in the population? Labor market conditions play an important role: the more unemployment risks are concentrated on the poor, the less high-income workers expect to rely on social transfers aimed at protecting against income shocks (Rehm, Hacker and Schlesinger 2012). Policy design matters in two ways. First, it can exclude middle and high-income groups *a priori*. By definition, means-tested public transfers are limited to the worse off (Korpi and Palme 1998). Second, policy design affects middle and high-income individuals' expectations of one day relying on social benefits, especially benefits targeted to those facing temporary job loss. A key factor is replacement rates, defined as the percentage of past income replaced by social transfers on average. If replacement rates are low, social transfers have income-smoothing properties only for the poor. Middle and high-income individuals who want to insure against the risk of catastrophic income loss (Moene and Wallerstein 2001) will more likely self-insure through the private market or private savings. In contrast, in countries with high replacement rates, high-income individuals will positively value the income smoothing properties of public unemployment insurance and social programs.<sup>9</sup> In other words, because of differences in labor market conditions and policy design, reliance on publicly-funded transfers will be more or less concentrated on the bottom of the income ladder.

---

<sup>9</sup> In addition, overall benefit generosity – which includes benefit duration in addition to replacement rates – directly affect the likelihood of *becoming* a recipient, with more generous unemployment transfers being associated with more time spent unemployed (Chetty 2005; Borghans, Gielen and Luttmer 2014).

To sum up, based on my argument, the small income gradients documented in Figure 2 for the tax/spend item has at least two sources: 1) a large proportion of high-income respondents who trust that social spending is done fairly and consequently oppose cuts and/or support increases out of fairness concerns, 2) self-interested opposition to cuts and/or support for increases among people with pessimistic beliefs about the prevalence of free riding. Cross-national differences in the size of this income gradient can be, at least partially, explained by differences in benefit concentration. Benefit concentration has the most implications for the attitudes of high-income individuals who hold pessimistic free riding beliefs: as benefit concentration decreases, so does their support for cuts and so does the income gradient.

## **Evidence**

To test these expectations, I rely on the 2008 wave of the European Social Survey (ESS) used in Figure 2. My main outcome of interest is the tax and spend item plotted in Figure 1. To measure free riding beliefs, I use the same survey items introduced in Table 1. First, I examine whether reliance on fairness reasoning, i.e., the extent to which tax/spend attitudes correlate with free riding beliefs, increases with income. Specifically, I regress the tax/spend variable on free riding beliefs interacted with income. I run this analysis on the pooled data using a hierarchical linear model with individuals nested in countries. Throughout, I use multilevel regression models with random effects, meaning that, unless noted, the coefficients and intercepts are modeled as random variables.<sup>10</sup>

Column (1) in Table 2 reproduces the basic pattern observed in Figure 2: on average, income does not predict attitudinal differences in tax/spend attitudes. Column (2) shows that individuals who differ in their free riding beliefs by 2 SD also disagree substantively in terms of their tax/spend attitudes: the average difference between the two is 0.43 SD. As shown in Appendix A.5, the

---

<sup>10</sup> Specifically, model parameter are allowed to differ across the twenty countries included in the analysis. Tables report the mean intercepts and coefficients, information on standard deviations are available on request.

**Table 2: Predicting Tax/Spend Attitudes**

	(1)	(2)	(3)
Income [1 to 10]	0.10 (0.01)	0.02* (0.01)	0.02* (0.01)
Free riding beliefs		0.43*** (0.04)	0.22** (0.07)
Free riding Beliefs * Income			0.04** (0.01)
Constant	- 0.08* (0.04)	- 0.11** (0.04)	- 0.10** (0.04)
N	30142	29908	29908

Results from multilevel linear regressions. Outcome variable is standardized using **country-specific** mean and SD. Free riding belief scores are demeaned and divided by 2 standard deviations using **country-specific** mean and SD. Coefficients can consequently be interpreted as a  $\beta$  SD(s) change in support for cuts following a 2 SDs change in free riding beliefs. Income variable ranges from 1 (1st decile) to 10 (10th decile).

Source: ESS 2008.

coefficient on free riding beliefs is not affected by the introduction of socio-economic controls, confirming the prior finding that free riding beliefs are orthogonal to the likelihood of being a recipient of benefit transfers. The model reported in column (3) interacts free riding beliefs with income levels. The coefficient on the interaction term indicates that the correlation between free riding beliefs and support for cuts is more than twice as large among top decile respondents ( $0.22 + 10 * 0.04 = 0.62$ ) than among bottom decile respondents ( $0.22 + 1 * 0.04 = 0.26$ ). Among people with pessimistic free riding beliefs (1 SD above country mean), the difference between top and bottom decile respondents is equal to  $[0.02 + 0.5 * 0.04][10 - 1] = 0.36$  SD, with the former more supportive of cuts than the latter. Among people with optimistic free riding beliefs (1 SD below country mean), there are no attitudinal differences between the rich and the poor.<sup>11</sup>

Figure 4 presents predicted attitudinal differences by free riding beliefs and income group, focusing on the case of Great Britain.<sup>12</sup> As a robustness check, I use a different measure of income that allows me to relax the assumption that income's effect on policy attitudes is linear.<sup>13</sup> Specif-

<sup>11</sup> Specifically,  $[0.02 - 0.5 * 0.04][10 - 1] = 0$ .

<sup>12</sup> Results are similar for all countries in the sample. See Appendix A.6 for the same analysis using Danish data.

<sup>13</sup> Note, however, that predicted attitudes plotted in Figure 4 still rely on the assumption of a linear relationship between tax/spend attitudes and free riding beliefs. This assumption seems plausible in light of empirical patterns plotted in Figure 1.

ically, I recode the income measure in order to better distinguish people in the top and bottom income *quintiles*.<sup>14</sup> In the analysis, the resulting trichotomous variable is dummied out and interacted with the free riding beliefs variable. To facilitate interpretation, I switch to two dichotomous measures of tax/spend preferences. The first one (left panel) identifies individuals who explicitly support tax and spending cuts over increases (i.e, a response category strictly inferior to 5 on the 0/10 scale, see notes in Figure 1 for wording). The other measure (right panel) identifies individuals who explicitly support tax and spending increases over cuts (i.e, a response category strictly superior to 5 on the 0/10 scale).

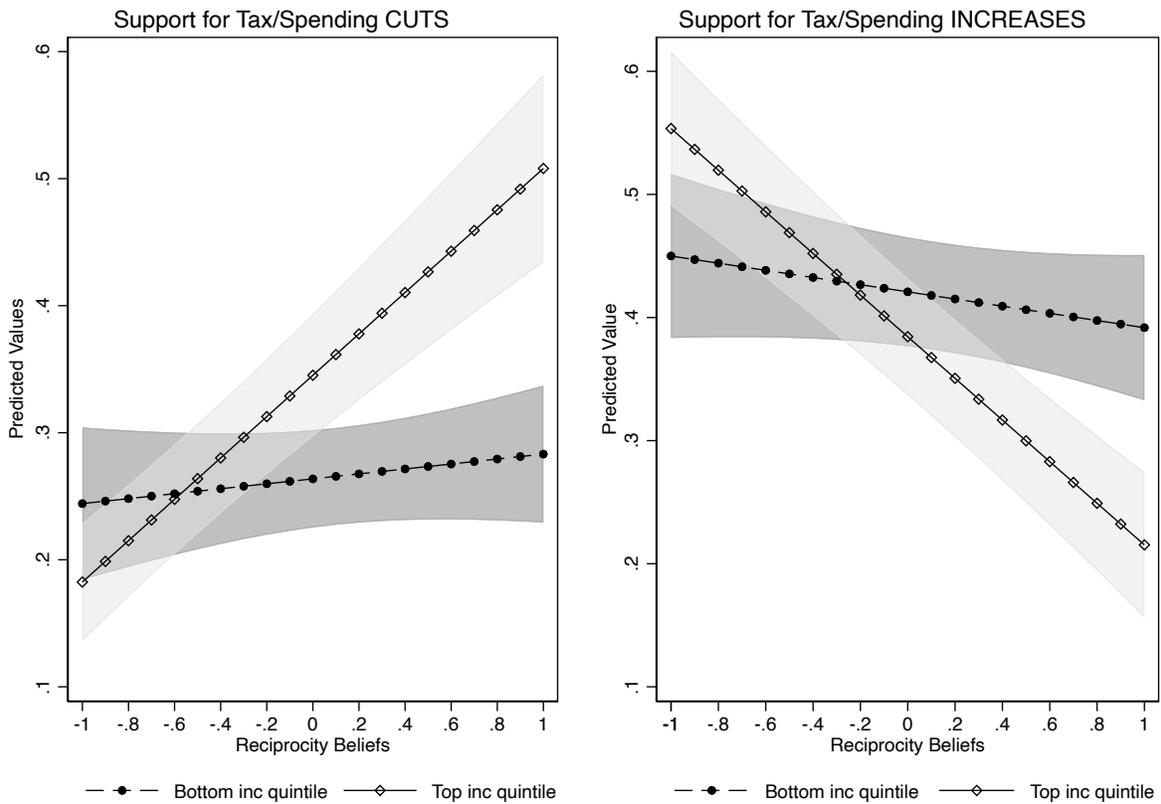
Irrespective of the outcome used, free riding beliefs correlate with tax/spend attitudes more so among the top quintile than among the bottom quintile. This means that income differences align with economic expectations only among people who find the status quo unfair because it benefits free riders (right-side of the figures). Among people who find the system fair (i.e., moral hazard is not a concern and recipients are deserving), there are no income differences (left-side of the figures). This is due to higher than expected – based on their “objective” material self-interest – opposition to tax and spending cuts, alongside higher than expected support for tax and spending increases, on the part of high-income individuals who find the system fair. In other words, it is this group that exerts a downward pressure on the income gradient, explaining the weak correlation documented in Figure 2.

Figure 5 examines how this pattern varies by country. Specifically, it plots the relationship between income and tax/spend attitudes, limiting the sample to individuals whose free riding beliefs score place them in their country’s top quartile (i.e., the subset most likely to find recipients undeserving and be concerned about free riding). For reference, I also plot the original income coefficients plotted in Figure 2 and computed using the full sample. If my framework has enough explanatory power, then the income gradient for this specific subset of individuals should be larger

---

<sup>14</sup> To identify income cut-offs – the P20 and the P80 – I use estimates from the LIS database. See Appendix A.1 for more detail on this variable.

**Figure 4:** Explaining Tax/Spend Attitudes in Great Britain

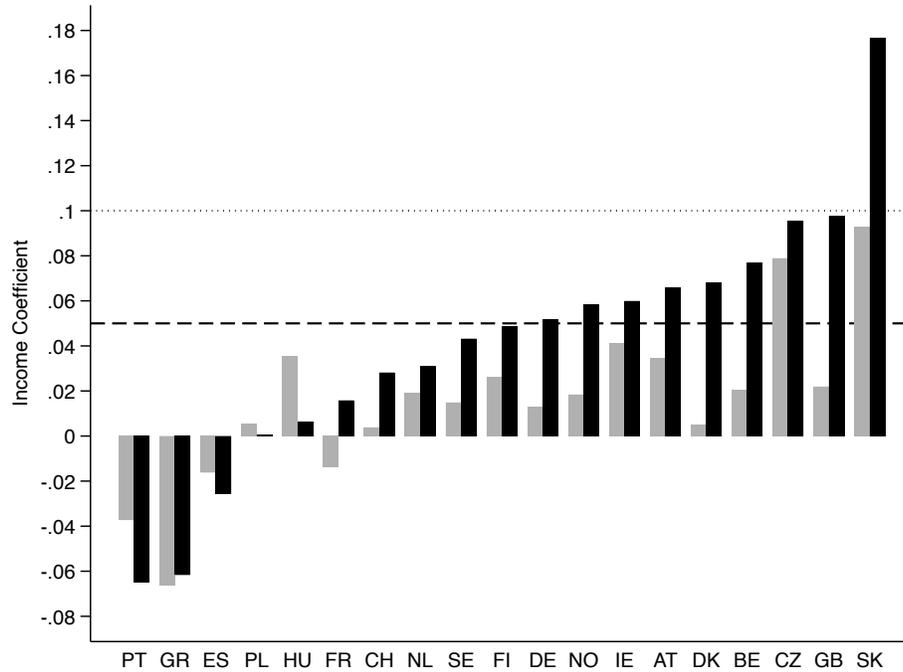


*Left panel:* plots predicted share of respondents, alongside 95% confidence intervals, choosing cuts over increases (i.e. a response category inferior to 5 on the 0/10 scale). *Right panel:* plots predicted share of respondents choosing increases over cuts (i.e. a response category superior to 5 on the 0/10 scale). Binary outcomes were regressed on free riding beliefs and interacted with income dummies (top quintile, bottom quintile, the rest). See Figure 1 for outcome wording. Predicted values are generated using free riding scores running from 1 SD below to 1 SD above the mean.

Source: ESS round 4, weighted

than the income gradient for the full country sample. This is the case for 14 of the countries under study. For 10 countries, attitudinal differences in this subgroup reach or surpass the 0.5 SD benchmark, meaning that tax/spend attitudes between the top and bottom deciles differ by at least half a standard deviation.

**Figure 5: In Search of The Income Gradient**



Coefficients and standard errors are obtained by regressing the tax-spend item described in Figure 1 on income interacted with country fixed effects. Gray bars reproduce the income coefficients plotted in Figure 2. The dark bars represent the same coefficients but limiting the sample to individuals in the top quartile of their country’s reciprocity score distribution. Source: ESS round 4, OECD (2008).

Still, for a little over a third of the countries in this sample, even within the subset of people with pessimistic free riding beliefs, the income gradient remains small or negative. A possible reason, I argued, is that social transfers reach higher up the income ladder in some countries than in others. This would suggest that, in countries to the left in Figure 5, income remains a poor predictor because even high-income individuals who are concerned about free riding benefit from social spending and adjust their preferences self-servingly.

To test this argument, I rely on data provided by the OECD (2008) on benefit concentration. The OECD’s research team has computed a measure similar to a Gini coefficient, which captures

**Table 3: Benefit Concentration and the Size of the Income Gradient**

Sample used depending on free riding beliefs score quartile	(1) Full sample	(2) Score $\leq 20^{th}$ perct.	(3) Score $\leq$ country av.	(4) Score $>$ country av.	(5) Score $\geq 80^{th}$ perct.
Income (10 deciles)	0.03*** (0.00)	0.01 (0.01)	0.02** (0.01)	0.05*** (0.01)	0.07*** (0.01)
Benef. concentration <small>High (0) to low (1)</small>	0.20*** (0.05)	0.14 (0.14)	0.27** (0.09)	0.22* (0.09)	0.34** (0.13)
Income * Benef. con.	-0.04*** (0.01)	-0.01 (0.02)	-0.04** (0.01)	-0.06*** (0.01)	-0.09*** (0.02)
Constant	-0.18*** (0.02)	-0.40*** (0.06)	-0.34*** (0.04)	-0.09* (0.04)	-0.09 (0.06)
N	27885	7158	13958	13707	6946

Results from multilevel linear regressions. Outcome variable is standardized using country-specific mean and SD. Income variable categorizes respondents by income deciles. Coefficient on income can be interpreted as a  $\beta_I * 10$  standard deviation(s) change in in tax/spend attitudes following a change from the bottom to the top decile. Benefit concentration variable is normalized: coefficient can be interpreted as a  $\beta_{BC}$  standard deviation(s) change in tax/spend attitudes following a change from the highest to the lowest benefit concentration value.

Source: ESS round 4, unweighted

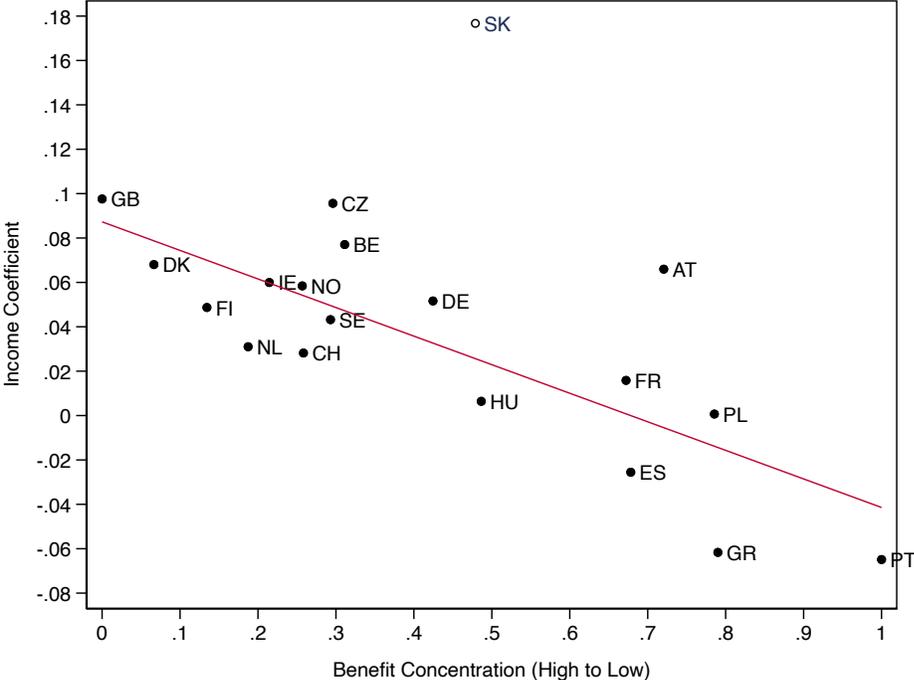
the differences between a group's share of the population and its share of all the cash transfers going to individuals of working age in a given year. A value of zero indicates that all income groups (ranked according to their disposable income) receive an equal share of all cash transfers. A negative coefficient indicates that lower income groups receive a higher share of transfers than their share of the population.<sup>15</sup> I normalize this variable such that 0 captures the smallest possible value (high concentration) and 1 the highest possible value (low concentration).

I examine whether the income gradient among individuals most concerned about free riding decreases as benefit concentration decreases. Table 3 reports regression results from a hierarchical linear model with a two-level interaction between income (individual-level) and benefit concentration (country-level). Column (1) reports results for the full sample. Column (2) through (5) report results for different sub-samples of the data, starting with people who hold the most optimistic free riding beliefs (with scores below the 20th percentile in their country) and ending with people who hold the most pessimistic free riding beliefs (with scores above the 80th percentile in their country). Notice how the absolute value of the interaction term between income and benefit

<sup>15</sup> In Appendix A.4, I detail a set of checks run to assess the quality of this measure for capturing cross-national differences in benefit concentration.

concentration increases from left (column (2)) to right (column (5)). This suggests that changes in the income gradient tied to cross-national differences in benefit concentration are mostly due to changes in the tax/spend preferences of people with the most pessimistic free riding beliefs. Figure 6 presents this evidence visually. First, for each country, I compute the income gradient for people with a free riding score equal or above the 80th percentile in their country (the dark bars in Figure 5). I then plot these estimates against the benefit concentration measures. As Figure 6 shows, when benefit concentration decreases, so does the size of the income gradient, with Slovakia as a notable outlier. This aligns with results in column (5).

**Figure 6: Income Gradient and Benefit Concentration**



Plots income coefficients (the dark bars in Figure 5) against benefit concentration measures.  
 Source: ESS round 4, OECD (2008).

My argument emphasizes the relationship between benefit concentration and the extent to which high-income individuals rely on fairness reasoning when asked about increasing or decreasing existing levels of taxation and social spending. This implies that, underpinning the income results described above, is a positive relationship between benefit concentration and the correlation between beliefs and preferences: as benefit concentration decreases, so does the correlation between

free riding beliefs and tax/spend attitudes. This decline should be due to a decline in support for cuts among high-income individuals with pessimist free riding beliefs. As shown in Appendix A.4, this is indeed the case. Note that my argument does not imply that benefit concentration affect average free riding beliefs: as I also show in Appendix A.4, countries with more concentrated benefits are not be countries with more pessimistic free riding beliefs.

## **Robustness Checks**

In Appendix A.7, I address two important confounders, namely cognitive capacity and ideological consistency. One might hypothesize that the relationship between income on the one hand, and the correlation between free riding and tax/spend attitudes on the other, is an artifact of high income individuals being more likely to associate the correct beliefs to a given set of social policy preferences. To address this issue, I look for systematic differences in patterns of answers to questions about free riding among the rich and among the poor. Lower inter-item correlation would indicate lower levels of “coherence” with implications for the results presented above. As documented in the appendix, I find not evidence of systematic differences, be it in the pooled data or in a country-by-country analysis.

Next, I examine whether the results discussed above might be an artifact of compositional differences in ideological consistency, with high-income individuals being better able to connect the correct proto-ideological belief to the correct policy preference due to a better knowledge of partisan ideological constructs. To probe this, I turn to a subset of respondents who, based on this alternative hypothesis, should be the least likely to exhibit a response pattern consistent with self-serving adjustment, namely people who hold pessimistic free riding beliefs (1 SD above country mean) and identify as right-wing (choose 7 or more on a 1 to 10 left-right scale). Despite members of this group being ideologically right-wing, attitudinal differences between top and bottom quintile respondents remain large at 0.4 SD. This pattern suggest, that even within this group of “ideologues”, self-serving adjustment is occurring.

More generally, if we assume that lower cognitive capacity and ideological consistency explain income differences in reliance on fairness reasoning, then one would expect the mismatch between free riding beliefs and tax-spend preferences to be evenly distributed among all poor. Specifically, one would expect the “error rate” to be the same among low-income individuals with optimistic beliefs and low-income individuals with pessimistic beliefs. Instead, I find that, in line with my argument, it is larger among the latter than among the former group. Furthermore, arguments emphasizing differences in cognitive capacity and ideological consistency have little to say about the cross-national results presented in the previous analysis.

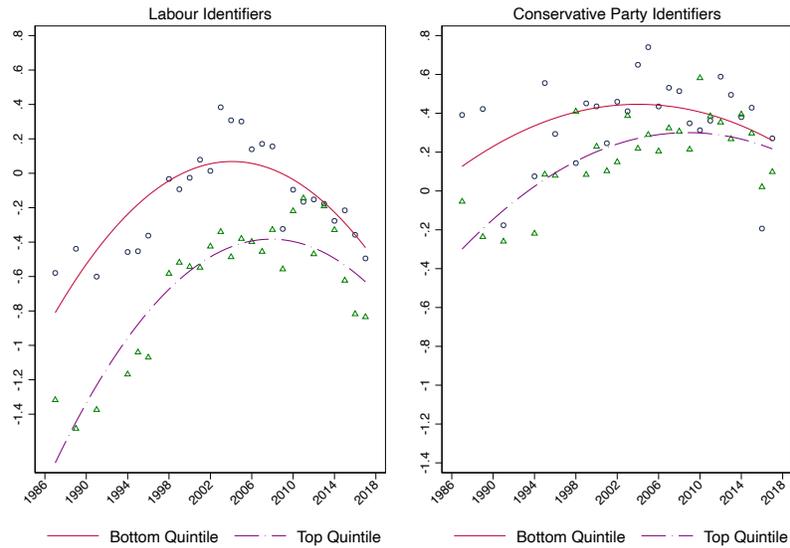
Next, I discuss the framework’s implications for understanding over time changes in attitudinal differences between the rich and the poor. I focus on the case of Great Britain, a country where mass beliefs about the prevalence of free riding have famously taken a pessimistic turn (Cavaille and Trump 2015; O’Grady 2021; Larsen 2013).

## **Explaining Over Time Differences in The Income Gradient**

In Great Britain, mass beliefs regarding the prevalence of free riding have changed dramatically since the 1980s. Figure 7 plots average free riding beliefs according to one’s income and partisan identification. To compute free riding beliefs, I use survey items similar to those listed in Table 1 (see Appendix B.1 for a full list of the items). As shown in Figure 7, the pessimistic shift in free riding beliefs has affect all income groups and partisan families. Notice the absence of income differences with regards to free riding beliefs. If anything, low-income individuals are more negative than high-income individuals when it comes to expressing concerns that generous income replacement policies unfairly reward free riders.

As discussed in the previous sections, income should help predict the extent to which a pessimistic shift in free riding beliefs will affect tax/spending attitudes. This should be especially true when asked specifically about spending on policies that provide income support to the poor and

**Figure 7: Free Riding Beliefs By Partisanship and Income**

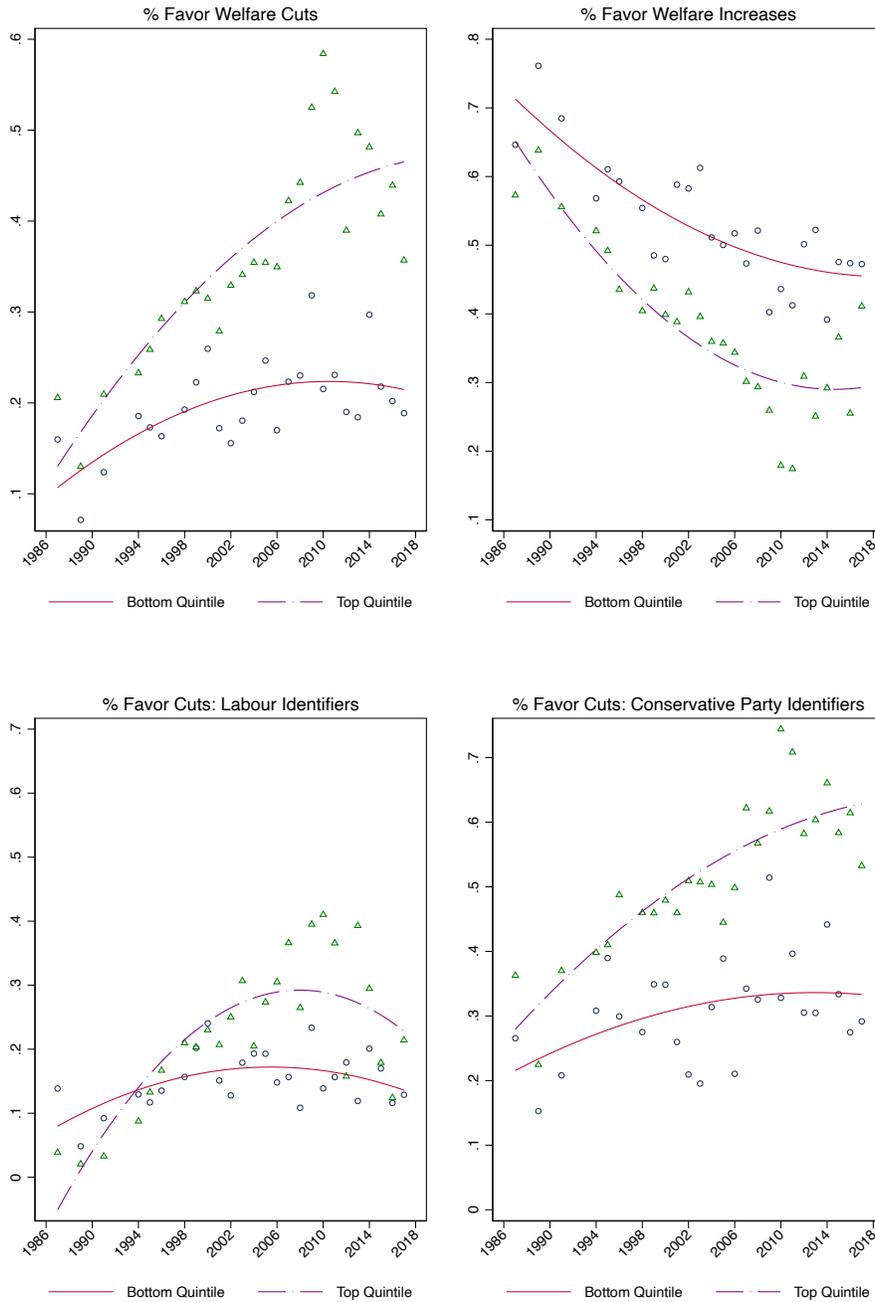


Plots changes in free riding beliefs by partisanship and income level. Free riding items were combined into a free riding score using weights recovered from an exploratory factor analysis. Higher scores indicate more pessimistic / right-wing free riding beliefs. See Appendices B.1 and B.2 for more information on item wording and on the income variable. Source: BSAS longitudinal dataset, weighted.

the unemployed. Indeed, as I argued above, the more benefits are concentrated on the worse off, the more high-income individuals are likely to rely on fairness reasoning. I consequently, expect low-income individuals to maintain high self-interested opposition to cuts in targeted social benefits, this *despite* changing free riding beliefs. In contrast, among high-income individuals, support for cuts should increase in line with changing free riding beliefs. The result should be a growing divergence between the attitudes of the rich and the poor on the issue of spending on welfare benefits.

Figure 8 examines this expectation. It plots support for cuts to welfare benefits, as well as support for increases by income groups (top panel). To better compare with Figure 7, I also split the analysis of support for cuts according to partisanship (bottom panel). In line with expectations, low-income individuals, despite a pessimistic shift in free riding beliefs, do not increase their support for welfare cuts. The observed decline in support for welfare increases is smaller among low-income individuals than high-income ones. In contrast, the shift in free riding beliefs is fully reflected in the welfare attitudes of high-income individuals, irrespective of the response category

**Figure 8: Attitudes Toward Welfare Cuts By Income**



Plots changes in tax/spend welfare preferences by partisanship and income level. Item wording: “Government should spend more money on welfare benefits for the poor, even if it leads to higher taxes.” *Top-left* and *bottom* figures plot the share of respondents who chose “Disagree strongly/Disagree.” *Top-right* figure plots the share who choose “Agree strongly/agree.” Item wordings: See Appendix B.2 for more information on the income variable.

Source: BSAS longitudinal dataset, weighted.

under scrutiny (cuts or increases). One consequence is the emergence of an income gradient (top panel): while there were no income differences in the late 1980s, today, the attitudinal gap in support for cuts is equal to 30 percentage points. In other words, while changes in fairness beliefs explain increasing support for benefit cuts and declining support for benefit increases, material self-interest still matters. In this case, it manifests itself as a deviation from what fairness reasoning prescribes, resulting in a significant income gap, especially with regards to support for cuts to welfare.<sup>16</sup>

---

<sup>16</sup> Splitting the analysis between support for cuts on the one hand, and support for increases on the other, reveals an interesting asymmetry: for low-income individuals, the impact of belief change appears stronger for increases (a decline in 20% points, versus 40% points for high-income individuals) than for cuts (an increase in 10% points only, versus 40% points for high-income individuals). Based on my argument, one might hypothesize that the threat of an income loss provokes a larger self-serving adjustment than the promise of an income gain. Such asymmetry might also emerge from a context in which fiscal stress makes cuts more credible than an increase. Absent experimental manipulations of the latter, this interpretation remains tentative.

## Discussion

For most people, most of the time, redistributive social policies' implications for one's own bottom line are unclear, making fairness reasoning the default (and anchoring) mode of reasoning. Yet, there are exceptions depending on policy design and its interaction with one's own socioeconomic conditions. In line with expectations, I found that the income gradient is the largest among people with pessimistic free riding beliefs, especially so in countries where social benefits are most concentrated on the worse off. Using this simple framework, I was also able to show that moral reasoning tied to free riding concerns on the one hand, and material self-interest on the other, combine in predictable ways to shape changes in mass social policy preferences. Simply put, the extent to which belief change implies attitudinal change will vary across individuals depending on how much is at stake for them and how readily available this information is. One observable implication is a growing income gradient in Great Britain, especially with regards to welfare cuts, which is a low-stakes moral issue for high-income individuals and a high-stakes pocketbook issue for low-income respondents.

Scholars have repeatedly argued that, without close attention to moral reasoning, important aspects of mass social policy preferences will remain unaccounted for. For example, according to Fong, Bowles and Gintis (2006) , “understanding egalitarian politics today requires a reconsideration of *Homo Economicus*, the unremittingly self-regarding actor of economic theory (...). conditional cooperation and punishment better explains the motivations behind support for the welfare state” (page 6). This paper provides one of the first systematic cross-national extension of the mainly US-centric research that provides the empirical bases for this claim. As this paper reveals, free riding concerns play an important structuring role and their omission comes at a cost. Still, I have also shown that, without material self-interest, researchers cannot explain why the predictive power of free riding beliefs varies across socio-economic groups and countries.

The framework presented in this paper has explanatory power beyond mass attitudes. It also sheds a new light on the coalitions behind social policy retrenchment in the “Age of Austerity”

(Pierson 2001). In many European countries, budgetary constraints and high levels of public debt following the Great Recession have rekindled the debate over the size of the welfare state. Mainstream models predict that high-income groups will be more likely to turn against a bankrupt welfare state for fear of having to foot the bill. The framework presented here predicts strong heterogeneity among the rich: because of uncertainty over who exactly will be footing this bill, high-income “bleeding heart” liberals, who believe the system to be fair and recipients deserving, are central to coalitions opposing retrenchment. The pivotality of bleeding heart liberals to anti-retrenchment coalitions varies with benefit concentration. The more benefits are evenly spread across the income distribution, the more high-income individuals concerned about free riding will perceive cuts as effecting them personally and oppose retrenchment. When transfers are targeted to the least well-off, high-income individuals who are concerned about free riding are very likely to defect and support retrenchment, increasing the importance of bleeding heart liberals in limiting cuts to social spending. In such a context, beliefs regarding the ubiquity of free riding are more likely to permeate public debates on welfare state reform.

Differences in the politics of health care reform in Great Britain and the United States help illustrate this point. Both countries are known for an extensive use of the deservingness and moral hazard rhetoric. Their populations hold, comparatively to other western countries, negative priors about the deservingness of the poor (Svallfors 2012). In Great Britain, debates over the privatization of segments of the National Health Service (NHS), have made no references to deservingness or moral hazard. The universal design of the NHS results in a strong self-interested support for the status quo. In contrast, the coalition behind Obamacare is a mix of self-interested low-income voters hoping to get access to year-around health insurance and high-income individuals who perceive the lack of insurance among the poor as a form of injustice. Opposition is the highest among people who are not directly affected by the reform and worry it mostly benefits incentivizes free riding (Skocpol and Williamson 2011).

I have highlighted the role of bleeding heart liberals in supporting social transfers to the worse off. This support is partly an artifact of uncertainty over the costs and benefits of their preferred

policies. This uncertainty is not a fixed attribute of the policy landscape. Take for example tax increases following a fiscal adjustment. When governments implement such taxes, they make it easier for high-income respondents to connect generous social spending to their own tax bill. In this case, one would expect high-income respondents with optimistic free riding beliefs to be more likely to self-servingly adjust and break away from fairness-induced opposition to tax and spending cuts. While low-income voters, believed to vote against their interest to “punish” undeserving recipients, have attracted most of the attention, the framework and findings presented in this paper suggest that bleeding heart liberals require as much, if not more, attention from researchers.

Beyond the study of redistributive politics, my argument speaks to studies on opinion formation in American Politics, which argue that most people form policy preferences using value-based or socio-tropic reasoning that is orthogonal to pocketbook reasoning (Sears et al. 1980; Sears and Funk 1990; Achen and Bartels 2006; Berinsky 2011). These studies have repeatedly shown that “respondents in mass opinion surveys” seem to “care less about their own personal (and material) stakes in policies than about whether those policies promote national welfare or serve longstanding values” (Chong, Citrin and Conley 2001*b*). In the words of Sears and Funk: “the personal and the political exist independently of each other in two very different cognitive worlds.” The argument presented in this paper advances this literature in at least two ways.

First, while I take stock of the well-documented disjuncture between policy preferences and pocketbook reasoning, I do not assume it is a fixed feature of opinion formation. Instead, I conceive of this disjuncture as a continuum in need of theorizing. Social spending, because it affects respondents’ income differently, represents higher stakes for some respondents than for others. To the extent that institutional contexts and policy design affect individual-level stakes, this disjuncture will not only vary systematically across socio-economic groups but also across countries. This approach, I showed, helps explain puzzling cross-national differences and longitudinal changes in the relationship between income and tax-spend attitudes.

Second, I draw from research in behavioral economics and evolutionary psychology to provide

a more precise definition of an important sub-set of “longstanding values” central to how people reason about social spending. Most people, I argued, agree that social spending should not reward free riding. Where people disagree is over the extent to which this is actually the case. As a result, differences in social policy preferences can be at least partially explained by people’s free riding beliefs. This definition of value-based reasoning is amenable to comparative research. Future studies, for example, will benefit from developing improved measures of free riding beliefs, which can then be used to leverage both within and cross-national variation to unpack how free riding beliefs form.

## References

- Achen, C.H. and L.M. Bartels. 2006. It feels like we are thinking: The rationalizing voter and electoral democracy. In *Annual Meeting of the American Political Science Association, Philadelphia*.
- Akerlof, G.A. 1970. "The market for "lemons": Quality uncertainty and the market mechanism." *The quarterly journal of economics* (1):488–500.
- Barr, Nicholas Adrian. 1998. *The economics of the welfare state*. Stanford university press.
- Beramendi, Pablo and Philipp Rehm. 2016. "Who gives, who gains? Progressivity and Preferences." *Comparative Political Studies* 49(4):529–563.
- Berinsky, Adam J. 2011. *New directions in public opinion*. Taylor & Francis.
- Borghans, Lex, Anne C Gielen and Erzo FP Luttmer. 2014. "Social support substitution and the earnings rebound: Evidence from a regression discontinuity in disability insurance reform." *American Economic Journal: Economic Policy* 6(4).
- Bowles, Samuel and Herbert Gintis. 2011. *A cooperative species: Human reciprocity and its evolution*. Princeton University Press.
- British Social Attitudes Survey, longitudinal survey. N.d. Cumulative file compiled by the authors NatCen Social Research [data collection]. UK Data Service [distributor]. <https://ukdataservice.ac.uk>.
- Cavaille, Charlotte. Forthcoming. *Fair Enough? Demand for Redistribution in the Age of Inequality*. Cambridge University Press.
- Cavaille, Charlotte and Kris-Stella Trump. 2015. "The Two Facets of Social Policy Preferences." *Journal of Politics* 77(1):146–160.
- Charness, Gary and Matthew Rabin. 2002. "Understanding social preferences with simple tests." *The Quarterly Journal of Economics* 117(3):817–869.
- Chetty, Raj. 2005. Why do unemployment benefits raise unemployment durations? Moral hazard vs. liquidity. Technical report National Bureau of Economic Research.
- Chong, Dennis, Jack Citrin and Patricia Conley. 2001a. "When Self-Interest Matters." *Political Psychology* 22(3):541–570.
- Chong, Dennis, Jack Citrin and Patricia Conley. 2001b. "When Self-Interest Matters." *Political Psychology*

22(3):541–570.

Cramer, Katherine J. 2016. *The Politics of Resentment: Rural Consciousness in Wisconsin and the Rise of Scott Walker*. University of Chicago Press.

Downs, Anthony. 1957. “An economic theory of democracy.”

European Social Survey, Round 4 Data. 2008. Data file edition 4.5. Sikt - Norwegian Agency for Shared Services in Education and Research, Norway – Data Archive and distributor of ESS data for ESS ERIC. <http://dx.doi.org/10.21338/NSD-ESS4-2008>.

European Social Survey, Round 8 Data. 2016. Data file edition 2.2. Sikt - Norwegian Agency for Shared Services in Education and Research, Norway – Data Archive and distributor of ESS data for ESS ERIC. <http://dx.doi.org/10.21338/NSD-ESS8-2016>.

Farber, Henry S, Daniel Herbst, Ilyana Kuziemko and Suresh Naidu. 2018. Unions and inequality over the twentieth century: New evidence from survey data. Technical report National Bureau of Economic Research.

Fischbacher, Urs and Simon Gächter. 2010. “Social preferences, beliefs, and the dynamics of free riding in public goods experiments.” *American economic review* 100(1):541–56.

Fischbacher, Urs, Simon Gächter and Ernst Fehr. 2001. “Are people conditionally cooperative? Evidence from a public goods experiment.” *Economics Letters* 71(3):397–404.

Fong, Christina. 2001. “Social preferences, self-interest, and the demand for redistribution.” *Journal of Public Economics* 82(2):225–246.

Fong, Christina M, Samuel Bowles and Herbert Gintis. 2006. “Strong reciprocity and the welfare state.” *Handbook of the Economics of Giving, Altruism and Reciprocity* 2:1439–1464.

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

Gilens, Martin. 1999. *Why Americans hate welfare: Race, media, and the politics of antipoverty policy*. University of Chicago Press.

Graham, Jesse, Jonathan Haidt and Brian A Nosek. 2009. “Liberals and conservatives rely on different sets of moral foundations.” *Journal of personality and social psychology* 96(5):1029.

- Häusermann, Silja and Hanspeter Kriesi. 2011. What do voters want? Dimensions and configurations in individual-level preferences and party choice. In *Conference on "The Future of Democratic Capitalism"*, Zurich.
- Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis and Richard McElreath. 2001. "In search of homo economicus: behavioral experiments in 15 small-scale societies." *American Economic Review* pp. 73–78.
- Holland, Alisha C. 2018. "Diminished expectations: Redistributive preferences in truncated welfare states." *World Politics* 70(4):555–594.
- Iversen, T. and D. Soskice. 2001. "An asset theory of social policy preferences." *American Political Science Review* 95(4):875–894.
- Korpi, Walter. 2006. "Power resources and employer-centered approaches in explanations of welfare states and varieties of capitalism: Protagonists, consenters, and antagonists." *World Politics* 58(02):167–206.
- Korpi, Walter and Joakim Palme. 1998. "The paradox of redistribution and strategies of equality: Welfare state institutions, inequality, and poverty in the Western countries." *American sociological review* pp. 661–687.
- Lamont, Michèle and Virag Molnar. 2002. "The study of boundaries in the social sciences." *Annual review of sociology* 28:167–195.
- Larsen, Christian Albrekt. 2008. "The institutional logic of welfare attitudes how welfare regimes influence public support." *Comparative Political Studies* 41(2):145–168.
- Larsen, Christian Albrekt. 2013. *The Rise and Fall of Social Cohesion: The Construction and Deconstruction of Social Trust in the US, UK, Sweden and Denmark*. Oxford University Press.
- Margalit, Yotam. 2013. "Explaining social policy preferences: Evidence from the Great Recession." *American Political Science Review* 107(01):80–103.
- Mau, Steffen. 2004. *The moral economy of welfare states: Britain and Germany compared*. Routledge.
- Moene, K.O. and M. Wallerstein. 2001. "Inequality, social insurance, and redistribution." *American Political Science Review* 95(04):859–874.
- OECD. 2008. *Growing Unequal?: Income Distribution and Poverty in OECD Countries*. Organization for

Economic Cooperation and Development.

- O'Grady, Thom. 2021. *The Transformation of British Welfare Policy: Politics, Discourse and Public Opinion*. Oxford University Press.
- Ostrom, Elinor. 1998. "A behavioral approach to the rational choice theory of collective action: Presidential address, American Political Science Association, 1997." *American Political Science Review* pp. 1–22.
- Ostrom, Elinor and James Walker. 2003. *Trust and reciprocity: Interdisciplinary lessons for experimental research*. Russell Sage Foundation.
- Petersen, Michael Bang. 2012. "Social welfare as Small-Scale help: Evolutionary psychology and the deservingness heuristic." *American Journal of Political Science* 56(1):1–16.
- Petersen, Michael Bang, Daniel Sznycer, Leda Cosmides and John Tooby. 2012. "Who deserves help? evolutionary psychology, social emotions, and public opinion about welfare." *Political psychology* 33(3):395–418.
- Petersen, Michael Bang, Rune Slothuus, Rune Stubager and Lise Togeby. 2011. "Deservingness versus values in public opinion on welfare: The automaticity of the deservingness heuristic." *European Journal of Political Research* 50(1):24–52.
- Pierson, Paul. 2001. *The new politics of the welfare state*. Oxford University Press, USA.
- Rehm, P., J.S. Hacker and M. Schlesinger. 2012. "Insecure alliances: Risk, inequality, and support for the welfare state." *American Political Science Review* 106(2):386–406.
- Roth, Christopher, Sonja Settele and Johannes Wohlfart. 2020. "Risk Exposure and Acquisition of Macroeconomic Information."
- Rueda, David and Daniel Stegmueller. 2015. "The externalities of inequality: Fear of crime and preferences for redistribution in Western Europe." *American Journal of Political Science* .
- Rueda, David and Daniel Stegmueller. 2019. *Who Wants What?: Redistribution Preferences in Comparative Perspective*. Cambridge University Press.
- Schumpeter, Joseph A. 1950. *Capitalism, Socialism, and Democracy*. 3d Ed. NY, Harper.
- Sears, David O, Richard R Lau, Tom R Tyler and Harris M Allen Jr. 1980. "Self-interest vs. symbolic politics in policy attitudes and presidential voting." *American Political Science Review* 74(3):670–684.

- Sears, D.O. and C.L. Funk. 1990. "The limited effect of economic self-interest on the political attitudes of the mass public." *Journal of Behavioral Economics* 19(3):247–271.
- Skitka, Linda J and Philip E Tetlock. 1993. "Providing public assistance: Cognitive and motivational processes underlying liberal and conservative policy preferences." *Journal of Personality and Social Psychology* 65(6):1205.
- Skocpol, Theda and Vanessa Williamson. 2011. *The Tea Party and the remaking of Republican conservatism*. Oxford University Press.
- Svallfors, Stefan. 2012. "Welfare Attitudes in Context." *Contested Welfare States: Welfare Attitudes in Europe and Beyond* p. 222.
- Tversky, Amos and Daniel Kahneman. 1991. "Loss aversion in riskless choice: A reference-dependent model." *The quarterly journal of economics* pp. 1039–1061.
- Wallerstein, Michael and Bruce Western. 2000. "Unions in decline? What has changed and why." *Annual Review of Political Science* 3(1):355–377.