

This is a section of [doi:10.7551/mitpress/14496.001.0001](https://doi.org/10.7551/mitpress/14496.001.0001)

Undue Hate

A Behavioral Economic Analysis of Hostile Polarization in US Politics and Beyond

By: Daniel F. Stone

Citation:

Undue Hate: A Behavioral Economic Analysis of Hostile Polarization in US Politics and Beyond

By: Daniel F. Stone

DOI: [10.7551/mitpress/14496.001.0001](https://doi.org/10.7551/mitpress/14496.001.0001)

ISBN (electronic): 9780262372367

Publisher: The MIT Press

Published: 2023



The MIT Press

4 Tastes and Truth

De gustibus non est disputandum.

(“There’s no disputing tastes,” Latin phrase)

And you say to me, friends, there is no disputing over tastes and tasting? But all of life is a dispute over taste and tasting!

—Friedrich Nietzsche, *Thus Spake Zarathustra*

According to the definition of affective polarization bias I’ve proposed, you can’t be wrong about *what* to judge people on. You can only be wrong about *how* you correspondingly judge them. Maybe there are criteria that we objectively should or shouldn’t use to make character judgments. But I don’t claim to know what these are. As a result, my definition of this bias could be “conservative” in the sense that it only captures some, and not all, ways that we may mistakenly dislike other people.

Distinguishing the subjective from the objective is often practically impossible. Perhaps that’s why we sometimes say “de gustibus non est disputandum” (“there’s no disputing tastes”). In the second epigraph quote above, Nietzsche points out that this phrase is empirically inaccurate—that we often do dispute tastes. But *de gustibus* is, I think, more usefully thought of as suggestion than description: when we disagree on matters of taste, we typically won’t be able to figure out whose tastes are “better,” so there’s often no sense in fighting about it. (My definition of affective polarization bias follows this advice.)

On the other hand, it’s certainly true, as Nietzsche notes, that many of us don’t shy away from disputing tastes. This chapter discusses two theories for why such disputes over tastes can in fact often cause undue hostility.

Moral Taste Buds

Jonathan Haidt's book on polarization, *The Righteous Mind*, is already widely considered seminal, though it was published just a decade ago. One of Haidt's key points is that Democrats and Republicans typically prioritize different moral values, or what he calls *moral foundations*. He discusses five main foundations in the book: fairness, care, loyalty, authority, and purity. Liberals value the first two higher than the others, and conservatives place approximately equal weight on all five. See figure 4.1 for a depiction of how the importance of each foundation corresponds to ideology.

Haidt famously compares moral foundations to taste buds. He argues that, in the same way that cuisines vary across cultures, preferences for different moral foundations vary systematically by culture. But just as cuisines must all please some of the same taste receptors that all humans share, moral codes must please at least some of our shared moral foundations.

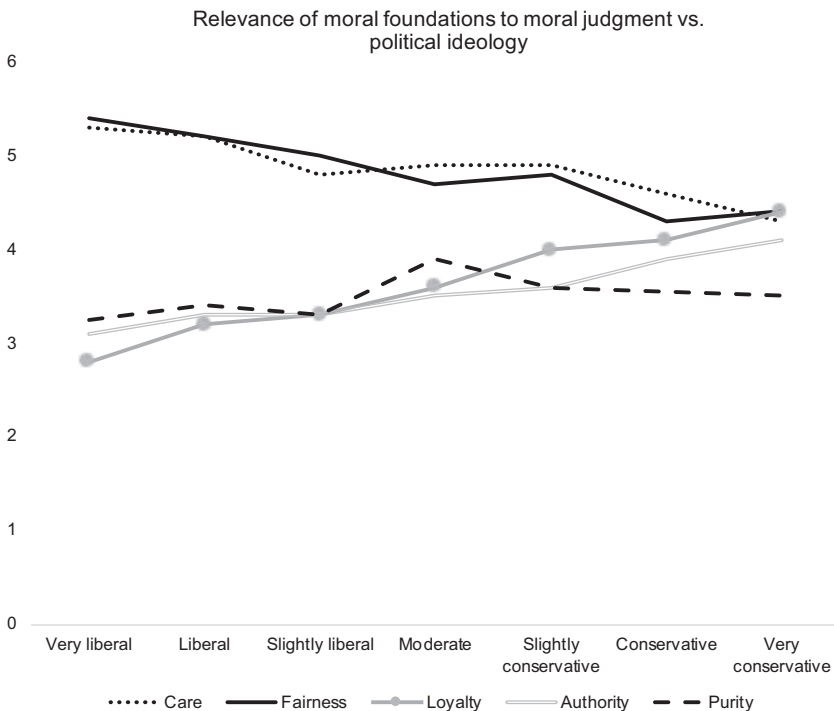


Figure 4.1

Moral foundations versus ideology. Source: adapted from Haidt (2007).

Moral foundations theory has received a lot of attention and a fair amount of criticism since Haidt's book was published (see, e.g., Hatemi et al., 2019; Curry, 2019). But the basic concept—that members of the two major parties tend to have different basic moral values and these values are as much a part of us as are our tastes in food and are thus nearly immutable for the near-term future—makes sense to most of us intuitively and has been supported by subsequent research (Kivikangas et al., 2021; Isler et al., 2021).

An important piece of support for moral foundations theory comes from behavioral economist Ben Enke (2020a) in his paper “Moral Values and Voting.” Enke argues that, consistent with Haidt's work, moral values are a key difference between Democrats and Republicans and this difference has indeed grown over recent decades. However, contra *The Righteous Mind*, Enke argues these differences largely boil down to a single moral dimension, universalism versus communitarianism. He defines universalism as “an ethic of universal human concern . . . irrespective of the context or identity of the people involved,” whereas communitarianism is “loyalty to the local community . . . tied to certain relationships or groups” (3680). Enke conducts a text analysis of congressional and presidential candidates' speeches since the 1960s and finds that Democrats use more universalist language than Republicans, to an increasingly greater degree over time. He also presents survey results indicating voters who embraced more universalist values were more likely to vote Democrat.

For the purposes of this chapter, the number of moral foundations doesn't matter as much as the metaphor comparing moral values to tastes. The expression “there's no disputing tastes,” interpreted as advice, then implies that we should tone down our culture wars and related political fights. But, again, the fact that we need to be *advised* not to dispute tastes underscores the fact that we so often *do* dispute them. Sometimes we even dispute literal matters of tastes—which types of food taste better than others. And these disputes can go beyond friendly debates, evoking real character judgments and even dislike. Why?

Naive Realism and False Consensus

Haidt doesn't explicitly answer this question in his book. Much of the book is ultimately about undue dislike—after all, its subtitle is “Why Good People Are Divided by Politics and Religion”—but he doesn't directly discuss mechanisms for why having different moral taste buds would make

us dislike other people more than we should. Similarly, some strands of the political science literature argue that policy disagreements and differences in moral convictions contribute to affective polarization, in addition to group identity differences. However, these papers also typically don't discuss underlying cognitive mechanisms.¹ It's very intuitive that such disagreements tend to be disagreeable (as I've already noted more than once in this book), so maybe that's why the topic isn't often analyzed explicitly. However, it's a fundamental underlying issue, so it's worth digging into to try to understand more clearly.

Before proceeding, it will be helpful to first provide a more precise definition of *subjective* versus *objective* for the context of this chapter. Definitions of these terms might be controversial in some disciplines (see, e.g., Gaukroger, 2012), but in economics there's a straightforward distinction between two closely related terms, *horizontal* and *vertical product characteristics*. Horizontal characteristics are those for which fully informed consumers have different preferences—some consumers like those characteristics in a good and some dislike them. Vertical characteristics are things that consumers all prefer more of. Location is a canonical example of a horizontal characteristic; consumers who live on the east side of town prefer firms located on the east side, all else equal, and consumers on the west side prefer firms on the west side. Durability is a vertical characteristic; we all agree that goods that last longer before breaking are better.

Like our moral taste buds, our horizontal preferences are part of who we are, at least for the immediate future. A disagreement between two consumers about which good is “better” driven solely by the two consumers having different horizontal preferences is thus not something that can be resolved via discussion or debate. There's no sense in arguing about which store is “better,” east or west, if location is the only difference between them (hypothetically) and store preferences are based just on location. No one is right or wrong; the consumers' preferences are just different. Therefore it's often wise to avoid disputing tastes—there's not really any point to it.

On the other hand, differences in opinion over vertical characteristics of goods (or other subjects) can *only* be driven by differences in beliefs about factual matters. These differences in opinion must be due to at least one of the two disputants being uninformed, misinformed, or downright wrong in some way. As a result, it very well might be productive to debate the point

and try to hash this out. Let's return to the example of durability as a vertical characteristic: in reality, we usually don't know for sure if, say, brand X or Y is more durable. If I think X is more durable and you think Y is, one of us is right and the other is wrong. It is reasonable for us to talk about this as discussion might allow us to improve the accuracy of one or both of our beliefs.

Thus, in this chapter when I say something is a "subjective" issue, I mean that differences in opinion for the issue are like differences in preferences for horizontal characteristics of a good, a matter of differing tastes. Analogously, when I say an issue is objective, I mean that we all share the same tastes for it, so it's akin to a vertical feature of a good. By the way, these definitions are consistent with how I've used the terms earlier in the book: we can have subjective preferences for different character traits in other people, while it's objectively better to have more accurate beliefs about other people's traits and future action probabilities. I'm just spelling things out a little more carefully here.

Now to return to undue hostility: this can arise in a few ways, all of which result in part, or entirely, from naive realism. Recall that naive realism, discussed briefly in chapter 3, is an overestimation of how objectively we see the world. Naive realism often refers to the overconfidence of the objectivity of our beliefs, causing overprecision. If others disagree with us on "vertical" issues and we're overconfident in our beliefs, we'll likely then overinfer that others' beliefs are mistaken and influenced by bias and poor judgment. Our mistaken inferences about others could then naturally lead to biased dislike toward them. For example, if you and I disagree on which headache medication works best, I might think you were brainwashed by ads for your favored brand or that you stubbornly ignore other valid evidence indicating you're wrong. Both beliefs could make me like you less—and dislike you too much if I'm overconfident in how correct my beliefs are.

Naive realism can also make us overestimate the objectivity of our *preferences*. If we think that if we like salads better than burgers, then salads must simply be (objectively) *better* than burgers, then we'll think that people who choose burgers are simply *wrong* and not just different. That is, naive realism can make us mistakenly think that horizontal issues are vertical ones. This type of naive realism in tastes contributes to why, for example, parents of each generation often think the music their kids like isn't as good as the

music they grew up with. (We typically like this music best mainly because we're more familiar with it and have positive nostalgic associations.²)

Naive realism in tastes also helps explain why we tend to debate matters of taste more than we should. We think our disagreements aren't intractable differences in taste buds but instead are actually simply matters of conflicting beliefs that can potentially be resolved via debate. This is, of course, *sometimes* true. I sometimes debate with my spouse what to make for dinner, and she turns out to be right—I'll later agree that what she wanted was better than what I had suggested. However, if we're biased toward overestimating our own objectivity (and we are!), then we'll overestimate how often these debates can be productive—how often our perceived preferences are driven by objective traits versus subjective ones.

Misperceiving our tastes as objective truths can then also naturally lead to undue dislike. The reasons are similar to why undue dislike can result from overconfidence in beliefs on truly vertical issues. For example, suppose feelings toward, say, some genre of music are driven largely by horizontal characteristics, but I think "I like this music; therefore it's high quality" (e.g., "jazz music from the 1920s is the *best* music of all time"). This would naturally lead me to have overly negative beliefs about people who disagree with me about this music. I'll overinfer that those who dislike it assessed its vertical characteristics incorrectly—that they simply don't realize how *good* it is, and so their judgment, and perhaps even intelligence, is worse than I would have thought otherwise. As a result, my confusion about matters of taste would cause me to hold unduly negative beliefs about their character traits (and corresponding action probabilities) and thus cause me to feel undue dislike toward them, that is, be subject to affective polarization bias ("If you don't like jazz from the 20s, there's something wrong with you"). See Ross (2018) for a more general discussion of these ideas and related research.

But why might disagreeing with other people on matters of tastes—even literal tastes for different types of food—affect, and even skew, our beliefs about other people's moral character? This is perhaps more puzzling. In my prior research (Stone, 2020a), I've proposed an explanation using an additional factor on the psychology of tastes: our tendency to overestimate the similarity of others' tastes and beliefs with our own. This phenomenon, called the *false consensus effect*, is also an extremely well-established empirical regularity in social psychology (see, e.g., Mullen et al., 1985). It's closely

Table 4.1

Examples of the false consensus effect (N = 122)

Item	Percentage of endorsers	Average endorser estimated percentage of endorsers	Average nonendorser estimated percentage of endorsers
1. I sweat very easily even on cool days	21%	44.5%	29.2%
2. My conduct is largely controlled by the behavior of those around me	28%	60.1%	49.4%
3. My hardest battles are with myself	73%	62.8%	46.2%
Total (for all 40 MMPI-2 items reported in the paper)	49.9%	55.6%	44.4%

Source: adapted from Krueger and Clement (1994).

related to naive realism and may often result from naive realism, but the two phenomena are distinct, as I'll explain. Table 4.1 shows some examples of false consensus: for statements from the Minnesota Multiphasic Personality Inventory (MMPI-2), people who endorse statements themselves tend to overestimate the percentage of others endorsing the statement, while people who don't endorse the statement tend to underestimate the total percentage of endorsers.

The false consensus effect is a simple and intuitive phenomenon. It makes sense that we'd use available information—our own tastes—to judge the prevalence of such tastes in others. The problem is we put too much weight on what we know about ourselves and systematically underestimate the degree to which other people differ from us in all sorts of ways. We are, again, too prone to think “what [we] see is all there is.” Overprecision exacerbates the problem by making us excessively confident in our (biased) beliefs about others' tastes and beliefs.

The false consensus effect doesn't always cause problems in relationships. Overestimating your similarity with your partner is usually a positive sign for the quality of the relationship (Montoya et al., 2008). But I argue that, perhaps counterintuitively, overestimating the extent to which we

share tastes with other people can also cause undue dislike. That's because if, say, Jill believes she shares the same tastes with Jane, then it's logical for Jill to interpret differences in her and Jane's *expressed opinions* as signals of differences in *moral character*. If Jill overestimates the degree to which she and Jane share the same true tastes, and sees Jill espousing an opinion inconsistent with these tastes, it's natural for Jill to think Jane has a hidden ulterior motive for what she says. Hidden motives are usually bad ones, and being disingenuous is usually a sign of bad character too. As a result, Jill would be subject to undue hostility toward Jane.

For instance, when Jill and Jane debate where to get dinner and Jane suggests the taco place because it's "the best" but Jill thinks the pizza place is the best, Jill might think Jane secretly agrees but has a hidden, self-serving motive for suggesting the taco place (perhaps it's more convenient for Jane). Moreover, if we observe others with different tastes from our own making such inferences and becoming more hostile toward us, and we don't understand the role of false consensus in their thinking, we may falsely attribute their behavior to poor character as well. (It's worth noting that the false consensus effect has not typically been found to apply to our own vertically differentiated characteristics, like competence and judgment—we don't overestimate how much we have these in common with other people.)

Naive realism and the false consensus effect can cause unduly negative character inferences in a wide range of settings. This may be most likely to happen when two people don't know each other well at all and so are especially likely to misunderstand each other's preferences. For example, consider a person who typically reclines his airplane seat sitting in front of, and reclining toward, someone who thinks reclining is rude. The recliner probably reclines without asking first because he thinks it's his right to do so and also thinks that this right is universally understood, due to overestimating his objectivity and falsely thinking there is a consensus about it ("That's why the buttons for reclining are there, right?").

The "antirecliner" thinks it's obvious that reclining is an invasion of another person's space and should only be done when you sit in front of an empty seat; if you must recline toward another person, "everyone knows" you should ask first ("and anyone who reclines without at least asking first is a jerk"). Do a web search for "airplane seat recline fight video" if you want to see examples of what can happen next. It's worth noting that this is a

situation where motivated reasoning would likely exacerbate the false consensus effect for both individuals—if I want to recline, I’ll be motivated to believe this is a widely accepted legitimate thing to do, and if I don’t want the person in front of me to recline, I’ll be motivated to think it’s widely accepted that one shouldn’t do this without at least asking first.

However, people who are intimately familiar with one another can be subject to a false consensus effect as well—and therefore also be subject to false consensus-driven undue conflict. Even friends, family, and spouses can be subject to false consensus with each other, and, again, while this can have positive implications, it can also cause disagreement that leads to undue conflict. Here’s an illustrative example. Spouse A loves flowers (and getting a gift of flowers), but spouse B doesn’t. Spouse B knows that A likes flowers better than B. But because of the false consensus effect, B underappreciates the difference in their tastes: B underestimates (“even after all these years”) how much more A likes flowers than B. So B decides it’s not worth bothering to get flowers often. Spouse A takes this as evidence that B does not care that much about A or is simply lazy or inconsiderate. You can see how undue hostility between A and B—and undue hostility given their misunderstanding of one another—might easily ensue.

Back to Politics

How do naive realism and the false consensus bias help connect Haidt’s moral foundations theory to political hostility? Well, of course we *don’t* have false consensus about support for many of the policies that we fight over. We more often have the opposite problem: we underestimate common ground in our views about specific policies. That’s the problem of false polarization discussed in chapter 2.

We’re well aware of real and perceived disagreements on many policy and related moral issues. When these disagreements are driven by different beliefs about the effectiveness of various policies, overconfidence in our own beliefs (overprecision, driven by naive realism among other factors) will make us overinfer others are mistaken in their beliefs and make us unduly judge others negatively.

Moreover, moral foundations theory implies many of these disagreements are ultimately subjective, at least for practical purposes. Again, naive

realism makes us confuse the subjective with the objective, and this confusion seems to in fact be especially pronounced for our beliefs about moral values. People indeed think their values are objective truths. For example, Skitka and colleagues (2021, 352) state this point clearly in their review article on the psychology of moral convictions: “People tend to perceive their morally convicted attitudes as objectively true facts that are grounded in fundamental truths about reality.”

It’s important to note that even if someone is *wrong* about the universality of their moral convictions, disliking another person because they fail to embrace the same morals does not necessarily cause affective polarization bias. Suppose you think fairness is always more important than other moral considerations and that everyone should think this. And suppose your brother Joe doesn’t care about fairness, and you know this about him and dislike him for this reason. Then you’re not misjudging him by your standards and are thus not subject to biased dislike. But if you think everyone, including Joe, *should* share your moral values, then you’ll probably dislike him further. You’ll think the fact that he doesn’t share the value signals something deeper that’s wrong with him.³ And naive realism will naturally lead us to misjudge how often our values should be shared.

For example, suppose (just for the sake of example!) that views on abortion are akin to Haidt’s moral tastes and are thus subjective as I’ve defined the term above. Suppose also, in the interest of simplicity, there are just two possible views on the issue, “pro-choice” or “pro-life,” and suppose I’m pro-choice and subject to naive realism. I thus think being pro-choice is objectively “correct.” If you’re pro-life and I know you’re pro-life and dislike you accordingly, that’s not affective polarization bias; I am judging you correctly by my standards. But if I incorrectly think we *should* share a view on this issue, and think you have the wrong view on the issue, I may then overinfer that you have other “bad” traits that caused you to have this wrong view. I’d then overestimate the probability of you taking other “bad” actions, leading me to dislike you more than I should (given my other preferences and values). Naive realism can make me think you’re a “bad” person in general, and not just for this one issue, more than I should.

Perhaps some (or even all) moral issues that moral foundations theory implies are “subjective” truly are “objective.” I’m not a moral philosopher and certainly won’t make any definitive claims here. But perhaps no one

can know for sure. And knowing that we perhaps can't know for sure, and that we have a general bias toward interpreting our personal tastes as universal truths, should temper our judgments. Even if we could somehow be sure that aspects of morality are "objective," knowing we tend to be overconfident in our beliefs on these issues should also temper our judgment of those who disagree. Moreover, if our differences on moral foundations are *practically* irreconcilable, at least for the immediate future, then our different opinions on these issues are essentially subjective, for practical purposes. If naive realism makes us overestimate the degree to which these differences can be reconciled, we'll still end up judging others with different moral values and different value-driven policy preferences too negatively.

Undue Suspicion

Politics is the art of making your selfish desires seem like the national interest.

—Thomas Sowell, "They're Baaack: Random Thoughts"

The false consensus bias can also cause political disagreement to yield undue dislike via suspicion that the other side is disingenuous about the reasons they support the policies they advocate. As Sowell alludes to in the quote just above, this suspicion is widespread. We often think others overstate the social benefits of policies they personally benefit from and sometimes even realize we do this ourselves. For instance, we might make arguments that lower or higher tax rates benefit society overall while truly being motivated by more personally relevant effects.

On the other hand, Sowell's quote is, of course, too cynical: often we do argue for policies that we feel are best for the country (due to a combination of our "tastes" broadly defined, beliefs, and some mixture of the two that we usually don't fully understand). The false consensus effect would then make us overestimate others' agreement that these policies are best. When we then observe out-partisans arguing for different policies, we'll logically infer their arguments are disingenuous or made in "bad faith."⁴ Or if we see someone take a political action (with "political" broadly defined) that "they know isn't right," we'll assume they have a bad motive. Bad faith and bad motives signal bad character, and, again, excessive inference of bad character yields affective polarization bias.

For example, suppose Democrats overestimate the similarity in their true views on the benefits of higher tax rates and level of government spending with Republicans: Democrats assume Republicans (silently) agree about the societal benefits of taxing and spending more than they really do. Suppose Democrats then observe Republicans vocally objecting to increased tax and spending. Democrats would then logically look for an alternative explanation for this vocal opposition by Republicans (such as wealthy members of the party benefiting being personally worse off from tax hikes) and suspect Republicans of being disingenuous when they make claims about the general benefits of tax cuts and harms of tax increases.

Or to return to abortion, suppose you're pro-life and subject to false consensus about pro-life being widely understood to be morally correct. Then if you see someone else support pro-choice policies, or even get an abortion themselves, you'll naturally assume they must have a bad motive. Perhaps you think they benefit from pro-choice policies leading to other political gains, or you think their choice to get an abortion reflects on some moral failing that they're aware of but don't have the strength of will to overcome. Again, judging them for taking an action that you think is "bad" would not imply affective polarization bias. The bias would come from your false assumption that "they know it's bad and they're doing it anyway," and false consensus is likely to make you think this.

Since the false consensus effect applies to both tastes and beliefs, it can cause undue suspicion regardless of whether policy disagreements are driven by tastes or beliefs. Democrats' and Republicans' disagreements on optimal tax policy might be due to different tastes for the "size" of government and different beliefs about the effects of tax policy on economic growth and the distribution of income. Either way, overestimating the underlying agreement logically causes overinference of the chance of the other side having an ulterior motive for claiming to disagree.

For another example, consider hostile disagreement over Trump's "America First" foreign policy. (For simplicity, let's ignore the phrase's unsavory historical associations.) This disagreement could be driven by differences in "moral tastes" (we may disagree over to what extent Americans should make sacrifices to help people around the world) and by differences in beliefs over how to best achieve the shared value of taking actions that serve America's interests (we may also disagree over to what extent supporting allies ultimately benefits Americans). Overestimating the similarity

of either our values or beliefs, or both, implies excessive inference of bad motives for those with different views. If I think that supporting allies does not benefit Americans much and I see that you do favor more support for allies (or so-called allies), I'm likely to think you might be disloyal, naive, or perhaps even personally or politically benefit from helping those other nations.

A Formal Example

Here's a simplified version of the model of false consensus-driven undue hostility from Stone (2020a) that I think provides additional clarity. Suppose there are two parties, left and right, and you're on the right. You want to know how "unselfish" or not someone on the left is. Suppose this character trait is denoted by C , and assume $0 < C < 1$, and you think each value is equally likely. So, on average, you expect C to be 0.5. Suppose you dislike selfishness, so you dislike someone more when you think their C value is lower.

You don't observe the out-party's (average) C directly, but you do observe a signal of it: a public policy that they support, which can be represented by a number greater than 0. This policy, call it X , affects society overall, and can provide "partisan benefits"—it can make members of the out-party better off, at society's expense. Now suppose that X represents the "size of government" and higher values of X yield greater partisan benefits for the left—for example, provide more government services, subsidies, and jobs for their constituents. So higher values of X might signal higher selfishness for the left. But the left also has a "taste parameter," T : a value of X that they feel is best for society, which is any number greater than or equal to zero. (This parameter could be driven by either direct preferences or beliefs about the extent to which higher values of X are good for society since false consensus can apply to either.)

Assume the left chooses X based on a combination of tastes and selfishness in a simple way:

$$X = T + \frac{1}{C} - 1.$$

If the left is fully unselfish and $C = 1$, they'd choose a value of X equal to T . If $C = 1/2$, then $X = T + 1$. If the left is more selfish with, say, $C = 1/3$, then X becomes larger ($X = T + 2$).

So, if you are on the right, you can infer the C value of someone on the left directly from X if you know their taste for the size of government, T . Rearranging the equation for X , the correct inference is $C = 1/(X + 1 - T)$. (Note that by assumption, $X \geq T$.)

Now suppose you think you know their T , but you are influenced by false consensus. Since you lean right, your taste is lower than theirs. Suppose you think their taste is t , with $t = T - b$, so b is your bias (the amount you underestimate how high their value of T is, because your value is smaller and you are subject to false consensus). Then your inferred C from observing their X , C' , is determined as follows:

$$C' = \frac{1}{X+1-t} = \frac{1}{X+1-T+b}.$$

Using the original definition of C to substitute $1/C$ for $X + 1 - T$ in the right-hand-side denominator, we can then show that

$$C' = C / (1 + bC).$$

This is strictly less than C if $b > 0$. In other words, the false consensus bias makes you underestimate the other side's character.

Moreover, the larger the bias (b), the more you underestimate their C . For example, suppose that since you lean right, your taste for government size is low, say, a 2, and suppose the left's taste is symmetric (around the midpoint of 5), so it is equal to 8. And suppose you know they lean left, but you're subject to a false consensus effect, so you perceive their taste is less than 8; let's say 6 (so your $b = 2$).

Then, if you see them choose $X = 8$, you should infer their $C = 1$, but you'll actually infer $C = 1/3$. If you see them choose $X = 10$, you should infer $C = 1/3$, but you'll infer $C = 1/5$. See figure 4.2 for an illustration; the gaps between inferred and actual character (affective polarization bias) are greater for larger values of character and larger values of b .

The details don't matter; as long as you experience false consensus bias (and your taste parameter is less than theirs), you'll always underestimate their C after observing their choice of X and thus experience biased dislike. If you observe them defending a high value of X that they chose by arguing that large X is optimal (i.e., large T), you'll think they're being disingenuous—that they actually agree with you that small T is better.

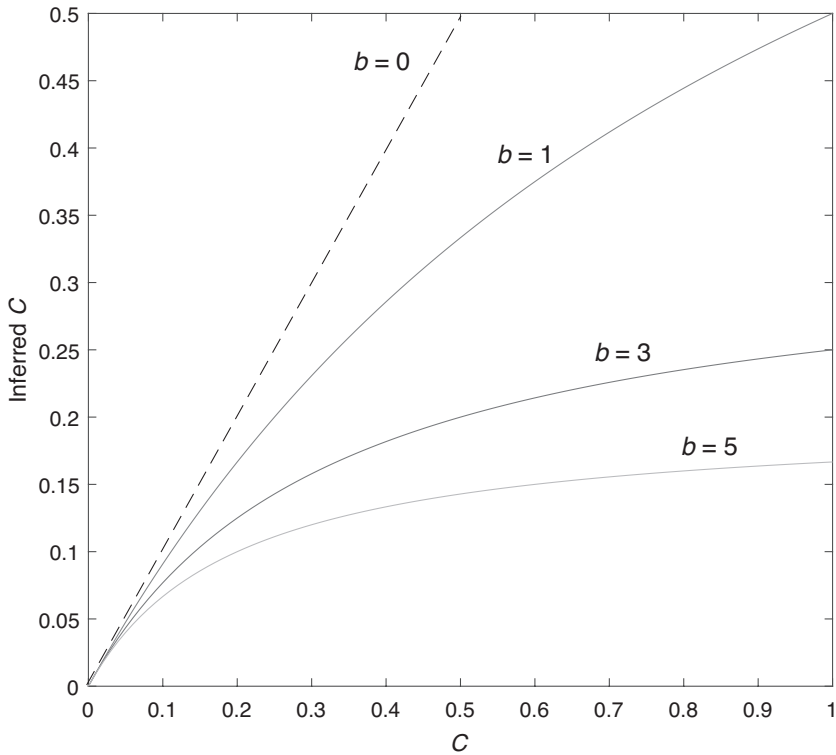


Figure 4.2

Inferred versus actual character (C) for different levels of false consensus bias (b). C minus inferred C is the underestimation of character due to false consensus bias.

Concluding Remarks

The worst offence . . . which can be committed by a polemic is to stigmatize those who hold the contrary opinion as bad and immoral men.

—J. S. Mill, *On Liberty*

Tough love or comforting attachment? Free-range or sheltered? Parents make these types of choices every day when deciding how much autonomy to grant their children and how to react when they struggle. There's usually no obviously optimal choice, just tradeoffs: more freedom generally means

more learning but also more painful mistakes.⁵ The good news is that there is also usually a wide range of reasonable options for how to balance these objectives. Even if it's unclear which option is precisely best, it's also clear most options are basically fine—they're good enough.

Many parents generally lean toward a single “parenting style”—and I've observed that many of us also, at least quietly, judge those whose styles differ from our own. This judgment is unlikely to lead to outright conflict between parents of different children but can do so. (My spouse is still upset about a stranger who once gave her advice on how to deal with a crying baby in a grocery store.) Explicit conflict over parenting styles is probably more likely to occur between members of different generations within a family: teens and parents most obviously but also parents and their grown children (over how the grown children were raised) and parents and grandparents (over how to raise the new generation of children).

Many disputes in politics today boil down to a similar question: to what extent should the government protect or push its citizens? Government is, of course, not literally a parental figure, and voters aren't children. Still, we look to it to play many of a parent's roles—to provide safety, protection, and guidance. It's possible for government to err in the direction of either too much or too little “protection,” broadly defined. This chapter is about why we unduly infer bad traits in those whose political opinions differ from our own. Just as we tend to dislike parents who we think are too strict—and those we think are too lax—we do the same for our fellow citizens who have different beliefs about various aspects of public policy.

I argue that we overinfer negative character traits in those who disagree with us due to naive realism and the related false consensus effect. A summary of these biases, and the mechanisms by which they can lead to biased dislike, is as follows:

- Naive realism is the overestimation of objective truth in our tastes or beliefs. This causes affective polarization bias by making us overinfer that actions or opinions that differ from our own ideals are flawed, and therefore that people who take those actions or have those opinions have “bad” character traits.
- The false consensus bias, which can result from naive realism, is the overestimation of the similarity of others' tastes or beliefs to our own. This causes affective polarization bias by making us falsely infer that people

who take actions or make statements inconsistent with our own tastes or beliefs have “bad” motives for those actions or statements, which we interpret as signals for generally “bad” character traits.

The arguments made in this chapter for why these general biases would cause biased dislike are applications of what’s sometimes called economic signaling theory. But the basic logic is straightforward: if we observe people taking opinions that differ from our own and underestimate the legitimate causes of differences in our opinions, we’ll overattribute these differences to negative factors.

It’s also straightforward to see how the mechanisms discussed here could have contributed to growth in biased dislike between the parties during the period partisan affective polarization has grown in the US. As the parties have sorted, they’ve grown to disagree more, in principle and in practice, and expressed moral values have diverged. Since larger disagreements imply larger errors due to the mechanisms discussed in this chapter, growth in disagreement implies growth in biased dislike.

The theories discussed in this chapter are logical extensions of the empirically well-established naive realism and false consensus effects. Prior literature has discussed naive realism as a cause of social conflict and polarization, though I would say to a surprisingly limited extent. Since the ideas of biased dislike and affective polarization bias are newly proposed in this book, there hasn’t been prior work explicitly arguing that these are implications of naive realism. My own prior theory work on false consensus and affective polarization is the only work that I’m aware of making this argument. And I’m not aware of direct evidence of either of these specific biases leading to biased character judgments. Constructing studies that directly examine the effects of naive realism and false consensus on undue dislike would be difficult but I’m sure not impossible. I hope future researchers are up for this challenge.

© 2023 Daniel F. Stone

This work is subject to a Creative Commons CC BY-NC-ND license.

Subject to such license, all rights are reserved.



The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in ITC Stone Serif Std and ITC Stone Sans Std by New Best-set Typesetters Ltd.

Library of Congress Cataloging-in-Publication Data

Names: Stone, Daniel F., author.

Title: Undue hate : a behavioral economic analysis of hostile polarization in US politics and beyond / Daniel F. Stone.

Description: Cambridge, Massachusetts : The MIT Press, [2023] | Includes bibliographical references and index.

Identifiers: LCCN 2022018587 (print) | LCCN 2022018588 (ebook) | ISBN 9780262047500 (paperback) | ISBN 9780262372374 (epub) | ISBN 9780262372367 (pdf)

Subjects: LCSH: Polarization (Social sciences)—Economic aspects—United States. | Polarization (Social sciences)—Political aspects—United States. | Economics—Political aspects—United States. | Economics—Psychological aspects. | Political culture—United States.

Classification: LCC HN90.P57 S766 2023 (print) | LCC HN90.P57 (ebook) | DDC 306.0973—dc23/eng/20221228

LC record available at <https://lcn.loc.gov/2022018587>

LC ebook record available at <https://lcn.loc.gov/2022018588>

10 9 8 7 6 5 4 3 2 1