Is paternalism wrong if we really don’t know what’s good for us? Research in behavioral economics and psychology has identified common flaws in people’s decision-making as well as interventions to correct or circumvent those flaws.¹ These findings, popularized by authors like Cass Sunstein and recent Nobel Laureate Richard Thaler, have inspired new defenses of paternalism that aim to reap the potential benefits of paternalist policies while overcoming the familiar objections to them. In particular, a family of views that I will refer to as welfare paternalism denies that paternalism needs to be either perfectionist or coercive. Rather, paternalism can aim more narrowly at improvements in people’s well-being and can do so through a range of techniques: we can improve (or improve on) people’s choices by altering the incentives they face or the “choice architecture” that shapes their decision-making, rather than (or in addition to) coercing them. Welfare paternalist arguments respond directly to worries that paternalism is a threat to autonomy, emanating chiefly (though not only) from government. In the views of their advocates, empirically-grounded welfare paternalist policies need not be understood as expressing a normatively important kind of insult to agents’ autonomy.

¹ Draft. Please do not cite without permission.

This paper offers a different way of understanding how paternalism can insult or wrong the people paternalized. Even if one accepts both that paternalist policies can avoid coercion and that having more choices is not always better from the point of view of individual welfare, one can object to paternalism as a threat to equality, emanating from private as well as public actors. Egalitarian objections provide a way of vindicating many people’s intuitive responses to cases of paternalism where choices are not significantly constrained, or where the choices on which the paternalist intervenes are relatively trivial. In such cases, the distribution of the paternalist policy or the hierarchical structure through which it is enacted may explain the objection.

The most worked-out defenses of welfare paternalism recognize equality as one kind of objection to paternalism and endeavor to show that some versions of paternalism can avoid expressing disrespect for people’s status as equals. If they succeed, then the new behavioral science-inspired paternalism will have given us something historically anomalous and normatively important: a defense of paternalism that avoids explicitly or implicitly expressing a claim about the superior judgment of some people over others (and the entitlement of the former to take decisions on behalf of the latter).

In this paper, I raise doubts about welfare paternalism’s ability to avoid such expressions of superiority and inferiority. Setting aside the question (still disputed) of whether new defenses of paternalism overcome autonomy-based objections, I consider what would be required for them to meet equality-based objections. Section one offers a characterization of the new welfare paternalism. Section two discusses objections to paternalism in general and outlines the reasons that welfare paternalists have for thinking that their brand of paternalism avoids those objections. Section three argues that the
obstacles to a paternalism that is applied in egalitarian ways are greater than welfare paternalists usually recognize. Section four argues that the structures through which paternalist policies are enacted can also give rise to egalitarian objections and considers the special problems that attend paternalism by private actors (for example, by employers and philanthropists). Section five concludes.

I. Defining the New Welfare Paternalism

I use the term paternalism to refer to attempts to restrict, circumvent, manipulate, or influence an agent’s choices, where those attempts express the judgment that the agent’s ability to choose or act well on her own behalf is deficient, or inferior to that of the paternalist, in some relevant respect or domain. This definition includes but is not

---

2 For similar definitions in the philosophical literature on paternalism, see Shiffrin (whose motive-based definition focuses on the paternalist’s attempt to substitute her judgment for that of the person paternalized); Jonathan Quong’s “judgmental definition” in Liberalism Without Perfection, 80; and, especially, Nicolas Cornell, “A Third Theory of Paternalism,” Michigan Law Review 113 (June 2015), 1295-1336. Cornell’s account is closest to my own because he focuses on the external meaning of paternalist actions rather than on the motive of the paternalist (which is Shiffrin and Quong’s focus). See Cornell, 1308fn43. The definition offered here is in one way narrower than my preferred account: like Shiffrin and Cornell, I believe that actions whose primary motive is to benefit on someone other than the person paternalized can count as paternalistic (provided they express the requisite negative judgment about the competence of the person paternalized). My definition could be amended (i.e. by eliminating the “on her own behalf”) to specifically cover interventions that aim to benefit someone other than the agent.
limited to cases of coercive paternalism; it will therefore not be accepted by people who argue for narrower definitions of paternalism that require interference with the liberty of the person paternalized.\(^\text{3}\) However, the defenses of paternalism that I am exploring in this paper generally accept that non-coercive actions can be paternalistic. For my purposes, the interesting point of controversy is not taxonomical (“Can non-coercive attempts at influence be paternalistic?”) but normative (“Can non-coercive attempts at influence be objectionable, for similar reasons that (other) cases of paternalism are objectionable?”).

I use the term welfare paternalism to distinguish paternalistic attempts to improve a target person’s or population’s health, wealth, or well-being from paternalism that aims at the moral perfection of the person or people paternalized. The label picks out an important point of agreement between the positions that I am interested in exploring: that paternalism is impermissible when perfectionist or moralistic but (sometimes) permissible when more narrowly tailored to promote well-being. Sometimes, proponents of welfare paternalism express this as a distinction between “paternalism about ends” (impermissible because a person should have freedom to choose her ends for herself) and “paternalism about means” (permissible to help people promote their own chosen ends).\(^\text{4}\)

In practice, paternalists have focused on interventions to improve people’s physical

\(^3\) See for e.g. the influential account in Gerald Dworkin, “Paternalism,” *The Monist* 56:1 (1972), 70-76.

health and economic security, both of which are understood as all-purpose means that will help people realize whatever (other) ends they value. In Rawlsian terms, welfare paternalism attempts to confer primary goods and not to promote a controversial conception of morality.\textsuperscript{5} In this paper, I accept for the sake of argument the claim that it is possible for paternalist interventions to avoid perfectionism. However, my own argument does not depend on accepting this claim, since the egalitarian objection to paternalism that I develop below holds whether the paternalism is welfarist or perfectionist in character.

The focus on welfare, and on promoting the objective interests of agents rather than aiming at their moral reform, is the central normative commitment linking recent defenses of paternalism. The new welfare paternalists are also united in accepting an empirical claim: that because of cognitive errors or weakness of will, people often fail to recognize or take the actions that would promote (what they themselves would recognize as) their own good. Recent defenses of paternalism are grounded in research in the behavioral sciences and rely heavily on evidence of people’s fallibility in decision-making to show that promoting people’s welfare often requires improving on or (where that is impossible or too cumbersome) circumventing their own decision-making.

The welfare constraint on paternalism and the claim that human fallibility in decision-making creates at least a prima facie case for paternalistic interventions

\textsuperscript{5} However, it may be that even welfare paternalism is unable to avoid controversy over conceptions of the good. For example, smoking may be inextricably tied to a person’s considered conception of the good life, so that anti-smoking “nudges” necessarily implicate that conception.
characterize all welfare paternalist positions. Two other features, while not necessary elements of a welfare paternalist position, recur often enough to be worth noting.

First, the commitment to securing better outcomes (measured in terms of well-being) allows welfare paternalists to elide the distinction between public and private actors. Paternalist interventions by either kind of actor should be judged by the same standard: the effects that they have on the welfare of individuals. Often, the differences between governmental and private paternalism are literally glossed over in parentheses. So, for example, Banerjee and Duflo write that to improve the lives of poor people, “The government (or a well-meaning NGO) should make the option that it thinks is the best for most people the default choice.” Similarly, Ben-Porath notes that “through instituting default rules, the state (or employer, or other regulator) can affect levels of participation in different programs, thus advancing what seems to be preferable . . . without imposing a change of motivation or preference on individuals, and without infringing upon their freedom.” Welfare paternalists often assume the portability of their arguments across different kinds of social, economic, and political relationships.

Second, many welfare paternalists express a preference for non-coercive paternalism, although this has been the subject of more internal debate. Welfare paternalism is primarily a view about good reasons for practicing paternalism and about the conditions under which paternalism is in principle permissible and desirable. It spans

---


a number of different positions regarding the techniques that paternalists may use. So-called “libertarian paternalism” rejects coercive paternalism entirely, limiting justified paternalism to the use of “nudging” techniques that influence people’s choices without actually restricting the choice set. But many other welfare paternalists accept some forms of coercive paternalism (e.g. seatbelt laws). Some are willing to go much farther, proposing forms of coercive paternalism well outside what public opinion would be likely to support (e.g. banning smoking).  

I refer to welfare paternalism rather than to libertarian paternalism or to nudging because these categories do not always overlap. Libertarian paternalism is a narrower category than what I call welfare paternalism, which can include forms of coercive paternalism. Nudging, on the other hand, is a much broader category and covers many non-paternalist interventions: nudges use behavioral science insights to influence people’s decision-making but can do so for purposes ranging from profit-seeking (as in commercial advertising) to environmental protection (as in the use of “green” default options on computer printers) to public health (as in efforts to increase rates of organ donation by changing the default option to require an opt-out rather than an opt-in to donation). Some of the concerns that I raise about paternalism may apply to such

---

8 Sarah Conly, Against Autonomy: Justifying Coercive Paternalism (Cambridge, UK: Cambridge University Press, 2013), 169-172.

9 Though the organ donation case might still count as paternalism on some understandings, e.g. Shiffrin’s claim that the paternalizer’s substitution of her judgment for that of the person paternalized (regarding matters that properly fall within the latter’s domain of autonomy) constitutes the core of what is morally troubling about paternalism, and that this can hold even when the usurpation of judgment aims to benefit someone other that the person paternalized.
interventions as well (either because they prove on closer inspection to be paternalistic or because they raise different but neighboring concerns about acceptable and unacceptable ways of using people—concerns about exploitation, for example). People will differ on the questions of whether paternalist or non-paternalist nudging faces a higher burden of proof and whether it is more disrespectful to use, manipulate, or influence a person for her own benefit or for the benefit of other people. Without taking a position on those broader issues, I will focus on cases where behavioral interventions express a concern for the welfare of the person whose choices they intend to alter or shape.10

Like other authors,11 I believe that the distinctive normative significance of paternalism in general comes from the negative judgment or insult that paternalism

Most welfare paternalists do not engage directly with this last possibility, and assume that paternalism aims to benefit the person paternalized. See Seana Valentine Shiffrin, “Paternalism, Unconscionability Doctrine, and Accommodation,” *Philosophy & Public Affairs* 29:3 (Summer 2000).

10 The word “express” here is intended to focus attention on the external meaning of an action rather than on the intentions of the actor. An intervention can express concern for someone’s welfare without being primarily motivated by such a concern, and on my understanding expressive content can make an action paternalistic independent of its motivation: for example, it is paternalistic when employers introduce weight-loss incentives for their employees, even if they do so in an attempt to save on health care costs rather than out of genuine concern for their employees’ well-being.

11 See for e.g. Shiffrin’s claim that “paternalist doctrines and policies convey a special, generally impermissible, insult to autonomous agents” (207) and Cornell’s argument that “instances of paternalism are objectionable because of their expressive content. Paternalism is suspect because
expresses toward the person(s) paternalized. This wrong can exist independently of coercion, of defects in the consent of the person paternalized, and of impermissible motivations on the part of the paternalizer.\textsuperscript{12} To see this, consider an example of parents who leave a gift of money in trust for their adult children, subject to conditions on the children’s behavior (e.g. that they refrain from smoking). Supposing that the children are not morally entitled to the money in question and have the option to refuse it, the trust is not coercive; and supposing that the parents have a conscientious objection to supporting tobacco companies, their motivations may not be paternalistic. Nevertheless, I take it that such conditional trusts are in some respect paternalizing and insulting to adult agents, because of the relational dynamic that they assert or attempt to preserve, in which parents exercise judgment about the personal choices of their adult children.

For both interpersonal and broader political examples of paternalism, there are different ways of explaining when and why interventions on an agent’s choosing express the relevant kind of insult. In the following section, I will suggest that we can divide accounts of the kind of insult paternalism involves—and thus objections to paternalism more generally—into two principal categories.

\textsuperscript{12} For arguments that paternalism need not be coercive or liberty-restricting (by both supporters and critics of paternalism), see Sunstein, Sunstein and Thaler, and Shiffrin. For an argument that the motivation of the putative paternalist is not the key factor in assessing paternalism, see Cornell, “A Third Theory of Paternalism.”
II. Paternalism, Autonomy, and Equality

The two principal objections to paternalism, and to the kind of disrespect that paternalism expresses, can be understood as two different specifications of the familiar claim that paternalism is objectionable for treating people like children. The first and more common claim is that paternalism insults or violates a person’s autonomy. The second is that paternalism insults or violates people’s equality. I take them in turn.

The Autonomy Objection to Paternalism

Many people object to paternalism as kind of impermissible coercion; hence the libertarian paternalist strategy of trying to reconcile some opponents to paternalism by showing that paternalism can avoid coercion. But of course, to object to paternalism as a case of impermissible coercion is question-begging: *why* is it impermissible to coerce, manipulate, or otherwise intervene on a person for her own good, when it would be permissible to make similar interventions for the good of or (at a minimum) to prevent harm to other people? And might an objection to paternalism be grounded in something other than an objection to coercion?

Many answers focus on paternalism’s affront to people’s autonomy and claim that paternalism is objectionable primarily as “a failure of respect, a failure to recognize the authority that persons have to demand, within certain limits, that they be allowed to make their own choices for themselves.”\(^\text{13}\) Consider also, for example, Shiffrin’s claim that “paternalist doctrines and policies convey a special, generally impermissible, insult to people’s autonomy.”

autonomous agents.”14 The importance of respecting autonomy has been cited as a reason for objecting even to non-coercive paternalism. Attempts to influence, manipulate, or “shape” people’s choices, as well as straightforward attempts to constrain the choice set, can threaten “the control that an individual has over his or her own evaluations and choices.”15 Even if paternalist nudges do not constrain choices in the narrow sense of “closing off alternatives or rendering them more costly,” they diminish liberty in the broader sense of autonomous deliberation and control over choices.16

Welfare paternalists emphasize the myriad ways that individuals fail to make the choices or take the actions that would best promote their own welfare. Nevertheless, they would agree that some choices should be left to individuals, with minimal outside intervention or manipulation. There is no difficulty explaining, in welfarist terms, why it is generally a good idea to let people decide for themselves whom to marry or what career to pursue. We need not think that this rule is required by respect for autonomy; we might simply think that, for some kinds of choice, attempting to shape an individual’s preferences or decisions will be an unreliable way of promoting her welfare. Welfare paternalists can incorporate epistemic arguments against some kinds of paternalistic intervention without agreeing that these arguments apply to all cases, and they can thus leave significant space for welfare paternalist policies.17 Often, welfare paternalists

14 Shiffrin 207.


16 Ibid.

17 See for e.g. Sunstein, *Why Nudge*, 91.
position themselves as criticizing or qualifying Mill’s argument that individuals are usually the best judges of what is good for themselves: they generally argue that this might be true in some areas but is definitely false in others.  

After all, reason welfare paternalists, not all choices are alike. Many choices are trivial or even burdensome rather than being closely linked to one’s sense of self-authorship. The Food and Drug Administration gives a federal agency (rather than individual consumers) authority to make decisions about what ingredients are safe to consume; but for many people it will be counterintuitive to see this outsourcing of cognitive labor as compromising individual autonomy in any meaningful way. So while welfare paternalists generally take the objection from autonomy or choice to be the most important obstacle to the more widespread endorsement of paternalist policies, they also believe that they have a persuasive answer to that objection. The kinds of paternalist policies that welfare paternalists are interested in promoting seem to them to be vindicated once we recognize that the value of choice for its own sake is easily overstated. The fact that having authority over some decisions—what career to choose, whether and whom to marry—is vitally important to a person’s ability to live a good life should not lead us to infer that all choices are equally important from the point of view of individual freedom or, especially, individual happiness. The welfare paternalist project aims to liberate people from burdensome choices, to make welfare-promoting choices

---


19 As suggested by titles like Conly’s *Against Autonomy* and Sunstein’s *Choosing Not To Choose*. 
easy and attractive, and so to leave people more free to express their individuality in the choices that really matter.

Welfare paternalists responding to autonomy-based objections generally emphasize either the possibility of non-coercive paternalism or the risk of overstating the generic value of autonomous choice. I agree with welfare paternalists that many of the interventions they recommend do not seem to raise significant worries about threats to autonomy. But welfare paternalist responses to autonomy objections are largely orthogonal to another category of objection to paternalism: objections from equality. (“Libertarian paternalism” is, after all, a label much more likely to reassure autonomy theorists than egalitarians.)

The Egalitarian Objection to Paternalism

To sustain an autonomy objection to paternalism, it will generally be sufficient to point to features of a dyadic interaction (e.g. in diagnosing the paternalistic intervention of A over B: the intervention is or is not of a kind that compromises B’s autonomy). But I am not convinced that we can assess whether an intervention is paternalistic in the first place, much less whether the paternalism in question is justifiable, solely from the dyadic

---

20 However, this is not to say that no welfare paternalist interventions raise significant worries of this kind (and non-welfarist paternalism may be even more likely to raise such concerns). My intention in what follows is not to rebut or dismiss autonomy as an objection to some kinds of paternalism, but rather to show that paternalism can be objectionable even when paternalistic behavior does not seem meaningfully to compromise autonomy.
point of view. Questions about background conditions, the distribution of influence, and the treatment that other agents encounter, will bear on the diagnosis and justifiability of paternalism: we need to assess paternalism through a social and relational lens.

Critics of paternalism (including ones who appeal to the value of autonomy) often observe that paternalism is insulting or disrespectful because it treats a person’s judgment about her own good as defective. But the insult this expresses to an agent’s competence might take one of two forms: absolute or relative. Paternalism can treat agents as incompetent (or imperfectly or insufficiently competent) full stop. It can also treat agents as less competent—less competent than other people not deemed in need of paternalist help, and/or less competent than the paternalist. The autonomy objection to paternalism can be understood as a response to absolute judgments about an agent’s competence. In the face of an (explicit or implicit) claim that an agent is incompetent to decide for herself, it demands that we accord each person the appropriate measure of control over her own choices and authority over her own life. (Usually, this means insisting that the agent is competent to decide for herself.)

The egalitarian objection to paternalism can be understood as a response to relative judgments about agents’ competence and incompetence. It rejects the claim (or implication) that differentially competent agents should encounter or practice paternalist treatment to different degrees. On the contrary, it holds that treating the judgment of some class of persons as inferior is objectionable amongst persons who are entitled to social and political equality.
A focus on autonomy explains why some authors have struggled to see how non-coercive paternalism could count as (objectionably) paternalistic and why many more deny that rational persuasion (to take an example I address in detail below) can count as paternalistic. But, if the egalitarian objection is (at least sometimes) warranted, our sense that paternalism is morally troubling need not come only from a sense that paternalism violates freedom or autonomy. An action or policy that might not in general represent a disrespectful way of treating persons can become disrespectful when applied or enacted in ways that violate equality.

Even paradigmatically respectful modes of engagement such as persuasion or advice giving can be objectionably paternalistic if distributed or managed in a way that expresses disrespect for some people’s decision-making capacities: for example, if laws compel the provision of more information and advice to women seeking medical treatment than to men. Florida’s “informed consent” rule requires that women seeking

________________________________________

21 See for e.g. Hausman and Welch, 127, arguing that many “nudges” are actually instances of rational persuasion, which ought not be counted as paternalistic. For an autonomy-focused argument against the widely accepted view that “one cannot act paternalistically by rationally persuading someone to do something,” see George Tsai, “Rational Persuasion as Paternalism,” Philosophy & Public Affairs 42:1 (2014). Tsai argues that, while rational persuasion is generally a respectful way to seek to influence someone, it can count as disrespectful and paternalistic under some conditions: “Rational persuasion may express, and be guided by, the motive of distrust in the other’s capacity to gather or weigh evidence, and may intrude on the other’s deliberative activities in ways that conflict with respecting her agency” (79).

22 Note that I have different reasons from Tsai for arguing that rational persuasion can sometimes count as paternalistic: he argues that persuasion can be paternalizing even (strictly) on autonomy
abortions receive (at least 24 hours before the procedure is performed) information about the age and stage of development of the fetus, a “list of entities that offer alternatives to terminating the pregnancy,” and that they be offered the opportunity to view live ultrasound images of the fetus.\(^{23}\) (The latter condition is waived if the woman presents documentation “which provides evidence that the woman is obtaining the abortion because the woman is a victim of rape, incest, domestic violence, or human trafficking.”) Similar laws exist in other U.S. states. The objection to “informed consent” abortion laws need not be that offers of information and advice in general violate autonomy, or even (as may in fact be the case) that the information provided is misleading or manipulative. Rather, we can object to the judgment such laws express about the decision-making competence and entitlements of women seeking abortions relative to that of people seeking other kinds of medical treatment. (In the Florida case, the negative judgment is directed more specifically at abortion-seekers whose pregnancies resulted from consensual sex.) Leaving aside debate about the specifics of this case, the point is that we can draw the line regarding what constitutes objectionable paternalism with reference to equality rather than (only) to autonomy.\(^{24}\)

\(^{23}\)For the full statute, see

\(^{24}\)I introduce the informed consent example as a case of paternalism broadly-speaking, and not necessarily of welfare paternalism. On definitions of paternalism that require a particular motive on the part of the paternalist (e.g. to benefit primarily or exclusively the person paternalized), the
Problems around unequal treatment and unequal power are central to both the theory and practice of paternalism, so much so that some authors interpret paternalism as oppressive by definition. This is probably too restrictive: we should leave open the possibility that paternalism can be appropriate and non-oppressive in some contexts (e.g. relations between parents and children). But while we need not regard paternalism as necessarily oppressive, I do think that we should regard it as unavoidably asymmetrical. Examples of hierarchical but non-oppressive paternalism (e.g. ordinary parent/child cases, where the authority is limited, justified by special features of the relationship, and temporary) look less strange qua examples of paternalism than do cases of “reciprocal” paternalism.

---

25 Ben-Ishai, for example, argues that “Paternalism is best understood as referring to forms of intervention that serve to perpetuate oppressive social relations by way of their coercive and stigmatizing tactics.” Elizabeth Ben-Ishai, “The New Paternalism: An Analysis of Power, State Intervention, and Autonomy,” *Political Research Quarterly* 65:1 (March 2012), 163.

26 Imagine, for example, a married couple composing New Year’s resolutions. Because both members of the couple are aware of their vulnerability to procrastination, temptation, and to prioritizing immediate benefits over future ones, they agree to monitor each other’s progress (on quitting smoking, on meeting writing goals, on donating to charity, or whatever) and occasionally to implement behavioral “nudges” in order to help each other achieve their own (self-defined) goals. Now imagine a second couple, where again both members compose resolutions, but where
Egalitarian objections to paternalism are not new; they appear in Mill’s own critique of paternalism, which focuses not just on the general wrongfulness of coercing a person for his own good but on the inappropriateness of relations of “dependence and protection” for a society of equals.27 Recent variants of the egalitarian criticism have been advanced against examples of governmental paternalism, notably “new paternalist” welfare policies that involve the coercion, surveillance, and stigmatization of the poor.28 These policies are important in practical terms but, from the point of view of egalitarian critique, relatively easy targets. What about versions of paternalism that look more

by agreement only one member monitors and attempts to shape the choices of the other. My suggestion is that the nudging behavior in the second case is more paternalistic than the first. I am inclined to describe the first case as non-paternalistic because of the reciprocity of the relationship and the acknowledgement that both people are equally in need of help (and equally competent to help each other). Note that, in both cases, the paternalistic behavior occurs with the consent of the person paternalized and is directed narrowly at meeting the self-defined goals of the person paternalized: that is, it is neither coercive nor perfectionist. Nevertheless, it seems to me that the asymmetry in the second relationship makes an important difference to our judgments about whether the relationship is paternalistic, and to the normative importance of the paternalistic behavior.


neutral—that do not necessarily involve coercion and that might be stigmatizing in practice but not so obviously by design?

In theory, at least, welfare paternalism gives us something historically anomalous: a paternalism grounded in people’s equal fallibility to weakness and error. Paternalism has generally been defended on the basis of the *differential* competence of the paternalizer and the person paternalized. This is true in cases where many people would regard paternalism as fully justified (e.g. of parents toward their children) as well as in more obviously objectionable cases. Defenses of male dominance pointed to women’s naturally inferior cognitive capacities to justify the authority of fathers and husbands over wives and daughters: women needed patriarchal guidance for their own good. A similar structure can be seen in defenses of feudal authority or of white supremacy. Of course, in each of these cases the paternalist defense of hierarchical relations is largely spurious: these relations were not *really* managed so as to promote the good of the subordinate party, whatever the people involved in the relationships may have believed. I give the examples not in the interest of considering whether these relationships were meaningfully paternalistic but rather to note the link between paternalistic *arguments* and claims of differential competence. It must for a long time have seemed puzzling how one could make a paternalistic argument *without* relying on claims about a hierarchy of cognitive capacities or capacities for self-control.

But in welfare paternalism, as Conly puts it, “it is not an assumption of superiority, but of shared fallibility, that moves us to paternalism, and no assumption is made about the superiority of one group of people over another.”29 The case for

29 Conly 37.
paternalism rests not on some people’s or kinds of people’s fitness to decide for others but on the benefits—for all people—of outside help in making decisions, especially decisions where the temptation to procrastinate or to indulge in pleasant but detrimental behavior is high.

But we should be wary of conflating the behavioral science findings that move us to paternalism with the structure of existing or proposed welfare paternalist interventions. In the remainder of this paper, I argue that (contrary to what its proponents suggest) welfare paternalism does treat some people’s judgment as superior to that of others. It does so in two different (though often related) ways. Most obviously, paternalist interventions may be applied in an uneven or discriminatory way. (Of course, this is often bad in cases of non-paternalist policies as well, but with paternalist policies it expresses a particular negative judgment about the competence of the people singled out for paternalist treatment.) Second, (putatively) paternalist policies can be enacted through more or less egalitarian and democratic processes. To answer the egalitarian objection to paternalism, welfare paternalism must be able to avoid both kinds of problem; I argue that, under existing and foreseeable conditions, welfare paternalist interventions cannot avoid inequalities of (at least) one kind or the other.

III. Inequalities in the application of paternalist policies

Welfare paternalists defend the egalitarian credentials of their proposals in two principal (and related) ways. First, they point to the egalitarian moral foundations of their
theories and emphasize that all people need paternalist help. Second, some explicitly attach egalitarian riders mandating the universal application of paternalist policies (and not just their universalist foundations). In this section, I argue that egalitarianism and paternalism are not so easily reconciled by either move.

It is worth noting that welfare paternalists who even recognize the challenge of reconciling paternalism and equality are the minority. Welfare paternalists sometimes acknowledge cases where interventions were discriminatory or stigmatizing but they usually treat these as unfortunate anomalies that do not affect the core of the welfare paternalist theoretical position. But the issue is not only one of an egalitarian theory being misapplied in practice. For welfare paternalists, the question of whether a paternalist intervention is justified turns primarily on its expected net welfare benefit. Most welfare paternalists do not require universal or symmetrical application as conditions for an intervention’s justifiability, and they sometimes go so far as to make distinctions between people who need paternalistic help and those who do not. One approach (which Thaler and Sunstein endorse in *Nudge*) recommends “asymmetric paternalism:” a regulation that is asymmetrically paternalistic “creates large benefits for those who make errors, while imposing little or no harm on those who are fully rational.” The suggestion seems to be

---

30 Welfare paternalist authors are notable for their modesty; they frequently self-identify as procrastinators, akratics, and bad choosers.

31 Colin Camerer et. al., “Regulation for Conservatives: Behavioral Economics and the Case for ‘Asymmetric Paternalism,’” *University of Pennsylvania Law Review* 151:3 (Jan. 2003), 1212. See also Thaler and Sunstein, *Nudge*, 252, where the principle is paraphrased as follows: “we should design policies that help the least sophisticated people in society while imposing the smallest
that paternalist policies are to be designed with two different kinds of person in mind. This need not mean that the “fully rational” and “those who make errors” encounter differential treatment. In this case, proponents of paternalism are rather noting the diverging effects of the same treatment on rational and irrational people (or, in Thaler and Sunstein’s modified version, the “most sophisticated” and “least sophisticated”). But notice too the diverging standards of justification invoked here: paternalist policies are to be endorsed when they confer significant welfare benefits to poor decision-makers and condemned when they impose excessive costs on good decision-makers. Thaler and Sunstein give no reason to rule out differential treatment, if that would magnify the benefits to one group or reduce constraints on the other.

Given their reliance on behavioral science evidence, it should not be surprising that welfare paternalists are often tempted to recommend asymmetrical treatment (and not only treatment that produces asymmetrical effects). While welfare paternalists interpret the capacity for good decision-making as largely contextual—a function of how one relates to the decision in question, and of how detached one can be from immediate temptations—they also posit situational factors that might compromise a person’s decision-making in a more general sense. Research on “ego depletion” and “decision fatigue” suggests that self-control is a limited resource, whose overuse can compromise decision-making competence. And so welfare paternalists sometimes attribute greater possible costs on the most sophisticated. (Libertarian paternalism is a form of asymmetric paternalism in which the costs imposed on the sophisticated are kept close to zero.)

fallibility in decision making to poor and disadvantaged people, from the expectation that such people’s cognitive and motivational capacities are more likely to be overtaxed by an excess of high-stakes decisions. This is different from the kind of moral judgment of poor and disadvantaged people that “new paternalist” welfare state policies are sometimes criticized for making. But it also supplies an empirical basis for differential treatment. If poverty negatively impacts a person’s decision-making, one obvious conclusion is that paternalist regulations or nudges should focus on the poor, who are most in need of paternalist help.

Libertarian paternalists in particular—while committed to practicing non-coercive paternalism—often treat coercion as the only serious normative objection that a person paternalized might have. As such, they fail to recognize other grounds (e.g. stigmatization or unfairness) for objecting to paternalist policies that target subsets (especially already-disadvantaged subsets) of the population.

To the degree that they recognize egalitarian objections to paternalism, welfare paternalists claim that their theories’ egalitarian normative foundations solve the

---


33 See for e.g. Soss, *Disciplining the Poor*, and Ben-Ishai, *Fostering Autonomy*.

34 Of course, an equally obvious inference from the same behavioral science evidence is the need for economic redistribution, and many welfare paternalists do endorse a more egalitarian distribution of wealth. That they nevertheless recommend (and focus on) paternalist interventions suggests either that they take redistribution to be an insufficient condition for better decision-making or that they regard large-scale redistribution as less politically feasible than paternalist policies that target the poor. On the latter point, they are probably correct.
Among the welfare paternalists, Conly gives the clearest articulation of and response to egalitarian criticisms. She acknowledges that “paternalist policies operate on the assumption that the person in the throes of making the decision about himself is less capable than those who have formulated the paternalistic policy on this issue,” and that paternalism on a political scale “is often taken to posit significant inequality among humans, inequality of the sort that could justify a class or caste system.” The fear underlying egalitarian criticisms is “that there will be one class of persons, self-styled experts, who make the rules, and another class, the supposedly cognitively impaired, who obey them,” which would be “undemocratic in a deep sense.”

But while Conly takes the egalitarian objection seriously, she believes that contemporary welfare paternalism overcomes it. The important move that welfare paternalists make is to interpret people’s decision-making competence as a matter of

35 This strategy of responding to objections from equality in the case of paternalism is similar to the strategy that David Estlund takes in justifying relations of unequal authority. Estlund often argues that unequal power among persons who are entitled to social and political equality can be justified if claims to authority make no “invidious comparisons” between persons.
36 Conly, 37.
37 Conly, 37.
38 While it is true that some pictures of paternalistic government have suggested that the more able will be in charge of the welfare of the less able, our present understanding of cognitive bias doesn’t support the view that one group is entitled to that kinds of authority over others. Not only does it not support a class division, it positively undercuts the grounds for such a division.” (Conly 37-38)
They claim that all people are prone to cognitive errors and weakness of will when deciding and acting for themselves: welfare paternalism addresses errors that are “a function of circumstance rather than the kind of person doing the thinking.” An appropriately-situated paternalizer might be better placed to take good decisions and help elicit good behavior not because he is smarter, stronger-willed, or more virtuous: he simply does not have the same vulnerability to cognitive biases and to immediate temptations when acting on behalf of other people as he would when deciding for himself.

Because the case for welfare paternalism rests on an understanding of our shared vulnerability to cognitive bias and weakness of will, the kinds of differential treatment it recommends need not (the argument goes) rest on invidious distinctions. A policy that I design to apply to the poor need not be problematically insulting so long as I hold the conviction that, were I to become poor, I would want this policy applied to me in the same way. This line of thought holds that, although some people and not others may be subject to pervasive paternalist regulation, welfare paternalism is egalitarian to the extent that it recognizes differences in decision-making competence to be a matter of context.

So far, that response stresses the point of view of the paternalizer and not the expressive content of paternalistic interventions. Recognizing, perhaps, that even well-intentioned policies can convey insults they do not intend, some welfare paternalists supplement emphasis on the contextual character of decision-making competence with provisos calling for symmetrical application in practice as well as in theory.

---

39 Conly 36-39. (See also Sunstein and Duflo.)

40 Conly 38.
Conly is friendlier than other welfare paternalists to paternalism in the form of (some) coercive government regulation; however, she is also more sensitive than most to egalitarian objections. As I argue in the following section, there is good reason to think that this conjunction is not coincidental. In any case, it leads her to condemn some real-world examples of paternalism, including some of Michael Bloomberg’s efforts to eradicate sugary beverages during his tenure as mayor of New York City. Why would Conly, who endorses wholesale bans on some unhealthy products, baulk at Bloomberg’s proposal to limit the use of food stamps to purchase sugary drinks? Because, she says, the proposal targeted only people on food stamps and therefore invited the interpretation that the policy was a statement about their (lack of) entitlements rather than about the health hazards of soda consumption. The possible health benefits of the proposal were not, in Conly’s view, likely to outweigh the harms caused by such a message.

This strategy of ruling out discriminatory paternalism in favor of policies that apply symmetrically to all citizens may seem promising, and other authors seeking to reconcile paternalism and equality have charted a similar course. Ben-Porath’s “structured paternalism” aims to promote civic equality as well as individual welfare, but it shares many features of the welfare paternalism that I have been discussing. Like Conly, she attaches an egalitarian rider to her position: paternalism “should not be directed toward the ‘other’—children, the mentally ill, those unable to realize their professed goals, the poor—but toward all citizens equitably. The decision about when

---

41 Conly, 159.
paternalism is justified should mostly be based on generalized rather than specific contexts.”

This apparent solution raises its own problems. Most importantly, the strength of the egalitarian rider and the kinds of paternalism that it rules out are left vague.\textsuperscript{43} The limitation imposed by the rider could be translated in different ways. It might mean, “Do not practice paternalism unless the policy or intervention applies to everyone equitably.” Or it could mean something much weaker: “When practicing paternalism, strive for equitable application.” The relationship between the call for egalitarian application and the justifiability of a paternalist policy needs more specification: to what extent may we trade off symmetrical application in favor of welfare benefits or increased efficiency?

Related to this is the practical difficulty of avoiding stigmatizing or expressing disrespect for the very groups on whom one is attempting to confer a welfare benefit. Many public and private programs are targeted to low-income people (and for good reasons) and so we need more guidance on how the negative messaging involved in

\textsuperscript{42} Ben-Porath, 38. Later (40), she includes paternalism directed specifically at immigrants and the elderly as further examples of “paternalist policies that are discriminatory against the ‘other,’” which fail to support civic equality and which her approach rejects.

\textsuperscript{43} Why should we treat unevenly-distributed paternalist interventions as potentially insulting or stigmatizing, when we would not have the same reaction to other kinds of unevenly-distributed beneficence? Here, it is important to emphasize that paternalism aims not only to confer a welfare benefit but to do so \textit{by} affecting the choices of the people the paternalist is trying to help. In cases where the same or a similar welfare benefit could be conferred without shaping some (and not other) people’s choices, the decision to promote welfare through paternalism rather than through other means expresses a negative judgment about the competence of the recipient.
Conly’s soda ban case may be avoided. If there is an objection to imposing restrictions on welfare recipients that are not imposed on other citizens, perhaps a fallback position is to rely instead on extending different incentives. This reply understands the objection to the food stamp case as primarily about coercion: if the state ought to provide people with food as a matter of justice, then attaching conditions or restrictions on those entitlements is coercive. But to the degree that the problem is insult and not only coercion, the egalitarian objection to extending paternalist incentives seems to remain. (Suppose, for example, that people below a specified tax bracket received tax credits in exchange for weight loss.) If this is right, then a great many of the programs that welfare paternalists propose will similarly risk stigmatizing their beneficiaries: some of the most celebrated welfarist interventions involve structuring the incentives of poor people to steer them toward better outcomes.⁴⁴ Among other things, this calls for a clearer statement by welfare paternalists of when the availability of non-paternalist options for producing the relevant welfare benefits undermines or extinguishes the justification for a paternalist intervention.

Imposing an egalitarian rider also complicates the justification for paternalism by private actors. Ben-Porath’s solution of applying polities to “all citizens equitably” is of limited help in guiding non-governmental forms of paternalism, which Ben-Porath nonetheless seems to endorse. She writes that, “through instituting default rules, the state (or employer, or other regulator) can affect levels of participation in different programs, thus advancing what seems to be preferable.”⁴⁵ But how is an employer to practice

⁴⁴ See generally Banerjee and Duflo.

⁴⁵ Ben-Porath, Tough Choices, 11. Emphasis added.
paternalism toward all people equitably when he lacks the control over other people’s choice environments that he enjoys in the case of his employees? How is a philanthropist to direct welfare paternalist interventions at “all citizens equitably?” Some kinds of paternalists lack the capacity to influence choices beyond a determinate group or finite number of people. Greater difficulties attend private and philanthropic nudging once the egalitarian objection to paternalism is recognized. To the extent that our concern is the symmetrical application of paternalist policies, governmental paternalism (with universal application) may prove a more easily or even uniquely justifiable option. If this is counterintuitive, it should underscore the point that the autonomy and equality objections to paternalism may sometimes conflict with as well as complement each other.

IV. Inequalities in the enactment of paternalist policies

Egalitarian objections to paternalism are not exclusively objections about the distribution of paternalist policies. We can imagine examples of paternalist policies (say, some bans on recreational drugs) that do not in themselves communicate an insulting or stigmatizing message about any particular groups (e.g. because the use of the relevant drugs is distributed, and known to be distributed, randomly across the population). To the extent that these conditions obtain, there is not yet any egalitarian objection to the paternalism (and I continue to bracket the question of whether there is a different, autonomy-based objection). But a program need not be targeted explicitly to a subset of citizens for inequality to be a worry: the problem can also arise on the input end, if
policies are *enacted* by a subset of citizens in ways that others do not have the opportunity to influence.\textsuperscript{46}

Welfare paternalists sometimes move from the argument that competence is a matter of where and how one is situated to the implication that paternalist policies are imposed by people who only temporarily occupy the relevant decision-making positions, and who will be affected by the policies in question just as other people are. But this reciprocity is often illusory. Even in the case of paternalism by public officials, the facts of unequal electoral influence and of money’s influence over officials once in office complicate the expectation that, in matters of paternalist policy, citizens nudge and are nudged in turn. Moreover, welfare paternalism has informed the design not only of public policy but also of corporate and philanthropic programs. In the latter cases, there is not even the expectation of reciprocity between the management personnel who decide on policies and the employees or populations affected by them.

Welfare paternalists often argue that paternalist policies should be enacted by governmental and non-governmental authorities alike. But there are important normative

\textsuperscript{46} As with the complaint against the uneven distribution of paternalist policies, it is worth clarifying why the undemocratic or nondemocratic imposition of *paternalist* policies is any more objectionable than the many other ways that people can use undemocratic or nondemocratic means of affecting each other. The specific problem in the paternalist case (whether or not it is a *worse* problem than other kinds of unequal influence) is again that the desired effects are brought about by intervening on people’s choices, with the explicit or implicit judgment that those people are not competent to make the right choices for themselves. In many economic and political interactions, one agent tries to manipulate or influence others in the belief that *her interests are different*; in the paternalist case, she does so from the view that *her judgment is better*.
differences between public and private power. In a democratic context, public power is supposed to afford people equal opportunities for influence and contestation. These equal ex ante opportunities can go some way toward justifying differential treatment ex post. But in the private sector and civil society there is not even the expectation of such opportunities for influence: employees of private firms and recipients of private charity have no meaningful opportunities to influence the paternalist interventions to which they might be subjected. It is therefore unclear why subjects of private paternalism should not regard such interventions as objectionable expressions of subordination.

Welfare paternalists often write as though a defense of public paternalism will suffice as a defense of (presumptively less normatively troubling) paternalism by private actors. The greater salience of liberty or autonomy-based objections to paternalism may explain why the normative problems with private paternalism are often overlooked. On a view that worries most about the power (including the paternalist power) of the state as a threat to freedom, the most threatening manifestations of welfare paternalism (and behavioral “nudges” more broadly) will be public. This orientation encourages close scrutiny of actions like President Obama’s Executive Order of September 15, 2015, which was titled “Using Behavioral Sciences to Better Serve the American People” and formally established a Social and Behavioral Sciences Team to help implement the Behavioral Insights Policy Directive. Welfare paternalists may focus on government paternalism in the belief that they are tackling a hard case: if we can justify paternalistic interventions even when they are backed by the coercive power or dominant, homogenizing influence of the state, surely parallel interventions will be easier to justify when deployed by less threatening actors (employers and philanthropists, for example).
On this interpretation, the decision to convert the UK government’s Behavioural Insights Team (popularly known as the “Nudge Unit”) to a social purpose company independent of the government looks like a step in the right direction. I think that this interpretation is wrong.

Consider one fairly representative argument for the greater danger posed by government paternalism. In developing their autonomy-based criticism of nudging as a threat to people’s control over their own choosing, Hausman and Welch focus on the limits that should be imposed on nudging by government. They justify this focus both on pragmatic grounds (because “it is difficult to see how to implement limits on persuasive techniques employed by non-governmental agents”) and because of the lesser threat posed by non-governmental nudging: “What makes the cacophony of invocations of irrational responses by non-governmental agents tolerable (to the extent that it is tolerable) are, we suggest, the limits to its effectiveness and the extent to which these invocations conflict with one another and cancel one another out.”47 This reasoning appears to have something like commercial advertising in mind, and may provide some comfort on that score.48 But it ignores the outsize influence and concentrated impact that some non-governmental nudgers can have over some people.

I am not convinced that even the autonomy objection justifies our worrying more about the state than about other paternalist actors. Insofar as our concern is people’s

47 Hausman and Welch, 131.

48 Although Hausman and Welch understate the degree to which collection of user data and the possibility of individual tailoring by online advertisers may permit more effective and consistent invocation of irrational responses by non-governmental agents.
autonomous control over their choosing, we should regard it as at least an open question whether that control is more threatened by government manipulation (in the form of default options in social policy and government advertising) or by employers who control the choice environment in which people may spend more than half their waking hours. It is very likely that government has the largest overall impact on people’s choices; but this is compatible with the possibility that each individual’s choices are more often influenced by nudges from employers or other non-governmental actors. (Similarly, government may have the single largest ability to shape the educational environment of the nation’s children, but individual children’s educational outcomes may depend more on decisions by parents.) The effects of non-governmental nudging may cancel each other out in aggregate without doing so at the level of individual control over choice (Hausman and Welch’s stated object of concern).

Once we take account of the egalitarian objection to paternalism, it should be still clearer that we should worry about private as well as governmental paternalism. Governmental paternalism often comes about in undemocratic ways. But private paternalism depends for its success on some people’s power over others. Hausman and Welch themselves gesture at this concern about private power:

Even when unshaped choices would have been just as strongly influenced by deliberative flaws, calculated shaping of choices still imposes the will of one agent on another. Suppose, for example, that, with the help of a consulting behavioral economist, an employer is able to structure the defaults, the contribution timing, and the framing of a retirement plan so as to achieve a very high contribution rate from a large majority of employees. When the employer
and the consultant get together to congratulate themselves on engineering the situation so that the employees chose in just the way that the employer had planned for them to choose, they are celebrating their power over the employees.\textsuperscript{49}

They do not pursue this line of thought, treating this celebration of unequal power mainly as an interesting feature to note, or perhaps flagging it as something a little unseemly. They do not address the possibility that important normative differences separate the unequal power exercised by the employer over her employees and the power exercised by a public official who was selected through meaningfully democratic procedures. If any forms of welfare paternalism can survive egalitarian objections, we should expect them to be ones enacted by democratic governments and applied symmetrically to all residents. This goes against the grain of most “libertarian paternalist” argument, which imports into welfare paternalism not just a rejection of coercion but also a distrust of government and greater sanguinity about private power. It is this strain in welfare paternalism that egalitarian objections should challenge going forward.

This is not to say that governmental paternalism is always or even generally justifiable, or that the same responses to paternalistic actions are appropriate whether the paternalist is a public or private actor. The considerations discussed here provide both public and private actors with reasons to avoid acting paternalistically, because of the insult that paternalistic actions and relationships express. I have further suggested that that insult can be understood as a violation of\textit{equality} as well as of autonomy. But here some more caution is needed regarding the moral\textit{force} of a charge of paternalism. Just as

\textsuperscript{49} Hausman and Welch, 133.
the democratic character of some interventions may help to rebut or mitigate charges of paternalism, there may also be cases where public actors should be held to a higher standard. The obligation that the state has to treat all citizens as equals is something that should be secured legally to a greater extent than the (moral) obligation of private actors and citizens to treat each other as equals. While it may be wrong to fail to treat friends, employees, and fellow citizens as one’s moral, social, and political equals, these ethical judgments do not automatically carry over into judgments about the appropriate regulatory structures. Private actors may have the right to act paternalistically, just as they may have the right to insult people or express disrespect in other ways. For this reason—as well as from respect for the person paternalized in cases where there is no defect in consent—the appropriate response to private paternalism will often not be legal or punitive. Rather, public criticism and perhaps other kinds of incentives are likely to represent a better way of reconciling commitment to social equality with respect for rights and for choice. My focus here has been on the normative weight of paternalism and the reasons that agents have to avoid acting paternalistically: arguments about paternalism’s normative force, and about the kinds of reasons that a charge of paternalism gives other agents or public authorities, will depend on contextual details.50 My concern here is to show how taking egalitarian objections to paternalism seriously lets us see that the differences between public and private paternalism are more significant and complicated than welfare paternalists often suggest.

Conclusion

Whether or not one thinks that autonomy presents a convincing objection to paternalist interventions, or whether or not one thinks that welfare paternalism has convincing answers to objections from autonomy, I have suggested that a second standard of evaluation should be applied to recent articulations of welfare paternalism. I have tried to articulate a distinct egalitarian objection to paternalism; to explain the welfare paternalist response to that objection; and to show some important ways that welfare paternalism falls short of meeting the egalitarian objection. We might be tempted to regard these latter cases as anomalies, or as cases of permissible techniques used with “impermissible motivations.” I have instead tried to show that the asymmetric application of paternalist policies, and the uneven distributions of private power that permit some people and not others successfully to enact paternalist policies, are not incidental features of some paternalist interventions. They are the things that make paternalist interventions possible, efficacious, and politically feasible.

The normative importance of paternalism is linked not only to questions of moral status but to ones of social and political status as well. Because of this, the tension between paternalism and egalitarian commitments goes deeper than is generally recognized. Paternalistic behavior occurs against the background of hierarchical and often unjustly unequal relationships. In acknowledging this, the egalitarian objection differs from autonomy objections to paternalism, which generally point to behavior that is always wrong (independent of contextual factors or background unfairness) and call simply for paternalistic behavior to be avoided. If we are committed to the importance

---

both of welfare improvements and of egalitarian social and political relations, then the problem is not resolvable simply by abandoning paternalistic behavior, without further changes to background social and political institutions. In addition to calling for the avoidance of paternalism where possible, and calling attention to the trade-offs that welfare paternalist policies involve, we should recognize that interventions on individual choices are inadequate substitutes for the egalitarian social and political relations that would make influence without insult a possibility.