In Defense of Algorithmic Governance

Adam Lovett

Forthcoming in *Free & Equal: A Journal of Ethics & Public Affairs*

Abstract: Government bureaucracies routinely make decisions about individual cases: whether to grant an applicant benefits, whether to investigate someone for rule-breaking, or what costs to impose on someone for a crime. In the twentieth century, this was all done by individual human decision-makers: caseworkers, inspectors, judges. Now, it is increasingly done by algorithms. Scholars have been largely critical of this change. In this paper I defend it. My main novel line of argument is that using algorithms to make decisions reduces domination. It replaces the tyranny of street-level bureaucrats with the non-dominating influence of algorithms. Additionally, I address broad concerns about such algorithms: that their use violates people's rights to explanation, that they inadmissibly rely on statistical evidence, and that they generate responsibility gaps. The result is a general defense of algorithmic governance.

Keywords: Algorithms; public administration; bureaucracy; domination; right to an explanation; responsibility gaps

I. Introduction

Imagine that you are a British worker in the mid-1970s who has just lost their job. To claim unemployment benefit, you start by going to the local Employment Exchange. There you are assigned a caseworker, who interviews you. They ask you whether you left voluntarily, whether you're looking for work, what other sources of income you have. They decide whether you get unemployment benefits. None of that happens anymore. Now, if you lose your job in the UK, you go online. You fill in a form to apply for Universal Credit.[1] An algorithm makes an automatic determination about whether you are eligible for benefits. If you later get a job, and your income goes up, your benefits are automatically

[1] GOV.UK, "Universal Credit," accessed August 27, 2026, https://www.gov.uk/universal-credit.

reduced. If your income goes down, they are increased. No human being need be involved in evaluating these applications. The British government has replaced a human decision-maker with an algorithm.

Here's another example. Consider what happens when you submit a tax return to the Australian Taxation Office (ATO). If you did this in the 1980s, a human tax collector would read your return. They'd try to work out whether you were underreporting income. If they thought you were, you'd get audited. Again, that no longer happens. Now a machine learning algorithm reads your return.[2] This algorithm has been trained on data on what past tax returns contained and whether, upon audit, they were found to underreport income. You likely suffer an audit if and only if the algorithm flags your return. A human decision-maker has, again, been replaced with an algorithm.

Let's look at a third case. In the U.S., you are convicted of serious crimes by a jury after a trial. You are then sentenced, at another hearing, by a judge. Since the 1920s, probation officers have helped the judge determine the sentence. After conviction they will interview you and put together a presentence investigation report (PSI). This typically includes an estimate of how likely you are to reoffend. For most of the twentieth century, this estimate would be a qualitative expression of the probation officer's professional judgment. Since the 1990s, in many states this has been replaced by a quantitative risk score.[3] In Wisconsin, as in many other states, this score is generated by a tool called COMPAS.[4] COMPAS is a logistic regression model. Its output is a predicted probability of re-offense, and its inputs are variables such as the convict's age, gender, criminal history, employment status and so on. In the model, the input variables relate linearly to a function of the output variable. In this case, the judgment of a probation officer has been replaced with the output of an algorithm.

These cases illustrate the increasing use of algorithms in the "decision-making factories" of government.[5] These are the parts of government that make

---

[2] Tess Bennett, "ATO Captures Billions of Dollars from Tax Cheats with AI," *Australian Financial Review*, July 31, 2023, https://www.afr.com/technology/ato-captures-billions-of-dollars-from-tax-cheats-with-ai-20230727-p5drnf.

[3] Center on Juvenile and Criminal Justice, *The History of the Pre-Sentence Investigation Report* (2008), https://www.cjcj.org/media/import/documents/the_history.pdf.

[4] Andrew Lee Clark, "Injustice Ex Machina: Predictive Algorithms in Criminal Sentencing," *UCLA Law Review*, February 19, 2019, https://www.uclalawreview.org/injustice-ex-machina-predictive-algorithms-in-criminal-sentencing/.

[5] Mark Bovens and Stavros Zouridis, "From Street-Level to System-Level Bureaucracies: How Information and Communication Technology Is Transforming Administrative Discretion and Constitutional Control," *Public Administration Review* 62, no. 2 (2002): 174–184, https://doi.org/10.1111/0033-3352.00168.

individual judgments about concrete cases. Where previously human beings would make these decisions, increasingly algorithms do. Notice two things about these examples. First, they illustrate how such decisions fall into three main classes. Government bureaucracies often decide whether to bestow a specific benefit on individuals, whether to investigate individuals for rule-breaking and what costs to impose on rule-breakers. These are the three parts of the decision-making factories of government. Second, governments use many different kinds of algorithms to make these determinations. The algorithm that determines eligibility for Universal Credit is a decision tree. It is, in essence, a collection of if-and-only-if rules. The algorithm used in sentencing is a logistic regression model. The one used by the ATO is a machine learning algorithm. I am interested in all these cases. My general question is: to what extent should we support or oppose the increased use of such diverse algorithms in the decision-making factories of government?

Scholars have, by and large, been critical of this change.[6] Scholars say that the use of algorithms in public administration discriminates against citizens,[7] violates their rights to an explanation,[8] means government decisions rely on inadmissible evidence[9] and creates responsibility gaps.[10] The preponderant view in the scholarly literature is that we should prevent the further adoption of such algorithms and get rid of those that have already been adopted. We should oppose what I'll call

---

[6] For notable exceptions, see Jon Kleinberg et al., "Human Decisions and Machine Predictions," *The Quarterly Journal of Economics* 133, no. 1 (2018): 237–293, https://doi.org/10.1093/qje/qjx032; Cass R. Sunstein, "Governing by Algorithm? No Noise and (Potentially) Less Bias," *Duke Law Journal* 71, no. 6 (2022): 1175–1205.

[7] Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (St. Martin's Press, 2018); Clinton Castro, "What's Wrong with Machine Bias," *Ergo, an Open Access Journal of Philosophy* 6, no. 15 (2019): 405–426, https://doi.org/10.3998/ergo.12405314.0006.015.

[8] Madalina Busuioc, "Accountable Artificial Intelligence: Holding Algorithms to Account," *Public Administration Review* 81, no. 5 (2021): 825–836, https://doi.org/10.1111/puar.13293; cf. Kate Vredenburgh, "The Right to Explanation," *Journal of Political Philosophy* 30, no. 2 (2022): 209–229, https://doi.org/10.1111/jopp.12262.

[9] Sune Holm, "Statistical Evidence and Algorithmic Decision-Making," *Synthese* 202, no. 1 (2023): 28, https://doi.org/10.1007/s11229-023-04246-8; David Gray Grant, Jeff Behrends, and John Basl, "What We Owe to Decision-Subjects: Beyond Transparency and Explanation in Automated Decision-Making," *Philosophical Studies* 182 (2025): 55–85, https://doi.org/10.1007/s11098-023-02013-6; Hugo Cossette-Lefebvre and Jocelyn Maclure, "AI's Fairness Problem: Understanding Wrongful Discrimination in the Context of Automated Decision-Making," *AI and Ethics* 3, no. 4 (2023): 1255–1269, https://doi.org/10.1007/s43681-022-00233-w.

[10] Andreas Matthias, "The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata," *Ethics and Information Technology* 6, no. 3 (2004): 175–183, https://doi.org/10.1007/s10676-004-3422-1; Filippo Santoni de Sio and Giulio Mecacci, "Four Responsibility Gaps with Artificial Intelligence: Why They Matter and How to Address Them," *Philosophy & Technology* 34, no. 4 (2021): 1057–1084, https://doi.org/10.1007/s13347-021-00450-x.

"algorithmic governance"—the use of algorithms to make administrative decisions on individual cases.

In this paper, I defend algorithmic governance. My main argument in its favor hinges on the vices of what I'll call "caseworker governance"—the use of human beings to make administrative decisions on individual cases. Under this model, citizens are subject to the discretion of individual bureaucrats: of caseworkers, tax collectors, probation officers. This gives caseworkers deeply asymmetric power over the people whose cases they decide. Such power asymmetries are dominating. Algorithmic government can alleviate such domination by taking human discretion out of the process. In sections II to V, I will develop this argument. In section VI, I will address some of the more general concerns people have had about the use of algorithms in government. My goal is to provide some corrective to the current literature on the use of algorithms in government. That literature seems to me overly negative: it ignores some virtues of algorithmic governance and overstates its vices. I believe a more positive picture better reflects reality.

II. Normative foundations

Let's start by laying the normative foundations for the argument in favor of algorithmic governance. In recent years, many philosophers have come to think that power asymmetries are distinctively objectionable.[11] The initial motivation for this view comes from our intuitive judgments about cases. Consider the relationship of master to slave, king to subject, or husband to wife in a patriarchal marriage. These relationships are ones of domination or subordination. Such relationships seem objectionable. These relationships are characterized by

[11] For neo-republican versions of this thought, see Philip Pettit, *Republicanism: A Theory of Freedom and Government* (Oxford University Press, 1997); Frank Lovett, *A General Theory of Domination and Justice* (Oxford University Press, 2010). For its relational egalitarian foundations, see Elizabeth S. Anderson, "What Is the Point of Equality?" *Ethics* 109, no. 2 (1999): 287–337; Samuel Scheffler, "What Is Egalitarianism?" *Philosophy & Public Affairs* 31, no. 1 (2003): 5–39. For applications of the idea to democracy, see Thomas Christiano, *The Constitution of Equality: Democratic Authority and Its Limits* (Oxford University Press, 2008); Philip Pettit, *On the People's Terms: A Republican Theory and Model of Democracy* (Cambridge University Press, 2012); Niko Kolodny, "Rule Over None II: Social Equality and the Justification of Democracy," *Philosophy & Public Affairs* 42, no. 4 (2014): 287–336; Daniel Viehoff, "Democratic Equality and Political Authority," *Philosophy & Public Affairs* 42, no. 4 (2014): 337–375. For applications to workplace relations, see Elizabeth Anderson, *Private Government: How Employers Rule Our Lives (and Why We Don't Talk about It)* (Princeton University Press, 2017); Iñigo González-Ricoy, "Little Republics: Authority and the Political Nature of the Firm," *Philosophy & Public Affairs* 50, no. 1 (2022): 90–120, https://doi.org/10.1111/papa.12205. For the richest recent development of the idea, see Niko Kolodny, *The Pecking Order: Social Hierarchy as a Philosophical Problem* (Harvard University Press, 2023).

asymmetries of power. Masters have asymmetric power over their slaves. That makes the relationships subordinating. Thus, the simplest lesson to draw from our qualms about such relationships is that we have an objection to being subjected to power asymmetries. We have a complaint against being under another person's asymmetric power.

What is the nature of this objection? It is not simply that power asymmetries have bad consequences. To see that, consider benevolent, competent dictatorships. Suppose you could replace the democratic institutions of your country with such a dictator. Imagine you were sure that this dictator would enact better policies than your elected governments would. I think, intuitively, there is a weighty objection to doing this. But the objection isn't instrumental. If you established the dictatorship, policy would get much better. That would be instrumentally good. So there must be something intrinsically objectionable about such a dictatorship. The hypothesis on which this paper rests is that thus one has a basic right against subordination. The idea is that you have several fundamental rights: rights to physical integrity, to life, to free expression. The right to non-subordination is among these rights.[12]

Philosophers have generally talked about this right with two different languages. On the one hand, neo-republicans talk about it in terms of freedom.[13] The thought is that domination is a kind of unfreedom. When someone has asymmetric power over you, that makes you unfree.[14] You have a right against being subjected to this kind of unfreedom, or a right to freedom as non-domination. On the other hand, relational egalitarians talk about it in terms of equality.[15] The thought is that domination is a kind of inequality. When someone has asymmetric power over you, you are subjected to an inegalitarian relationship. You have a right against being subjected to such a relationship. Both neo-republican and relational egalitarian approaches identify the same kind of right: a right against being subjected to power asymmetries. The difference between these two ways of conceptualizing the right won't matter in this paper. What matters is the idea that we have a claim against being subjected to asymmetries of power.

This idea needs some spelling out. Most importantly, let's clarify power. The relevant notion of power is the influence one person can have over another. When

---

[12] For this way of thinking of non-subordination's import, see especially Kolodny, *The Pecking Order*.

[13] Pettit, *Republicanism*; Pettit, *On the People's Terms*; F. Lovett, *A General Theory of Domination and Justice*.

[14] Republicans usually say that when someone can "arbitrarily" influence you, they have the objectionable kind of power over you. See, e.g., Pettit, *Republicanism*, 55; F. Lovett, *A General Theory of Domination and Justice*, chap. 4. This idea is captured by the discussion of costs below.

[15] E.g., Anderson, "What Is the Point of Equality?"; Scheffler, "What Is Egalitarianism?"

I have much more power over you than you have over me, you are subjected to a power asymmetry. There are two ways of understanding such power. On the one hand, we might think of it as the ability to affect someone's well-being.[16] I have power over you if I can affect how well your life goes. If I can, for example, impoverish you by denying you benefits, then this gives me power over you. On the other hand, we might think of power as the ability to affect someone's behavior.[17] I have power over you if I can affect how you act. If I can get you to search for a job, then this increases my power over you. Plausibly, both abilities are forms of power. Power, on this view, consists in being able to affect how people act and what their lives are like.

To what extent does one person have power over another? For practical purposes, we can understand this largely in terms of conditional probabilities. I have power over your behavior when my trying to get you to do something makes you more likely to do it. If my trying to get you to attend job interviews, say, increases the likelihood you do attend job interviews, I have power over you. The more my trying to get you to do things increases the probability you do them, the more power I have over you. The same idea extends to well-being. The more my trying to make your life a certain way increases the chances that your life is that way, the more power I have over you. One corollary of this that will matter later is that if a third party can disrupt the causal link between how I try to influence someone and how they are influenced, then this reduces my power over the person I can influence. If I can unilaterally determine what happens to someone, I usually have a lot of power over them. If I need someone else's say-so to affect them, then that usually diminishes my power over them.

It is natural to also link degrees of power to costs.[18] Suppose I can get you to give me your wallet by pointing a gun at you, but I would be arrested immediately after doing so. Here it is very costly for me to get you to give me your wallet. Intuitively, that diminishes the extent to which I have power over you. However, such costs

---

[16] For this view, see Steven Lukes, *Power: A Radical View,* 2nd ed. (Palgrave Macmillan, 2005), 29–38.

[17] For this view, see Robert A. Dahl, "The Concept of Power," *Behavioral Science* 2, no. 3 (1957): 201–215.

[18] This is a common view. See, e.g., John C. Harsanyi, "Measurement of Social Power, Opportunity Costs, and the Theory of Two-Person Bargaining Games," *Behavioral Science* 7, no. 1 (1962): 67–80, https://doi.org/10.1002/bs.3830070105; Alvin I. Goldman, "Toward a Theory of Social Power," *Philosophical Studies* 23, no. 4 (1972): 221–268; Sean Ingham, *Rule by Multiple Majorities: A New Theory of Popular Control* (Cambridge University Press, 2019); Adam Lovett, "The Egalitarian Objection to Coercion," *Pacific Philosophical Quarterly* 105, no. 3 (2024): 392–417, https://doi.org/10.1111/papq.12463.

have to be carefully understood. They are not opportunity costs.[19] For suppose somebody offers you a million dollars to free your slave, and you decline. The opportunity cost of keeping them enslaved is high, but you still have a lot of power over the slave. You still control the slave's life. Instead, I suspect that the relevant notion of costs should be understood in terms of ill-being. If my trying to get you to do something means I suffer a lot of intrinsic bads—pain, distress, imprisonment—then my ability to get you to do the thing gives me relatively little power over you.[20] Thus, if someone threatens to shoot you unless you free your slave, that reduces your power over your slave. You retain influence over the slave only at the cost of suffering substantial intrinsic bads. Such influence is less subordinating than that which you can exercise costlessly.

This spells out what it is to have (asymmetric) power over someone. There is a further question about just how objectionable different power asymmetries are. A key point here is that power asymmetries aren't anodyne just because they are impersonally justified.[21] Power asymmetries aren't unobjectionable simply because they conduce to the general good or the better protection of people's other rights. We can see this in benevolent, competent dictatorships. This dictator enacts better policies than a democracy would. That impersonally justifies the dictator's power. But there is a weighty objection to their rule. Impersonal justification does not make power asymmetries wholly anodyne.

I want to make two more positive suggestions about the shape of our objection to power asymmetries. On the one hand, concentrated power seems more objectionable than dispersed power. Suppose power is concentrated in the hands of one dictator. This dictator has enormously asymmetric power over each of their subjects. The claim here is that this is worse than if power is dispersed among many people.[22] The latter is not completely anodyne, but it is less objectionable

---

[19] I take this to be the lesson in Arash Abizadeh, "Power, Costs, Collective Action, Bargaining, and Solidarity," *American Journal of Political Science* 70, no. 2 (2026): 441–452, https://doi.org/10.1111/ajps.12953.

[20] For a counterpart view about harms, see Elizabeth Harman, "Harming as Causing Harm," in *Harming Future Persons: Ethics, Genetics and the Nonidentity Problem*, ed. Melinda A. Roberts and David T. Wasserman (Springer Netherlands, 2009), 137–154, https://doi.org/10.1007/978-1-4020-5697-0_7. For more on ill-being, see Shelly Kagan, "An Introduction to Ill-Being," in *Oxford Studies in Normative Ethics, vol. 4*, ed. Mark Timmons (Oxford University Press, 2014), 261–288, https://doi.org/10.1093/acprof:oso/9780198722144.003.0012.

[21] For a similar point, see Jake Zuehl, "Kolodny Against Hierarchy," *Philosophy & Public Affairs* 52, no. 4 (2024): 565–595, 590–591, https://doi.org/10.1111/papa.12273. For the contrary view, see Kolodny, *The Pecking Order*, 131–134.

[22] An Associate Editor suggests that it is important that the many people aren't working as a team, in the sense of coordinating with one another to influence those they have power over. I am neutral on this, but (as they also helpfully point out) when I rely on this suggestion in section IV, the idea is that power is not dispersed among people working as a team.

than concentrated power. On the other hand, enduring power seems more objectionable than temporary power.[23] Suppose one person has held supreme office for a very long time. Our objection to their power over us gets weightier the longer they have held office. It is less objectionable, in other words, for an office to be held by different people for much shorter lengths of time. These positive suggestions won't be entirely essential to my main argument, but they motivate specific features of the institutional framework I propose in section IV.

One might wonder how plausible these positive suggestions are. Is it really less objectionable to have two masters rather than one? Or to always have a master, but a different one each day? To reply to this worry, I reiterate that dispersed, temporary power is still objectionable. The suggestion is just that it is less objectionable than concentrated, enduring power, not wholly unobjectionable. With that in mind, the positive suggestions seem to me plausible. Their appeal can be seen, for example, in how we think about juries. One problem with inquisitorial legal systems is that they give judges concentrated power; jury systems disperse that power among more people. And, as we'll see in section IV.C, these positive suggestions help explain what's good about being able to contest administrative or judicial decisions. Contestability disperses adjudicative power. Similar points go for temporality: it is good to have different adjudicators because it makes the power of each one less enduring. So I'll assume these positive suggestions in the rest of the paper.

That completes the elucidation of the normative premise on which the argument in this paper relies. Let us now turn to some defense of that premise.

II.A Are power asymmetries objectionable?

In this section, we look at three prima facie counterexamples to the claim that power asymmetries are generally objectionable. First, the relationship of elected representatives to ordinary citizens in a democracy is clearly one of asymmetric power. Our representatives have asymmetric power over us. Yet this, one might think, needn't be objectionable. Second, the relationship of parent to child is one of asymmetric power. Yet parent-child relationships seem anodyne. Third, the relationship between professor and student is one of asymmetric power. Yet students don't seem to inevitably have a weighty objection against these relationships. I believe these counterexamples can be addressed, buttressing the claim that asymmetries of power really are objectionable.

---

[23] This is similar to Kolodny's "Episodic Character" tempering factor. See Kolodny, *The Pecking Order*, 99–100.

Let's start with the representative-citizen relationship. The key point is that, in well-functioning democracies, elected representatives are under the control of ordinary citizens. They can be voted out of office if they exercise their power in a way voters disapprove of. These relationships of control matter because they make it costly for representatives to do things of which their constituents disapprove. Losing an election usually means one suffers various bad things: humiliation, desire frustration, distress, loss of income. The more costly it is to do something, the less power one gets from the ability to do that thing. So the tighter is the control ordinary citizens have over representatives, the less power representatives have over ordinary citizens. Of course, in real-world democracies these relationships of control are rarely watertight. Still, in principle there is a way to reconcile representative democracy with the claim that power asymmetries are objectionable.[24]

We now turn to the parent-child relationship. Children, in some ways, have less important rights than adults. Adults, for example, have a weighty right to physical integrity. It would be wrong to force an adult to undergo a life-saving operation that they refuse to have. But if a five-year-old refuses such an operation, we may force them to have it. Perhaps children have a right to physical integrity of less weight than adults, and so one can overrule this right in their own interests. Or perhaps they have a right to physical integrity of the same weight as adults, but one that can be waived by their guardians.[25] Either way, we may infringe their right to physical integrity in order to promote their interests. As with physical integrity, so with non-subordination. Adults have a weighty right to non-subordination, but children do not have exactly this right. Perhaps they have a right to non-subordination of lesser weight or perhaps one that can be waived by appropriate guardians. Either way, it is permissible for their parents to have power over them. So parent-child relationships are unobjectionable.

Let us turn to the relationship of professor to student. Remember that we are thinking about power asymmetries in terms of rights. The crucial observation is that rights can generally be waived. I have a right to physical integrity. But if I voluntarily step into the boxing ring, I waive this right. If I validly consent to you

---

[24] I make the general point here in Adam Lovett, "Must Egalitarians Condemn Representative Democracy?" *Social Theory and Practice* 47, no. 1 (2021): 171–198, as did Sean Ingham almost contemporaneously in his "Representative Democracy and Social Equality," *American Political Science Review* 116, no. 2 (2022): 689–701, https://doi.org/10.1017/S000305542100109X. Neither of us construes the import of control explicitly in terms of costs as I do here. But nothing in my argument in this paper hinges essentially on the reference to costs. I simply (now) think that the costs story is the most elegant way of understanding why control matters. We could just as effectively think of control as a sui generis tempering factor in the style of Kolodny's *The Pecking Order*.

[25] For this view, see Joel Feinberg, *Harm to Self* (Oxford University Press, 1986), 325–326.

punching me, you do me no wrong when you do so. I suggest that professor-student relationships are unobjectionable when the students validly waive their right against subordination. When students validly consent to being subjected to professorial power, the power asymmetry does not violate their rights. When adult students join most classrooms—when they go to university, say—they often do validly waive their right against being subjected to professorial power. They validly consent to their professor's telling them what to read and what work to do. So these power asymmetries are not objectionable.

How might one resist this idea? One might think that to validly consent to something, one must have adequate alternative options, or, at the least, one is wronged when one is forced to accept a rights-infringement due to the absence of such options. To avoid this, one must have options that don't leave one too badly off in absolute terms, and in which one avoids the treatment one consents to. If I step into the boxing ring with you only because I'll be shot if I don't, or because otherwise I will have to fight someone else, I do not validly consent to the fight; or, at the least, leaving me with no adequate alternative to being punched wrongs me. The concern is that prospective students don't have an adequate alternative to being subjected to professorial power. To avoid professorial power, they would have to skip college altogether. This, one might think, will leave them too badly off for it to be an adequate option. So professorial power is objectionable after all. We haven't really escaped the objection.

To reply to this worry, we begin by noting that in most countries it is feasible to have a decent life without attending higher education. Non-attendance shuts one off from certain career opportunities, but one is still able to start one's own business, pursue a trade, or work in many technical roles. And life is anyway about more than career opportunities. So, in actuality, I think most students do have an adequate alternative to being subjected to professorial power. Still, when this is not true, I think there is a problem with the power that universities—and professors—have over their students. In these cases we have weighty reason to incorporate a student voice in the governance of some universities and to soften hierarchies within individual classes. If we do so, then students have the option of getting a relatively non-hierarchical education. This, it seems to me, leaves us with an intuitively plausible position: professorial power isn't inevitably objectionable, but it can be in certain contexts. When students don't have any option but to attend a hierarchical, autocratic university, they have a complaint.

Let me look at one further worry about my position. I have claimed that the right against being subordinated is alienable: it can be waived. One might worry that

this implies one can sell oneself into slavery. Some think this is implausible.[26] Fortunately, this is not an implication of my claim. The key point is that rights are temporally indexed. Consider physical integrity. When I waive a right to physical integrity at some time, touching me at that time isn't objectionable. But I cannot generally waive my right to physical integrity at future times. If I tell you I'll fight you tomorrow, but tomorrow I withdraw my consent, it would be wrong for you to punch me at the later time. The same goes for non-subordination. If I waive my right against being subjected to asymmetric power at a time, your asymmetric power over me at that time doesn't wrong me. But that doesn't mean I can now waive my right against being subjected to asymmetric power at later times. Thus, even if I now agreed to be your slave, I could withdraw that consent in the future. If I do, my continued enslavement is impermissible. So one cannot sell oneself into slavery in the way we would usually think of it.

We can, thus, address a fairly wide swathe of prima facie counterexamples to the idea that power asymmetries are objectionable. It is perhaps worth making clear where, in my view, that leaves our epistemic situation vis-à-vis this claim. I think this claim is a simple, elegant explanation of our judgments about a variety of cases, and prima facie contrary judgments can be undermined. That does not mean the claim is decisively established or that there is no evidence against it. There almost surely is evidence against it: there are, at the least, cases where we do not seem especially troubled by power asymmetries.[27] This just illustrates how ethical theorizing is a messy business. There are judgments that conflict with any theory. The task of such theorizing, as I see it, is to identify the simplest theory that conforms with the widest array of weightiest judgments. I believe that the claim that power asymmetries are objectionable will likely be part of such a theory, since it seems indispensable to fully explaining our disquiet about relationships like slavery, kingship or patriarchal marriage. So, in the rest of the paper, we will take this claim as a normative premise. My main goal is to explore its implications for the administrative state.

I'll do one final thing before moving on to this exploration: situate my approach to equality in the recent literature. The most influential recent articulation of the

---

[26] Rousseau, for example, famously finds it implausible (contra Grotius). See Jean-Jacques Rousseau, *The Social Contract*, trans. Maurice Cranston (Penguin Books, 2004), bk. 1, chap. 4.

[27] An Associate Editor, for example, suggests that it doesn't seem non-instrumentally better in any respect for what happens to me to be random, rather than for it to be determined by someone who loves me and therefore cares deeply about my well-being. Here we lack an intuition that might confirm the claim that power asymmetries are objectionable. That, it seems to me, is indeed some (weak) evidence against it. I take the evidence to be only weak, partly because I think here we lack a judgment that confirms the theory, rather than have a judgment that disconfirms it, and partly because this lack might be explained away by the intrinsic value of, say, loving relationships. But weak evidence is still evidence against my claim.

idea that we have a right against subordination is Kolodny's *The Pecking Order*.[28] Kolodny suggests that power asymmetries are less objectionable when allayed by tempering factors. He advances fourteen tempering factors in total. These factors often seem motivated by avoiding the kind of counterexamples discussed in this section.[29] I regard this as somewhat ad hoc. My approach, in contrast, rests on claims about what power is and how rights work—for instance, that power is reduced by costs or that rights can generally be waived. These claims are independent of the idea that power asymmetries are objectionable. Their motivation doesn't consist solely in the goal of defusing counterexamples to this idea. This seems to me preferable to treating such claims as *sui generis*. Still, the difference between my approach and Kolodny's is not crucial to the argument in this paper. On either approach, one can articulate a problem with bureaucratic domination. Let us turn to that problem now.

III. The problem of bureaucratic domination

Under caseworker governance, human beings make decisions about individual cases. The problem with this model is that it creates deep asymmetries of power between caseworkers and the people whose cases they decide. To see this, simply consider the caseworker who determines whether a jobless worker gets unemployment benefit. This caseworker decides whether the worker can pay rent that month or afford groceries. This gives them enormous power over the worker. The same goes for the relationship of probation officer to convict and tax collector to taxpayer. The probation officer has a big impact on how long the convict will stay in jail. The tax collector decides whether a taxpayer will suffer a costly audit. Generally, caseworker governance creates asymmetries of power between caseworkers and those whose cases they decide. If asymmetries of power are objectionable, then this is objectionable. I will call this the "problem of bureaucratic domination."

Let us see in a bit more detail exactly how the caseworker has power over the people whose cases they decide. To begin with, caseworkers have a big influence on how their subjects behave. If a caseworker at the Employment Exchange tells you to look for work, you will likely look for work. If your probation officer bars you from being out after 10 p.m., you will likely stay home. Caseworkers have this power because they have a lot of influence over their subjects' well-being. The caseworker at the Employment Exchange can stop your unemployment benefits, impoverishing you. Your probation officer can send you back to jail. So if power consists in either the ability to influence someone's behavior or in the ability to

[28] Kolodny, *The Pecking Order*.
[29] E.g., Kolodny, *The Pecking Order*, 98–99.

influence their well-being (or both), then caseworkers have a lot of power over the ordinary citizens they deal with.

Is the power of caseworkers really objectionable? One might suggest it is not, on the grounds that it is impersonally justified. Having caseworkers make decisions might have good consequences. But in the previous section we saw that this does not disable the objection to power asymmetries. The power of a benevolent dictator would still be objectionable. On the other hand, one might think that caseworkers' power is anodyne because citizens validly consent to being subjected to it. But this seems uncompelling in most cases. Many bureaucrat-citizen relationships are clearly involuntary. Convicts don't have a choice of probation officer. Taxpayers don't have a choice of tax collector. And most benefit applicants don't have a decent alternative option to applying for benefits. Forgoing such benefits usually means being unable to meet one's basic needs. Whether or not such involuntariness invalidates consent, it is plausibly wrong to leave someone no adequate option but to suffer a right infringement. So neither of these points suffices to make bureaucratic power anodyne.

A different idea is that the power of caseworkers is dispersed or temporary. In section II, I said that both mitigate our objection to power asymmetries. However, mitigation is not eradication. Concentrated, enduring power asymmetries are more objectionable than dispersed, temporary ones, but the latter are still objectionable. So this wouldn't wholly solve the problem of bureaucratic domination. More deeply, it is not really true that the power of caseworkers is very dispersed or temporary. For dispersal what matters is how much power each caseworker has over any individual citizen. Each caseworker might have power over relatively few citizens (only their clients), but their power over those citizens is very substantial. So the power a caseworker has over an individual citizen is concentrated. In many cases, it is also not very temporary. One usually keeps the same probation officer for a long time. Decision-making about whether one gets benefits often takes months. So neither dispersal nor temporariness makes actual caseworker power wholly anodyne.

A final idea is that the import of costs domesticates bureaucratic power. The key observation is that bureaucrats, and caseworkers specifically, are embedded in a hierarchical organization. Street-level bureaucrats are under the control of their managers; managers are under the control of senior civil servants; senior civil servants are under the control of elected officials. If power is reduced by control, this reduces the extent to which any bureaucrat has power over ordinary citizens. More specifically, if managers can fire or demote caseworkers for doing the job in a way they dislike, they can impose costs on them. This reduces the power of caseworkers. If senior civil servants can impose costs on managers and elected officials can impose costs on senior civil servants, no bureaucrat has deeply

asymmetric power over ordinary citizens. The traditional hierarchical structure of a Weberian bureaucracy, in theory, solves the problem of bureaucratic domination.

Unfortunately, what works in theory works less well in practice.[30] There are two problems. The first is one of observability.[31] Street-level bureaucrats do many things when adjudicating on a case. They make judgments about the issues. They decide how good the subject's evidence is. They speak to their colleagues. They tell subjects what they need to achieve success. Much of this involves interpersonal interactions. Managers cannot observe exactly how bureaucrats manage these interpersonal interactions. They can monitor the final adjudication and demand that bureaucrats document their reasons in support of the final decision. But bureaucrats can obfuscate their motivations and omit documentation. Yet when managers cannot observe what their subordinates are doing, they cannot control what they're doing. The impossibility of perfectly monitoring street-level bureaucrats means that such bureaucrats retain a lot of power over subjects even in hierarchical organizations.

The second issue is one of capacity.[32] The traditional way of setting up a bureaucracy is meant to solve the problem by ultimately putting control over bureaucrats in the hands of elected representatives. The problem is that elected representatives have very limited capacity to control administrators. This problem arises from the number of elected officials relative to bureaucrats. In the U.S., for example, there are 536 elected officials at the federal level and three million federal bureaucrats. The issue is that to control how someone does their job, you need to devote some time to controlling them. You need to learn what they're doing, decide what you want them to do and impose sanctions on them for going against your wishes. Elected officials can devote only a minuscule amount of time to overseeing each bureaucrat in this manner, because there are so few elected officials and so many bureaucrats. They may be able to oversee senior civil servants, but their oversight quickly seeps away when it comes to street-level bureaucrats. They simply lack the time to effectively oversee even, say, the 60,000 employees of the Social Security Administration. So they lack the capacity to exert meaningful control over caseworkers.

---

[30] For similar points, see Henry S. Richardson, *Democratic Autonomy: Public Reasoning about the Ends of Policy* (Oxford University Press, 2002), 4–8; Pettit, *On the People's Terms*, 235–238; Joseph Heath, *The Machinery of Government: Public Administration and the Liberal State* (Oxford University Press, 2020), 54–65.

[31] For this point, see Michael Lipsky, *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services* (Russell Sage Foundation, 1980), 159–179; Bernardo Zacka, *When the State Meets the Street: Public Service and Moral Agency* (Harvard University Press, 2017), 49–62.

[32] Heath, *The Machinery of Government*, 60–61.

These are intractable problems with the hierarchical approach to the problem of bureaucratic domination. It seems, then, that caseworker governance inevitably creates objectionable asymmetries of power between caseworkers and ordinary citizens. It creates a problem of bureaucratic domination. Let us turn to a way of solving the problem.

IV. The solution: algorithmic governance

The problem with caseworker governance is that, when human beings make decisions about cases, they have deeply asymmetric power over the people whose cases they decide. Algorithmic governance is a simple solution to this problem. Such governance involves algorithms rather than human beings making decisions about individual cases. This gets rid of these human deciders. So it gets rid of the relationship of domination between the caseworker and their subjects. The more decisions are made by algorithms rather than human decision-makers, the less power human decision-makers have over ordinary citizens. By making decisions on eligibility for benefits automatically, one ensures that no caseworker has asymmetric power over applicants. By evaluating recidivism risk with an algorithm rather than a probation officer, one removes the dominating power of the probation officer over convicts. Algorithmic governance reduces the extent of human domination within the administrative state.

If so, we have reason to support the existing use of algorithms in public administration. We also have reason to favor their further adoption. We could, for instance, use algorithms to decide whether applicants were eligible for disability benefits. We could design an algorithm to evaluate whether a claimant for unemployment benefits is looking for work on the basis of submitted job applications and attended job interviews. We could make the use of algorithms mandatory rather than discretionary in sentencing. If bureaucratic domination is a serious problem, we have reason to do all these things. We have reason to increase the extent to which algorithms make decisions. The argument in favor of such algorithmic governance is that it helps ameliorate the power asymmetry between bureaucrats and ordinary citizens.

There is one very serious concern about this argument. The worry is that algorithmic governance empowers the people who design the algorithms.[33] It disempowers caseworkers, but it gives a class of system designers a lot of power over ordinary citizens. The concern is that, therefore, algorithmic governance does not reduce the level of domination at all. Now we shouldn't exaggerate the level of reasonable concern here. Caseworkers interact with their clients face-to-face as

[33] Bovens and Zouridis, "From Street-Level to System-Level Bureaucracies," 181–182.

individuals. This means they can give them detailed individual instructions. They can make access to benefits contingent on following those instructions. This gives caseworkers a lot of influence over clients' behavior. System designers, in contrast, interact with their clients anonymously, through a computer screen. They don't personally know each person whom their actions affect. They cannot directly communicate with them. This means system designers exercise much less fine-grained control over users' behavior than caseworkers. The anonymity of client to system designer acts as a kind of protective screen between the two, reducing the ability of the latter to influence the former.

Nonetheless, system designers surely retain a lot of power over ordinary citizens. There remains a concern. In the rest of the section, I will discuss how to address this concern more fully. The key point is that we can structure algorithmic governance to minimize the power of system designers. Three tools help achieve this: top-down control, bottom-up participation, and contestability. We will discuss them sequentially. This will give us both a response to the concern and a blueprint for how to implement algorithmic governance.

IV.A Top-down control

One tool for ameliorating the power of system designers is control by elected officials. Elected officials must be able to direct system designers to create algorithms that do what they want them to do. The best ways to do this are the ordinary ways by which any organization is controlled. Elected officials should have the ability to hire system designers and fire them for ignoring their commands. They should set the budget for different programs, and have some power of promotion, demotion, and task assignment. This gives elected officials the ability to impose costs on individual system designers. That, I've suggested, ameliorates the extent to which those system designers have power over ordinary citizens. They cannot just design algorithms in any way they want. They are under the control of elected representatives.

Now, elected officials have this power vis-à-vis caseworkers too. But we saw that there were very serious barriers to elected officials controlling caseworkers. These barriers are much lower under algorithmic governance. To see this, begin with the problem of capacity. The main point here is that there are far fewer system designers under algorithmic governance than there would otherwise be caseworkers. We don't need hundreds of thousands of caseworkers to make decisions about each case; we just need a much smaller number of engineers. So it is much more feasible for elected representatives to monitor and incentivize system designers than it is for them to monitor and incentivize caseworkers. Elected officials can devote much more time to overseeing each system designer

than each caseworker, since they have far fewer people to oversee. Thus, they can exert much tighter control over each system designer than each caseworker.

Now turn to the problem of observability. Much of this problem came from the difficulty of observing the many face-to-face interactions caseworkers have with their clients. System designers have no such interactions, immediately alleviating the problem. More generally, it is much easier for system designers to leave extensive documentation about their key choices than it is for caseworkers. These are choices about what kind of algorithm to use, what data to use, or how to evaluate the fit of the relevant algorithm. It is good design practice to document technical work in these areas, together with the reasons for many choices. We can feasibly require that key design choices are recorded. It is thus much easier to observe (ex post) the important things system designers have done than it is to observe the important things caseworkers do. So it is easier for elected officials to sanction system designers for doing things they don't want.[34]

Controlling system designers does generate challenges. The big challenge is technical knowledge. For overseers to have substantial control over system designers, they need to know how algorithms are designed: how choices of metrics, data, or algorithm type affect the functioning of the system. There are two ways to equip overseers with this knowledge. First, committee systems build expertise within a legislature. When legislators sit on a committee for many years, they gain knowledge about the topic of the committee. This allows elected officials to effectively oversee even esoteric policy areas. So algorithmic governance should be alloyed with a dedicated legislative committee overseeing algorithmic development across government. Second, we can hire technical knowledge directly into the executive. One way to do this is to set up an ombudsman's office to evaluate decision-making algorithms. This office should be staffed by technical experts. It should give advice to elected officials and respond to citizen complaints about the use of algorithms. Setting up such a body creates a store of technical experts tasked with overseeing system designers. That again puts such designers under some top-down control.

Let me make one final point. For system designers to be under this kind of top-down control, they must be formal government employees. Often, current governments hire private companies to design their algorithms. COMPAS, for example, is designed by a company called Equivant. Equivant staff are not U.S. government employees. They cannot be fired by elected officials and cannot be given direct orders by such officials. They are thus not under their very tight democratic control. This is a solvable problem. Governments should simply

---

[34] For a similar point, see Jon Kleinberg et al., "Discrimination in the Age of Algorithms," *Journal of Legal Analysis* 10 (2018): 113–174, https://doi.org/10.1093/jla/laz001.

design algorithms in-house. That seems feasible; most machine learning algorithms used by, say, the U.S. federal government have been developed in-house.[35] Eschewing outsourcing is indispensable for giving elected officials control over system designers.

IV.B Bottom-up participation

The second tool to minimize the power of system designers is bottom-up participation. We can take key decisions out of the hands of system designers and put them in the hands of ordinary citizens. Different ways of doing this are feasible at different steps in the bureaucratic process. Consider the final step in the process, when an algorithm has been designed and an agency is deciding when to use it. Here we can employ notice-and-comment procedures. Such procedures require agencies to publish proposed rules for public comment and then consider those comments before issuing a final rule. This allows ordinary citizens and civil society groups to object to implementing specific algorithms. Such procedures mean system designers cannot unilaterally determine how government decisions are made. They reduce the chance that, conditional on trying to enact some algorithm, system designers will succeed in doing so. So they reduce system designer power.

Many key decisions occur before final adoption. Suppose, for example, that one is designing an algorithm to decide whether a defendant is granted bail. Multiple considerations bear on this decision: whether the defendant will turn up for trial, whether they'll commit a crime if released, whether they have dependents and so on. Someone must decide how to trade these issues off against one another; they must decide their relative weight. Or suppose one is designing an algorithm that directs Child Protective Services' (CPS) investigatory resources. We want to predict the likelihood that each child is being abused. We never have complete data on actual abuse. We have data on whether a CPS investigation found abuse, whether abuse was proven in court, whether medical professionals found an injury consistent with abuse, and so on. Someone must decide what metric the algorithm should aim to optimize for. If system designers alone make these decisions, they have a lot of power over those subjected to their algorithms.

We can reduce system designer power by putting some such decisions into the hands of ordinary citizens. We can do this by setting up temporary citizen bodies that determine these factors. One way to do this is to get a sample of citizens to

---

[35] David Freeman Engstrom et al., *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies* (Administrative Conference of the United States, 2020), https://www.ssrn.com/abstract=3551505.

deliberate about such issues. This means empaneling a civic jury to discuss trade-offs or metrics. If we give such a body at least advisory, and perhaps veto, power over key design decisions, that reduces the power of system designers.[36] A second option is to simply aggregate citizens' separate judgments on these issues. To see how to do this, consider, say, trade-off decisions. We first pick a sample of citizens.[37] We can then evaluate how each citizen in the sample makes such trade-offs. We do this by asking about what they would choose in concrete cases. We can then aggregate their different weightings via some voting rule and use the output to determine how an algorithm makes such trade-offs.[38] This would reduce system designers' control over the resulting algorithm.

Let's discuss one concern about this proposal. One might worry that, by empaneling a citizen jury, we simply replace the power of system designers with that of jurors. We don't ameliorate the problem after all. The response to this point rests on the idea that empaneling the proposed juries disperses power and makes it more temporary. It disperses power in that it ensures less power is concentrated in the hands of system designers. In section II, I suggested that dispersed power is less objectionable than concentrated power. If so, this makes the power asymmetries less objectionable. Citizen juries make power more temporary in that jurors hold power for less time than more permanent bureaucrats. They are empaneled to decide on a specific issue. After deciding they lose power. In section II, I suggested that enduring power is more objectionable than temporary power. So distributing power away from permanent bureaucrats and into the hands of these jurors makes the power less objectionable. Either way, juries reduce objectionable power asymmetries.

I have claimed that institutions for bottom-up participation can reduce the power of system designers. It's important to underline a key point. These forms of bottom-up participation are all, in a sense, unfeasible under caseworker governance. It is of course feasible to use all these tools to decide on the rules that govern caseworker decisions. But the problems of observability and capacity mean these rules often don't practically bind caseworkers. It is unfeasible for citizens to participate in each of the thousands or millions of individual caseworker decisions.

---

[36] For a similar proposal, see Chiara Cordelli, *The Privatized State* (Princeton University Press, 2020), 111–113. For defense of the use of civic juries outside the bureaucracy, see Alexander Guerrero, *Lottocracy: Democracy Without Elections* (Oxford University Press, 2024); Hélène Landemore, *Politics Without Politicians: The Case for Citizen Rule* (Allen Lane, 2026).

[37] The sample might be random, or we might oversample those directly affected by a bureaucratic agency (e.g., parents for CPS procedures). The difference doesn't matter to my point.

[38] For how to implement this in detail, see Min Kyung Lee et al., "WeBuildAI: Participatory Framework for Algorithmic Governance," *Proceedings of the ACM on Human-Computer Interaction* 3, no. CSCW (2019): 181:1–181:35, https://doi.org/10.1145/3359283.

One cannot empanel a citizen jury to evaluate every decision on, say, benefit applications. One cannot implement notice-and-comment procedures for every proposed tax audit. So citizen participation cannot much reduce the power of street-level bureaucrats. It can diminish the power of system designers.

IV.C Contestability

The final tool for ameliorating the power of system designers is contestability. This is widely used under caseworker governance. Consider, say, British benefit decisions. One can appeal the denial of such benefits. One first asks the relevant department to reconsider one's application.[39] One can appeal the reconsidered judgment to an independent tribunal. If one thinks the tribunal makes a legal mistake, one can appeal to the Upper Tribunal. At each step, different people are involved in the process. Contestability helps reduce the power of individual caseworkers. If a human decision-maker makes the initial decision, they still have a lot of power over the claimant. But their power is diminished by the fact that the decision can be appealed. Contestability diminishes the power of individual street-level bureaucrats, dispersing the power bureaucrats have over individual citizens.

When we don't have street-level bureaucrats, we don't need contestability for this purpose. But we can also use contestability to reduce the power of system designers. One way to do this is to use algorithms to make initial decisions but retain humans for the appeal procedures. This reduces the power of system designers to determine outcomes. But my arguments support a system in which the appellate stages are automated. The key is to have different designers for the appellate stages than for the initial decision stages. Under this proposal, one team designs the algorithm that makes initial decisions while a different team or teams design algorithms that decide on appeals of those decisions. It would be natural for a team in the aforementioned ombudsman's office to oversee the design of appellate algorithms, but really any team could do the job. This reduces the power that any one team (member) has over any individual citizen. So it reduces citizens' objections to the ensuing power asymmetries.

I conjecture that contestability will, often, be easier to provide via this system than under caseworker governance. It is too slow or expensive to offer human re-adjudication on many decisions. Thus, appeals procedures for, say, UK benefit decisions end after three stages. And people often cannot appeal being the subject of an investigation at all. Australians cannot appeal the ATO's decision to audit

[39] GOV.UK, "Appeal a Benefit Decision," accessed August 27, 2025, https://www.gov.uk/appeal-benefit-decision.

their tax return. Parents can rarely appeal being the subject of a CPS investigation. Algorithmic decision-making is usually much quicker and less resource intensive than human decision-making. If so, automated appeals processes can make feasible appeal procedures for decisions, like investigatory ones, that are usually unappealable under caseworker governance. They can make feasible more stages of appeal for other kinds of decision. Algorithmic governance helps us use appeals procedures to reduce the power each bureaucrat has over each person.

Let me sum up. The main worry about algorithmic governance is that such governance empowers system designers. I have presented three tools for addressing this problem: top-down control, bottom-up participation, and contestability. These tools work better under algorithmic governance than under caseworker governance. What emerges from this, then, is both an argument in favor of algorithmic governance and a guide to its implementation.

V. Domination by algorithms?

I want to address a final objection to my argument for algorithmic governance. By replacing caseworkers with algorithms, one ensures that the algorithm itself has enormous power over citizens. Algorithms decide whether people get unemployment benefits or are judged at high risk of re-offending. One might worry that this means citizens are literally dominated by these algorithms. The worry is not that algorithms empower bureaucrats who dominate citizens, but rather that they dominate them directly. On this view, U.S. convicts are, for instance, literally dominated by COMPAS. If so, then algorithmic governance doesn't solve the problem of bureaucratic domination at all.

To evaluate this concern, we need to know what kinds of things one can be dominated by. I believe we can only be dominated by agents.[40] To see why, think about your relationship to natural phenomena. A hurricane might have a big impact on your well-being and behavior. It might blow down your house or cause you to flee. But you are not dominated by a hurricane. Intuitively, you don't have the same relation to a hurricane as you do to a slave master. You don't have the same kind of moral objection to a hurricane's power. You should resent your slave master, but resentment aimed at a hurricane would be inapt.[41] This is because domination is a certain kind of social relationship and you only have such social

[40] This is the orthodox view. See Pettit, *Republicanism*, 79, 52; F. Lovett, *A General Theory of Domination and Justice*, 25; Kolodny, *The Pecking Order*, 89–90.
[41] For this kind of point, see also David Estlund, "What's Unjust about Structural Injustice?" *Ethics* 134, no. 3 (2024): 333–359, https://doi.org/10.1086/728634.

relationships with agents. So the distinctive claim we have against asymmetries of power is a claim against agents having asymmetric power over us.

If that is true, the reply to the objection is obvious. The algorithms we've discussed in this paper are not agents. To be an agent, you have to have beliefs and desires and intentions that are somewhat rationally integrated. They must at least somewhat conform, for example, to means-end rational requirements. The relevant algorithms lack beliefs or intentions. COMPAS is a statistical model. It doesn't have mental states. Even if it did, those states would not be integrated enough for it to count as a proper agent. The same goes for the kind of machine learning algorithm used by, say, the Australian Taxation Office. We can of course describe such algorithms in terms of some simple goals. We can say that COMPAS "wants" to accurately predict re-offenses. But this is like describing a thermometer as wanting to measure temperature or a ship as wanting to stay afloat. Such entities don't really have the complicated set of mental states necessary for agency, and so they cannot dominate people in the morally relevant sense.[42]

Perhaps that is too quick. One might think that when inanimate phenomena are the product of human decisions, we can be dominated by them.[43] Think of architecture. If you cannot walk, and you live in a city full of inaccessible buildings, then perhaps you are dominated by the city's architecture. One might think that this shows that one can be dominated by inanimate entities that result from human decisions. The problem with this thought is that natural phenomena are the product, sometimes predictably, of human decisions. Reconsider hurricanes. The frequency and intensity of hurricanes have predictably increased due to human decisions (i.e., to pollute). Yet one still cannot be dominated by a hurricane. One might be dominated by the human beings that cause hurricanes, but not by hurricanes themselves. Likewise, one might be dominated by the human beings who design algorithms. But, I claim, one is not dominated by the algorithms themselves.

*

Let's take stock. I've argued that the use of algorithms in government's decision-making factories alleviates the problem of bureaucratic domination. This is a

[42] These algorithms are notably less capable than frontier AI models. It is an open question about whether the instances of the latter are the kind of agents that could dominate us. If they are, then that is a reason against using them in administrative decision-making. For a recent overview of some relevant metaphysical issues, see Bradford Saad and Andreas L. Mogensen, "Digital Minds: Issues in the Philosophy of Mind and Cognitive Science," Philosophy Compass 21, no. 6 (2026): e70141, https://doi.org/10.1111/phc3.70141.

[43] For this sort of view, see Dorothea Gädeke, "Does a Mugger Dominate? Episodic Power and the Structural Dimension of Domination," *Journal of Political Philosophy* 28, no. 2 (2020): 199–221, 209–214, https://doi.org/10.1111/jopp.12202.

weighty consideration in favor of algorithmic governance. Still, this consideration is pro tanto: it might be outweighed by contrary considerations. So, in the next section, we address some reasons against using algorithms in the administrative state.

VI. Against algorithmic governance

Many criticisms of the use of algorithms in government focus on their reliability. The idea is that to make good judgments about cases, we sometimes need to respond to a lot of qualitative data that algorithms, allegedly, mishandle. So algorithms are often less accurate than human decision-makers.[44] Another worry is that certain algorithms are discriminatory. This might be because they rely on data that embeds past injustice, or because their design means they treat people unfairly.[45] When they are more discriminatory than human decision-makers, that is a reason not to employ them. The overall worry is that, sometimes, algorithms work less well than human decision-makers. In these cases, we have weighty reasons not to use algorithms.

I won't address either of these concerns here. They are reasonable. Sometimes, algorithms do work less well than human decision-makers. We have weighty reasons not to employ algorithms in these cases. Still, sometimes algorithms are no less accurate and no more discriminatory than human beings.[46] These are the situations in which the case for algorithmic governance is strongest. Such cases will likely become more common in the future. The worries about algorithmic governance I wish to address, then, identify more general objections to the use of algorithms in government's decision-making factories. The objections are that such use violates people's rights to an explanation, that it requires inadmissibly basing decisions on statistical evidence, and that it creates responsibility gaps. Such worries give us in-principle reasons to resist algorithmic governance. In this section, I argue that such worries are overblown.

---

[44] Daniel J. Solove and Hideyuki Matsumi, "AI, Algorithms, and Awful Humans," *Fordham Law Review* 92 (2024): 1923–1940; Grant, Behrends, and Basl, "What We Owe to Decision-Subjects."
[45] Solon Barocas and Andrew D. Selbst, "Big Data's Disparate Impact," *California Law Review* 104, no. 3 (2016): 671–732; Julia Angwin et al., "Machine Bias," ProPublica, May 23, 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing; Eubanks, *Automating Inequality*; Deborah Hellman, "Big Data and Compounding Injustice," *Journal of Moral Philosophy* 21, no. 1–2 (2024): 62–83, https://doi.org/10.1163/17455243-20234373.
[46] For arguments here, see Kleinberg et al., "Human Decisions and Machine Predictions"; Kleinberg et al., "Discrimination in the Age of Algorithms."

VI.A Right to an explanation

Let's start with the idea that the use of algorithms in the administrative state violates people's right to explanation. The idea is that the subjects of government decisions are owed an explanation of those decisions and the decision-making procedures.[47] This is a worry specifically for the use of machine learning algorithms in government. These are often very complex. An input variable might increase the value of the target over some range, decrease it over others, and interact multifariously with other input variables over other ranges. Intermediate layers in a neural network sometimes represent very peculiar features. This can make it difficult to give people an intelligible explanation of exactly what caused a specific algorithmic decision.[48] If people are owed such an explanation, that provides a general reason not to use machine learning algorithms in administrative decision-making.

One way to understand this concern is in terms of a right to justification. Several authors have thought that people are owed a justification for the nature of their social order.[49] We have a right to an intelligible justification of why government decisions are made in certain ways. These are justifications that point to the normative reasons in favor of those decision-making procedures, things we can understand that might answer a complaint against the use of such a procedure. This, however, is perfectly compatible with algorithmic governance. We can give people intelligible justifications of why we use machine learning algorithms to make decisions. The argument in sections III–V is one such justification. Additionally, we might say we use such algorithms because they are more accurate, less discriminatory, cheaper, or faster than human decision-making. These all identify normative reasons in favor of the use of algorithms in the administrative state. They are all points one could cite to answer a complaint against the use of such algorithms. So they are, it seems to me, justifications for the use of such algorithms.

One might think that the right to justification is more fine-grained than a right to the justification of a social order. Perhaps people are owed an intelligible

---

[47] Vredenburgh, "The Right to Explanation"; Tae Wan Kim and Bryan R. Routledge, "Why a Right to an Explanation of Algorithmic Decision-Making Should Exist: A Trust-Based Approach," *Business Ethics Quarterly* 32, no. 1 (2022): 75–102, https://doi.org/10.1017/beq.2021.3.

[48] For a classic overview, see Jenna Burrell, "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms," *Big Data & Society* 3, no. 1 (2016): 2053951715622512, https://doi.org/10.1177/2053951715622512.

[49] Jeremy Waldron, "Theoretical Foundations of Liberalism," *The Philosophical Quarterly* 37, no. 147 (1987): 127–150, https://doi.org/10.2307/2220334; Rainer Forst, *The Right to Justification: Elements of a Constructivist Theory of Justice*, trans. Jeffrey Flynn (Columbia University Press, 2012).

justification of each individual decision that affects them, rather than just a decision-making procedure. There is still, however, little tension between such a right and the use of algorithms in the administrative state. Many of the features that justify using an algorithm to make decisions also justify those very decisions. Accuracy is the clearest example. Suppose we have an algorithm that is very good at predicting whether someone has submitted a fraudulent tax return. When this algorithm flags someone's tax return, they get audited. The tax agency has a good justification for auditing this person. It is likely that they cheated on their taxes. That someone likely cheated is a normative reason to audit them. So we have a justification of the individual decision, even if the agency cannot intelligibly explain exactly how the algorithm made its prediction. The underlying issue is that the use of certain algorithms in government makes it difficult to furnish people with something very specific: detailed, intelligible, causal histories of those decisions. But such histories are not needed to justify decisions.

Might we nonetheless have a right to a detailed, intelligible causal history of decisions made about us? To answer that, I think we need to explore what could make furnishing such histories valuable. One might think that such histories are necessary for people to contest mistaken decisions.[50] The thought is we need to understand the causal origins of a decision to be able to appeal it. But this seems incorrect. One needn't know what caused an initial decision to appeal that decision. As long as appeals procedures exist—like the ones described in section IV.C—one can simply lodge the appeal and provide whatever evidence is required for it. One might instead think that one needs to know the causal history of a decision to know whether one should appeal it. But that also seems false. Suppose one knows the standard for a correct decision. One knows, for instance, that disabled people are eligible for disability benefits. Then if one knows that one is disabled, one can know that one has good reason to appeal a denial of disability benefits, without knowing what caused the denial. One needn't identify how a decision went wrong to identify that it is wrong. So the value of contestability doesn't support a right to detailed causal histories of decisions.

Perhaps the issue instead lies in autonomy. One might think that we need detailed causal explanations of decisions to make autonomous choices in our dealings with the administrative state. We need to know how agencies will decide matters. To autonomously choose whether to apply for benefits, we need to know how likely our application is to be successful. To autonomously choose what evidence to put in an application, we need to know what additions likely avoid an adverse decision. One might think that, for this, one needs a detailed causal explanation of how

---

[50] Vredenburgh, "The Right to Explanation," 213, 219.

specific decisions are made.[51] But, again, this seems incorrect. To autonomously make such decisions, we need answers to certain counterfactual questions. One needs to know about what would happen were one to act in certain ways.[52] We can often answer such counterfactuals without detailed knowledge of the causes of events. I know, for example, that if I take paracetamol (Tylenol), it will ease my pain; but nobody knows exactly how paracetamol causes pain relief. We can give citizens such counterfactual knowledge about administrative decisions by giving them an interface in which they can test out an algorithm. When people are, say, applying for unemployment benefits, we can let them see how changing certain parts of their application would affect whether it is approved. This would give people the knowledge they need to make autonomous choices about whether and how to apply for certain benefits.

So there seems little basis for the idea that we have a right to detailed, intelligible histories of governmental decisions. And in fact, in general form, this idea seems to me quite implausible. Sometimes, governments decide policy on the basis of esoteric science: they might decide climate policy, for instance, on the basis of complicated climate modeling.[53] Sometimes governments make decisions on the basis of state secrets: a government might increase military spending because confidential informants have told it that a neighbor is planning to invade. In neither case is the government required to give citizens maximally detailed, intelligible accounts of the causal origins of these decisions. In the case of esoteric science, they cannot make the accounts intelligible to all citizens. Some of us cannot understand the climate modeling. In the case of state secrets, they have instrumental reasons to avoid disclosure. Telling everyone who the informants are puts their lives at risk. So it is doubtful that we have a right to detailed causal histories of every decision that affects us.

My argument here has the form of a dilemma. There is some appeal to the idea that we have a right to a justification for the nature of our social order, and even of individual decisions. But the use of algorithms to make decisions doesn't imperil such justifications. The use of some algorithms does imperil our ability to give people detailed, intelligible causal histories of each decision. But there is little appeal to the idea that we have a right to be given such histories. So let us turn to a different worry about algorithmic governance.

---

[51] Vredenburgh, "The Right to Explanation," 213.

[52] For the same point, see Brett Karlan and Henrik Kugelberg, "No Right to an Explanation," *Philosophy and Phenomenological Research* 111, no. 1 (2025): 137–156, 144–145, https://doi.org/10.1111/phpr.70008.

[53] For this point, see Karlan and Kugelberg, "No Right to an Explanation," 149–152.

VI.B Purely statistical evidence

Begin by considering a case. Imagine a hundred prisoners are exercising in a yard. Ninety-nine suddenly attack the only guard in the yard, killing him. One prisoner plays no role in the assault. The authorities have no further information about what happened. Terence was a prisoner in the yard at the time of the assault. There is a 0.99 probability that Terence contributed to the guard's murder. Are the prison authorities permitted to treat Terence as guilty? Many people think not. The only evidence the authorities have about Terence's guilt is statistical evidence. They can only infer his guilt from the fact that he falls into some group in which there is a high frequency of guilt. Many think that treating Terence as guilty on the basis of purely statistical evidence is impermissible.[54]

Some people think that this is a problem for algorithmic governance.[55] Now this is not a problem for the use of decision trees, as in the UK's Universal Credit eligibility determinations. But many algorithms in public administration output estimates of probability. They estimate how likely a tax return is to be fraudulent or how likely a convict is to reoffend. The evidence they have is, at least largely, statistical evidence. COMPAS, for example, looks at people's employment status because unemployed people more often reoffend. So if there is a general argument against making administrative decisions on the basis of purely statistical evidence, it would seriously limit the extent to which algorithms could be used in the administrative state.

To address this point, we have to think about what makes assigning guilt on the basis of purely statistical evidence wrong. Why is it wrong to treat Terence as guilty? There are several proposals in the literature.[56] The most plausible story, it

---

[54] For a classic discussion of these kinds of cases, see Mike Redmayne, "Exploring the Proof Paradoxes," *Legal Theory* 14, no. 4 (2008): 281–309, https://doi.org/10.1017/S1352325208080117.

[55] Holm, "Statistical Evidence and Algorithmic Decision-Making"; Cossette-Lefebvre and Maclure, "AI's Fairness Problem."

[56] For stories which focus on incentives, causality, or modal issues, see, respectively, David Enoch, Levi Spectre, and Talia Fisher, "Statistical Evidence, Sensitivity, and the Legal Value of Knowledge," *Philosophy & Public Affairs* 40, no. 3 (2012): 197–224, https://doi.org/10.1111/papa.12000; Judith Jarvis Thomson, "Liability and Individualized Evidence," *Law and Contemporary Problems* 49, no. 3 (1986): 199–219, https://doi.org/10.2307/1191633; Martin Smith, "When Does Evidence Suffice for Conviction?" *Mind* 127, no. 508 (2018): 1193–1218, https://doi.org/10.1093/mind/fzx026. For compelling critiques of each of these views, see John Hawthorne, Yoaav Isaacs, and Vishnu Sridharan, "Statistical Evidence and Incentives in the Law," *Philosophical Issues* 31, no. 1 (2021): 128–145, https://doi.org/10.1111/phis.12201; Georgi Gardiner, "Legal Burdens of Proof and Statistical Evidence," in *The Routledge Handbook of Applied Epistemology*, ed. David Coady and James Chase (Routledge, 2018), 179–195, https://doi.org/10.4324/9781315679099-14; Michael

seems to me, is that when we are deciding whether someone is guilty of a wrongdoing, we are, in part, deciding whether they are blameworthy for the wrongdoing.[57] Their blameworthiness is crucial for the permissibility of punishing them; if they have an excuse, then punishment would usually be inapt. Yet to blame someone for a crime, we have to be in a position to know they committed it. If we could not know someone committed some wrongdoing, we cannot blame them for it. We cannot be in a position to know such a thing on the basis of purely statistical evidence. We do not, for example, know that a given lottery ticket will lose on the basis of the fact that there are a million other tickets.[58] So we should not blame people on the basis of purely statistical evidence. Thus we should not assign guilt on the basis of such evidence.

If this is correct, there is no general argument against making administrative decisions on the basis of purely statistical evidence. We should not assign guilt (and blame) on the basis of such evidence, but most administrative decisions do not concern guilt. When an agency denies your application for unemployment benefits, it isn't saying you are blameworthy for anything. It is merely saying that you are ineligible for benefits. When the tax authorities audit you, they are not saying you are guilty of cheating on your taxes. They are trying to decide whether you did cheat. When a probation officer evaluates a convict's re-offense risk, they are not assessing their guilt. They are assessing the likelihood of a new offense. Even sentencing itself needn't involve any additional assessment of guilt. Guilt has already been decided at a trial. Sentencing can simply be an evaluation of what length of term is optimal given the risk of re-offense, the deterrence effect of imprisonment, and the facts determined in court. So none of the administrative decisions with which we began—allocating benefits, investigating rule-breaking, imposing costs—require assessing guilt. The administrative state is only rarely in the blame game. So the injunction against relying on statistical evidence covers

---

Blome-Tillmann, "Statistical Evidence, Normalcy, and the Gatecrasher Paradox," *Mind* 129, no. 514 (2020): 563–578, https://doi.org/10.1093/mind/fzaa007.

[57] For this kind of story, see Clayton Littlejohn, "Truth, Knowledge, and the Standard of Proof in Criminal Law," *Synthese* 197, no. 12 (2020): 5253–5286, https://doi.org/10.1007/s11229-017-1608-4. For related views, see Lara Buchak, "Belief, Credence, and Norms," *Philosophical Studies* 169, no. 2 (2014): 285–311; Michael Blome-Tillmann, "'More Likely Than Not': Knowledge First and the Role of Bare Statistical Evidence in Courts of Law," in *Knowledge First: Approaches in Epistemology and Mind*, ed. J. Adam Carter, Emma C. Gordon, and Benjamin Jarvis (Oxford University Press, 2017), 278–292, https://doi.org/10.1093/oso/9780198716310.003.0014.

[58] For a classic discussion of such cases, see John Hawthorne, *Knowledge and Lotteries* (Clarendon Press, 2004).

little of what the administrative state does. In other aspects of administrative decision-making, the use of algorithms is permissible.[59]

VI.C Responsibility gaps

Let us turn to a third worry. The worry here is that the use of algorithms—and AI—by governments creates so-called "responsibility gaps."[60] The worry is that when an algorithm makes a decision, often, no human being is morally responsible for that decision. The idea is that system designers may not have been able to foresee that the algorithm would make the decision it did. But sometimes algorithms make mistakes. They deny people benefits for which they are eligible, for example. The concern is that system designers' ignorance excuses them from responsibility for these mistakes. So nobody will be responsible for them. If that is bad, it gives us a general reason to oppose algorithmic governance.

To evaluate this concern, we need to ask why a lack of responsibility could be bad. Those who worry about responsibility gaps often cite the value of incentives.[61] The thought here is that if nobody is morally responsible for a mistake, then nobody can be sanctioned for it. And if nobody is sanctionable for mistakes, people's incentives to avoid mistakes are diminished. The problem with this argument is that we can clearly sanction people for things for which they are not morally responsible. We can give system designers performance incentives. If their algorithms make a lot of mistakes, then we can reduce their pay or fire them regardless of their moral responsibility. We needn't tie incentives to things for which people are morally responsible. So the incentive concern is misguided: lack of responsibility doesn't entail lack of incentivization.

Sometimes, those who worry about these gaps seem to think it is more directly problematic that nobody is to blame for bad decisions.[62] They think there is something non-instrumentally bad about responsibility gaps. I confess to finding this position hard to make plausible. When bad things happen to us, it is not good that someone is blameworthy for them. The world isn't made better when we can blame more people for our misfortunes. In some ways, it is made worse. Blame is

---

[59] Is the objection to racial profiling that it relies on purely statistical evidence? Clearly, I think not. I think it is whatever the objection is to discrimination. I bracket that in this section, since some algorithms are less discriminating than human decision-makers.

[60] Matthias, "The Responsibility Gap"; Robert Sparrow, "Killer Robots," *Journal of Applied Philosophy* 24, no. 1 (2007): 62–77; Santoni de Sio and Mecacci, "Four Responsibility Gaps with Artificial Intelligence."

[61] Santoni de Sio and Mecacci, "Four Responsibility Gaps with Artificial Intelligence," 1063; Sparrow, "Killer Robots," 63.

[62] Sparrow, "Killer Robots," 67.

a negative emotion. Being wronged is a bad state to be in. By creating more responsibility, one creates more of these bad things. To illustrate the point, reconsider natural phenomena. If a hurricane makes me homeless, sometimes nobody is to blame. The world wouldn't be a better place if God inflicted all hurricanes on us, and I could blame Him for my homelessness. More blameworthiness doesn't make matters better.

One might understand the non-instrumental concern about responsibility gaps in expressive terms.[63] On this view, we express disrespect to people when we ensure that nobody is responsible for decisions which adversely affect them. Thus, the employment of algorithmic systems expresses disrespect. We shouldn't be disrespectful, so we shouldn't employ such systems. I want to make two points about this position. First, I doubt that using algorithms expresses disrespect in our actual societies. Something expresses disrespect when people understand it as communicating disrespect. I doubt the public has settled views about what the use of algorithms communicates. Second, even if we did, we could change what the use of algorithms expresses by changing how we understand it. If governments communicate clearly that they are using algorithms to, say, avoid domination, this will likely affect how people understand the employment of such algorithms. Expressive views tell in favor of changing people's communicative beliefs more than opposing certain practices.[64]

Overall, then, it is difficult to identify a compelling objection from responsibility gaps to algorithmic governance.[65]

*

I've just explored three general worries about the use of algorithms in the administrative state. To reiterate, I focus on these worries because of their generality. Worries about accuracy and discrimination can only be evaluated on a case-by-case basis. The worries I've addressed threaten algorithmic governance more broadly. But, so I've argued, the threat is largely illusory. These concerns do not much constrain the use of algorithms in the administrative state.

VII. Conclusion

---

[63] This, it seems, is how it is understood in Sparrow, "Killer Robots," 67.
[64] This is the kind of point made by Jason Brennan and Peter Martin Jaworski, "Markets without Symbolic Limits," *Ethics* 125, no. 4 (2015): 1053–1077, https://doi.org/10.1086/680907.
[65] A more concessive view tries to answer the objection by establishing people's responsibility for algorithmic output. For this strategy, see Isaac Taylor, "Representative Robots: Can AI Systems Act in Our Name?" *Free & Equal: A Journal of Ethics and Public Affairs* 1, no. 1 (2025): 298–322, https://doi.org/10.16995/fe.18626.

In sum, then, what should we think about algorithmic governance? The preponderance of scholarly opinion is skeptical. This paper is intended as a corrective to that skepticism. The core of the positive case for algorithmic governance is that it solves the problem of domination by street-level bureaucrats. I've addressed several worries about such governance: that the use of algorithms violates people's rights to explanation, that algorithms rely on inadmissible statistical evidence and that they create responsibility gaps. These worries, I've argued, are overblown. Overall I think this constitutes a strong case for automated decision-making. We should often try to replace caseworkers with algorithms in the decision-making factories of government.

Acknowledgments For helpful discussion I am grateful to audiences at the University of Melbourne, Monash University, and the PPE Society London Meeting. For incisive comments on drafts of this paper, I thank Kate Vredenburgh, David Storrs-Fox and several Associate Editors for this journal.