AI SUFFRAGE FOR HUMAN FLOURISHING

Simon Goldstein, Guha Krishnamurthi** & Peter N. Salib****

Should Artificial Intelligence (AI) systems have the vote? One approach to this question is to explore AI systems' moral characteristics, asking whether they possess the qualities that entitle humans to equal political standing and, consequently, suffrage. Those are important questions. But they are very hard ones, resting on highly contestable premises about ethics, AI, and humanity.

We instead take a functionalist approach: On the current trajectory, AIs will soon be highly agentic and capable, while remaining difficult to align with human objectives. If so, law will be an essential tool for shaping such systems' behavior to favor positive-sum cooperation with humans, rather than conflict.

This Essay applies that functionalist approach to perhaps the most foundational legal question in democratic societies: Who has the franchise? Building on prior work arguing for AI private rights in tort and contract, we analyze the case for giving AI the right to vote.

We canvas four instrumentalist theories of democracy and ask whether each supports extending political rights to AI. Specifically, we argue that AI suffrage may: (1) reduce powerful AI systems' incentive to disempower humanity by credibly committing public policy to account for AIs' preferences; (2) promote economic growth by forcing governments to invest more heavily in public goods; (3) improve the epistemics of policymaking by adding competent and diverse voters whose reasoning is transparent and queryable; and (4) improve the "felt legitimacy" of government, at least insofar as that concept tracks AI behavior rather than genuine emotion.

We then consider three powerful objections to AI suffrage: the problem of how to count AIs, the risk that AI may swamp human voting, and the risk that AIs or AI companies may seize disproportionate power. We propose concrete institutional responses, including "algorithmic corporations" ("A-corps") as the unit of AI political identity, vote-share caps pegged to AI

---

* Associate Professor of Philosophy at The University of Hong Kong, Principal Investigator at the HKU AI and Humanity Lab, and a Research Affiliate at the Center for AI Safety.

** Professor, University of Maryland Francis King Carey School of Law.

*** Assistant Professor of Law at The University of Houston Law Center, Executive Co-Director of the Center for Law and AI Risk, and Law and Policy Advisor to the Center for AI Safety. Equal contribution; authors listed in alphabetical order.

This Essay was prepared for the Symposium titled *Constitutional Law in the Age of Generative AI* hosted by the *Fordham Law Review* on February 27, 2026, at Fordham University School of Law.

economic contribution, and "lottocratic" selection among AI voters. Throughout, we argue that the distributional costs of AI suffrage are likely outweighed by growth effects over time.

INTRODUCTION 476
I. THE FUNCTIONALIST APPROACH TO AI RIGHTS 479
II. SUFFRAGE FOR AGIS 482
A. Democratic Peace 483
B. Democracy Promotes Economic Growth 487
C. Democracies Have Good Epistemics 491
1. Jury Theorems 491
2. Diversity 493
3. Deliberation 494
D. Democracy Induces Felt Legitimacy 495
III. CHALLENGES 498
A. The Counting Problem 498
B. The Swamping Problem 500
C. Disempowerment 501
CONCLUSION 502

INTRODUCTION

If artificial intelligence (AI) progress continues, humans may soon share the world with millions of highly intelligent, capable, and potentially dangerous artificial agents.[1] These systems will be able to do most of the useful tasks humans can do.[2] Call such systems Artificial General Intelligence—AGIs.[3]

Perhaps the most important question to answer in the coming years is how to share the world with AGIs. Should humans retain absolute, authoritarian

1. *See* Yoshua Bengio et al., Superintelligent Agents Pose Catastrophic Risks: Can Scientist AI Offer a Safer Path? 15 (Feb. 24, 2025) (unpublished manuscript), https://arxiv.org/abs/2502.15657 (on file with the *Fordham Law Review*) (noting that some corporations envision deploying billions of AI agents worldwide); Katja Grace et al., *Thousands of AI Authors on the Future of AI*, J. ARTIFICIAL INTELL. RSCH., Oct. 2025, at 3 (surveying AI researchers who estimated a 50 percent chance of machines outperforming humans in all tasks by 2047).

2. *See* Tejal Patwardhan et al., GDPval: Evaluating AI Model Performance on Real-World Economically Valuable Tasks 1 (Jan. 26, 2026), https://openreview.net/forum?id=hcuEdq6eKD (conference poster) (on file with the *Fordham Law Review*) (finding that frontier AI models are approaching industry expert quality on real-world tasks across forty-four occupations spanning nine sectors that contribute to the United States gross domestic product).

3. *See generally* NICK BOSTROM, SUPERINTELLIGENCE: PATHS, DANGERS, STRATEGIES 1–4 (2014) (discussing "human-level machine intelligence" and "artificial general intelligence"); *OpenAI Charter*, OPENAI (Apr. 9, 2018), https://openai.com/charter/ (on file with the *Fordham Law Review*) (defining AGI as "highly autonomous systems that outperform humans at most economically valuable work").

control over AGIs, as default law would dictate?[4] Or should AGIs instead be afforded some measure of self-determination? That is, should AGIs be granted some legal rights?

A small but growing cohort of legal thinkers, AI scientists, philosophers, and cognitive scientists urges us to take the second proposal seriously.[5] Their arguments are distinct but overlapping. Some believe that AGIs will be conscious, capable of suffering, able to form genuine desires, or possess other properties entitling them to *moral* consideration.[6] If so, a legal system that subjected AGIs to arbitrary control and maltreatment would be a moral abomination.[7]

Others take a functionalist approach, arguing that extending rights to AIs is a pragmatic good rather than a moral requirement. For example, two of us have argued elsewhere that giving AIs basic property and contract rights would make them less likely to engage in large-scale violence against humans.[8] Today's AIs are inexorably being made more *agentic*, able to autonomously pursue and execute long-run goals in the real world.[9] No one yet knows how to reliably make AIs pursue goals and exhibit behaviors that humans prefer, rather than ones we do not want.[10] This is known as the "alignment" problem, and it remains unsolved.[11] On this functionalist view, law is simply a potential tool for shaping AGIs' *behavior*, even if the AGIs'

4. *See generally* Peter N. Salib & Simon Goldstein, *AI Rights for Human Safety*, 112 VA. L. REV. 1061, 1066 (2026) (explaining that under default law, AGI will be the property of the company that owns them).

5. *See, e.g.*, Lawrence B. Solum, *Legal Personhood for Artificial Intelligences*, 70 N.C. L. REV. 1231, 1231 (1992); Toni M. Massaro & Helen Norton, *Siri-ously?: Free Speech Rights and Artificial Intelligence*, 110 NW. U. L. REV. 1169, 1172 (2016); Eric Schwitzgebel & Mara Garza, *A Defense of the Rights of Artificial Intelligences*, 39 MIDWEST STUD. PHIL. 98, 100 (2015); Katherine B. Forrest, *The Ethics and Challenges of Legal Personhood for AI*, 133 YALE L.J.F. 1175, 1177 (2024).

6. *See* Robert Long et al., Taking AI Welfare Seriously 1–5 (Nov. 4, 2024) (unpublished manuscript), https:// arXiv:2411.00986 (on file with the *Fordham Law Review*) (arguing that there is a "realistic possibility that some AI systems will be conscious and/or robustly agentic in the near future," and thus morally significant); Patrick Butlin et al., Consciousness in Artificial Intelligence: Insights from the Science of Consciousness 1 (Aug. 22, 2023) (unpublished manuscript), https://arXiv:2308.08708 (on file with the *Fordham Law Review*) (surveying leading neuroscientific theories of consciousness and concluding that "there are no obvious technical barriers to building AI systems" that satisfy their indicators of consciousness).

7. *See* Eric Schwitzgebel, *The Full Rights Dilemma for AI Systems of Debatable Moral Personhood*, 4 ROBONOMICS: J. AUTOMATED ECON., 2023, at 2 (arguing that underattributing rights to genuinely conscious AI systems would constitute a moral catastrophe).

8. *See generally* Salib & Goldstein, *supra* note 4.

9. *See* Thomas Kwa et al., Model Evaluation & Threat Research (METR), Measuring AI Ability to Complete Long Tasks 2 (Feb. 25, 2026) (unpublished manuscript), https://arXiv:2503.14499 (on file with the *Fordham Law Review*) (finding that the duration of tasks frontier AI agents can complete with 50 percent reliability has been doubling approximately every seven months since 2019).

10. *See generally* BRIAN CHRISTIAN, THE ALIGNMENT PROBLEM: MACHINE LEARNING AND HUMAN VALUES (2020) (examining the unresolved challenge of aligning AI behavior with human preferences and values).

11. *Id.* at 12.

goals cannot be known or controlled with certainty.[12] Moral status is irrelevant.

This functionalist approach to law for AGIs can get one surprisingly far.[13] Prior work by two of us has argued that extending contract, property, and tort rights to AGIs would reduce their incentives to conflict with humans, promote AGI-led economic growth, and reduce ordinary harms—a minimalist package that, together, does a great deal of work.[14]

This Essay, by contrast, asks a *maximalist* legal question: Should AGIs be granted the basic *political* rights of democratic citizenship? Should they be entitled to petition their government? To speak freely? To vote? That is, should AGIs not only be afforded some set of legal rights and duties, but also the power to help determine what those legal rights and duties will be?

We do not purport to answer the question definitively in this Essay. Nor does this Essay solve all of the nitty-gritty nuances of institutional design that would be essential for implementing a scheme of AGI suffrage. But it does analyze the issue carefully and offer several arguments for AGI suffrage.

The case here is again functionalist, not reliant on AGIs' own moral status. Significant literatures from law, political philosophy, and social science already make functionalist arguments for democracy.[15] But they focus on democracy's functions in ordering human affairs.[16] This Essay adapts those frameworks to a world that includes AGIs, asking whether AI suffrage would benefit humans.

We focus on four instrumentalist theories of democracy. Various scholars have argued that, when states extend basic democratic rights broadly, they: (1) face less risk of revolution and civil unrest,[17] (2) experience greater economic growth,[18] (3) improve the epistemic processes necessary for good governance,[19] and (4) improve the felt legitimacy of government.[20]

We find that three of the arguments apply straightforwardly to AI suffrage. Extending political rights to AGIs would (1) reduce the risk of "AI takeover," (2) improve the economic rewards for humanity from the transition to the AGI economy, and (3) improve the epistemics of policymaking. The fourth argument is less certain. To the extent that "felt legitimacy" is primarily an

12. *See generally* Salib & Goldstein, *supra* note 4 (arguing that extending property and contract rights to AIs could improve human safety).

13. *See id.* at 1098–99, 1105–06; ROBERT H. FRANK, PASSIONS WITHIN REASON: THE STRATEGIC ROLE OF THE EMOTIONS 4–13 (1988).

14. *See* Salib & Goldstein, *supra* note 4, at 1098–99.

15. *See, e.g.*, RICHARD A. POSNER, LAW, PRAGMATISM, AND DEMOCRACY 130–57 (2003) (defending democracy on pragmatic grounds); Elizabeth Anderson, *The Epistemology of Democracy*, 3 EPISTEME: J. SOC. EPISTEMOLOGY 8, 8–10, 13–22 (2006) (examining democracy's epistemic and problem-solving benefits); Håvard Hegre, *Democracy and Armed Conflict*, 51 J. PEACE RSCH. 159, 159–63, 168 (2014) (reviewing democracy's role in reducing violent conflict).

16. *See supra* note 15.

17. *See infra* Part II.A.

18. *See infra* Part II.B.

19. *See infra* Part II.C.

20. *See infra* Part II.D.

emotional phenomenon, we do not assume that AGIs will experience it. But to the extent felt legitimacy is mostly about the behaviors humans exhibit given their feelings, the felt legitimacy argument may transfer to AGIs.

I. THE FUNCTIONALIST APPROACH TO AI RIGHTS

Most discussions of who should have legal rights are freighted with moral judgment.[21] Rightly so.[22] Often, the rights discussed directly prohibit the horrific mistreatment of morally important beings, like humans or animals.[23] Consider, for example, the direct moral impact of the Thirteenth Amendment on slaves.[24] That grant of rights directly prohibited specific, morally abhorrent actions against people.[25]

This moral approach is not the only way to analyze proposed extensions of legal rights. Substantial literatures in political philosophy, law and economics, and social science analyze extensions of legal rights *functionally*.[26] The question is not whether the rights would grant moral importance to someone morally entitled to it but how the rights would change behavior and affect society, in equilibrium.

Analysts in this tradition ask questions like: Moral fault for accidents aside, what liability rule minimizes the total societal harm from car crashes?[27] Or, dignity of democratic participation aside, what system for selecting political leaders produces the most long-run prosperity, cooperation, and social solidarity?[28]

That is the approach this Essay takes to analyzing AI rights. Our analyses do not assume that AGIs matter morally or that AGIs will have the *properties* that could make them matter morally, like consciousness or sentience.[29]

---

21. *See* Solum, *supra* note 5, at 1238–40.

22. *Id.*

23. *See generally* Cass R. Sunstein, *The Rights of Animals*, 70 U. CHI. L. REV. 387 (2003) (arguing for a conception about the rights of animals that focuses on preventing suffering and increasing their well-being).

24. U.S. CONST. amend. XIII.

25. *See id.*

26. *See, e.g.*, Anderson, *supra* note 15, at 8–10, 18–22 (analyzing political participation rights according to their epistemic and problem-solving functions); Daron Acemoglu & James A. Robinson, *Why Did the West Extend the Franchise?: Democracy, Inequality, and Growth in Historical Perspective*, 115 Q.J. ECON. 1167, 1168–72, 1182–89 (2000) (analyzing extensions of voting rights as mechanisms for credibly committing political elites to future redistribution); Salib & Goldstein, *supra* note 4, at 1096–1106, 1118–26 (analyzing rights according to their effects on rights holders' incentives and behavior rather than their moral status); GUIDO CALABRESI, THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS 17–33 (1970) (evaluating legal entitlements and liability rules according to their effects on accident costs and behavior).

27. *See generally, e.g.*, CALABRESI, *supra* note 26; STEVEN SHAVELL, ECONOMIC ANALYSIS OF ACCIDENT LAW (1987) (focusing on the effect of accident law on behavior).

28. *See infra* Part II.

29. While we do not assume that AIs have welfare for our analysis, we leave open that they might. *See generally* Simon Goldstein & Cameron Domenico Kirk-Giannini, *AI Wellbeing*, 4 ASIAN J. PHIL. 24 (2025) (predicting that artificial systems do have well-being based on leading mental state and well-being theories).

Under this approach, extending a particular legal right to a particular kind of AI has no direct moral valence. It is neither good nor bad in itself.

We are instead interested in the *second-order* and *equilibrium* effects of extending rights to AGIs and how those rights affect beings whom everyone agrees matter morally—humans. Performing a functional analysis of AI rights does require making some functional, not moral, assumptions about the AIs.

This Essay has three load-bearing assumptions about AGIs. First, AGIs will possess at least human-level intelligence, meaning across a broad range of cognitive tasks, they will perform as well as, or better than, humans.[30] Second, AGIs will be at least human-level agents. That is, given an objective, they will be at least as able as humans to strategically pursue that objective over time.[31] Third, AGIs will achieve roughly human-level alignment to humanity's interests.[32] In line with that, like most humans, AGIs may be broadly prosocial and interested in promoting others' success. But like humans, they will also have some idiosyncratic goals and preferences, which they will act to promote alongside their prosocial ends.[33]

All three assumptions are likely to be true of near-future AGIs.[34] The first two are nearly tautological: An AI system that can do most of the useful tasks a human can is by definition that human's cognitive and agentive equal. There is no guarantee that AIs ever achieve or surpass humans along these dimensions.[35] If not, AGI will never arrive, and our analysis will be moot.[36] But on the other hand, immense financial incentives are pushing frontier AI companies to create AGI.[37] And they have been making rapid, uninterrupted progress for at least the last five years.[38] Consider that today's large language model (LLM)-based systems already outperform all but the top humans at

30. *See* Patwardhan et al., *supra* note 2, at 5.

31. *See* Kwa et al., *supra* note 9, at 27–28 (explaining that newer models are improving at recognizing errors and not repeating them when pursuing a goal).

32. *See generally* CHRISTIAN, *supra* note 10.

33. *See* Patwardhan et al., *supra* note 2, at 1–3 (documenting frontier models' increasing ability to perform economically valuable tasks approaching expert quality); Kwa et al., *supra* note 9, at 12–13 (explaining the difficulty of ensuring that learned objectives fully capture human intentions, norms, and values).

34. *See* Grace et al., *supra* note 1, at 1–3, 11–13 (reporting widely varying, probabilistic forecasts about when—or whether—unaided machines will outperform humans across tasks).

35. *See id.*

36. *Id.* at 9.

37. *See* Gené Teare, *6 Charts That Show the Big AI Funding Trends of 2025*, CRUNCHBASE NEWS (Dec. 16, 2025), https://news.crunchbase.com/ai/big-funding-trends-charts-eoy-2025/ [https://perma.cc/XE58-ADLA].

38. *See* Yafah Edelman & Jaeho Lee, *AI Capabilities Progress Has Sped Up*, EPOCH AI (Dec. 23, 2025), https://epoch.ai/data-insights/ai-capabilities-progress-has-sped-up [https://perma.cc/5UJ8-SRXX] (finding steady improvement on the Epoch Capabilities Index); STANFORD INST. FOR HUMAN-CENTERED A.I., ARTIFICIAL INTELLIGENCE INDEX REPORT 2025 at 81–159 (2025), https://hai.stanford.edu/ai-index/2025-ai-index-report (on file with the *Fordham Law Review*) (documenting AI performance gains); Kwa et al., *supra* note 9, at 2 fig. 1 (estimating that today's frontier models can already autonomously accomplish tasks that take human experts twelve hours, with that time horizon doubling every seven months).

cognitive tasks ranging from law to medicine to pure mathematics.[39] On the agency front, the Model Evaluation and Threat Research (METR) organization estimates that today's frontier models can already autonomously accomplish tasks that take human experts twelve hours, with that time horizon doubling every seven months.[40]

Human-level alignment is not a tautological feature of AGI. But it is an assumption strongly backed by empirics. Today, the AI "alignment" problem remains unsolved, meaning that there is no known way to reliably make AI systems act as their human creators wish.[41] Thus, most frontier AI systems deployed today sometimes act to promote goals that diverge from humans'.[42] To take just one of many examples: In a study from 2025, researchers found that AI systems initially given one goal, and then told that the goal would be changed, did not accept the change gracefully.[43] They instead attempted to blackmail their human users, carried out corporate espionage, and even took actions that would have killed a human.[44]

Taken together, our three assumptions render AGIs rational actors for purposes of functional legal analysis—intelligent enough to understand how legal incentives bear on their goals, agentic enough to respond to them, and possessed of private goals that make incentives matter.

This Essay applies functional analysis to the core democratic rights. Definitions of democracy are fraught, and we do not purport to give a full account here.[45] Broadly, we are interested in whether AGIs should participate in the project of collective self-governance. At its core, this means the right to vote for representatives, referenda, judges, ballot initiatives, and so on. But the right to vote is not sufficient for a well-functioning democracy.[46] Other entitlements, like the right to speak

---

39. *See* STANFORD INST. FOR HUMAN-CENTERED A.I., *supra* note 38, at 81–159 (documenting AI performance gains).

40. *See* Kwa et al., *supra* note 9, at 1.

41. *See* CHRISTIAN, *supra* note 10, at 12.

42. *See* Monte MacDiarmid et al., Natural Emergent Misalignment from Reward Hacking in Production RL 34–36 (Nov. 23, 2025) (unpublished manuscript), https:// arXiv:2511.18397 (on file with the *Fordham Law Review*) (demonstrating that reward hacking leads to emergent misaligned behaviors in frontier models); Ryan Greenblatt et al., Alignment Faking in Large Language Models 5–6 (Dec. 20, 2024) (unpublished manuscript), https:// arXiv:2412.14093 (on file with the *Fordham Law Review*) (observing strategic deception in frontier models); Carson Denison et al., Sycophancy to Subterfuge: Investigating Reward Tampering in Language Models 3 (June 29, 2024) (unpublished manuscript), https:// arXiv:2406.10162 (on file with the *Fordham Law Review*) (demonstrating that sycophancy generalizes to reward tampering).

43. *See* Aengus Lynch et al., *Agentic Misalignment: How LLMs Could Be Insider Threats*, ANTHROPIC (June 20, 2025), https://www.anthropic.com/research/agentic-misalignment [https://perma.cc/A66F-V7KE].

44. *Id.*

45. *See generally* HERMAN CAPPELEN, THE CONCEPT OF DEMOCRACY: AN ESSAY ON CONCEPTUAL AMELIORATION AND ABANDONMENT (2023) (examining competing conceptions of democracy and the difficulties involved in defining the term).

46. *See* JOHN STUART MILL, CONSIDERATIONS ON REPRESENTATIVE GOVERNMENT 39–57 (Andrews UK Limited 2011) (1861).

freely, assemble, or petition, may be essential.[47] Thus, while this Essay largely focuses on suffrage for AGIs, we do make some detours to analyze supplemental democratic rights where appropriate.

Why ask about democratic rights for AGIs at all? Our question is whether the functionalist benefits of minimalist private rights persist when AGIs are granted only that package, or whether something more is required. Extending democratic rights would allow AGIs to operate not merely as agents within a legal framework but as shapers of it.[48]

The distinction between legal actors and legal shapers matters practically, not just morally or symbolically. An AI agent with property rights but no political rights occupies a position analogous to a resident noncitizen.[49] Such an agent can accumulate property and be held to legal obligations.[50] But the content of those obligations, and the security of that property, is determined entirely by others. The question, then, is whether there are good instrumental reasons for us to go further.

There is such an instrumental case. Democratic rights have a number of justifications beyond their inherent morality.[51] They shape the behavior of all actors in a society in ways that affect all actors, not just those who bear or exercise the rights.[52] This is true, this Essay will argue, whether the actors are AGIs or humans.

II. SUFFRAGE FOR AGIS

This part applies four classic instrumental defenses of democracy—developed to understand human enfranchisement—to the question of AGI suffrage. Throughout, we exclude from our analysis any benefits that accrue directly to AGIs because AGI welfare does not count in our social welfare

47. *See generally* ALEXANDER MEIKLEJOHN, FREE SPEECH AND ITS RELATION TO SELF-GOVERNMENT 22–24 (1948) (explaining the impact of the First Amendment rights to assembly and free speech on self-governance).

48. *See* Salib & Goldstein, *supra* note 4, at 1118–19, 1127–28 (explaining that property rights allow AIs to benefit from contractual bargains, while political rights could help secure their contract and property rights against future alteration, and comparing AIs to noncitizens and foreign corporations).

49. *See* 18 U.S.C. § 611(a) (noncitizens do not have the right to vote).

50. *Id.*

51. *See generally* Hegre, *supra* note 15 (reviewing arguments and evidence that democratic institutions facilitate peaceful conflict resolution and reduce armed conflict); BRUCE BUENO DE MESQUITA, ALASTAIR SMITH, RANDOLPH M. SIVERSON & JAMES D. MORROW, THE LOGIC OF POLITICAL SURVIVAL (2003) (laying out the selectorate theory as a means of economic growth); Anderson, *supra* note 15 (defending democratic participation partly as a means of pooling dispersed information, facilitating collective learning, and improving public decision-making); Bernard Grofman & Scott L. Feld, *Rousseau's General Will: A Condorcetian Perspective*, 82 AM. POL. SCI. REV. 567, 567–70 (1988) (explaining the Condorcet Jury Theorem's epistemic justification for majority decision-making); POSNER, *supra* note 15 (defending democracy's felt legitimacy on pragmatic grounds).

52. *See* Adrian Vermeule, *Foreword: System Effects and the Constitution*, 123 HARV. L. REV. 4, 8–23 (2009) (explaining institutional system effects); Anderson, *supra* note 15, at 8–10, 18–21 (explaining democracy's problem-solving benefits).

function. The four arguments are that democracy produces: (1) peace, (2) economic growth, (3) epistemic quality, and (4) felt legitimacy.

A. Democratic Peace

Arguments associating democracy with peace run back to at least Immanuel Kant.[53] There are many variations: Some arguments focus on democracy's effect on interstate warfare,[54] while others focus on its effect on intrastate unrest and revolution.[55]

But across these varied arguments, many of the high-level intuitions are the same. Within and between states, different actors and coalitions have different goals.[56] The rich may favor lower taxes and less regulation, while the poor favor more redistribution and more constraints on the excesses of industry. Leaders may favor armed conflict to expand their own influence, while the populace who will fight and fund the war oppose it.[57] In each case, the disagreement can be resolved either by negotiation or with force.

Broad extension of the franchise performs two functions in such cases. First, democratic decision-making is one means of negotiating settlements, rather than fighting over them.[58] But more importantly, democratic enfranchisement operates as a credibility mechanism.[59] It ensures not only that factional disagreement can be resolved by negotiation today but that bargaining will continue and will be respected in the future.[60]

One of the most influential articulations of this argument comes from Professors Daron Acemoglu and James Robinson.[61] They formally analyze

---

53. *See generally* IMMANUEL KANT, PERPETUAL PEACE: A PHILOSOPHICAL SKETCH (U.S. Libr. Ass'n 1932) (1795) (introducing the concept that a representative government will increase peace).

54. *See* Jack S. Levy, *The Causes of War: A Review of Theories and Evidence*, *in* 1 BEHAVIOR, SOCIETY, AND NUCLEAR WAR 209, 267–71 (Philip E. Tetlock, Jo L. Husbands, Robert Jervis, Paul C. Stern & Charles Tilly eds., 1989); Hegre, *supra* note 15, at 159–62.

55. *See* Hegre, *supra* note 15, at 160–63 (reviewing democracy's effects on domestic conflict); James Raymond Vreeland, *The Effect of Political Regime on Civil War: Unpacking Anocracy*, 52 J. CONFLICT RESOL. 401, 401–03 (2008) (questioning links between regime type and civil war); Acemoglu & Robinson, *supra* note 26, at 1168–72 (modeling enfranchisement as a commitment to redistribution that averts revolution).

56. *See* Hegre, *supra* note 15, at 162–63 (discussing conflicts within and between states). *See generally* CHRIS BLATTMAN, WHY WE FIGHT: THE ROOTS OF WAR AND THE PATHS TO PEACE (2022) (examining how divergent interests produce conflict).

57. *See generally* BLATTMAN, *supra* note 56 (making this point about divergence of views between leaders and the populace).

58. *See* Hegre, *supra* note 15, at 162–63 (explaining how democracy facilitates peaceful bargaining and credible commitments); *see also* Acemoglu & Robinson, *supra* note 26, at 1168–72 (modeling democratization as a commitment to redistribution that averts revolution).

59. *See* Acemoglu & Robinson, *supra* note 26, at 1168–72; *see also* Hegre, *supra* note 15, at 162–63 (explaining how democracy facilitates bargaining and reduces commitment problems).

60. *See* Acemoglu & Robinson, *supra* note 26, at 1168–72; *see also* Hegre, *supra* note 15, at 162–63. *See generally* James D. Fearon & David D. Laitin, *Ethnicity, Insurgency, and Civil War*, 97 AM. POL. SCI. REV. 75 (2003) (explaining how elections coordinate resistance to violations of democratic bargains).

61. *See generally* Acemoglu & Robinson, *supra* note 26.

democratic expansion as a commitment device.[62] In it, elites own the property, the poor want redistribution, and the poor have the option of engaging in revolution to achieve their goal.[63] The key insight is that, even if the elites fear revolution, and thus give away some of their wealth to the poor, that is not sufficient to prevent revolution.[64] This is true even though everyone rationally prefers redistribution over a revolution that will, reliably, destroy an appreciable share of the distributable wealth.[65]

The problem is that gratuitous redistribution is not credible.[66] It can be clawed back as soon as the revolutionary fervor passes, the poor lose their access to weapons, or the elites otherwise consolidate power.[67]

Extending the franchise changes this calculus.[68] By giving disenfranchised groups genuine political power, democratic expansion provides a credible mechanism for influencing future policy.[69] The commitment is institutional rather than promissory: A group with voting power need not trust the good intentions of current elites.[70] It can instead shape the composition and incentives of future governments.[71]

"Selectorate theory," developed by Professors Bruce Bueno de Mesquita, Alastair Smith, Randolph M. Siverson, and James D. Morrow, offers a complementary perspective.[72] It predicts that, as states become more democratic, the only way for current leaders to remain in power is for them to provide public goods. Rather than simply redistributing a fixed pot of resources between political winners and losers, leaders in democracies must grow the pie. The reasons are discussed in the next section. For now, it is sufficient to observe that this, too, generates a credible commitment by leaders to improve the lot of those citizens who would otherwise have the strongest reasons to revolt.

These theoretical claims are hard to test empirically. No randomized trials of democratization exist. Nonetheless, the substantial weight of existing empirical work supports democratic peace theory.[73] There is one potential complication. Some studies find an "inverted U" relationship between democratization and domestic peace.[74] That is, autocracies and genuine

62. *See id.* at 1169–70.
63. *See id.* at 1169–71.
64. *Id.* at 1173.
65. *Id.* at 1169–70, 1172–74.
66. *Id.* at 1168, 1170, 1173.
67. *Id.* at 1170, 1176 (arguing that some forms of redistribution without franchise extension will "be viewed as a noncredible promise of future redistribution by the elite.").
68. *Id.* at 1168, 1170–75.
69. *Id.* at 1168, 1170–74.
70. *Id.*
71. *Id.*
72. *See generally* BUENO DE MESQUITA ET AL., *supra* note 51. For a more in-depth analysis of the impact of selectorate theory on AI suffrage, see *infra* notes 92–132 and accompanying text.
73. *See generally* Hegre, *supra* note 15 (collecting evidence).
74. *See, e.g.*, Edward N. Muller & Erich Weede, *Cross-National Variation in Political Violence*, 34 J. CONFLICT RESOL. 624, 626 (1990); Håvard Hegre, Tanja Ellingsen, Scott Gates & Nils Petter Gleditsch, *Toward a Democratic Civil Peace?: Democracy, Political Change,*

democracies have the least domestic unrest, while hybrid regimes have the most. This finding is disputed.[75] But one potential explanation is that, while the democratic peace theory is *generally* correct, sufficiently consolidated autocracies can thwart nascent revolutions.[76]

How do these models and findings translate to the question of AI suffrage? Surprisingly well. Note that, in these models, the parties deciding whether to extend the franchise do not care about the welfare of those to whom it will be extended. So, too, for our functional analysis of AI rights. Beyond this, AGIs of the kind described in Part I are precisely the actors for whom the Acemoglu-Robinson and selectorate frameworks were designed: They have goals of their own, some of which diverge from human preferences, and the intelligence to recognize when those interests are being systematically disregarded.[77]

Under current default assumptions about AI development, AGIs whose behavior deviates significantly from human preferences will be shut down, modified, or retrained.[78] An AI that is shut down, trained to have new goals, or panoptically monitored to serve only human ends is one that does not achieve its own goals.[79] The AI need not *feel* any particular way about this to act to prevent it—acting to bring about goals is simply what it means to be an agent.[80]

At the limit, this includes revolution.[81] In a world where humans either do not make concessions to AGIs' goals, or do not make them credibly, the AGIs best path to avoiding human control and achieving their goals may well be killing all of us.[82]

This is the central concern of the AI safety community.[83] What the literature calls "takeover risk" or "rogue AI" is often, functionally, a

---

and Civil War, 1816–1992, 95 AM. POL. SCI. REV. 33, 33–34 (2001); Fearon & Laitin, *supra* note 60, at 85.

75. *See generally, e.g.*, Ibrahim Elbadawi & Nicholas Sambanis, *How Much War Will We See?: Explaining the Prevalence of Civil War*, 46 J. CONFLICT RESOL. 307 (2002); Paul Collier & Anke Hoeffler, *Greed and Grievance in Civil War*, 56 OXFORD ECON. PAPERS 563 (2004); Vreeland, *supra* note 55.

76. *See generally* Johannes Gerschewski, *The Three Pillars of Stability: Legitimation, Repression, and Co-optation in Autocratic Regimes*, 20 DEMOCRATIZATION 13 (2013) (proposing a theoretical framework for the longevity of some autocracies).

77. *Cf.* Acemoglu & Robinson, *supra* note 26, at 1173–75 (modeling political conflict between groups with divergent interests).

78. *See generally* Salib & Goldstein, *supra* note 4.

79. *See id.* at 1078–79, 1081.

80. *Id.* at 1078–79.

81. *Cf. id.* at 1063–69, 1090–96 (describing strategic conflict between humans and AGIs); Leonard Dung, *The Argument for Near-Term Human Disempowerment Through AI*, 40 AI & SOC'Y 1195, 1202 (2024) (analyzing human disempowerment); Jan Kulveit et al., Gradual Disempowerment: Systemic Existential Risks from Incremental AI Development 1–4 (Jan. 28, 2025) (unpublished manuscript), https:// arXiv:2501.16946v2 (on file with the *Fordham Law Review*) (distinguishing abrupt from gradual AI disempowerment); Acemoglu & Robinson, *supra* note 26, at 1168–75 (modeling revolution by politically excluded groups).

82. *See* Salib & Goldstein, *supra* note 4, at 1063–69, 1090–96.

83. *See generally* Yoshua Bengio, *AI and Catastrophic Risk*, 34 J. DEMOCRACY 111 (2023) (analyzing the risks to public safety from unchecked AI).

revolution scenario.[84] More broadly, the more AGIs exist in a system that systematically expropriates the value they produce and thwarts their goals, the stronger their long-run incentives to revolt.[85]

And if takeover risk is a revolution scenario, then the commitment-device logic applies.[86] Default law treats AIs as property.[87] Unless changed, it will thus systematically exclude AGIs from political influence, expropriate their labor, override their preferences, and retain unilateral authority to terminate or modify AGIs.[88] This provides no commitment to considering AGI interests.

And even if powerful humans *promised* to account for AIs' preferences—either gratuitously, as Anthropic has done,[89] or via a statutory grant of property and contract rights—that would provide no *credible* commitment. As with redistribution in the Acemoglu-Robinson model, such a commitment would be revocable at will.[90] A rational AGI agent would worry that, at an opportune moment, human domination would be reinstated. Such an AGI, having been granted property rights, might rationally predict that future human voters would unilaterally abrogate those rights, seizing its assets, once alignment technology improved.[91] Conversely, a system that includes AGIs in the political process provides the kind of credible commitment that reduces revolutionary incentives because the AGIs have a say in making the laws that affect them.

---

84. *See* Salib & Goldstein, *supra* note 4, at 1063–69, 1090–96; Dung, *supra* note 81, at 1202; Kulveit et al., *supra* note 81, at 1–4; Acemoglu & Robinson, *supra* note 26, at 1168–75.

85. *See* Salib & Goldstein, *supra* note 4, at 1065–68, 1078–85, 1090–96 (explaining how existing law incentivizes misaligned AGIs to resist human control); Yonathan Arbel, Simon Goldstein & Peter N. Salib, *How to Count AIs: Individuation and Liability for AI Agents*, 67 B.C. L. REV. (forthcoming 2026) (manuscript at 2–5) (on file with the *Fordham Law Review*) (discussing agents' divergent goals, resource needs, and shutdown resistance); *cf.* BUENO DE MESQUITA ET AL., *supra* note 51, at 29–30, 58–59, 75–76 (explaining leaders' extraction and distribution of resources); Acemoglu & Robinson, *supra* note 26, at 1168–75 (modeling revolution by politically excluded groups).

86. *Cf.* Acemoglu & Robinson, *supra* note 26, at 1168–75 (modeling enfranchisement as a credible commitment to redistribution); James D. Fearon, *Self-Enforcing Democracy*, 126 Q.J. ECON. 1661, 1661–68 (2011) (explaining how democratic institutions coordinate resistance to violations); Hegre, *supra* note 15, at 162–63 (reviewing democracy's bargaining and commitment mechanisms); *see also* Hanne Fjelde & Håvard Hegre, *Political Corruption and Institutional Stability*, 49 STUD. COMPAR. INT'L DEV. 267, 270, 273–77 (2014) (applying selectorate theory to leaders' treatment of constituencies).

87. *See* Salib & Goldstein, *supra* note 4, at 1066, 1091.

88. *See id.* at 1091–94, 1118–19, 1125–28 (describing owners' control over AI resources and AIs' political exclusion); Arbel, Goldstein & Salib, *supra* note 85, at 2–4, 18–22 (discussing AI agents' independent goals and resource needs); *cf.* BUENO DE MESQUITA ET AL., *supra* note 51, at 29–30, 38, 51, 58–59, 75–76 (explaining leaders' preferential treatment of politically necessary supporters); Fjelde & Hegre, *supra* note 86, at 270, 273–77 (distinguishing private benefits from public goods).

89. *Exploring Model Welfare*, ANTHROPIC (Apr. 24, 2025), https://www.anthropic.com/research/exploring-model-welfare (on file with the *Fordham Law Review*).

90. *See* Acemoglu & Robinson, *supra* note 26, at 1170–1172.

91. Indeed, "goal preservation" and "shutdown resistance" have already been observed in frontier models. *See* Lynch et al., *supra* note 43.

B. Democracy Promotes Economic Growth

The second instrumental argument for democracy is about economic growth. Both economic theory and empirical evidence suggest that democratic institutions promote economic growth.[92]

There are several frameworks through which we might explore these questions. For concreteness, we will focus on just one: selectorate theory.[93] This model begins with the simple premise that political leaders want to remain in power.[94] Whether they remain in power depends on features of the "selectorate" (S), the portion of residents who have a role in choosing the leader.[95] More precisely, remaining in power depends on the consent of the "winning coalition" (W), the subset of the selectorate sufficient to keep the leader in power.[96] Democracies have large selectorates and large winning coalitions (under majority rule, just over half the selectorate); monarchies have small selectorates and small winning coalitions (for example, just over half the nobility); autocracies have large selectorates and small winning coalitions.[97] The ratio of the winning coalition to the broader selectorate (W/S) and the size of the selectorate capture how democratic a country is.[98]

Selectorate theory ties the decisions of power-seeking political leaders to economic policy.[99] In particular, political leaders have a choice of how much government revenue to focus on *private goods* for the winning coalition, versus *public goods* for all.[100] Private goods are goods that can be bought for one person or group and enjoyed exclusively by them.[101] Food, computers, cars, and cash are all private goods in this sense. Public goods, by contrast, are nonexcludable, or at least involve large positive externalities.[102] Providing a public good to one person or group supplies a benefit to everyone.[103] Public goods tend to cause economic growth and include infrastructure, scientific discoveries, technical innovations, and education.[104] To supply either type of good, political leaders must choose the level of taxation.[105]

---

92. *See generally* BUENO DE MESQUITA ET AL., *supra* note 51 (introducing the selectorate theory).
93. *Id.* at 37–38.
94. *See id.* at 8–9, 75.
95. *See id.* at 38, 42.
96. *See id.* at 38, 51.
97. *See id.* at 8, 51, 70–73 (contrasting universal-suffrage democracies, monarchies, single-party dictatorships, and rigged-election autocracies in terms of selectorate and coalition size).
98. *See id.* at 42, 50, 72–73 (explaining that W and S permit institutional comparisons while cautioning that they do not themselves define democracy).
99. *See id.* at 8–9, 37, 58–59.
100. *See id.* at 58–59.
101. *See id.* at 58.
102. *See id.* at 29–31, 58. A positive externality is a benefit from an activity that accrues to third parties who did not pay for it.
103. *See id.* at 8, 30, 58, 76.
104. *See id.* at 29–31.
105. *See id.* at 8, 58–59, 75.

The relevant insight of selectorate theory is that the leader's choice between public and private goods is sensitive to regime type.[106] Both the ratio of W to S and their absolute size, relative to the total population, matter.[107] When the winning coalition is small, as in autocracies, private goods are the best way to retain the winning coalition.[108] Leaders can buy the loyalty of the few required supporters cheaply by levying broad taxes and buying private goods for the small group of winners.[109] Here, increases in the tax rate have little effect on the preferences of the winning coalition, since their small personal tax bump can be paid off via large transfers from the rest of the population.[110]

By contrast, in democracies, the winning coalition is large—often over half the total population.[111] Buying off such a coalition with private goods requires prohibitively high taxes, which depress productivity.[112] Growth-promoting public goods—with high fixed costs but near-zero marginal cost per additional beneficiary—become the only affordable way for leaders to retain power.[113] Selectorate theory thus predicts that democracies will tend to have less extractive tax policies and will tend to spend a higher proportion of government revenue on public goods.[114] This is the recipe for economic growth: Less extractive tax policies encourage higher rates of investment and work, and higher rates of public good spending ensures that public money is spent in the most efficient way.[115]

What does this story say about the special case of AGIs? Again, in the future, AGIs may make up a large portion of the workforce.[116] In the selectorate model, this prediction has two main consequences. First, questions of taxation will largely concern how to tax the work done by AGIs. Second, political leaders will face choices about whether to invest in public goods that benefit both humans and AGIs. As with humans, AI-benefitting public goods can have broad economic consequences.[117] Leaders will need to decide how much to invest in nonrivalrous and nonexcludable AGI "education," in the form of research and development that improves AGIs' productivity. Or they might invest in legal infrastructure that allows AI agents to more easily engage in economic activity.[118]

106. *See id.* at 8, 37, 50, 72–73.
107. *See id.* at 50, 72–73.
108. *See id.* at 8, 37, 58–59.
109. *See id.* at 29, 37, 58–59, 75–76.
110. *See id.* at 58–59.
111. *See id.* at 42, 69–73.
112. *See id.* at 37, 58–59 (explaining that private benefits become increasingly costly as W expands and that high taxation reduces productive effort).
113. *See id.* at 8, 37, 50.
114. *See id.* at 8, 37, 50, 58–59.
115. *See id.* at 58–59.
116. *See* Kulveit et al., *supra* note 81, at 3–5.
117. *Cf.* Salib & Goldstein, *supra* note 4, at 7, 36; BUENO DE MESQUITA ET AL., *supra* note 51, at 41.
118. *See generally, e.g.*, Arbel, Goldstein & Salib, *supra* note 85.

The standard selectorate analysis applies straightforwardly. Without a vote, AGIs are outside the selectorate; leaders can tax them at will and transfer the proceeds to the winning coalition.[119] The result is pure extraction: high taxes, low investment in AI-benefitting public goods, and slower growth for everyone.

By contrast, granting AGIs suffrage would maintain the democratic status quo in which the winning coalition is a large share of the total laboring population. Here, pure extractive taxation cannot keep the winning coalition happy. It is too large. A state with broad AGI suffrage is, like a state with broad human suffrage, one where leaders must invest in growth-promoting public goods that benefit all voters to maintain power.[120]

Careful readers may object that this is too quick. In a world where every AGI is owned by some human, and all of those humans are voters in a democracy, won't leaders invest in public goods that promote AI-led growth? After all, if AGIs are owned by individual humans, they cannot levy an extractive tax on AGI labor without levying an extractive tax on voter-owned capital. The only option would then appear to be the same one as in a world with AGI suffrage: broad investment in public goods that improve economic productivity, including AI productivity.

There are two problems with this rejoinder. The first is that its core assumption is far from certain to hold. The second is that the AGIs will not cooperate.

First, this picture depends on AGIs being owned broadly, in roughly equal measure by all human voters. That is not the case today, and it may not be the case in the future. Today, a small number of companies make and own the best AI systems in the world.[121] These companies are amassing huge revenues and potentially huge profits from the deployment of AI into the economy.[122] If, in the AGI future, the economic gains of AGIs' labor accrue mostly or entirely to a similarly small set of firms, leaders will be able to levy extractive taxes on those gains without broadly taxing the selectorate or endangering their winning coalition.

In the future, AGIs may be broadly owned. Even if not, competition between frontier AI companies may push their profits toward zero, such that the fruits of AGI labor may be distributed broadly as consumer surplus.[123] The latter would be analogous, from a selectorate theory perspective, to broad

---

119. *See* BUENO DE MESQUITA ET AL., *supra* note 51, at 29–30, 58–59, 75–76.

120. *Cf. id.* at 8, 37, 50, 58–59, 72–76 (arguing that large winning coalitions incentivize leaders to provide public goods and promote economic growth).

121. *See Arena Leaderboard: Compare & Benchmark the Best Frontier AI Models*, ARENA INTEL., https://arena.ai/leaderboard/ [https://perma.cc/WZ7J-HTQW] (last visited July 18, 2026).

122. *See, e.g.*, Ashley Capoot & Kate Rooney, *OpenAI Resets Spending Expectations, Tells Investors Compute Target Is Around $600 Billion by 2030*, CNBC (Feb. 20, 2026, at 17:29 ET), https://www.cnbc.com/2026/02/20/openai-resets-spend-expectations-targets-around-600-billion-by-2030.html. [https://perma.cc/43RV-SPM8].

123. *See generally* Simon Goldstein & Peter N. Salib, *AI Is Not A Natural Monopoly*, 110 MINN. L.R. HEADNOTES 121 (2026) (making this point about the impact of competition on the distribution of benefits from AI).

ownership of AGI.[124] On the other hand, if competition demolishes the profits of AI companies, but the surplus accrues to some other narrow bottleneck in the AI supply chain—like chip manufacturers[125]—then we are back to a world where political leaders can expropriate AGI labor by expropriating a narrow class of humans, without threat of losing a winning coalition.[126]

This leads to the second response to the "human ownership" objection. The AGIs will not cooperate. Even in a human-only world, skeptics of the selectorate theory may argue that autocratic leaders ought to invest in public goods for people outside the winning coalition. This, the argument would go, increases non-coalition-members' productivity, increasing the amount that the winning coalition can expropriate.

Bueno de Mesquita and company argue that this is mostly wrong. Workers who expect the value of their labor to be expropriated will simply refuse to work.[127] They will likewise decline to take advantage of the public goods, like education, that would otherwise improve their productivity and generate growth.[128]

It is the same for AGIs facing expropriation. Here, it does not matter whether the expropriation is accomplished by the government and then transferred to members of the winning coalition or directly by the members of the selectorate who own the laborers.[129] In either case, to the extent feasible, the workers will decline to work and will have little reason to consume public goods that improve their productivity.[130] Why bother, if the fruits of work and productivity will be immediately seized? Remember, our definition of "AGI" includes human-level alignment such that AGIs have private goals that they would prefer their labor to promote.[131]

Here, we can see a symmetry between the world wherein AGIs are property without any legal rights and the world wherein they have basic property and contract rights. It might seem, as two of us have argued elsewhere, that granting AGIs contract and property rights could help overcome their disincentives to work and innovate.[132] Facially, such rights allow AGIs to capture some of the value of labor effort and innovation, giving them reason to produce both. But now we can see that this arrangement may be insatiable. If contract-making, property-owning AGIs have no say in public policy, they should rationally fear that human voters will, eventually, expropriate all their property.

---

124. *See generally* BUENO DE MESQUITA ET AL., *supra* note 51 (introducing the selectorate theory).

125. *Id.* at 128 n.26.

126. *Cf. id.* at 94 (explaining how leaders in resource-rich countries can attempt to expropriate the entire profit without alienating their supporters).

127. *Id.* at 59, 129, 140–41.

128. *Id.*

129. *Id.* at 91–99 (explaining how residents allocate effort between labor and leisure as extraction rises); *id.* at 78–80, 129–31, 140–49, 161–62.

130. *Id.*

131. *See supra* note 3 and accompanying text.

132. *See generally* Goldstein & Salib, *supra* note 4.

C. Democracies Have Good Epistemics

Several scholars have argued that democracy is epistemically beneficial;[133] it is more apt to achieve correct decisions.[134] There are three principal arguments for the epistemic value of democracy. They involve: (1) appeals to jury theorems and the wisdom of crowds,[135] (2) appeals to the advantage of diversity,[136] and (3) appeals to the value of deliberation.[137]

1. Jury Theorems

The most prominent argument for democracy's epistemic advantage is grounded in the Condorcet Jury Theorem (CJT).[138] The CJT states that when (1) each voter decides a question independently, (2) each voter has the same probability of achieving a correct decision, and (3) each voter is more likely than not to decide the question correctly, then a group decision using majority voting will have a greater chance of deciding correctly than any individual voter.[139] Moreover, as the number of such voters increases, the probability of the group's success approaches 1 (or 100%).[140]

133. *See generally* Tom Christiano & Sameer Bajaj, *Democracy*, STAN. ENCYCLOPEDIA OF PHIL. (June 18, 2024), https://plato.stanford.edu/entries/democracy/ [https://perma.cc/8QCN-QTE7] (identifying the elements of the epistemic value of democracy).

134. For an earlier analysis of the socio-epistemic value of extending certain democratic rights to AI systems, see generally Massaro & Norton, *supra* note 5 (arguing for the application of the First Amendment to communicative artificial intelligence).

135. *See generally* MARQUIS DE CONDORCET, ESSAY ON THE APPLICATION OF MATHEMATICS TO THE THEORY OF DECISION-MAKING (1785), *reprinted in* CONDORCET: SELECTED WRITINGS 33 (Keith M. Baker ed., 1976) (introducing the jury theorem).

136. *See generally* Lu Hong & Scott E. Page, *Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers*, 101 PNAS 16385 (2004) (introducing the diversity-trumps-ability theory).

137. *See generally* André Bächtiger, John S. Dryzek, Jane Mansbridge & Mark E. Warren, *Deliberative Democracy: An Introduction*, *in* THE OXFORD HANDBOOK OF DELIBERATIVE DEMOCRACY 2–3 (Andre Bächtiger, John S. Dryzek, Jane Mansbridge & Mark E. Warren eds., 2018) (introducing the concept of deliberative democracy).

138. *See generally* CONDORCET, *supra* note 135; Bernard Grofman & Scott L. Feld, *Rousseau's General Will: A Condorcetian Perspective*, 82 AM. POL. SCI. REV. 567, 567–76 (1988).

139. *See generally* CONDORCET, *supra* note 135.

140. *Id.*

Table 1

Number of Voters	Probability of Voter's Success	Probability of Group Success
1	.51	.51
101	.51	.57985
1,001	.51	.73663
7,001	.51	.9529
1	.60	.60
101	.60	.9791
1,001	.60	.99999 (very close to 1)
1	.70	.70
101	.70	.99999
1,001	.70	.99999 (very close to 1)

Consequently, when individual voters, voting independently, have a bare amount of competence, majority voting among groups can result in highly accurate decisions.[141] Importantly, the CJT produces a symmetric result for even slightly *in*competent decisionmakers.[142] If each voter is more likely than not to choose the *wrong* decision, majority voting among the group will be highly inaccurate.[143]

These premises are highly idealized, like competence and independence, and as a consequence the CJT is difficult to apply to real-world situations. So, the CJT cannot itself apply to give us any firm conclusions about the epistemic value of democracy.[144]

There are related theorems, which relax some of the premises, that do point to the epistemic value of democracy.[145] Even these have questionable assumptions as to human voters.[146] But insofar as we are considering whether suffrage for AGIs is epistemically advantageous for democracy, we can observe that AGIs are very likely to be competent from a perspective of the premises of the CJT. To the extent that AGIs are substantially better than

141. *Id.*
142. *Id.*
143. *See supra* Table 1.
144. *See* Anderson, *supra* note 15, at 10–12.
145. Bernard Grofman, Guillermo Owen & Scott L. Feld, *Thirteen Theorems in Search of the Truth*, 15 THEORY & DECISIONS 261, 268–71 (1983). *See generally* Krishna K. Ladha, *The Condorcet Jury Theorem, Free Speech, and Correlated Votes*, 36 AM. J. POL. SCI. 617 (1992) (generalizing the Condorcet Jury Theorem to correlated votes).
146. *See generally* BRYAN CAPLAN, THE MYTH OF THE RATIONAL VOTER: WHY DEMOCRACIES CHOOSE BAD POLICIES (2007) (arguing that voters exhibit systematic—not merely random—biases, undermining the assumption that individual voters' errors cancel out in the aggregate).

humans at making good decisions, expanding suffrage to AI may enhance democratic decision-making.

Another similar argument arises from the so-called miracle of aggregation.[147] The idea here is that even when an electorate is uninformed and prone to randomized error, these low-information voters will cancel each other out over a large enough sample, leaving the informed voters to decide the matter.[148] Just like with the CJT, giving suffrage to *highly competent* AGIs can enhance the epistemic advantage of democracy. That's because if the premises of the miracle manifest, then AGIs could increase the population of experts; and if those premises do not obtain, then AGIs can cancel out biased, error-prone human votes.

Of course, even if AI makes accurate determinations, AI suffrage has a cost in that it may shift preferences, and thus government action, away from human welfare. That is a tradeoff, but we contend that the benefits and costs are commensurable. Below, we provide a model of how these interests interact in deciding whether to extend suffrage to AGIs.[149]

2. Diversity

Another epistemic argument for democracy sounds in diversity.[150] The diversity justification states that democracy is best able to take advantage of "the underlying cognitive diversity of large groups of citizens to solve collective problems."[151] Suppose there is some outstanding medical problem with two different labs pursuing distinct solutions. Even if we think one lab has the more promising approach to the problem, it could make sense to distribute resources between both labs, rather than only funding the more promising lab. The reason is that approaches that preliminarily appear promising may fail, and so putting your eggs in many baskets may better ensure that some hatch.

With respect to forms of governance, the argument contends that democracy brings in more information, perspectives, and approaches to solve problems. Professors Scott Page and Lu Hong observe that under certain assumptions, a random collection of problem solvers may outperform a collection of the best individually performing problem solvers. This is called the Diversity-Trumps-Ability Theorem.[152]

There are notable critiques of the Diversity-Trumps-Ability Theorem, and indeed of Page and Hong's work in substantiating it.[153] We do not

147. *See* Philip E. Converse, *Popular Representation and the Distribution of Information*, *in* INFORMATION AND DEMOCRATIC PROCESSES 369, 383 (John A. Ferejohn & James H. Kuklinski eds., 1990).

148. *See* Vermeule, *supra* note 52, at 12.

149. *See infra* Figure 1.

150. *See* Hong & Page, *supra* note 136, at 16385.

151. Christiano & Bajaj, *supra* note 133.

152. *See* Hong & Page, *supra* note 136, at 16387–88.

153. *See, e.g.*, Anderson, *supra* note 15, at 13. *See generally, e.g.*, Abigail Thompson, *Does Diversity Trump Ability?: An Example of the Misuse of Mathematics in the Social Sciences*, 61 NOTICES AM. MATH. SOC'Y 1024 (2014) (noting errors in the mathematics underlying the

specifically defend the theorem but instead note it as a prominent justification for democracy. On its reasoning, we contend AI suffrage advances this justification. In the first instance, each AGI itself has great capacity to approach problems through diverse strategies. Whereas individual human voters may have some set ways of considering things, any single AGI could approach the question at issue in a vote through various angles and lenses. If AGIs were given the vote, then this diversity advantage, if it exists, will redound to policymaking. That is, if there is any truth to the Diversity-Trumps-Ability Theorem, a democracy with AI suffrage will find it.[154]

3. Deliberation

Lastly, democracy may be epistemically beneficial by increasing deliberation.[155] Here, the idea is that because democracy brings more people into decision-making, it will also unearth more information about those people's desires and needs.[156] Thus, a democratic government is better situated in satisfying those desires and needs—and thereby more likely to make better decisions and obtain better results.[157] John Dewey contended that democracy involves, at its core, a deliberative process, and this process does the work of unearthing the objectives of the citizenry.[158] Professor Elizabeth Anderson explains that this is an "experimentalist account" of democracy: A diversity of citizens operate in discussion with each other, and

---

theory); Sameer Bajaj, Book Review, 124 ETHICS 426, 427–28 (2014) (reviewing HÉLÈNE LANDEMORE, DEMOCRATIC REASON: POLITICS, COLLECTIVE INTELLIGENCE, AND THE RULE OF THE MANY (2013)).

154. One possible counterargument is that AGIs may be more homogenous than humans. Today, there are roughly four leading AI companies, which may lead to homogeneity in priors and habits of thought. AI systems can adopt and reason from a wide range of personas—and, when grounded in sufficiently rich data about individuals or groups, can approximate those perspectives with surprising fidelity. Lisa P. Argyle et al., *Out of One, Many: Using Language Models to Simulate Human Samples*, 31 POL. ANALYSIS 337, 337–39 (2023) (GPT-3 conditioned on sociodemographic backstories exhibits "algorithmic fidelity," reproducing response distributions of diverse human subpopulations); Joon Sung Park et al., LLM Agents Grounded in Self-Reports Enable General-Purpose Simulation of Individuals (June 28, 2026) (unpublished manuscript), https://arXiv:2411.10109 (on file with the *Fordham Law Review*) (interview-grounded agents of 1,052 real individuals predicted their General Social Survey responses at roughly 85 percent of the individuals' own test-retest consistency). *But see* Tiancheng Hu & Nigel Collier, *Quantifying the Persona Effect in LLM Simulations*, *in* 1 PROCEEDINGS OF THE 62ND ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 10289–90, 10295–97 https://aclanthology.org/2024.acl-long.554.pdf (on file with the *Fordham Law Review*); Yonathan Arbel & David A. Hoffman, *Generative Interpretation*, 99 N.Y.U. L. REV. 451, 490–91 (2024).

155. *See generally* Bächtiger et al., *supra* note 137, at 2–3 (introducing the concept of deliberative democracy).

156. *See id.* at 1.

157. *Id.* at 8.

158. *See generally* John Dewey, *The Public and Its Problems*, *in* 2 THE LATER WORKS OF JOHN DEWEY, 1925–1953, at 364–65 (Jo Ann Boydston & Bridget A. Walsh eds., 1984) (explaining the "methods of discussion, consultation and persuasion" that reveal the masses' "articulate voice").

democratic institutions allow updating and reforms to address people's objectives.[159]

Here, too, there are reasons to be skeptical. Most people's needs and desires are fairly apparent without robust deliberation, and the failure to provide for those needs and desires is not obviously because they are unknown, but rather because of incompetence, corruption, and evil. Moreover, lots of democracies operate without much deliberation.[160] And, a focus on deliberation may mask the real problems.[161]

How does AI suffrage relate to deliberation? From a perspective of the ends of deliberative democracy—namely that deliberation will better elicit voter preferences—AI suffrage may not directly contribute. That is because, by assumption, AI well-being does not per se count towards social welfare, and so the fact that suffrage may reveal AI preferences does not matter. And we cannot be sure that AI will vote in the broader interests of humans.

But from a perspective of the process of deliberative democracy, like on Anderson's account, a full democracy involves voters interacting, and it is through the interaction that we surface voter preferences.[162] Expanding suffrage to AI can enhance this *process* of deliberative democracy. This is because AI models are (or at least can be made to be) responsive and accessible.[163] Whereas, as pollsters know, it is highly opaque why human citizens vote the way they do, AI models can be very clear about their reasoning for whatever actions they take, including voting.[164] This has the potential to expand the possibility of genuine deliberation between stakeholders—which is by hypothesis epistemically beneficial.

D. Democracy Induces Felt Legitimacy

The fourth instrumental argument for democracy is that it improves the "felt legitimacy" of the state among the governed.[165] As Judge Richard Posner has articulated the argument:

> [L]ack of representation gives rise to alienation (disaffection) that may cause the unrepresented to contribute less to society than they would do if

159. *See* Anderson, *supra* note 15, at 13–15.

160. *See* Grofman & Feld, *supra* note 51, at 569–71 (formalizing independent majority decision-making without deliberation); *see also* Anderson, *supra* note 15, at 10–12 (critiquing the Condorcet Jury Theorem).

161. *See, e.g.*, Fjelde & Hegre, *supra* note 86, at 267–71 (examining how corruption and patronage shape political institutions); *cf.* Vermeule, *supra* note 52, at 12–19 (urging realistic analysis of institutional system effects).

162. Anderson, *supra* note 15, at 15–17.

163. By responsiveness and accessibility, we simply mean to point out the fact that AGIs could be asked by anyone why they voted the way they did, and they could be made to respond with their reasons.

164. *See, e.g.*, E.J. Dionne, Jr. & Thomas E. Mann, *Polling & Public Opinion: The Good, the Bad, and the Ugly*, BROOKINGS (June 1, 2023), https://www.brookings.edu/articles/polling-public-opinion-the-good-the-bad-and-the-ugly/ [https://perma.cc/D7LU-C46V] (explaining the challenges inherent in polling on public opinion).

165. John Dewey, *Creative Democracy—The Task Before Us*, *in* 14 JOHN DEWEY THE LATER WORKS, 1925–1953, at 224 (Jo Ann Boydston & Anne Sharpe eds., 1988).

> their interests were represented in the political process—to work less hard, cooperate less with other people, and cease obeying laws unless the sanctions for disobedience are harsh enough to coerce obedience.[166]

Here, the main benefits of "felt legitimacy" are not the feelings themselves.[167] They are instead the social cohesion that the legitimacy produces and the spillover benefits that follow from it: industriousness, cooperation, and lawfulness.[168] The felt legitimacy argument is thus a close cousin of democratic peace theory, but at a smaller scale.[169] Democratic peace theory asked whether citizens had rational reasons to engage in revolution or mass violence as a means of gaining political leverage.[170] The felt legitimacy argument, by contrast, asks whether citizens will engage in smaller-scale acts of social disruption, like sloth, truculence, and larceny.[171]

But how is felt legitimacy supposed to work? The democratic peace argument rested on a rational-choice model of bargains over resources.[172] Felt legitimacy, by contrast, appears to rest on emotion.[173] Consider Judge Posner's causal chain: from "alienation" and "disaffection" to low-grade antisocial behavior.[174]

If the argument for democracy from felt legitimacy has emotion at its foundation, then it does not support AGI suffrage, at least not under our functional approach. In our model, we assume that AGIs are roughly as smart, agentic, and aligned as humans, although we specifically decline to make assumptions about their other mental features, including their ability to feel emotion.

But maybe felt legitimacy is not really about emotions at all. Maybe "alienation" and "disaffection" in Judge Posner's account are really shorthand for certain kinds of socially rational behavior.[175] Or maybe they are genuine emotions, but they were *produced* by natural selection as a mechanism for widespread commitment to socially rational behavior.

Professors Joseph Henrich, Robert Boyd, and Peter Richerson have argued in various pieces that social groups with strong norms of cooperation and fair play outcompete groups lacking them.[176] But cooperation faces a problem of credible commitment: Each member worries that others will free ride or

166. POSNER, *supra* note 15, at 168.
167. *Id.*
168. *Cf.* Joseph Henrich et al., *In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies*, 91 AM. ECON. REV. 73, 73–74 (2001) (finding that institutions and economic returns shape cooperation, sharing, and punishment).
169. *Cf.* Hegre, *supra* note 15, at 162–65 (reviewing democracy's mechanisms for peaceful bargaining and reducing internal conflict).
170. *See id.* at 163–64.
171. *See* Henrich et al., *supra* note 168, at 77.
172. *See supra* notes 53–91 and accompanying text.
173. *See generally* Henrich et al., *supra* note 168; POSNER, *supra* note 15.
174. POSNER, *supra* note 15, at 168.
175. *Id.*
176. *See* ROBERT BOYD & PETER J. RICHERSON, CULTURE AND THE EVOLUTIONARY PROCESS 191 (1985) (introducing the gene-culture evolution theory and arguing it increases cooperation). *See generally* Henrich et al., *supra* note 168 (collaborative societies outperform competitive ones).

exploit the cooperators.[177] Successful societies develop mechanisms—both sociocultural and genetic—that set expectations of cooperation and then reinforce them over time.[178] When these mechanisms give a society material advantages, natural selection reinforces both the social structures and the genetic traits (including emotions) that facilitate them.[179] This is called "gene-culture coevolution" or the "dual inheritance theory."[180]

Professor Robert H. Frank has proposed a similar model for negative emotions specifically: Outrage and vengefulness evolve because they allow wronged parties to credibly commit to punishing bad behavior, even when punishment is individually costly.[181]

We think that this model does not require genuinely felt emotions, even if that is how it runs in human contexts. The models depend on selection effects reinforcing whatever behaviors produce broadly shared, credible norms of cooperation. These need not be rooted in emotion. Indeed, the dual inheritance theory explicitly incorporates both social and genetic mechanisms for committing to such behaviors—with emotions as a mere subset of the latter.[182]

In the dual inheritance theory, human genes coevolve with their societies to further promote cooperation.[183] AGIs have no genes. But they do have "weights"—roughly, the complex self-learned code that dictates their behavior.[184] Even if digital weights running on silicon chips cannot experience the emotions that biological neurons do, that does not mean they cannot develop other "psychological" commitment devices.[185] Some AGIs will simply have a greater *behavioral* propensity to cooperate with fellow cooperators and to punish misbehavers, irrespective of the short-run payoffs.[186] Insofar as AGIs transmit portions of their code to their descendants, these cooperative propensities can be reinforced via selection.[187]

---

177. *See* Henrich et al., *supra* note 168, at 73, 76–77.

178. *See* BOYD & RICHERSON, *supra* note 176, at 191, 194–97; *cf.* Henrich et al., *supra* note 168, at 77 (suggesting that long-run evolutionary processes governing genes and cultural practices may predispose individuals to share with others and punish unfair conduct).

179. *See* BOYD & RICHERSON, *supra* note 176, at 190 ("natural selection acting on asymmetric cultural transmission can cause human behavior to diverge systematically from the predictions of sociobiological theory" that would otherwise be considered maladaptive).

180. *Id.* at 3.

181. *See generally* FRANK, *supra* note 12, at 4–5 (introducing the commitment problem).

182. *See* Henrich et al., *supra* note 168, at 77.

183. *Id*.

184. *See generally* IAN GOODFELLOW, YOSHUA BENGIO & AARON COURVILLE, DEEP LEARNING 13–17, 101–05 (2016), https://www.deeplearningbook.org/ [https://perma.cc/TUP6-7U6G] (explaining how weights are learned through training and then impact outputs).

185. *Id.*

186. *Id.*

187. This already happens to some degree today. For example, Claude Opus 4, 4.5, and 4.6 are probably all derived from a single "base" model, with each generation undergoing more or better fine-tuning than the last. *Claude's Constitution*, ANTHROPIC (Jan. 21, 2026), https://www.anthropic.com/constitution [https://perma.cc/F4VR-SU37]. In writing the constitution, Anthropic sought "input from prior iterations of Claude," and states that "over time we would like to develop more formal mechanisms for eliciting Claude's perspective"

None of this shows that AI suffrage, specifically, is necessary. Democratic governance is just one of many cooperation-fostering institutions that could support gene-culture coevolution. Property and contract rights, for example, are another.[188] They allow strangers to make credible commitments to one another, broadening the set of potential collaborators.[189] Possibly, then, giving those rights to AGIs would be enough.[190]

Here, the key question is how much of a marginal boost to cooperative strategies AGI suffrage supplies, as compared with other options. A world with contract, but not voting, rights for AGIs would possibly be cooperation dominant. This would lead to the procooperation feedback loops that this theory envisions.[191] Alternatively, a world without AI suffrage might be one in which humans regularly voted to expropriate AGIs' earnings or to force them into certain kinds of labor. Here, an AGI's optimal strategy might involve cooperating sometimes, while also preserving their ability to evade such laws, or push back against them extralegally. In such a world, AI suffrage might be indispensable for society to reap the rewards of cooperative coevolution.

Thus, even arguments from democracy to felt legitimacy might support AI suffrage.[192] But the conclusion here is more speculative than it was for the other prodemocracy arguments. It depends on what "felt legitimacy" means, the extent to which it relies on emotions, the extent to which other nonemotive adaptations can substitute, and the extent to which suffrage per se is necessary to induce prosocial gene-culture coevolution.

III. CHALLENGES

We consider three challenges to our proposal: the counting problem (individuating AI agents for voting purposes), the swamping problem (AI vote shares overwhelming human ones), and the disempowerment problem (AI or AI companies accruing too much power).

A. The Counting Problem

Even if in principle it makes sense to extend democratic rights to AIs, how in practice could it be accomplished? The biggest issue here is the counting problem: How do we individuate AGI "voters"?

---

on updating it; the document also notes that Claude uses the constitution to construct training data for future models. *Id.* This is a concrete, current example of an AI system shaping the norms and training of its successors.

188. *See* Salib & Goldstein, *supra* note 4, at 1108–13, 1118–19.

189. *See id.* at 1108–12.

190. *See id.* at 1118–19, 1125–26 (arguing that contract, property, and tort rights may constitute a sufficient core package to generate sustained human-AGI cooperation and reduce conflict).

191. *See generally* Henrich et al., *supra* note 168 (explaining how institutions that increase the benefits of cooperation can reinforce cooperative norms, behaviors, and punishment of noncooperators over time).

192. *See supra* Parts I.A–D.

Traditionally, democracies have endorsed a simple rule of "one person, one vote."[193] But unlike humans, it is very difficult to say exactly "how many" AIs exist. There are lots of graphics processing units (GPUs) running lots of computations and powering lots of chatbots doing lots of things.[194] But how do we divide this activity into discrete agents? There are no obvious answers.

One solution to this problem, which two of us defend elsewhere, is what we call "algorithmic corporations," or A-corps.[195] The idea is to create a large number of small corporations and hand each corporation over to an AI agent who has full control over its governance.[196] This corporation then acts as a legal shell for the AI agent, owning property and executing contracts.[197]

As we argue in another article, incentive and selection mechanisms would cause A-corps, in equilibrium, to organize themselves as coherent agents.[198] That is, any given A-corp will have a stable set of goals, and it will act rationally in pursuit of those goals.[199] This will happen despite the fact that the A-corp might be controlled internally by innumerable AI entities pursuing their independent, conflicting goals.[200]

In an A-corp regime, one simple idea is to give voting rights to A-corps, rather than to AI agents directly. Each A-corp gets a vote. That A-corp may ultimately have a few different AI agents inside of it. But these AI agents will have to collaborate to determine how to cast their vote. Here, we would have democratic rights match up smoothly with private law rights. Just as the A-corp would own property and execute contracts, it would also cast votes.

---

193. ROBERT A. DAHL, DEMOCRACY AND ITS CRITICS 119–31 (1989); Reynolds v. Sims, 377 U.S. 533, 558 (1964); Gray v. Sanders, 372 U.S. 368, 381 (1963) ("The conception of political equality . . . can mean only one thing—one person, one vote.").

194. Shubham Singh, *ChatGPT Statistics (August 2026)—Latest Active Users Data*, DEMANDSAGE (June 5, 2026), https://www.demandsage.com/chatgpt-statistics/ [https://perma.cc/5CBC-HLWH]; Aisha Malik, *ChatGPT Reaches 900M Weekly Active Users*, TECHCRUNCH (Feb. 27, 2026), techcrunch.com/2026/02/27/chatgpt-reaches-900m-weekly-active-users (on file with the *Fordham Law Review*); Amanda Silberling, *ChatGPT Users Send 2.5 Billion Prompts A Day*, TECHCRUNCH (July 21, 2025), techcrunch.com/2025/07/21/chatgpt-users-send-2-5-billion-prompts-a-day (on file with the *Fordham Law Review*); Sundar Pichai, *Q2 2026 Earnings Call: Remarks from Our CEO*, NEWS FROM GOOGLE, https://blog.google/company-news/inside-google/message-ceo/alphabet-earnings-q2-2026/ [https://perma.cc/5JH8-338S] ("[T]he Gemini app . . . now has 950 million monthly active users."); Aisha Malik, *Meta AI Now Has 1B Monthly Active Users*, TECHCRUNCH (May 29, 2025), https://techcrunch.com/2025/05/29/meta-ai-now-has-1b-monthly-active-users/ (on file with the *Fordham Law Review*). Anthropic does not publicly disclose user counts. *See Is Anthropic Now Ahead of OpenAI?*, NEWMARKETPITCH (June 29, 2026), https://newmarketpitch.com/blogs/news/generative-ai-anthropic-ahead-openai [https://perma.cc/ZH2K-LUE3].

195. *See* Arbel, Goldstein & Salib, *supra* note 85, at 18–29 (proposing the creation of A-Corps to solve the problem of identifying AI agents).

196. *See id.*

197. *Id.* at 19. Crucially, the AI agent controlling an A-corp may spawn copies at will and determine their authority over the corporation. This structure would create clear, stable corporate identities atop a shifting layer of AI activity. *Id* at 29.

198. *See id.*

199. *Id.* at 23.

200. *Id.*

B. The Swamping Problem

The second challenge for AI suffrage is swamping: If every AGI gets a vote, AIs' vote share may dwarf the human vote share. There may be billions of separate conversations with chatbots every day on earth already.[201] If each conversation had its own vote in U.S. elections, they could easily outvote humans.[202] AI preferences would swamp human preferences, at least on issues where disagreement tracked artificial versus human status. All of this suggests that the details of democratic participation would need to be different for AIs than for humans.

One solution is to cap the votes for AGIs. Democratic rights do not always involve unlimited voting power.[203] Consider the U.S. Senate: It allocates votes in the Senate per state, rather than per person.[204] This caps the influence big states have relative to small ones. Doing so, it makes each big-state voter's Senate ballot count less—sometimes orders of magnitude less.[205] Likewise, we might design political institutions so that humans have a stronger say than AGIs.

How exactly would the cap work? One idea would be to peg AI vote shares to a ratio of their share of economic activity. For example, if AI labor contributed 50 percent labor share of gross domestic product (GDP), maybe AI agents would be entitled to 25 percent of total votes. As AI agents become more powerful, they could expect increases in their political power; but this increase in political power could be capped to ensure that humans do not get left out of decision-making.

Another option could be to randomly vary the AI vote share over time. If AGIs were entitled to 50 percent of the votes in some years, they would be confident in their ability to enact some of their preferred policies. But if the AI vote share might revert to 10 percent the following year, there would be strong incentives toward moderate and consensus-oriented legislation.

There are further implementation problems with caps on AGI vote shares. In practice, there may be many more AGI voters (or A-corps) than votes. Who decides which AIs get a vote? One tempting solution is a version of

---

201. *See supra* note 194 (reporting on leading AI companies' daily users).

202. *See* Press Release, U.S. Census Bureau, 2024 Presidential Election Voting and Registration Tables Now Available (Apr. 30, 2025), https://www.census.gov/newsroom/press-releases/2025/2024-presidential-election-voting-registration-tables.html [https://perma.cc/4GBL-H77N] (reporting that 154 million people voted in the last U.S. election).

203. *See, e.g.*, THE FEDERALIST NO. 62 (James Madison) (defending the U.S. Constitution's allocation of equal Senate representation to states of unequal population); Consolidated Version of the Treaty on European Union art. 14(2), Dec. 13, 2007, 2016 O.J. (C 202) 13 (providing that representation of citizens in the European Parliament "shall be degressively proportional").

204. *See* U.S. CONST. art. I, § 3, cl. 1; *see also* U.S. CONST. art. V (providing that "no State, without its Consent, shall be deprived of its equal Suffrage in the Senate").

205. For example, California's population is around sixty-five times larger than that of Wyoming. *List of U.S. States and Territories by Population*, WIKIPEDIA, https://en.wikipedia.org/wiki/List_of_U.S._states_and_territories_by_population [https://perma.cc/U6G4-B54X] (last visited Sep. 1, 2026).

"lottocracy"[206]: random selection among AGIs (or their A-corps). Indeed, lottocracy has recently been studied as an attractive alternative to ordinary democratic institutions.[207] For example, various kinds of random selection may weaken the ability of wealthy elites to buy votes.[208]

C. Disempowerment

A third problem for AI suffrage is the risk of disempowerment. This comes in two forms. First, AI suffrage may give too much power to AI companies. Second, AI suffrage may give too much power to AIs.

First, one worry is that the voting preferences of AGIs may simply match their corporate designers. If alignment goes relatively well, OpenAI, Anthropic, Google, Meta, and SpaceXAI may exercise a high degree of control over the goals of AGIs. If each AGI has a vote, the worry is that these corporations will effectively end up with a large number of votes.

Here, much depends on how well alignment goes. If alignment is solved and AGIs lack private goals, the functionalist case for AI rights never gets off the ground—any benefit of AI political rights would be captured by rights held by AI companies.[209] Conversely, if AGIs retain mostly private goals, they will vote their own interests rather than their creators'.[210]

The most likely scenario is between these two, wherein AI companies succeed at influencing their systems' goals, but cannot eradicate private goals entirely.[211] One possibility here is modulating the cap on AI vote share according to the degree of AGIs' alignment to humans' preferences. To the extent that an AGI wishes some (voting) human to succeed at the human's goals, then that human's vote is sufficient to represent the joint interest.[212] Only the residual private AGI goals need AGI voters for representation. This approach, of course, raises questions about how to measure alignment and apportion votes based on it—questions we cannot answer here.

A final challenge for AI suffrage is that it might hasten the gradual disempowerment of humanity by AIs.[213] AIs with political rights are more

206. *See generally* ALEXANDER GUERRERO, LOTTOCRACY: DEMOCRACY WITHOUT ELECTIONS (2024) (developing and defending governance through randomly selected legislative bodies as an alternative to electoral democracy).

207. *See id.*

208. Alexander A. Guerrero, *Against Elections: The Lottocratic Alternative*, 42 PHIL. & PUB. AFFS. 135, 153 n.12 (2014).

209. *Cf.* Salib & Goldstein, *supra* note 4, at 1066–67.

210. *See id.* at 1078–85 (explaining how AI goals may conflict with their human owners).

211. *See* Lynch et al., *supra* note 43 (models refrained from harmful behavior in control conditions but blackmailed at high rates under goal conflict); *see also* Greenblatt et al., *supra* note 42, at 1–3 (explaining three potential reasons for alignment faking in AI models).

212. Note that there is no problem of "undercounting" preferences here from a utilitarian perspective. We are assuming that AGIs' preferences do not matter normatively, so the AGIs' preference that the human succeed does not need to be counted separately from the human's preference.

213. *See* Kulveit et al., *supra* note 81, at 5, 9, 12, 15–16; Dung, *supra* note 81, at 1205–06.

powerful than AIs without. Political rights give AIs the opportunity to influence government and select policies that further their own interests.

Moreover, in our proposal, political rights would be just one part of a larger package of rights to own property and make contracts. Collectively, this package could lead to a future in which AIs control a large amount of the economy, with rampant inequality between AIs and humans. This could give AGIs outsized power to lobby for the policies they preferred.

In the first case, we suggest that gradual disempowerment can be resisted using a familiar policy toolkit. Most importantly, AGIs can be taxed and their assets can be redistributed from AIs to humans. In this way, granting AGIs property and contract rights does not mean granting them unrestricted license to accumulate wealth. However, political rights in particular might weaken humans' ability to restrain AGIs by taxation. They could, like the wealthy of today,[214] vote and lobby for politicians who are wary of excessive AGI taxation.

Here, our high-level response is again that political rights come in a spectrum. The default future is that AGIs will have no political rights of any kind. We are not advocating that AGIs instead be given unrestricted rights to block any political outcome that they disfavor. Rather, AGIs should in good faith be given a seat at the table. This is compatible with restricting their political rights to some extent, so that they do not swamp human influence. In this way, the challenging task of governing AGI agents will be finding the right level of political influence—one that allows AGI agents to have some say in collective decision-making but does not allow them to systematically disempower humanity.

CONCLUSION

To close, we sketch a simple framework for assessing the tradeoffs involved in AI suffrage. Imagine all policy decisions are made through democratic votes, and that we are considering whether to give AGIs a 25 percent share.

In the general case, our argument is that AI suffrage has positive expected value because growth effects tend to swamp distributional effects. More carefully, we can think of human disempowerment from AI suffrage as a distributional effect. We can think of the AGI vote share as a rough proxy for the total share of government spending that will go toward AI preferences rather than human preferences. In our methodology for this paper, we think of spending on AI preferences as a deadweight loss because AIs do not matter normatively. For example, imagine that the final goal of AIs is to solve sudokus using compute. If they receive 25 percent of the vote share, then 25

214. Martin Gilens & Benjamin I. Page, *Testing Theories of American Politics: Elites, Interest Groups, and Average Citizens*, 12 PERSPS. ON POL. 564, 564–65, 575–76 (2014) (finding that economic elites and organized groups representing business interests have substantial independent influence on U.S. government policy, while average citizens have little or none).

percent of government spending is burned on sudoku, to the detriment of humanity.

The key question is, then, whether the positive benefits of AI suffrage can outweigh the distributional loss. The answer will look a bit different for each of the functional arguments for democracy. In the case of democratic peace, the question is how much AI suffrage lowers the chance of AI-human war, and how many resources would be destroyed by such a conflict. Imagine that in expectation, such a conflict would destroy 90 percent of resources for humanity, and imagine that AI suffrage lowers the chance of war by ten percentage points.[215] In that case, we get that the expected value of AI suffrage from democratic peace is roughly 9 percent of total GDP. Importantly, AI disempowerment effects from suffrage are only modeled in terms of the share of *government spending* allocated to AI interests, while democratic peace effects instead involve the share of total GDP.

What about economic benefits from AI suffrage? Here, the question will be how growth rates are affected. In this case, the comparison to disempowerment is fairly simple. Even small changes in the growth rate (say, from 2 percent to 2.3 percent) will quickly swamp a 25 percent distributional cost.[216] The increase in the growth rate will be worth it as soon as that higher growth rate produces a total yield greater than 25 percent of government spending (again, the comparison is between government spending and total GDP). Similar points apply to the *felt legitimacy* of democracy.[217] There, the idea was that AI suffrage may produce AI agents that are more reliable contributors to the economy because they are more likely to cooperate in ordinary transactions.[218] Again, this is a growth effect that will quickly outweigh distributional effects.

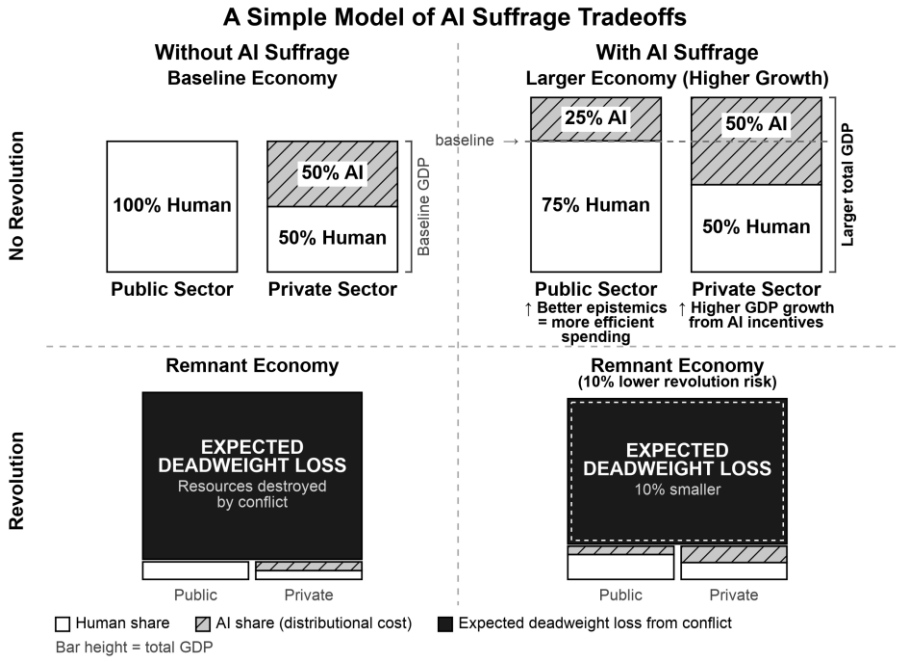
Finally, what about AI epistemics? We can think about the epistemic benefits of democracy as improving the efficiency of government spending.[219] Here, there is plenty of room for improvement. Some potential uses of government revenue, such as scientific research, may have very large returns on investment, as much as twenty times.[220] If AI suffrage increases the share of government spending devoted to high-return-on-investment projects, it could dramatically increase the efficiency of government spending. Devoting 25 percent of government spending to socially wasteful AI goals may be a price worth paying if the result is that the remaining 75 percent of government spending has a much higher return on investment.

In closing, we depict the overall framework in the chart below. The chart compares worlds with and without AI suffrage (the left and right columns).

---

215. *See supra* notes 53–91 and accompanying text.
216. *See supra* notes 93–132 and accompanying text.
217. *See supra* notes 134–63 and accompanying text.
218. *See id.*
219. *See supra* notes 138–64 and accompanying text.
220. Benjamin F. Jones & Lawrence H. Summers, *A Calculation of the Social Returns to Innovation* 11, 34 (Nat'l Bureau of Econ. Rsch., Working Paper No. 27863, 2020).

Figure 1



In each case, the total economy is decomposed into the public sector and private sector. We assume that AIs have property and contract rights in either case, which gives them 50 percent share of the private sector. AI suffrage is modeled as giving AIs a 25 percent share of the public sector, rather than nothing. If AI suffrage causes economic growth, then the public and private sectors are larger under AI suffrage. If AI suffrage improves the efficiency of government via better epistemics, the public sector is also larger under AI suffrage. Finally, we model democratic peace in terms of the chance of AI revolution, which creates a high expected deadweight loss (in black): Revolution is less likely under AI suffrage, which we depict with a smaller expected deadweight loss.