Governing AI

by
Doğuhan Sündal

Vermont senator Bernie Sanders's proposal to bring major AI companies under public ownership has gained considerable attention. The proposal itself should not come as a surprise. These companies, first, rely on information produced collectively; second, their tools become better as more users adopt them; and third, unlike industries where increasing returns to scale can often be addressed through price and quantity regulation, AI firms do not have a clear marginal-cost structure and many of their tools are offered for free. As I have discussed elsewhere, these features make a public utility-approach to Big Tech a plausible route toward public control.

Guastella and Burgis, in separate articles, have supported the nationalization approach. It is worth re-emphasizing some of their key observations. Guastella presents the popular direction of AI firms as a fair and straightforward way to address the problem of companies serving private interests while benefiting from public resources and potentially posing public risks through the innovation paths they choose. Burgis, meanwhile, emphasizes the justice of publicly owned AI firms even from within a Lockean framework. Bruenig similarly supports the public-ownership argument and, most significantly, highlights the contradictory character of many arguments against state-owned enterprises: "The state-as-owner is thus sometimes a rapacious capitalist and at other times a doddering central planner." Indeed, the case against public ownership often rests less on demonstrated impossibility than on speculative and inconsistent hypotheticals.

Let us cut to the chase. Rikap and McCarthy, in their respective articles, largely agree with the motivation behind these arguments: the public should be brought into the governance of AI firms. But they also argue that public ownership alone is insufficient. Rikap writes:

"As a stand-alone initiative, it risks backfiring. Just as sovereign oil funds can dampen popular enthusiasm for a renewable energy transition, a sovereign AI fund could bolster the adoption of today's AI, the world's largest control technology. A just society is not only one in which wealth is redistributed. Control, thus decision-making power, and knowledge should be equally shared."

This is a sharp point. Public ownership schemes may fail to bring genuine public control into the picture; worse, they may prevent an alternative AI agenda from emerging at all. McCarthy similarly emphasizes a concern familiar to many people who have worked in public institutions: administrations in these institutions often appear far too willing to

serve the private interests of a small group. Representative democracy, in his view, is "beset by principal-agent problems that have been easily corrupted by both money in politics and the structural power of capital."

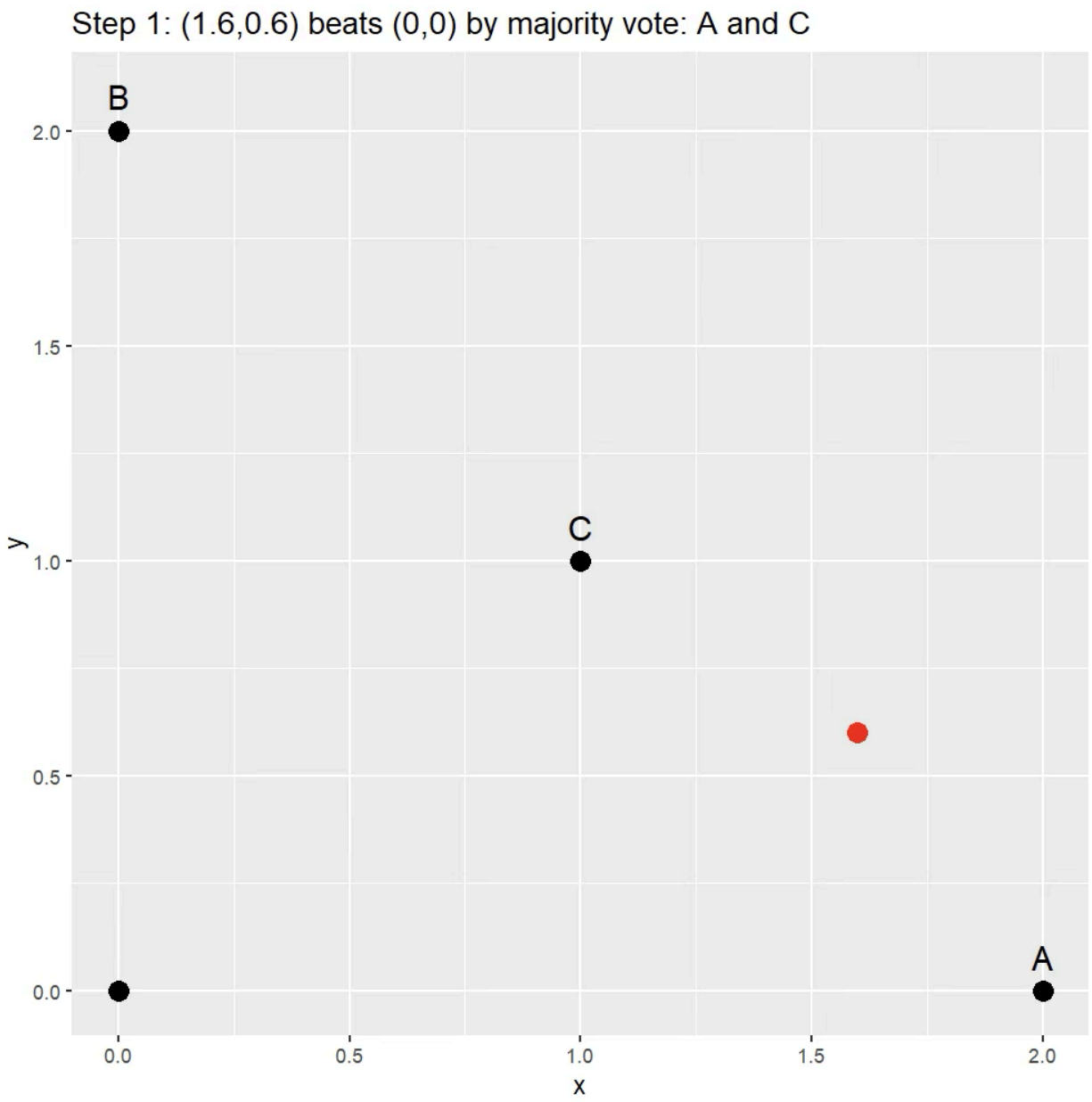
These are serious social-choice questions, even though they are often neglected in progressive debates. Amartya Sen once recalled that "Joan [Robinson] thought that my interest in welfare economics and social choice theory reflected a clear failure to grasp what was really important. Much later, while I was doing the Collective Choice book, she wrote me a letter asserting that I had told her that when I finished this book, I would come back and do some serious economics." I mention this because I see a similar tendency today: debates over public control often proceed without engaging the developments in social choice theory that bear directly on the problem. In this case, that literature is especially relevant to McCarthy's proposal.

McCarthy should be credited for offering a tangible institutional proposal. In progressive circles, one often encounters an imbalance between the amount of work devoted to critique and the amount of work devoted to concrete paths forward. His proposal has a real degree of rigor. It begins from the premise that "if an institution or organization makes decisions that deeply affect the interests of some groups in society, those groups should have a meaningful say in those decisions." Following this principle, McCarthy argues that the environmental effects of AI firms and their potential contribution to mass unemployment identify a set of affected people who should have decision-making power. This group can constitute a popular assembly, filled through random selection, or sortition, and empowered to make decisions for these firms. The goal would be not merely to distribute dividend payments, but to establish genuine control. McCarthy writes:

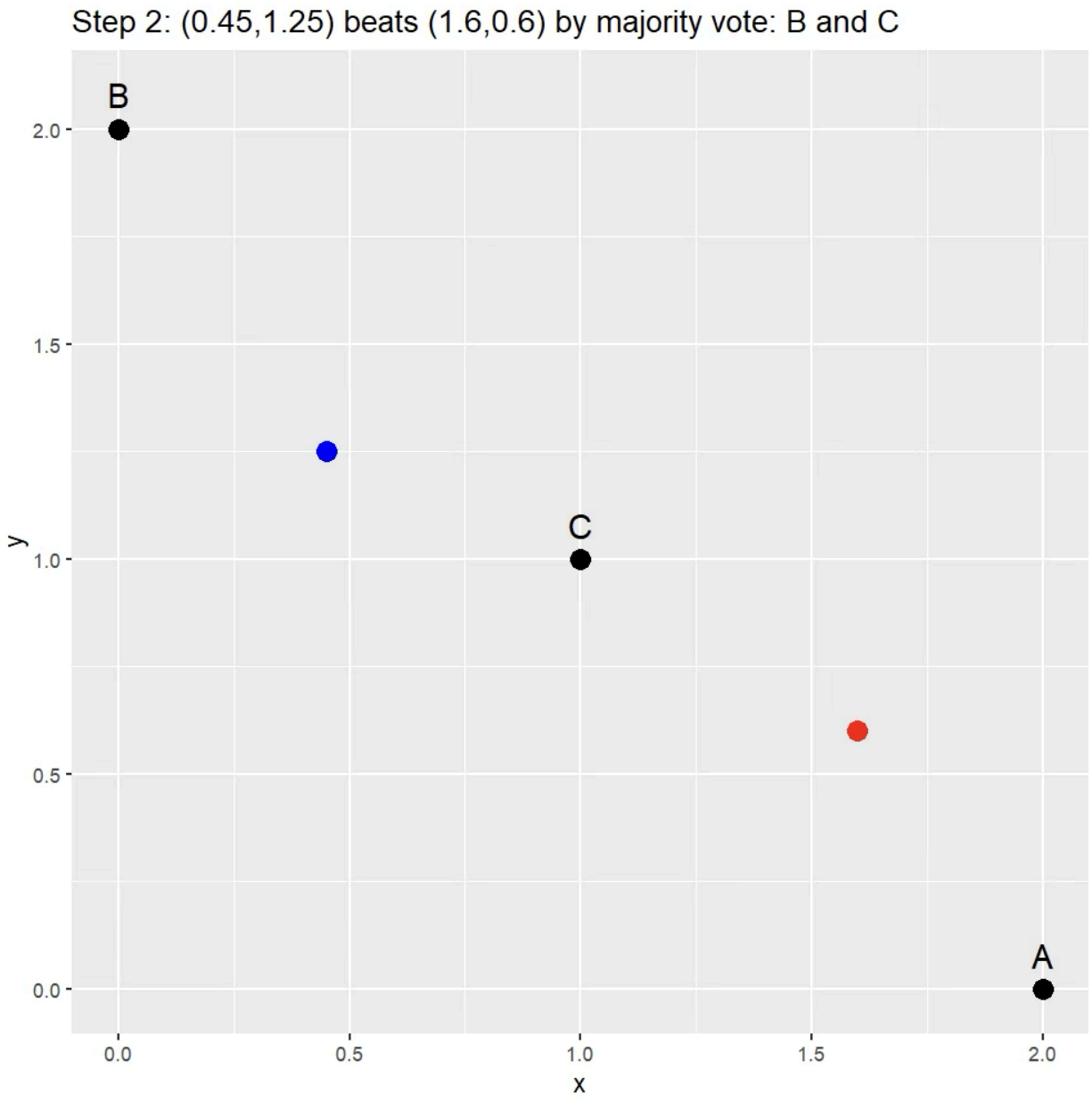
"Picture a worker at Amazon receiving a notice akin to a jury summons, informing her that she will sit on the assembly for a fixed term to deliberate over binding decisions about how these models are developed and deployed. The assembly meets on one Friday a month, and it is paid work. She is not asked to become an engineer or to master the mathematics of large language models any more than a juror is asked to fully comprehend forensics or the science of DNA evidence. Instead, those with technical expertise in artificial intelligence and its impacts act as consultants to the assembly and are marshaled by neutral facilitators in order to lay out risks, trade-offs, and options that inform the deliberations of the assembly. The participants would learn about an issue, deliberate over a set of options concerning it, and then make decisions about it."

My concern, based on the McKelvey-Schofield theorem, is the following. Imagine that the assembly is composed of three people who must decide how to allocate some amount of resources across two tasks, (x) and (y). Each individual has an ideal

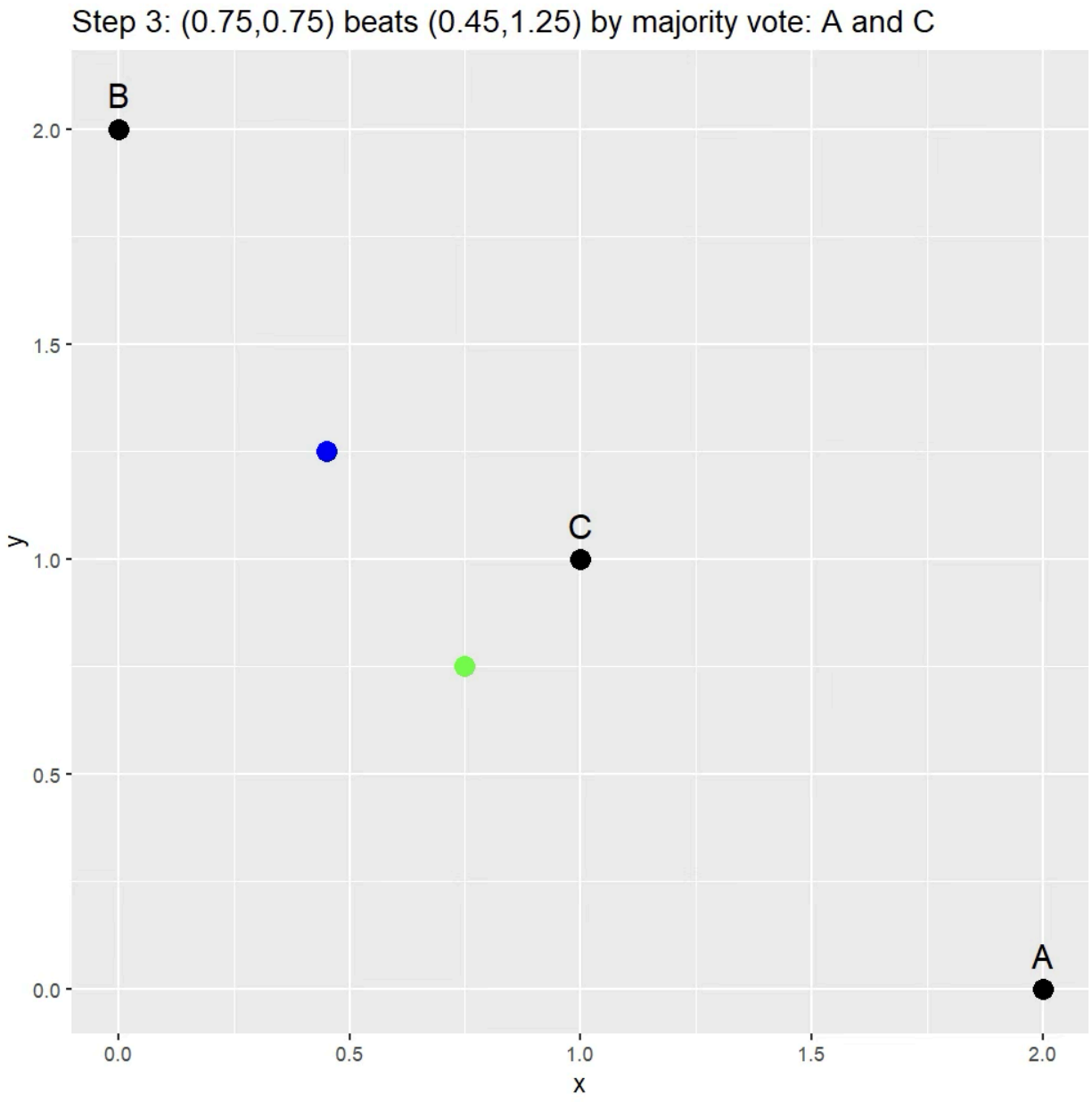
allocation, or bliss point, and evaluates alternatives by their Euclidean distance from that ideal. To illustrate the point, consider the first figure below.



Suppose the technicians first propose that, instead of spending nothing on either (x) or (y), the company should spend $1.6 billion on (x) and $600 million on (y). If one calculates the Euclidean distances, person B, who prefers spending everything on (y), finds the original point (0,0) closer than the red point (1.6,0.6). But the red point is closer to the bliss points of both A and C than the initial point (0,0) is. Thus, A and C form a majority in favor of (1.6,0.6).



Next, the technicians propose spending $450 million on (x) and $1.25 billion on (y). This combination is shown as the blue point. Compared with the red point, the blue point is closer to both B's and C's bliss points. B and C therefore vote for the blue point over the red point, and the blue point becomes the new majority choice.



Finally, the technicians propose a more balanced allocation: $750 million on each task. This is shown as the green point (0.75,0.75). Compared with the previous blue point (0.45,1.25), the green point is closer to the bliss points of both A and C. A and C therefore vote for the green point.

So what has happened? By proposing these combinations in a particular order, the technicians are able to guide the assembly toward an outcome that is worse for everyone than an available alternative. If the company instead spent $1 billion on both

(x) and (y), all three agents would prefer that allocation to the final green point. Yet through sequential majority voting and agenda control, the assembly ends up with an outcome that all three members would reject if it were placed directly against the obvious compromise.

There are two kinds of questions one might ask at this stage. The first is whether this model is realistic: does this ever happen, and how could it occur in practice? The second, which I find more interesting, is whether majority voting is a robust method for revealing public will under fairly weak assumptions. The answer is no. In multidimensional policy spaces, majority rule can fail to produce a stable collective choice. The danger is that a formally democratic assembly may still be governed by those who structure its choices. At this point, I will have to ask the reader to take my word for it: the manipulability of majority voting, and of many other voting rules, has been a major headache in social choice theory.

Finally, what appears to be a weakness of public ownership may, under some circumstances, also be a strength. It is true that, under certain conditions, the principal-agent problems of public institutions may not favor workers. But the presence of a principal-agent problem also means that there is an identifiable agent whose incentives can be studied, constrained, and redirected. With a public institution operating under a clear mandate, we may at least know what behavior to expect and how to discipline it. By contrast, a multidimensional assembly operating through majority voting may generate unstable or manipulable outcomes.

Footnotes:

1. I feel the urge to paraphrase Duncan Foley's warning here: there is no randomness, what appears random reflects our lack of information about the process generating the outcome.

References:

Sündal, Doğuhan. "Anti-monopoly and the Socialization Program." Available at SSRN 6683782 (2026). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6683782

https://jacobin.com/2026/07/ai-policy-nationalization-commons-work

https://jacobin.com/2026/07/ai-nationalization-sanders-libertarians-property

https://jacobin.com/2026/07/ai-sovereign-weath-fund-sanders-critics

https://jacobin.com/2026/07/ai-big-tech-global-ownership-control

https://jacobin.com/2026/07/ai-democratization-working-class-institutions

Klamer, Arjo. "A conversation with Amartya Sen." Journal of Economic Perspectives 3.1 (1989): 135-150.