

Introduction: Political Economy and Artificial Intelligence

By Ajay Agrawal, Joshua Gans, Avi Goldfarb, and Catherine Tucker

As the economic potential of artificial intelligence (AI) becomes clearer, a number of questions arise as to its impact on the political environment, including with respect to regulation, competition between nations, and political preferences. To provide some structure to future economics literature on these questions, we organized an NBER conference in September 2024 in Toronto with the aim, as set out in the invitation, “to set the research agenda for economics, emphasizing how AI might impact economic systems, their governance, and the relative power of individuals and nations.” We brought together a mix of economists and political scientists to identify open questions and provide frameworks for research going forward.

This volume is the third edited collection to emerge from the NBER conference series on the economics of artificial intelligence. The first volume, *The Economics of Artificial Intelligence: An Agenda* (based on the first such conference in 2017), laid out a broad research agenda spanning AI’s impacts on issues such as productivity, labor markets, scientific discovery, trade, and economic growth—several chapters have since become standard readings in many PhD curricula. The second volume, *The Economics of Artificial Intelligence: Health Care Challenges* (2022 conference), narrowed the focus to a single industry by examining economic issues associated with AI and healthcare. Building on these foundations, the present volume—born of the eighth NBER AI conference in September 2024—occupies a middle ground in scope, addressing the wide-ranging political economy of AI while remaining more focused than the inaugural collection and less narrowly sector-bound than the healthcare volume.

This volume contains a mix of invited articles and comments from that conference. It is organized into three key themes: the political economy of AI regulation, AI and military power, and AI and political identity. Consistent with NBER norms, our goal in putting this volume together was policy-relevant analysis without policy recommendations. We included papers with

an eclectic mix of approaches—empirical studies, economic theory, thought pieces, and book summaries—to illuminate different facets of AI's political economy. Given the field's early stage and the limited availability of data on key dimensions of AI and political economy, this combination enables contributions that both identify what are likely to be the most important questions and begin to explore ways of answering them. While AI's impact is already being felt in areas such as content creation and healthcare, its implications for politics and nationhood remain comparatively underexplored.

The political economy of AI regulation

Beraja and Yuchtman first discuss various ways that AI might be disruptive to existing economic structures, such as labor markets and the role of privacy in society. The chapter then points out a previously underappreciated aspect of AI's commercialization: that it will create new conflict over property rights for the inputs into AI applications. These include online content for generative AI, where current copyright policy does not provide clear guidance, and urban space for autonomous vehicles, including flying drones. The chapter, therefore, identifies important political forces that will shape the economic implications of AI, depending on how these property rights ultimately are allocated.

Farronato's chapter discusses one aspect of Beraja and Yuchtman's newly identified conflict over property rights: data. Building on the metaphor that Beraja and Yuchtman emphasized at the conference, that data is the new oil, Farronato notes that property rights over information are particularly complicated, as Arrow (1962) highlighted. Therefore, the underlying economic frameworks for developing property rights for information will need to be different from those developed for non-information goods such as oil.

In contrast to the chapters by Beraja, Yuchtman, and Farronato, which discuss the regulatory challenges in newly contested property rights, Callander's chapter examines how incentives affect relevant policy outcomes, particularly with respect to regulation. The chapter

discusses the role of interest groups and asymmetric information between policymakers and innovative firms. Politics affects the path of technology, and through interest groups, the anticipated path of technology can affect politics.

The following two chapters turn to the regulation of algorithms directly. Like Callander, Gillis et al. emphasize asymmetric information between firms and regulators. They provide a detailed framework for assessing the relative benefits of different regulatory regimes. If the regulator emphasizes inputs into the algorithm, that can limit the upside potential of AI. If the regulator emphasizes the outputs, that can punish firms for bad luck. They point out that regulation can also occur “ex-interim”, where the process can be scrutinized in ways that, under certain conditions, incentivize efficient deployment.

Gans’s chapter also highlights tradeoffs in regulating algorithms, noting that the economic approach to regulating through market design is generally more efficient than an approach that focuses on regulating the inputs to the algorithm or the algorithm itself.

Furman provides a broader discussion of the goals of AI regulation, emphasizing that there are meaningful differences between a “precautionary principle” approach and a cost-benefit approach. He identifies six principles to consider.

Together, these chapters suggest the contours of a research agenda for the political economy of AI regulation. How do alternative property-rights regimes for AI inputs—ranging from open-access licenses to exclusive copyrights—affect innovation incentives, market entry, and social welfare? Under what conditions do ex-interim process-based regulation (e.g., model-audit requirements) outperform input- or output-based regimes in fostering both safety and dynamism? What are the equilibrium choices of interest groups and regulators when asymmetric information about AI capabilities is high, and how do these choices shape the pace and direction of technological diffusion? How can we quantify the welfare trade-offs between a precautionary-principle approach versus a cost-benefit analysis in AI oversight, particularly in

high-stakes domains like finance or transportation? Finally, what empirical strategies can reveal the political feedback loops by which early regulatory decisions on data governance influence subsequent lobbying, public opinion, and legislative reform? These types of questions sketch an agenda for understanding the rich interplay between economic forces and the governance of AI.

AI and military power

Economic power and military power are interconnected. To emphasize this connection, we invited national security expert Paul Scharre to discuss his book, *Four Battlegrounds*. His chapter serves as a transcript of his remarks at the conference. He examines each of the “battlegrounds” for AI dominance in military conflict: data, hardware, talent, and institutions. Gross’s commentary on Scharre’s chapter underscores that these four battlegrounds are primarily economic in nature, suggesting it could be framed as a “resource-based view of military performance.” Consequently, the relative military power of nations relies heavily on a specific set of economic resources that countries apply to the military use of AI.

Another political scientist attending the conference, Jon Lindsay, argued that political science is probably the “More dismal science”, partly because the consequences of military conflict are severe and partly because the implications of technological change on military power are difficult to ascertain. Thus, seemingly good news is often mixed with bad. Tying military outcomes to the role of judgment in decision-making (Agrawal, Gans, and Goldfarb 2018), Lindsay emphasizes that conflicts are often decided by willingness to pay costs as much as by military firepower. AI might not have much impact on national resolve even if it changes the ability of nations to inflict harm on each other.

Together, these chapters point to a rich set of questions for the economics of AI and military power. How do marginal investments in the four battlegrounds—data acquisition, specialized hardware, human capital, and institutional frameworks—translate into shifts in national military effectiveness and bargaining power? What are the opportunity costs and

spillovers when governments divert private-sector talent toward defense-oriented AI research, and how does that reshape civilian innovation ecosystems? Can we estimate models to quantify the trade-offs between stockpiling AI-optimized hardware versus investing in adaptive talent pipelines when facing uncertain future conflicts? How does asymmetric information about AI capabilities between rival states influence arms-race dynamics and the probability of preemptive strikes? What empirical strategies can uncover the causal impact of AI-driven military expenditures on both conflict escalation and deterrence? Questions such as these frame a research agenda for the economics AI and military power.

AI and political identity

A third theme was the role of AI in shaping political identity. Fouka and Reich emphasize recent research on the role of education in shaping national identities. Nation building is a deliberate process, and it has long been recognized that education plays an important role in that process. Furthermore, there are reasons to expect that AI will lead to more effective and personalized education. Combining these ideas, they argue that AI could lead to more effective nation-building. They, therefore, connect disparate ideas on nation-building, education, and AI to suggest a novel pathway through which AI might affect political identity.

Petrova et al. discuss a different way that AI can shape identity: through its impact on politicians and voters. They document that the automation of factories likely increased the activities of right-wing populist politicians in affected areas and, in related work, note that this in turn likely influenced election outcomes in the United States. Colantone offers some suggestions to enhance this research. Mansell et al. highlight that support for AI is not necessarily stronger on the left. Using data from Canada, they show that both the far left and the far right are more skeptical of AI, while traditional conservatives appear to be the most supportive.

Acemoglu et al. develop separate models of how AI's use in social media and targeted political advertising affect political preferences and political outcomes. For social media, curated news feeds lead to polarization. For advertising, two political parties aim to reach voters with digital ads, which can be crafted and targeted by AI. Naive voters may be convinced by incorrect information, and this increases the effectiveness of digital ads for political purposes. Thus, their models suggest AI-powered media lead to polarization of voters, and that political parties respond to this by choosing polarizing policies. They note that competition between platforms will not solve these issues, and instead explain how regulation of the algorithms (enforcing social media feed diversification or weaker ad targeting) could reduce polarization. In addition to touching the theme of regulation, this concluding chapter provides specific mechanisms for an idea that is perhaps the central theme of this volume on AI and political economy, that politics affects technology and technology affects politics.

Political scientist Joshua Tucker commented on the presentation by Acemoglu and Siderius at the conference, which is related but distinct from the chapter those authors contributed to this volume. Tucker's comments highlight ideas similar to those of Beraja and Yuchtman, Farronato, and especially Callander, noting that regulation occurs in a contested environment. However, it remains unclear who would advocate for the policies in AI and online advertising that Acemoglu and Siderius's research emphasizes.

These contributions invite a host of empirical and theoretical questions at the intersection of AI and political identity. How does personalized, AI-driven education alter measures of national cohesion and civic engagement, and what are the long-run welfare implications of such shifts? Can we estimate the causal effect of AI-mediated job displacement on the rise of populist identities, perhaps using regional variation in automation adoption as an instrument? What structural models can capture the feedback between AI-curated news feeds and the evolution of individual ideological sorting, and how might interventions—such as enforced content diversity—mitigate polarization without stifling user engagement? How do AI-

powered targeted advertisements reshape voters' identity salience and turnout, and what are the comparative magnitudes of these effects across different electoral systems? Finally, what econometric strategies can uncover the role of social networks in amplifying AI-driven identity dynamics, and how might regulatory or platform-design changes alter these network effects? These types of questions set a research agenda for understanding how AI is reshaping the very fabric of political identity.

Conclusion

This volume introduces an exciting new frontier at the intersection of artificial intelligence and political economy, laying out a research agenda for empirical and theoretical investigation. Two overarching challenges stand out.

First, bringing ideas from other disciplines into economics. AI's rapid and broad diffusion requires economists to engage with insights from military studies, political science, and data science—applying rigorous economic models to frameworks presented in non-traditional formats. For example, Scharre's resource-based view of AI in defense and Lindsay's political-science lens on conflict underscore the payoff from importing novel hypotheses into general equilibrium and game-theoretic settings. By recasting these narratives in the language of incentives, information asymmetries, and strategic complementarities, economists can both sharpen the underlying theory and generate testable empirical predictions.

Second, novel data generation and measurement. Understanding AI's political-economic impact hinges on new data such as lobbying expenditures, algorithmic audits, and platform governance. High-frequency measures of data sharing between firms and regulators, granular geo-coded records of automation adoption, and micro-level assessments of targeted political ads are examples of data that will be instrumental for identifying causal channels. These datasets will allow researchers to quantify welfare trade-offs—such as safety versus

innovation—and to estimate structural parameters that govern interest-group dynamics, regulatory design, and the polarization of political identity.

Taken together, the contributions in this volume provide a conceptual scaffolding and methodological roadmap for researchers: from model construction and counterfactual simulation to field experiments and natural-experiment–based identification. We hope that this collection inspires the next generation of scholars to chart the dynamic interplay between AI, institutions, and social welfare to enhance our understanding of whether and how this rapidly advancing technology will impact the political economy.

SHORT SUMMARY/ABSTRACT

As the economic potential of artificial intelligence (AI) becomes clearer, a number of questions arise as to its impact on the political environment, including with respect to regulation, competition between nations, and political preferences. To provide some structure to future economics literature on these questions, we organized an NBER conference in September 2024 in Toronto, where the goal set out in the invitation was “to set the research agenda for economics, emphasizing how AI might impact economic systems, their governance, and the relative power of individuals and nations.” We brought together a mix of economists and political scientists to identify open questions and provide frameworks for research going forward.

This volume contains a mix of invited articles and comments from that conference. It is organized into three key themes: the political economy of AI regulation (with entries by Beraja and Yuchtman, Farronato, Callandar, Gillis et al, Gans, and Furman), AI and military power (with entries by Scharre, Gross, and Lindsay), and AI and political identity (with entries by Fouka and Reich, Petrova et al., Colantone, Mansell et al., Acemoglu et al., and Tucker).