Balancing the Tradeoff between Regulation and Innovation for Artificial Intelligence: An Analysis of Top-down Command and Control and Bottom-up Self-Regulatory Approaches

Keith Jin Deng Chan[1], Gleb Papyshev[2], Masaru Yarime[1,3]
[1]Division of Environment and Sustainability, The Hong Kong University of Science and Technology, Hong Kong SAR, China
[2]Division of Social Science, The Hong Kong University of Science and Technology, Hong Kong SAR, China
[3]Division of Public Policy, The Hong Kong University of Science and Technology, Hong Kong SAR, China

Abstract

In response to the rapid development of AI, several governments have established a variety of regulatory interventions for this technology. While some countries prioritize consumer protection through stringent regulation, others promote innovation by adopting a more hands-off approach. However, this tradeoff has not been analyzed systematically. We developed an economic theory on how the welfare-maximizing level of regulatory stringency for AI depends on various institutional parameters. Our game-theoretic model is motivated and built upon the comparison of regulatory documents for AI from the EU, the UK, the US, Russia, and China. The results show that if a government strives to find the right balance between innovation and consumer protection to maximize actual consumer welfare, stringent regulation is optimal when foreign competition is either high or low, whereas light-touch regulation is optimal when foreign competition is intermediate. Meanwhile, minimal regulation is rationalizable only if a government prioritizes other objectives in its agenda, such as maximizing innovation, domestic producer surplus, or perceived consumer welfare.

1 Introduction

The spectrum of regulatory proposals spans from hard laws and the prohibition of certain systems to industry self-regulation based on AI ethics. Both the top-down stringent approach and the bottom-up ethics-based approach have been criticized in the academic literature. While strict regulatory requirements may better protect society against the risks of AI technologies, they also tend to hinder the pace of innovation. It is unclear to policymakers and researchers which approach maximizes consumer welfare, and under what conditions. The conceptual difficulty in addressing this dichotomy partly stems from the lack of a common framework which incorporates both sides of the argument. In response to this gap in the literature, this paper has developed a model to address the following interrelated questions:

(1) What are the advantages and disadvantages of the two regulatory approaches?

(2) What institutional factors influence the outcomes of the two approaches?

(3) How should governments optimally balance the tradeoff between AI innovation and consumer protection in general?

In this paper, we look at the global landscape of AI regulation to highlight two major regulatory approaches: the top-down command and control approach and the bottom-up self-regulatory approach. By top-down command and control approach we understand regulatory approaches where the governments control access to the market for certain types of AI systems through legally-binding regulations. On the other hand, bottom-up self-regulatory approach is based on the idea of industry self-regulation through non-enforceable ethical principles for AI. We use the case studies of the EU, the US (if the Algorithmic Accountability Act is adopted), and China to illustrate proposals in the first group, and cases from Russia and the UK for the second group.

The regulatory stringency chosen by the government is modelled by the probability that the exploitative practices of local AI companies are revealed to the consumers. This modelling choice is motivated by a unique challenge facing AI regulators. One key aspect of regulating AI is the difficulty to interpret the workings of the black box systems, particularly what kinds of data are collected and what types of algorithms are used to extract valuable information by the companies. This fundamentally differs from industrial sectors where their social cost of production such as environmental pollution is relatively easily monitored and detected. In that sense, it is important that our model incorporates the possibility of revealing information to consumers, which will affect the behavior of consumers and their welfare in the end. After all, for unethical but lawful exploitation, it is consumers' knowledge of such practices rather than top-down prohibition that acts as a disciplining device. This modelling choice is therefore especially relevant to

consumer-facing AI systems. For concrete examples of policy mechanisms that strengthens disclosure and scrutiny of AI systems:

(1) In the EU, firms must first ensure the transparency and interpretability of their AI systems and file relevant technical documentation to the government before entering the market;

(2) In China, providers of recommendation algorithms must file self-assessment reports to the government and inform the users about their services and how these algorithmic recommendations are designed;

(3) In the US, the Algorithmic Accountability Act proposes a repository of publicly available information about the usage of automated decision-making systems.

The key contribution of our model is to link the optimal regulatory stringency to the degree of foreign competition (relative to domestic innovation capacity) faced by the local AI industry. Foreign competitive pressure is naturally a driving force behind firms' innovation and exploitation decisions; but the presence of competitive foreign substitutes also determines the extent of preferential treatment a government should show towards the local AI industry. Note that the degree of foreign competition is by no means solely driven by technological differences; economic policies play a key role as well. For instance, a government can influence the foreign competitive pressure facing the local AI firms by restricting or opening access to the domestic market for foreign companies, adjusting the subsidies for attracting foreign businesses, and enacting protectionist policies in the form of tariffs.

2. Results

The analysis highlights the subtle trade-off between AI innovation and consumer protection. The analysis shows that there are two distinctively different approaches to regulating AI: a top-down command and control approach adopted in the EU and China (and possibly in the US in the future) and a bottom-up self-regulatory approach adopted in Russia and the UK. However, there are debates regarding which of these approaches is more effective in facilitating innovation while protecting consumers. Previous studies had difficulties addressing this issue, as they did not have a systematic framework that addresses this trade-off under different institutional environments. Therefore, we constructed a game-theoretic model to clarify the conditions that affect the trade-off, which incorporates the incentive structures facing various key stakeholders and their resulting strategic interaction.

Based on our game-theoretic analysis, we have developed an economic theory of how the welfare-maximizing level of regulatory stringency for AI depends on various institutional parameters. Under high foreign competition, domestic innovation plays a relatively small role in serving consumers. On the other hand, consumers benefit most when they are not misled to underuse the highly competitive foreign AI systems. As a result, the prioritization of consumer protection should motivate a government to choose a high level of regulatory stringency under high foreign competition.

Meanwhile, under low foreign competition (for instance, due to strong protectionist policies), the domestic AI industry can effortlessly win over local consumers from their foreign competitors. This means domestic firms can derive high marginal benefits in terms of market share from improving their algorithms. As a result, the robustness of domestic firms' innovation incentives should motivate a government to also choose a high level of regulatory stringency under low foreign competition.

Interestingly, under intermediate foreign competition, the government faces a delicate trade-off between consumer protection and innovation. Too stringent regulation stifles the innovation incentive of the domestic AI industry, whereas minimal regulation subjects the consumers to excessive exploitation. To maximize the actual consumer welfare, the government may strategically lower its regulatory stringency and turn a blind eye on some occasions.

Across all institutional environments, however, minimal regulations are never compatible with maximizing actual consumer welfare. As such, the objectives of such regulatory design may be either rationalized by the prioritization of innovation, domestic producer surplus, or the perceived welfare of the consumers. In the latter, the government is primarily concerned with the image that this regulatory intervention produces without worrying too much about the actual protection of consumers – essentially using a loosely designed regulation as a PR tool. This suggests that further empirical studies should pay close attention to cases where governments are proposing very loosely defined.