Modeling the Market Impact of Artificial Intelligence

Wun Yung Shaney Sze
Barcelona School of Economics

December 13, 2022 – Draft I

Abstract

Artificial intelligence (AI) has been a hot topic among businesses and the global economy given its breakthrough results and the significant cuts in labor costs it has provided. Given its novelty in our society, academics are still looking for methods to properly model the economics of the technology. By summarizing academic and industry findings, I will discuss how AI has been able to reduce costs for firms and illustrate with an economic model. In the model, I analyze the economic forces behind AI and outline externalities it produces within the system. I will examine how the cost of prediction, automation, and risk analyses gets driven down by AI, and how such cuts in cost generates a larger impact on the market and beyond.

1 Introduction

First and foremost, we need to understand the importance of modeling the economic impact of AI. Over the years, there has been a growing interest in the economic reach of AI, but given its stage of diffusion, many discussions are necessarily speculative and theoretical [3]. As we move forward, we will need more public data sets on AI adoptions in order to apply practical economic theories on. An economic model is crucial for carrying out policy making decisions, both within the business as well as the larger macroeconomic context. Modeling AI in the industry context will be key to future regulatory and policy making decisions.

In my model, I will create the theoretical framework as to how AI will shift the cost in the following categories: (i) prediction, (ii) replication, and (iii) unit labor. As cost decreases, it manifests countervailing effects such as an increase in productivity and capital accumulation, creating a rich framework modeling AI.

The paper proceeds as follows. Section 2 provides a nontechnical background on the capabilities and applications of AI in our current society, and outlines some of the prominent literature surrounding the topic. In Section 3, I define the economic scope under consideration and introduce my economic model on the market impact of AI alongside its limitations, within this section, I also draw comparison from my model to several economic theories such as the constant elasticity of substitution and the Baumol effect. In Section 4, I will offer some insights for further applications of the model. Finally, Section 5 concludes.

2 Background

In December 2022, the global consulting agency McKinsey published a report on the state of AI [8]. They have found that AI adoption by firms has more than doubled in the past 5 years, as well as the average AI capabilities of firms. Meaning, more and more firms are leverage techniques such as computer vision and natural-language generation. Research and development investments towards

AI has increased and firms are reaching their arms towards more and more areas and functionalities to automate.

In the past few years, academics have theorized that we can model AI as a drop in the cost of prediction, defining AI has a substitute for human prediction and a complement to other skills such as human judgement [3]. With the ability to define the utility function for AI, we will be able to better understand externalities generated by firms using the technology and the technology itself. For every technological breakthrough–think electric motors, semiconductors, the internet–economists race to redefine the economy surrounding the new change. Here, by appending the technology of AI to the list above, I plan to continue the conversation in a more formal setting by introducing an economic model to explain variables driving the technology of AI and hope to generate productive discussions within the field both for academics and decision makers.

3 Model

3.1 Setup

For our interests, the players in consideration are the firms themselves. Since we are modeling the effects of cost reduction of AI, the primary endogenous variables are the components that contribute to such reductions: (i) prediction power, (ii) replication efficiency, and (iii) unit labor costs.

Exogenous variables include technological change, governmental restrictions, geopolitical forces, funding for artificial intelligence development, strategic cooperation, and creation and destruction of firms and jobs. When exogenous variables like governmental restrictions and geopolitical forces change, it places limits to the speed and feasibility of endogenous variables. For example, maybe a government feels that a company will have too much power so they set restrictions to the usage of quantum computing, which causes the company to lose opportunity cost on that task. Another example, if funding for AI development decreases, then AI capital will decrease, which leads to more more manual tasks, leads to lower productivity.

3.2 Framework

In this model, I will reference Brynjolfsson's model on computing productivity on a firm-level [7] and Acemoglu & Restrepo's task-based framework on AI and displacement [1].

Under a strategic setting with only two firms in a closed economy, we have a homogeneous and fully employed workforce. There are two types of firms, firms with (i) high level of AI capability (H) and firms with (ii) low level of AI capability (L). Within the model, we treat whether a firm have a high or low level of AI capability as a given.

The objective function of one firm will measure the profits of the firm, with the linear combination of (i) prediction power (P), (ii) replication efficiency (R), and (iii) unit labor costs (C). Prediction power and replication efficiency are negatively correlated with unit labor costs. For firms with higher prediction power, they will experience relatively lower unit labor costs, which leads to higher profits, vice versa. For trade-offs in the model, I introduce the cost of employing higher-skilled workers. Some companies may also not be in the appropriate sector, or their scale is too small to provide enough data for prediction purposes, so they do not act on the implementation.

Following the standard growth framework used for the study of productivity, I assume that the production process can be represented by a production function (F) that relates firm value added (Q) to the previously introduced variables. In addition, we assume that the production function is affected by time (t), and by the industry (j) in which a firm (i) operates. Thus, for each of the high and low level of AI capability firms, we can compute the value added by AI as

$$Q_{it} = F(P_{it}, R_{it}, C_{it}, i, j, t). \tag{1}$$

Assuming that this relationship can be approximated by a Cobb-Douglas production function.

We implement this function with the three inputs–prediction power, replication efficiency, and unit labor costs:

$$Q = A(i, j, t)P^{\beta_p} R^{\beta_r} C^{\beta_c}, \tag{2}$$

or

$$q = a(i, j, t) + \beta_k k + \beta_l l + \beta_c c. \tag{3}$$

Then, we can consider the production function to be

$$Y = AQ_1^{\alpha_1} Q_2^{\alpha_2} \cdot \ldots \cdot Q_n^{\alpha_n} \tag{4}$$

where

$$\sum \alpha_i = 1$$

with α denoting capital share. In order to distinguish firms with low level of AI capability and firms with high level of AI capability, we set

$$Q_{it} = \begin{cases} L_{it} & \text{if firm's AI capability is low} \\ H_{it} & \text{if firm's AI capability is high} \end{cases} \tag{5}$$

If we have the optimal number of low and high AI capability firms, the production function can be expressed as

$$Y_t = A_t H_t^{\alpha} L_t^{1-\alpha}. \tag{6}$$

3.2.1 Elasticity of Substitution

There also exists the elasticity of substitution between labor and *AI capital*, where AI capital is distinguished by the fact that it is generated by AI, rather than traditional human labor [9]. From existing literature, there are at least two school of thoughts in terms of incorporating this elasticity of substitution into the production function. Agrawal, McHale and Oettl (2017) incorporated AI into an innovation-based growth model to show how AI can speed up growth along a transition path [4]. When both streams of capital *substitute* each other, a competition between them arises. On the other hand, when both streams of capital *compliment* each other, AI can help humans complete more tasks in less time, saving labor and raising wages. Here, an elasticity of substitution shows how some tasks are better off without the technology of AI, in which case they are substitutes ($\alpha > 1$), and complimentary otherwise ($\alpha < 1$).

3.2.2 Diminishing Marginal Utility

Following the law of diminishing marginal utility, if AI is not being used effectively or if it is not being integrated into the firm's operations in a way that maximizes its potential benefits. However, in other cases, a firm may continue to see increasing returns from using AI as it becomes more advanced and is able to perform more complex tasks. Ultimately, the impact of AI on a firm's marginal utility would depend on a variety of factors.

Marginal utility is defined by the *change in total utility* over the change in *number of units consumed.* Here, I define the former term to be utility generated from implementing AI, and latter term to be the scale of AI functionalities within a firm. I speculate that as a firm decide to implement one AI functionality after the other, whether it be a form of natural language processing or computer vision, it will not grow as rapidly as the increase in total utility. This is because of two things: (1) the fruit of the labor needs *time* to come into fruition, and (2) the reach of AI functionality is limited, there are only so many areas that a firm can implement it on. The novelty will wear off over time and over each product line. In terms of diminishing marginal utility, an important note to keep in mind of is the coinciding nature between a firms' tenure within the industry and its growing number of AI functionality, which can cause potential spurious associations in the practical setting.

3.2.3 Externalities

Considering the broader economic environment, AI can generated economic effects via multiple channels.

First, it will increase economic gains from the increased global flows. Second, it will generate wealth. Third, it will encourage reinvestment. As AI increases the rates on data generation across digital platforms, firms are able to leverage such data to their advantage and generate more profits. AI can also improve supply chain efficiency and automate domestic and international compliance processes, producing a more efficient digital global economy. As economies become more efficient, the increased output from firms can be passed to workers in form of wages and stock options, leading to higher wealth, boosting domestic economy. Similarly, firms can also decide to reinvest their profit into more research and development work towards AI and other operations.

On the other hand, AI can also induce negative distributional externalities among firms and the wider economy. First, firms can suffer implementation costs. Second, it can remove manual laborers from the global workforce. Third, firms will have to provide severance and other compensations to such workers. As firms transition across systems of different technology levels, they will incur costs such as consulting fees and integration fees. In terms of human resources, they will also incur a cost as they train employees into the new technology, which would take human capital away from generating productive work. The economy can also experience the Baumol effect, where a decline in labor share is caused by the implementation of AI, which I will discuss further in the following section. As some workers can displaced due to AI, unemployment benefits and other provisions will have to come out of the firm's pocket, creating a negative externality.

3.3 Applications

3.3.1 Baumol Effect and Engels' Pause

One economic theory of relevance is the Baumol's effect or Baumol's cost disease. Upon observing a string quartet, Baumol (1967) noticed that while the productivity of a quartet has not increased over time, their salary has increased relatively dramatically. In the industry context, Baumol found that sectors with rapid productivity growth often see their share of GDP decrease, whereas sectors with relatively slow productivity growth see an increase in their share of GDP. An illustration of the Baumol effect can be shown in the model above. As a firm leverages more AI technology, they produce more AI capital accumulation, which improves productivity and leads to a decline in the share of GDP. This aligns to my hypothesis of the diminishing marginal utility of AI. An excessive utilization of AI triggers the Baumol effect and causes a firm's share in GDP to drop. It also causes firms to lose touch with the current pace of technology. To solve this issue, firms need to keep the productivity growth generated by AI at a steady pace such that the marginal utility is optimized.

From a macroeconomic point of view, we can relate the supply side's Baumol effect to demand side's Engels' Pause, which illustrates wage stagnation during during technological upheavals [5]. An increase in a firm's market share in the industry will decrease the labor share of income. As AI becomes more and more saturated, less salaries will need to be paid for laborers in relative to the firm's profit share.

3.3.2 Capital Accumulation

For our purposes, we will define the *capital* in capital accumulation as the added value generated from AI, as opposed to traditional capital generated by manual labor or other forms of labor but AI. This is denoted as α in our model. Note that α is an exponential variable for firms with high level of AI capability, and $1 - \alpha$ for firms with low levels of AI capability. In other words, capital will accumulate much faster for firms and high levels of AI capability and much slower for firms with low levels of AI capability. The demand for capital and labor is cyclical–high demand for capital leads to accumulation, which leads to demand for labor.

3.3.3 Productivity Effect

As costs of production decreases due to the implementation of AI, it raises the demand for non-automated tasks [2]. That is, the work of AI scientists and machine learning engineers. Here, we have a substitution of intelligence capital, as engineers automate part of their work have more time for innovative tasks, they will be able to generate more ideas for technologies that increase productivity. An example would be Google's 20% side project time, where the firm allows employees and leverage their down time to work on creative projects. Mathematically, we can denote the productivity affect as

$$\frac{\Delta \ln(\text{Profits/Labor})}{\Delta \text{AI Capability}} \tag{7}$$

The change in AI capability is positively correlated with the logarithmic change in labor, denoting the complimentary nature. The change in AI capability is negatively correlated with the in the logarithmic change in profits, supporting the theory that there is a counterproductive excess rate of AI capability that a firm may experience.

3.4 Limitations

It is important to note the model is limited by the novelty of the technology, of which the impact of may build up at an accelerating pace over time. Our firm-level model is limited to a certain subgroup, and I knowledge the approach has its potential limitations. First, the data is difficult to gather and generalize among firms. For example, on the offset it will be a challenge to separate *low* and *high* level AI firms, affecting the quality of the model. Second, the survey results may move towards early movers, which may lead us to overestimate the economic impact since such technologies may not be necessarily widely leveraged yet. In the model, we essentially generalizes the entirety of the subject of AI into three variables, which is most likely not enough to describe the complete consequences of AI.

4 Discussion

The technology of AI can bring about ethical questions and new unique risks to the industry's integrity and safety, of which the full extent is yet to be assessed. So far, I have outlined an economic model on the impact of AI on market dynamics and discussed its theoretical effects. Here, I will discuss how such effects can be translated to the larger global scale. Policy plays an important role to enable society to reap optimal benefits while technological change minimizes the disruptive effects. Company that leverage AI technologies can run into ethical and privacy concerns such as data scraping on individuals. AI can both improve human decision-making and exploit human weaknesses. It is up to how policymakers decide the line for firms are.

The impact of AI will depend on how the increased income will be distributed. When the internet was first introduced, it had led to an increase in inequality due to skill-bias [6]. Some important questions for economists to consider is whether the increase from productivity using AI is actually worth the cost of implementing it and the displacement costs in the short term. What is the cost disease? What is the optimal point for firms to implement the automation on? In countries with already-high unemployment rate, what are the consequences of generating more work and capital share towards a specific industry like tech?

5 Conclusion

In this paper, I summarize the conceptual framework for the implications for firms to use AI. I restricted the model to firms with high and low AI capability levels and emphasized how each firm can differ in their production functions, while generalizing AI into the linear combination of (i) prediction power, (ii) replication efficiency (P), and (iii) unit labor costs (C). Economic theories such as the diminishing marginal utility and Baumol Effects were considered, where I argued how AI can be excessive for some firms and that firms will optimize their AI levels in the long run. Finally, I discussed the limitations of the generalized model and discuss the role of firms and policymakers in the future role of AI. Nevertheless, AI implementation is a long process and economic models will undeniably differ and be modified over time. We cannot take the technology and the freedom to build on such technology for granted. We should encourage productive discussion between academics, industry leaders, and policy makers to optimize the shared gains from the impact of AI.

References

[1] Acemoglu, D., and Restrepo, P. Artificial Intelligence, Automation and Work. NBER Working Papers 24196, National Bureau of Economic Research, Inc, January 2018.

[2] Acemoglu, D., and Restrepo, P. Artificial intelligence, automation and work. NBER Working Papers 24196, National Bureau of Economic Research, Inc, 2018.

[3] Agrawal, A., Gans, J., and Goldfarb, A. *Introduction to Economics of Artificial Intelligence*. University of Chicago Press, September 2022.

[4] Agrawal, A., McHale, J., and Oettl, A. Finding needles in haystacks: Artificial intelligence and recombinant growth. Working Paper 24541, National Bureau of Economic Research, April 2018.

[5] Allen, R. C. Engels' pause: Technical change, capital accumulation, and inequality in the british industrial revolution. *Explorations in Economic History 46*, 4 (October 2009), 418–435.

[6] Autor, D. H., Katz, L. F., and Krueger, A. B. Computing Inequality: Have Computers Changed the Labor Market?*. *The Quarterly Journal of Economics 113*, 4 (11 1998), 1169–1213.

[7] Brynjolfsson, E., Ma, C., Hitt, L., Bartelsman, W., Berndt, E., Bresnahan, T., Griliches, Z., Kahn, J., Mairesse, J., and Malone, T. Computing productivity: Firm-level evidence. Tech. rep., October 2000.

[8] Chui, M. The state of AI in 2022—and a half decade in review. Tech. rep., April 2022.

[9] Huang, X., Hu, Y., and Dong, Z. The macroeconomic consequences of artificial intelligence: A theoretical framework. Economics Discussion Papers 2019-48, Kiel Institute for the World Economy, September 2019.