World Economy Brief

February 6, 2020 | Vol. 10 No. 4 | ISSN 2233-9140

Data- and AI-driven Economic Growth in a General Equilibrium Model

Kyu yub Lee Research Fellow, Trade and Investment Policy Team, Department of International Trade (kylee@kiep.go.kr)
Hyun Park Professor, Department of Economics, Kyunghee University (econhyun@khu.ac.kr)

I. Introduction

Data is the new oil (The Economist 2017). Data is the oxygen (Bowen 2017). Like these metaphors, we can easily find many others emphasizing the importance of data (Atkinson 2019). Such metaphors could highlight some aspects of data, but they may also mislead readers from understanding the nature and characteristics of data (see Table 1).

Table 1. Characteristics of data, oil, oxygen

	Data	Oil	Oxygen
Input factor	O	O	X
Non-rivalry	O	X	O
Non-excludability	△	X	O
Information/idea	O	X	X
Privacy violation	O	X	X

Note: 1) O: yes, 2) X: no, 3) △: partially true
Source: Authors

As discussed in the literature, data is non-rival (Jones and Tonetti 2019), data usage can affect consumers' welfare through breaching and identity theft and the like (Acquisti 2014). To build a growth model with data and artificial intelligence (AI), it is necessary to refer to the canonical growth model (Romer 1990), recently developed growth models with data and/or AI (Aghion et al. 2018, Gordon 2016, Brynjolfsson et al 2018, Jones and Tonetti 2019), and a model with robots and labor (Acemoglu and Restrepo 2016) among many others.

We attempt to characterize a data- and AI-driven economy and establish a general equilibrium growth model in order to describe the data economy and examine how data and AI can affect the economy in the long run.

The constructed growth model herein is closely related to Jones and Tonetti (2019). Unlike them, however, our model is distinct in four ways. First, privacy violation is captured by disutility in preferences and shows how privacy plays a role as negative externality in the long-run growth. Second, we add data dynamics that include creation and destruction. Third, we categorize data into different types and add the process how economic agents generate production and technology data using labor and raw data from consumers. Finally, we consider substitution/complementarity between labor and capital in output technology.

II. The model

The economy has two production sectors, final output and intermediate goods, and one R&D sector for AI technology. We consider three types of data: raw data, production data, technology data. It is important to note that we assume raw data is non-rival and both production and technology data are generated by using raw data and labor. There is no population growth and time is continuous (see Figure 1).

1. Household

The household maximizes intertemporal utility over time t, $t \in [0, \infty)$:

$$U = \int_0^{\infty} e^{-\rho t} u(c_t, z_t) dt$$

where ρ is time discounting factor, c_t is consumption level and z_t is amount of raw data at time t. Flow utility is

$$u(c_t, z_t) = \frac{1}{1-\theta}\left(\frac{c_t}{z_t^{\omega}}\right)^{1-\theta}$$

where θ is intertemporal substitution and ω is degree of disutility from privacy violation.

The stock of raw data is defined as

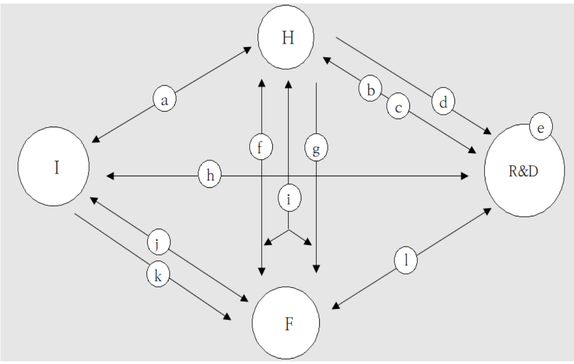
$$z_t = \varphi \int_{-\infty}^{t} \bar{c}_t e^{-\delta(s-t)} ds$$

where φ is intensity of raw data creation, $\delta \geq 0$ is rate of raw data destruction and $\bar{c}_t$ is average consumption at time t. In constructing raw data equation, we refer to data characteristics pointed out by Tucker (2018). The evolution of raw data shows

$$\dot{z}_t = \varphi \bar{c}_t - \delta z_t$$

where a dot denotes time derivative.

Figure 1. Structure of the model



Note: 1) H, I, F represent Household, Intermediate good producer, Final output producer, respectively.
2) ⓐ: H provides physical assets and receives returns.
ⓑ: H provides physical assets and receives returns.
ⓒ: H supplies labor to produce technology data and receives wages.
ⓓ: Raw data is used in the R&D sector.
ⓔ: AI technology, including machine learning and deep learning, also referred to as Invention of Method of Inventing (IMI)
ⓕ: F sells final goods or services and receives payment.
ⓖ: Raw data is used in final output sector.
ⓗ: Blueprint, recipe, idea
ⓘ: H supplies labor for production and for generating production data and receives wages.
ⓙ: I supplies intermediate goods to F and receives payment. Goods provided to F can be interpreted as robots or automation.
ⓚ: Variety, expansion
ⓛ: Learning-by-doing
Source: Authors.

The household's budget constraint is

$$\dot{a}_t = r_t a_t + w_t^y l_t^y + w_t^d l_t^d + w_t^A l_t^A - c_t$$

where a is asset and r is interest rate. Labor supplies are l^y, l^d, l^A and their corresponding wage rates are w^y, w^d, w^A for production, production data generation, and technology data generation, respectively.

Euler equation for consumption is derived as

$$\frac{\dot{c}_t}{c_t} = \frac{r_t - \rho}{\theta} + \omega\left(1 - \frac{1}{\theta}\right)\left[\varphi\frac{\bar{c}_t}{z_t} - \delta\right].$$

In deriving this equation, we treat that consumers take z_t as given (Varian 2018). When we ignore privacy ($\omega = 0$), consumption dynamic returns to the standard form.

2. Final output producer

Final output technology is

$$y_t = (l_t^y)^\alpha (d_t^y)^\beta \left(\int_0^{A_t} (x_t^i)^\nu di\right)^{\gamma/\nu}$$

where α is labor share, β is share of production data, and $1-\gamma=\alpha+\beta$. ν is elasticity of substitution among intermediate goods, x^i is intermediate good $i \in [0, A^t]$ and A^t represents variety of intermediate goods.

Production data is generated by employing labor and raw data,

$$d_t^y = \frac{1}{\zeta} l_t^d z_t$$

where ζ is data labor productivity in final output sector.

Final output producers choose l^y, l^d(or d^y), x^i to maximize profits by taking market prices as given. Necessary conditions are $\alpha y = w^y l^y$, $\beta y = w^d l^d$, and $\gamma y \frac{\partial X}{\partial x^i} = p^i X$ where p^i is price of intermediate good i and $X = \left(\int_0^{A_t} (x_t^i)^\nu di\right)^{1/\nu}$. The demand for intermediate good i is with price index P

$$x_t^i = X_t P_t^{\frac{1}{1-\nu}} (p_t^i)^{\frac{1}{\nu-1}}.$$

3. Intermediate good producer

Given demand for intermediate good i, intermediate good producer maximizes profits by setting a price,

$$\max_{p_t^i} \pi_t^x = p_t^i x_t^i - r_t \eta x_t^i$$

where $\eta > 0$ is constant. The optimal pricing strategy becomes

$$p_t^i = \frac{\eta}{\nu} r_t.$$

The price of intermediate goods becomes constant so that economy-wide capital can be defined as $k_t = \eta x_t A_t$. The profit is obtained $\pi_t^x = \eta r_t x_t^i (1-\nu)/\nu$. Supplying intermediate good i, the firm i gets compensated for sunk cost through retention of monopoly power over the commercial use of the invention and this monopoly power is supported by patents of infinite duration (Alvarez-Pelaez and Groth 2005).

4. R&D sector: AI technology

AI technology in the R&D sector evolves:

$$\dot{A}_t = \phi A_t \frac{d_t^A}{k_t}$$

where ϕ is productivity level and technology data for AI is generated by labor and raw data,

$$d_t^A = \frac{1}{\chi} l_t^d z_t$$

where χ is data labor productivity in the R&D sector. $1/k_t$ is introduced to adjust economic scale (Jones 1995, Gordon 2016). AI innovator (or inventor) maximizes $\pi_t^I = p_t^A \dot{A}_t - w_t^A l_t^A$. Optimization gives

$$p_t^A \phi A_t \frac{z_t}{\chi k_t} = w_t^A.$$

This condition means that the value of labor service is closely related to (1) the value of new technology p_t^A, (2) externality of existing technology A_t, (3) raw data generated from consumers z_t, and (4) capital showing current economy level k_t.

5. System of dynamic equations

In equilibrium, $\bar{c} = c$ (consumption), $w^y = w^d = w^A \equiv w$ (wage), $x^i = x^j \equiv x$ (intermediate good). Every agent optimizes her objective subject to constraints, no arbitrage condition (the return on a patent must be equal to the return on intermediate good) arises in the R&D sector, all markets are clear, and the transversality condition holds given the initial capital stock. Imposing equilibrium conditions, we can characterize the competitive equilibrium with data and AI technology by constructing the system of dynamic equations. Define u_t as a share of labor in the R&D sector (total labor is normalized as 1).

Consumption

$$\frac{\dot{c}_t}{c_t} = \frac{r_t - \rho}{\theta} + \omega\left(1 - \frac{1}{\theta}\right)\left[\varphi\frac{c_t}{z_t} - \delta\right]$$

Raw data

$$\frac{\dot{z}_t}{z_t} = \varphi\frac{c_t}{z_t} - \delta$$

Final output

$$\frac{\dot{y}_t}{y_t} = \frac{\gamma(1-\nu)}{\nu}\frac{\dot{A}_t}{A_t} - \frac{u_t(\alpha+\beta)}{(1-u_t)\alpha}\frac{\dot{u}_t}{u_t}$$

AI technology

$$\frac{\dot{A}_t}{A_t} = \phi\left(\frac{d_t^A}{k_t}\right)$$

National resource

$$\frac{\dot{k}_t}{k_t} = \frac{y_t}{k_t} - \frac{c_t}{k_t}$$

Labor

$$\frac{\dot{u}_t}{u_t} = -\frac{(1-u_t)\alpha}{u_t(\alpha+\beta)}\left[\gamma\nu\frac{y_t}{k_t} - \frac{\gamma(1-\nu)(1-u_t)}{(\alpha+\beta)}\frac{\dot{A}_t}{A_t} - \frac{\gamma(1-\nu)-\alpha\nu}{\alpha\nu}\frac{\dot{A}_t}{A_t}\right]$$

III. Simulation results

We are interested in how changes in variables relating to data policy impact economic outcomes in the data economy constructed in this article. Using the system of dynamic equations, we can calculate the long-run equilibrium and conduct simulation analysis focusing on privacy and utility (ω, θ), creation and destruction of raw data (φ, δ), AI productivity (ϕ, χ), technology substitution (ν), and factor complementarity in final output technology (α, β, γ). Due to space limitations, we focus on the first three cases and deliver concise results on each case. Initial parameter values imposed and benchmark equilibrium are summarized in Appendix.

1. Privacy and utility function

Due to the negative externality of privacy, a higher intensity of disutility from privacy violation (high ω) leads, through creation and destruction of raw data, to lower rates of technology and capital growth in the long run (see Figure 2). Accordingly, the growth rate of consumption falls. Privacy plays an important role in determining growth rates of economic variables, implying that the economic rationale to protect private information can be based on not only utility loss, but also changes in growth rates. The above results are robust whenever intertemporal substitution of consumption (θ) is larger than one.

2. Raw data

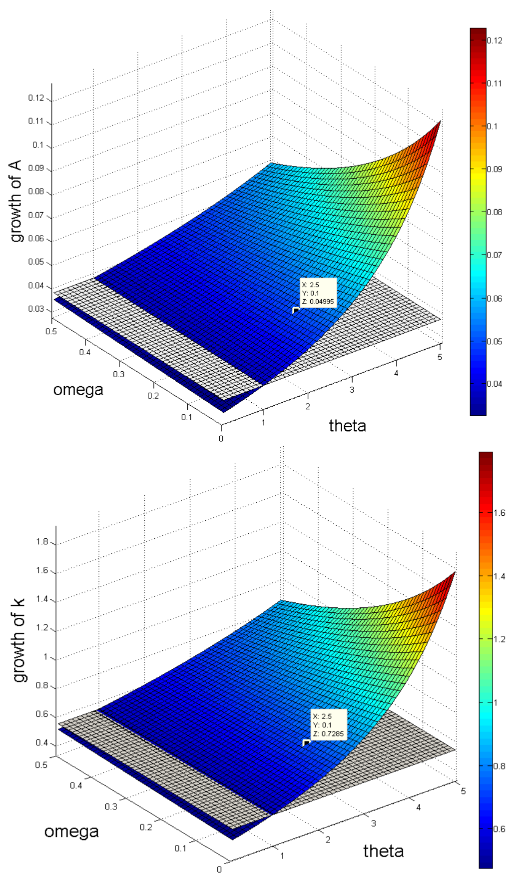
Creation and destruction of raw data can contribute to the usage of technology and production data. Raw data per capital and labor share in producing technology data increase as the economy has higher creation and lower destruction of raw data (high φ, low δ). It turns out that high raw data per capital leads to help enhancing productivity and technology growth in the long run.

3. Productivity of AI technology

Productivity changes in either AI technology (high ϕ) or labor of producing technology data in R&D (low χ) can directly affect shares of labor across sectors. It is worth noting that the economy-wide growth arises from AI-technology used as general purpose technology across sectors. We find that higher productivity of AI technology leads to higher growth rates of capital and technology, helping the economy grow in the long run.

Figure 2. Technology and capital growth: privacy violation and intertemporal substitution

Unit: %



Source: Authors' calculations.

IV. Concluding remarks

We build a general equilibrium model in which an economy grows by using data and AI technology. To the model, we add several characteristics of data including non-rivalry, privacy as generator of negative externality, information/idea as AI technology and others. By constructing a data- and AI-driven economy, we delineate how economic agents interact with one another through generating data and developing AI technology. Based on the constructed model, we also examine how changes in variables of data policy can affect long-run equilibrium.

This article provides three policy implications. First, the authority should have a balanced view between privacy protection and data usage in economy-wide technology in terms of long-run growth. Privacy should not be considered only as utility loss, but must be considered as a contributor to loss in growth rates. Second, economic growth can be achieved by using higher amounts of data as well as continuous development in AI technology. A caveat is that AI-technology can boost economic growth only when it applies to all industries as general purpose technology. Lastly, the authorities should keep considering how to deal with new issues that include data ownership, outlaw data sharing, data market, AI bias, and so forth. Our model can be used as a starting point to such examinations. KIEP

Appendix

Table 2. Initial parameter values

Parameter	Explanation	Value
α	Labor share in final output	0.6
β	Data labor share in final output	0.05
γ	Intermediate good share in final output	0.35
ν	(Inverse) elasticity of substitution among intermediate goods	0.8
ρ	Time discounting rate	0.03
δ	Destruction rate of raw data	0.6
φ	Intensity of raw data creation	0.1
ω	Degree of disutility from privacy violation	0.1
θ	(Inverse) intertemporal substitution	2.5
ϕ	Productivity in AI technology	1
χ	Labor productivity in R&D sector	0.03

Source: Authors.

Table 3. Long-run equilibrium as benchmark

Parameter	Explanation	Value
y/k	Output per capital	0.168
c/k	Consumption per capital	0.161
c/z	Consumption per raw data	6.073
z/k	Raw data per capital	0.027
u_A	Labor share in technology data	0.057
u_y	Labor share in final output	0.871
u_d	Labor share in production data	0.073
$\dot{A}/A$	Growth rate of AI technology	5.00%
$\dot{k}/k$	Growth rate of capital	0.73%

Source: Authors' calculations.