A Generalized Model of Growth in the Data Economy

Danxia Xie, PhD[1]; Longtian Zhang, PhD[2]
[1]Tsinghua University, [2]Central University of Finance and Economics

Abstract

In this paper, we develop a generalized model to describe the long-run effects of various data types. In addition to the conventional data types generated from consumption activities, we introduce two new types—producer data and nature data—generated from the production process and nature, respectively. To utilize these data, labor is required to collect and clean them; however, we assume no privacy concerns as in previous studies that solely focus on consumer data. Our investigation delves into the balanced growth path and find that the economy's growth rate is higher than that derived from the model using consumer data alone—a revelation previously unreported in existing studies. Moreover, we extend our analysis to incorporate all data types and present a comprehensive picture of the transformations of different growth regimes. Our framework has novel and crucial implications for understanding the developing trend of the data economy.

Introduction

In the data economy, a critical question arises is whether and how data will affect economic growth. Various theories have been proposed to describe and predict the developing trend of this new form of economy, especially its long-run growth. Most papers, such as Jones and Tonetti (2020), Cong et al. (2021), Cong et al. (2022), and Farboodi and Veldkamp (2021), consider data as an input factor generated directly or indirectly from consumer activities. However, recent observations show that a significant amount of data come not from consumers but from the production processes of industrial firms. For instance, the Industrial Internet of Things (IIoT) links sensors, instruments, and other equipment with computers, enabling data collection, exchange, and analysis, thereby potentially improving production efficiency. Moreover, we also observe that some data neither come from consumers nor the production process, but from nature, such as weather and geological data, which are critical for both production and consumers.

We propose a generalized model of growth in the data economy, highlighting the characteristics describing a spectrum of data types that can lead to different long-term effects. Our analyses incorporate the endogenous supply of new data types, the transition from previously studied models to our generalized model, and a new form of knowledge accumulation. Different combinations of the three data types also lead to various labor allocations among different sectors, complicating policy implications.

Model Settings

Generalized Model

- Representative consumer

$$U = \int_0^\infty e^{-(\rho-n)t}\left[\frac{c(t)^{1-\gamma}-1}{1-\gamma} - \varphi(t)^\sigma\right]\mathrm{d}t\,,$$

where $c(t)$ is the per capita consumption and $\varphi(t)$ is the usage of consumer data.

- Final good producer
  - The production function is

$$Y(t) = L_E(t)^\beta \int_0^{N(t)} x(v,t)^{1-\beta}\mathrm{d}v\,,$$

  where $L_E(t)$ is labor employment, $N(t)$ is variety, and $x(v,t)$ is intermediate goods.
  - Final good producer also generate producer data

$$\phi(t) = y(t)^{\alpha_y} L_D(t)^{\alpha_l},$$

  where $y(t)$ is per capita output and $L_D(t)$ is labor employment.
- Nature data firm

$$\phi^*(t) = A_N L_N(t)^{\beta_l}$$

- Intermediate good producer
  - Production phase, the profit function

$$\pi(t) = p_x(v,t)x(v,t) - \psi x(v,t)$$

  - Innovation phase, innovation possibility frontier

$$\dot{N}(t) = \eta N(t)^\zeta \Phi(t)^\xi L_R(t)^{1-\xi}$$

$$\Phi(t) = \left[a\big(\varphi(t)L(t)\big)^{\frac{\varepsilon-1}{\varepsilon}} + b\phi(t)^{\frac{\varepsilon-1}{\varepsilon}} + (1-a-b)\big(\phi^*(t)\big)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}$$

Contact

Longtian Zhang (张龙天)
Central University of Finance and Economics
Email: zhanglongtian@cufe.edu.cn
Website: https://longtianzhang.com
Phone: (+86) 13161693539

Results

- BGP growth rates are the same in both decentralized economy and the social planner's problem, but the variable levels are different.
- When the combination of data is lack of elasticity, $0 < \varepsilon < 1$,

$$g^* = \left[\frac{\sigma}{(1-\zeta)\sigma - \xi(1-\gamma)}\right]n, \qquad g_\varphi^* = \left[\frac{1-\gamma}{(1-\zeta)\sigma - \xi(1-\gamma)}\right]n < 0,$$

$$g_{l_D}^* = \left\{\frac{(1-\varepsilon)\left[(1-\alpha_l)\big(\sigma(1-\zeta) - \xi(1-\gamma)\big) - \sigma\alpha_y + 1 - \gamma\right]}{\alpha_l\left[(1-\zeta)\sigma - \xi(1-\gamma)\right]}\right\}n < 0.$$

- When the combination of data is unit elasticity, the BGP growth rates are a mix of the three types of data.
- When the combination of data is full of elasticity, $\varepsilon > 1$,

$$g^* = \left[\frac{1-(1-\alpha_l)\xi}{1-\zeta-\alpha_y\xi}\right]n,$$

$$g_\varphi^* = -\left\{\frac{\alpha_l[(\varepsilon-1)(1-\zeta) + \varepsilon\xi(\gamma-1)] + \varepsilon[(1-\xi)(\gamma-1) - (1-\zeta)] + \alpha_y(\varepsilon-1) + (1-\zeta)}{[1+\varepsilon(\sigma-1)]\big(1-\alpha_y\xi-\zeta\big)}\right\}n < 0,$$

$$g_{l_D}^* = 0.$$

- The BGP growth rate of producer data is

$$g_\phi^* = \alpha_y g^* + \alpha_l\big(g_{l_D}^* + n\big).$$

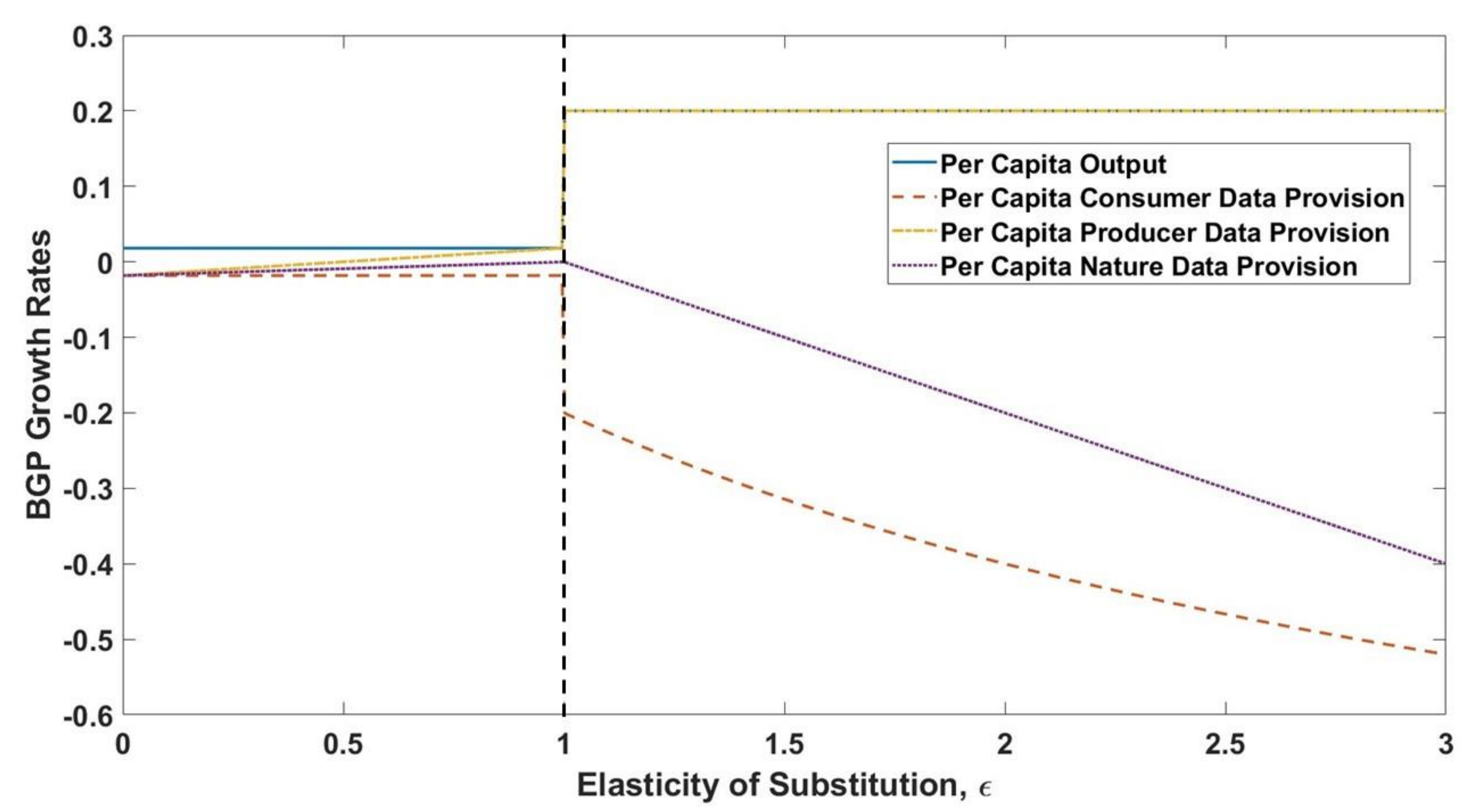

Figure 1. BGP growth rates with different elasticity of substitution, ε

Table 1. BGP growth rates in the three related models.

	Per capita output	Per capita data provision
Producer data model	$g_{\text{prod}}^* > g_{\text{com}}^*$	$g_{\text{prod},\phi}^* > g_{\text{com},\varphi}^*$
Nature data model	$g_{\text{nature}}^* > g_{\text{com}}^*$	$g_{\text{nature},\phi}^* > g_{\text{com},\varphi}^*$
Consumer data model	$g_{\text{com}}^* = \left[\frac{\sigma}{(1-\zeta)\sigma-\xi(1-\gamma)}\right]n$	$g_{\text{com},\varphi}^* = \left[\frac{1-\gamma}{(1-\zeta)\sigma-\xi(1-\gamma)}\right]n$

Table 2. Comparisons of different models.

	Generalized model		
	$0 < \varepsilon < 1$	$\varepsilon = 1$	$\varepsilon > 1$
Consumer data model	$g_{\text{com}}^* = g^*$, $y_{\text{com}}^* < y^*$	When $a = 1$, $g_{\text{com}}^* = g^*$	$g_{\text{com}}^* < g^*$
Nature data model	$g_{\text{nature}}^* > g^*$	$g^*(a=1) < g_{\text{nature}}^* < g^*(b=1)$	$g_{\text{nature}}^* < g^*$
Producer data model	$g_{\text{prod}}^* > g^*$	When $b = 1$, $g_{\text{prod}}^* = g^*$	$g_{\text{prod}}^* = g^*$ $y_{\text{prod}}^* > y^*$

Conclusions

We develop a generalized model to integrate all the possible data types and study the transformations of different growth regimes in a data economy. We study a generalized model and study changes among different models as we vary the elasticity of substitution between the three data types. We find a more effective way to push the data economy to a higher growth regime, which will further improve social welfare and provide a full picture of data economy from the initial stage to a high-growth regime.

References

1. Cong, L. W., W. Wei, D. Xie, and L. Zhang (2022). Endogenous Growth Under Multiple Uses of Data. *Journal of Economic Dynamics and Control*, 141, 104395.
2. Cong, L. W., D. Xie, and L. Zhang (2021). Knowledge Accumulation, Privacy, and Growth in a Data Economy. *Management Science*, 67(10), 6480-6492.
3. Farboodi, M. and L. Veldkamp (2021). A Model of the Data Economy. NBER Working Papers 28427, National Bureau of Economic Research.
4. Jones, C. I. and C. Tonetti (2020). Nonrivalry and the Economics of Data. *American Economic Review*, 110(9), 2819-2858.