Macroeconomic Dynamics (2024), **28**, pp. 1206–1230
doi:10.1017/S1365100523000445

ARTICLE

Creative destruction, human capital accumulation, and growth in a digital economy

Lewei Liao[1], Xuezheng Qin[2], Xiaolong Li[3], and Liutang Gong[4]

[1]Guanghua School of Management, Peking University, Beijing, China
[2]School of Economics, Peking University, Beijing, China
[3]School of Modern Post, Beijing University of Posts and Telecommunications, Beijing, China
[4]Guanghua School of Management, Peking University, Beijing, China
Corresponding author: Xiaolong Li; Email: tell714@gmail.com

Abstract

The rapid development of the digital economy has highlighted the crucial role of data in economic growth. This study investigates the impact of two types of innovation on long-term growth by incorporating data into a model of creative destruction and knowledge accumulation. Unlike traditional factors, data exhibit nonrivalry between the two research and development (R&D) sectors, thereby influencing the growth rate of economic outputs simultaneously without interference. Our findings reveal the existence of a balanced growth path (BGP) in both the decentralized economy and the social planner's economy. In horizontal innovation, data can be transformed into digital knowledge to promote the economic growth [Cong et al. (2021)]. In addition to horizontal innovation, the utilization of data in vertical innovation also enhances the success rate of innovation, with a gradual decrease in per capita data usage on the BGP. Moreover, as agents accumulate human capital, the economy achieves higher output levels, effectively addressing consumer privacy concerns. However, along the transitional path, insufficient data provision by both R&D sectors leads to lower economic growth rates or more intense economic fluctuations, necessitating policy interventions.

Keywords: Data; Vertical Innovation; Human Capital; Privacy; Nonrivalry

JEL classifications: J24; O33; O41

1. Introduction

With the continuous advancement of new generation information technologies, such as the mobile web and artificial intelligence, the effective stock of data is experiencing an explosive growth trend, showcasing its immense value-creating potential.[1] Various countries possess distinct comparative advantages in leveraging data to drive economic development.

Numerous recent studies have delved into the utilization of data as a novel factor directly or indirectly employed in production to foster the economic growth [Jones and Tonetti (2020), Cong et al. (2021), and Farboodi and Veldkamp (2021)]. Data have found extensive applications in the production of final goods as well as in knowledge creation. Besides, these studies acknowledge that in cases where producers have ownership over data, they may choose to hoard it as a preventive measure against creative destruction.[2] However, they fail to recognize creative destruction as a driving force behind economic growth.[3] We aim to explore the role played by data in both

We thank Xiaoyong Cui, Chan Wang, Danxia Xie, Longtian Zhang, Pengfei Sun, Junru Chen and the anonymous associate editor and referees for their feedback. All remaining errors are the authors.

vertical and horizontal innovation, thereby implying the existence of two distinct markets for data applications.

To address this research gap, this paper develops an endogenous growth model for the digital economy, building upon Howitt (1999) framework. In the benchmark model, agents generate data through their consumption activities and possess the right to profit from selling this data to two types of innovators. Meanwhile, they also experience disutility due to privacy concerns. Additionally, agents can supply labor to both the innovation sectors and the production sectors in exchange for wages. Notably, the nonrival nature of data allows for its utilization both as digital knowledge and as a means to improve the productivity parameter attached to intermediate goods. This latter aspect distinguishes the model from existing frameworks. We specifically focus on characterizing equilibria on the balanced growth path (BGP) and delves into exploring the role of data in vertical innovation.

The main findings of this paper highlight the positive impact of data usage in vertical innovation on the success rate of innovation. By employing data to replace labor in the innovation process, intermediate producers can allocate more labor to final goods production. Consequently, as the economy develops, the cost of labor increases while the price of data decreases. Ultimately, data usage in vertical innovation leads to greater monopoly profits for intermediate producers. The continuous spillover effect resulting from knowledge accumulation enables data to contribute to innovation success, while the higher productivity parameter adds complexity to achieving successful innovation outcomes.

In contrast to the decentralized economy, we show that the social planner's economy exhibits relatively higher growth rates on the BGP. Three market imperfections contribute to this disparity. First, monopolistic markups in the production of intermediate producers result in the crowding out of labor in both types of innovation [Jones (1995)]. Second, there exists an intertemporal spillover of intermediate products, as R&D sectors primarily consider the temporary stream of profits derived from innovation, while society as a whole enjoys the benefits of innovation. Lastly, in the decentralized economy, intermediate producers must purchase data at a corresponding price. Conversely, in the social planner's economy, data are allocated to intermediate producers without any cost. Consequently, fewer data and labor are employed in the two types of R&D sectors, leading to a weaker driving force of innovation in the decentralized economy compared to the social planner's economy.

Furthermore, previous studies have struggled to strike a balance between the extent of data usage and economic growth in the evolving digital economy. This paper reveals that achieving rapid economic development is feasible even with a limited amount of data, provided that the positive impact of digital knowledge on human capital accumulation is taken into consideration.

During the transition towards steady growth in the social planner's economy, data play distinct roles in the long run, particularly in relation to the two types of innovation. First, horizontal innovation influences the time it takes for the economy to reach the BGP. If data provision for horizontal innovation is limited, the economy will require more time to converge to the BGP and may experience more pronounced fluctuations. Second, vertical innovation determines the growth rate of the economy prior to reaching the BGP. If data usage in vertical innovation is constrained, it will adversely affect social welfare and consumer surplus during the transitional period.

We contribute to the extensive literature on economic growth by examining the role of data, complementing recent studies such as Jones and Tonetti (2020), Cong et al. (2021), and Farboodi and Veldkamp (2021). In line with these studies, we find that data exhibit a bounded effect on long-term economic growth, leveraging an innovative approach inspired by Howitt (1999) to indirectly incorporate data into production. Prior research, including studies by Aghion and Howitt (1992), Howitt and Aghion (1998), Howitt (1999), has explored the impact of various input factors on creative destruction and economic growth, such as labor, physical capital, and human capital. By extending this line of inquiry, we enhance the model's realism by incorporating data as an additional input factor. This approach allows for a more comprehensive understanding of the dynamics of economic growth, capturing the multifaceted nature of real-world economies.

In addition to its contributions to the literature on economic growth, we also make significant contributions to the emerging field of research on the digital economy, privacy, and information. Early studies, such as those by Hirshleifer (1971), Admati and Pfleiderer (1990), and Murphy (1996), have explored the social value, sales, and property rights of information. More recent research, such as Choi et al. (2019), Acemoglu et al. (2019), Ichihashi (2020, 2021a, b), focuses on the competition among digital platforms or intermediaries. Others, including Akçura and Srinivasan (2005), Casadesus-Masanell and Hervas-Drane (2015), Fainmesser et al. (2022), and Sun et al. (2022), delve into the connection between data and privacy concerns. Building upon the works of Jones and Tonetti (2020) and Cong et al. (2021), which propose that privacy concerns can be mitigated if consumers have the autonomy to determine the extent of data sharing, this paper sheds light on the underlying mechanism that promotes economic growth while mitigating privacy issues. It further reveals that the accumulation of human capital through digital knowledge can effectively alleviate privacy concerns without compromising overall welfare. By addressing these important topics, we expand the understanding of the intricate relationship between the digital economy, privacy concerns, and economic growth, offering valuable insights into the potential synergies and trade-offs in these domains.

The remaining sections of this paper are structured as follows. Section 2 presents the construction of an endogenous growth model within the framework of a decentralized economy, incorporating two research and development (R&D) sectors. Section 3 focuses on solving and analyzing the BGP of the endogenous growth model in the decentralized economy, using the benchmark model as a reference. Moving forward, Section 4 establishes an endogenous growth model within the context of a social planner's economy. The BGP of this model is solved and analyzed to understand its key properties. In Section 5, we investigate the scenario where innovators possess ownership of data and compares it with the benchmark model. Section 6 delves into the role of human capital in the digital economy, examining its implications for growth. Subsequently, Section 7 conducts a numerical analysis to study the characteristics of transitional dynamics. Finally, Section 8 concludes the paper, summarizing the key findings and offering insights for future research.

2. The baseline model

This paper incorporates key elements such as vertical innovation (creative destruction), horizontal innovation (digital knowledge accumulation), data production, and privacy concerns into a macroeconomic model within the framework of a decentralized economy. The aim is to capture the essential characteristics of economic growth in a dynamic digital economy. The dynamic digital economy model consists of various representative agents who fulfill multiple roles, acting as both consumers and workers, as well as horizontal and vertical innovators. Additionally, there are intermediate producers and a final goods producer within the model's structure. It is important to note that the model assumes a continuous and infinite time framework, providing a suitable context to examine the dynamics of the digital economy and its implications for economic growth.

2.1. Representative consumers

The numerical value of homogeneous representative consumer groups within the model exhibits a constant growth rate denoted by n and reaches a population size of $L(t)$ at time t. In addition to their consumption decisions, each consumer has the option to supply one unit of labor inelastically to either of the two types of R&D sectors or to final goods production.

Data, generated as a by-product of consumer activities, can be individually sold by consumers to the two types of R&D sectors, implying that consumers possess the property rights over their own data [Cong et al. (2021)]. For instance, when consumers engage in online shopping or use ride-hailing apps, their consumption patterns and behavior characteristics are recorded as data. R&D sectors can leverage this data to develop new business models, create innovative products, and enhance their ability to meet market demands.

However, when consumers sell their data to any R&D sector, their utility is impacted by concerns related to the leakage and misuse of personal information. As consumers can derive profits from selling their data, they strive to strike a balance between the financial gains and the associated privacy issues. Data subsidies can be viewed as additional bonuses beyond traditional income, enabling agents to optimize their utility based on their privacy sensitivity. Given the uncertainty surrounding how the data will be utilized, the disutility stemming from selling data to different R&D sectors varies among agents.

Each consumer's utility maximization problem can be represented as[4]

$$\max_{c(t),D_H(t),D_V(t)} \int_0^{\infty} e^{-(\rho-n)\cdot t}\cdot\left[\frac{c\,(t)^{1-\gamma}-1}{1-\gamma}-\chi_H\cdot D_H\,(t)^{\sigma}-\chi_V\cdot D_V\,(t)^{\sigma}\right]dt; \tag{1}$$

subject to

$$\dot{a}\,(t)=(r\,(t)-n)\cdot a\,(t)+w\,(t)+p_{D,H}\,(t)\cdot D_H\,(t)+p_{D,V}\,(t)\cdot D_V\,(t)-c\,(t)\,; \tag{2}$$

and

$$D_H\,(t)\leq\Theta_1(c(t)); \tag{3}$$

$$D_V\,(t)\leq\Theta_2(c(t)); \tag{4}$$

In the equations provided, $c(t)$ represents the per capita consumption level at time t, while $D_H(t)$ and $D_V(t)$ represent the quantities of data provided by a consumer to horizontal and vertical R&D sectors, respectively. The parameter ρ denotes the consumer's discount rate, and γ, which lies within the range $(1,\infty)$, represents the reciprocal of the elasticity of intertemporal substitution of consumption. The parameter σ parameterizes the disutility associated with data leakage, which depends on the quantity of data provided each time. The variables χ_H and χ_V represent the degrees of negative effect on utility when consumers sell data to horizontal and vertical R&D sectors, respectively [Cong et al. (2022)]. Furthermore, $\Theta_1(\cdot)$ and $\Theta_2(\cdot)$ are both increasing exogenous general functions that determine the generating process of data. The specific form of $\Theta_1(\cdot)$ and $\Theta_2(\cdot)$ reflects the constraints imposed on data generation, which can be influenced by factors such as the digital infrastructure, legal framework, and local privacy regulations. In the baseline model, it is assumed that $\Theta_1(c(t))$ and $\Theta_2(c(t))$ are sufficiently large, enabling consumers to have the autonomy to determine the quantity of data they wish to sell based on their preferences and constraints.

In the budget constraint equation (2), $a(t)$ represents the asset held by a consumer at time t, $r(t)$ represents the interest rate at time t, and $w(t)$ represents the wage for labor at time t. The variables $p_{D,H}(t)$ and $p_{D,V}(t)$ represent the time-t prices of data for horizontal and vertical R&D sectors, respectively. These prices are determined based on the nonrivalry and exclusivity of data. In order to simplify the model, it is assumed that consumers can directly sell the same data to both types of R&D sectors simultaneously, without the presence of competitive data intermediaries. This assumption aligns with previous studies such as Jones and Tonetti (2020) and Cong et al. (2021). The disutility of data leakage caused by selling data is not constant and varies depending on the value of the parameter σ. If σ is larger than 1, the privacy cost associated with selling data will increase. Therefore, the price of data purchased by the R&D sectors is influenced by the quantity of data purchased by each sector. Constraints (3) and (4) impose limitations on the

https://doi.org/10.1017/S1365100523000445 Published online by Cambridge University Press

growth rate of data provision, ensuring that it remains bounded by the corresponding growth rate of consumption. This constraint reflects the idea that data, being by-products of economic activities, cannot exceed a certain proportion of consumption activities.

To obverse the relevant variables for transitional dynamics, we derive the system's evolution in the form of Euler equations from the Hamilton:

$$\frac{\dot{c}(t)}{c(t)} = \frac{r(t) - \rho}{\gamma}; \tag{5}$$

$$\frac{\dot{p}_{D,H}(t)}{p_{D,H}(t)} - (\sigma - 1) \cdot \frac{\dot{D}_H(t)}{D_H(t)} = r(t) - \rho; \tag{6}$$

$$\frac{\dot{p}_{D,V}(t)}{p_{D,V}(t)} - (\sigma - 1) \cdot \frac{\dot{D}_V(t)}{D_V(t)} = r(t) - \rho; \tag{7}$$

2.2. The final good producer

A representative final good producer operates within a competitive environment, characterized by a production function:

$$Y(t) = L_E(t)^{\alpha} \cdot \int_0^{N(t)} A(v,t)^{\alpha} \cdot x(v,t)^{1-\alpha}\, dv; \tag{8}$$

In equation (8), $Y(t)$ represents the gross output, $L_E(t)$ denotes the amount of labor employed in the production of the final goods, $N(t)$ signifies the number of intermediate goods varieties utilized in the final goods production, and $A(v,t)$ represents the productivity parameter associated with the most recent version of intermediate good v at time t.Additionally, $x(v,t)$ refers to the total quantity of the intermediate good of variety v at time t, which can only be utilized in the production of final goods for a single period. The elasticity coefficient of labor in final goods production is denoted as α, and the rental fee for intermediate goods is denoted as $p_x(v,t)$. The first-order conditions derived from the profit maximization of the final goods producer, with respect to both labor employed and the quantity of each intermediate good, are as follows:

$$x(v,t) = \left[\frac{(1-\alpha)}{p_x(v,t)}\right]^{\frac{1}{\alpha}} \cdot A(v,t) \cdot L_E(t); \tag{9}$$

$$\alpha \cdot [L_E(t)]^{\alpha-1} \cdot \int_0^{N(t)} A(v,t)^{\alpha} \cdot x(v,t)^{1-\alpha}\, dv = w(t); \tag{10}$$

In equation (9), the quantity of intermediate good variety v at time t is determined by a combination of the rental fee, productivity parameter, and the amount of labor employed. Equation (10) represents the condition where the marginal productivity of labor in the final goods production is equal to the wage rate, denoted as $w(t)$.

2.3. Intermediate producers

The decision-making process of potential intermediate producers in conducting research is not limited, allowing them to determine whether to engage in research activities and the extent of their investment. Successful horizontal innovation grants these producers a monopoly over the intermediate product they have developed. Consequently, the market will witness the emergence of a certain number of monopoly intermediate producers. Conversely, if vertical innovation proves

https://doi.org/10.1017/S1365100523000445 Published online by Cambridge University Press

successful, an existing monopoly intermediate producer will be replaced. Until the next instance of vertical innovation occurs, new intermediate producers can continue to enjoy monopoly profits. It is important to note that we assume that once previous innovators exit the market, they are unable to reenter. Therefore, the latest innovator will never face competition from past innovators [Howitt and Aghion (1998)]. To address the dynamics of intermediate producers, we employ a backward induction methodology, allowing for a comprehensive analysis of the problem.

2.3.1. Final good production phase

For each monopoly intermediate producer entering the market, the profit obtained from the intermediate good of variety v in a single period, denoted as t is determined by the following profit-maximizing function:

$$\max_{p_x(v,t)} \left[p_x\left(v,t\right) \cdot x\left(v,t\right) - \psi \cdot x\left(v,t\right) \right]; \tag{11}$$

In equation (11), ψ represents the constant marginal cost associated with this production process in the economy.

Substituting equation (9) into equation (11) and taking the derivative with respect to $p_x(v, t)$ yields the optimal price for each variety of intermediate goods:

$$p_x\left(v,t\right) = \frac{\psi}{1-\alpha}; \tag{12}$$

It is evident that the optimal price of intermediate goods, $p_x(v, t)$, is independent of the specific intermediate good variety v and the time period.

Substituting equations (12) and (9) into equation (11), we can derive the quantity of intermediate good of variety v:

$$x\left(v,t\right) = \frac{\left(1-\alpha\right)^{\frac{2}{\alpha}}}{\psi} \cdot A\left(v,t\right) \cdot L_E\left(t\right); \tag{13}$$

By substituting equation (13) into equation (11), the maximum profit obtained from the intermediate good of variety v can be expressed as

$$\pi\ \left(\mathrm{v},\mathrm{t}\right) = \frac{\left(1-\alpha\right)^{\frac{2}{\alpha}-1}}{\psi^{\frac{1}{\alpha}-1}} \cdot \alpha \cdot A\left(v,t\right) \cdot L_E\left(t\right); \tag{14}$$

Furthermore, the gross output, $Y(t)$, and the wage rate, $w(t)$, can be expressed as

$$Y\left(t\right) = \left[\frac{\left(1-\alpha\right)^2}{\psi}\right]^{\frac{1}{\alpha}-1} \cdot L_E\left(t\right) \cdot \int_0^{N(t)} A\left(v,t\right) dv; \tag{15}$$

$$w\left(t\right) = \alpha \cdot \left[\frac{\left(1-\alpha\right)^2}{\psi}\right]^{\frac{1}{\alpha}-1} \cdot \int_0^{N(t)} A\left(v,t\right) dv; \tag{16}$$

In equations (15) and (16), the gross output, $Y(t)$, is determined by the integral over all varieties of intermediate goods, denoted by $N(t)$, and their respective productivity parameter $A(v, t)$. The wage rate, $w(t)$, is a function of the integral over the same range.

2.3.2. Vertical innovation phase

This paper examines the vertical innovations that result in quality improvements. Vertical R&D sectors are capable of conducting research and development (R&D) activities by utilizing both labor, denoted as $L_V(t)$, and data acquired from consumers, represented as $D_V(t) \cdot L(t)$. The

Poisson arrival rate of vertical innovations in any given sector is determined by $\Phi(t) = \mu \cdot \phi(t)$, $\mu > 0$, where μ is a parameter indicating the productivity of vertical R&D, and $\phi(t)$ represents the innovation success rate equation adjusted for productivity in vertical R&D within each sector [Howitt (1999)]. The productivity-adjusted innovation success rate equation can be formulated as follows:

$$\frac{L_V\,(t)^{\beta} \cdot [D_V\,(t) \cdot L\,(t)]^{1-\beta}}{N\,(t) \cdot A^{\max}\,(t)}; \tag{17}$$

Equation (17) illustrates that vertical R&D sectors can enhance the success rate of innovation by utilizing a substantial amount of data during the scientific research process. In this equation, the leading-edge productivity parameter, $A^{\max}(t)$, represents the maximum value among all productivity parameters associated with the latest version of intermediate products. The parameter β signifies the contribution of labor in the vertical innovation process. $L_V(t)$ denotes the labor employed in vertical R&D sectors, while $l_V(t)$ represents the fraction of labor allocated to vertical R&D sectors. It is evident that as technology advances and the number of intermediate goods expands, vertical innovation becomes more intricate and challenging.[5]

Upon the successful completion of R&D, a new intermediate producer enters the monopolistic market, replacing the previous producer. The present value of the intermediate good with the leading-edge productivity parameter in a single period, denoted as t, can be calculated using the following expression:

$$V\,(t) = \int_t^{\infty} e^{-\int_t^{\tau}[r(s)+\Phi(s)]ds} \cdot \tilde{\pi}\,(\tau)\,d\tau, \tag{18}$$

In equation (18), $\tilde{\pi}(\tau)$ represents the profit obtained from the intermediate good with the leading-edge productivity parameter in a single period, given by

$$\tilde{\pi}\,(t) = \frac{(1-\alpha)^{\frac{2}{\alpha}-1}}{\psi^{\frac{1}{\alpha}-1}} \cdot \alpha \cdot A^{\max}\,(t) \cdot L_E\,(t)\,; \tag{19}$$

The present value expression implies that higher investment in R&D leads to a greater success rate of innovation and a higher probability of becoming a monopoly producer of intermediate goods. However, it also implies that a higher success rate of innovation in the next phase results in a shorter cycle of replacement and a lower present value that can be obtained. This concept is rooted in the notion of creative destruction [Aghion and Howitt (1992)].

Vertical R&D sectors determine the optimal allocation of labor, $L_R(t) = l_R(t) \cdot L(t)$, and data, $D_V(t) \cdot L(t)$, to maximize the expected net profit. This decision can be formulated as follows[6]:

$$\max_{D_V(t), l_V(t)} \left\{\Phi\,(t) \cdot V\,(t) - w\,(t) \cdot L_V\,(t) - p_{D,V}\,(t) \cdot D_V\,(t) \cdot L\,(t)\right\}; \tag{20}$$

The first-order condition gives rise to two free-entry conditions:

$$\frac{\beta \cdot \mu \cdot l_V\,(t)^{\beta-1} \cdot D_V\,(t)^{1-\beta} \cdot \tilde{\pi}\,(t) \cdot [r\,(t) - n]}{w(t) \cdot [r\,(t) + \Phi\,(t) - n]^2 \cdot N(t) \cdot A^{\max}\,(t)} = 1; \tag{21}$$

and

$$\frac{(1-\beta) \cdot \mu \cdot l_V\,(t)^{\beta} \cdot D_V\,(t)^{-\beta} \cdot \tilde{\pi}\,(t) \cdot [r\,(t) - n]}{p_{D,V}\,(t) \cdot [r\,(t) + \Phi\,(t) - n]^2 \cdot N\,(t) \cdot A^{\max}\,(t)} = 1; \tag{22}$$

Equations (21) and (22) represent the two free-entry conditions derived from the first-order condition, which are integral to the decision-making process of vertical R&D sectors.

2.3.3. Horizontal innovation phase

Horizontal innovations are the result of R&D efforts aimed at creating new intermediate goods. In the horizontal R&D sector, both labor, $L_H(t)$, and data purchased from consumers, $D_H(t) \cdot L(t)$, are employed [Romer (1990); Jones (1995)]. The evolution of the aggregate innovation possibility frontier can be described by the following equation:

$$\dot{N}(t) = \frac{\eta \cdot N(t)^{\xi} \cdot (D_H(t) \cdot L(t))^{\theta} \cdot L_H(t)^{1-\theta}}{A^{\max}(t)}; \tag{23}$$

This equation can be simplified as

$$\dot{N}(t) = \frac{\eta \cdot N(t)^{\xi} \cdot l_H(t)^{1-\theta} \cdot D_H(t)^{\theta} \cdot L(t)}{A^{\max}(t)}; \tag{24}$$

From equation (24), we can observe that data can be transformed into digital knowledge in the process of horizontal innovation. Here, $\eta > 0$ is an efficiency term of innovation, $\theta \in (0, 1)$ represents the contribution of data in the horizontal innovation process, and $\xi \in (1, +\infty)$ represents the spillover effect of digital knowledge. $L_H(t)$ denotes the labor employed in the horizontal R&D sectors, and $l_H(t)$ represents the fraction of labor employed in the horizontal R&D sectors. Notably, digital knowledge is created through data spillovers to future periods by creating new varieties [Cong et al. (2021)]. The larger the leading-edge productivity parameter, the more complex the process of horizontal innovation becomes [Howitt (1999)].

Each horizontal innovation gives rise to a new intermediate product whose productivity parameter is randomly drawn from the distribution of existing intermediate products.

The expected value of a horizontal innovation can be expressed as

$$E\left[\frac{A(\nu, t)}{A^{\max}(t)}\right] \cdot V(t).$$

In the decision-making process, horizontal R&D sectors determine the optimal allocation of labor, $L_H(t) = l_H(t) \cdot L(t)$, and data, $D_H(t) \cdot L(t)$, to maximize the expected net profit. This decision can be formulated as follows:

$$\max_{l_H(t), D_H(t)} \left\{ E\left[\frac{A(\nu, t)}{A^{\max}(t)}\right] \cdot V(t) \cdot \dot{N}(t) - w(t) \cdot L_H(t) - p_{D,H}(t) \cdot D_H(t) \cdot L(t) \right\}; \tag{25}$$

The first-order condition leads to two free-entry conditions:

$$\eta \cdot (1-\theta) \cdot N(t)^{\xi} \cdot l_H(t)^{-\theta} \cdot D_H(t)^{\theta} \cdot V(t) \cdot E\left[\frac{A(\nu, t)}{A^{\max}(t)}\right] = w(t) \cdot A^{\max}(t); \tag{26}$$

and

$$\eta \cdot \theta \cdot N(t)^{\xi} \cdot l_H(t)^{1-\theta} \cdot D_H(t)^{\theta-1} \cdot V(t) \cdot E\left[\frac{A(\nu, t)}{A^{\max}(t)}\right] = p_{D,H}(t) \cdot A^{\max}(t); \tag{27}$$

Equations (26) and (27) represent the two free-entry conditions derived from the first-order condition, which play a crucial role in the decision-making process of horizontal R&D sectors.

2.3.4. Spillovers

The growth in the leading-edge parameter, $A^{\max}(t)$ occurs due to the knowledge spillovers produced by vertical innovations. $A^{\max}(t)$ always grows at a rate proportional to the aggregate rate of vertical innovations [Caballero and Jaffe (1993)]. The factor of proportionality, $\frac{\varphi}{N(t)}$, $N(t)$, by the flow of vertical innovations per sector. This relationship can be expressed as

$$\frac{\dot{A}^{\max}(t)}{A^{\max}(t)} = \varphi \cdot \mu \cdot \phi(t); \tag{28}$$

The distribution of productivity parameters among new intermediate goods mirrors the distribution across existing intermediate goods at time t. The distribution of relative productivity parameters, $b(v,t) = \frac{A(v,t)}{A^{\max}(t)}$, can be described by the following expression [Howitt (1999)][7]:

$$F\left(b(v,t) \leq b\right) = \left[b(v,t)\right]^{\frac{1}{\varphi}};$$

In the long term, it follows that:

$$E\left[\frac{A(v,t)}{A^{\max}(t)}\right] = \frac{1}{1+\varphi}; \tag{29}$$

Equation (29) represents the expected value of the ratio of the productivity parameter of a new intermediate good to the leading-edge parameter in the long term.

3. Decentralized economy on the balanced growth path

In the decentralized economy on the BGP, an equilibrium is achieved where the evolution of the variable $\{N(t)\}_{t=0}^{\infty}$ is determined by free entry. Intermediate producers optimize their choices of $\{p_x(v,t), D_H(t), D_V(t),$

$L_H(t), L_V(t)\}_{t=0}^{\infty}$ to maximize the discounted value of profits. The evolution of $\{r(t), w(t), p_V(t), p_H(t)\}_{t=0}^{\infty}$ is consistent with market clearing, and the evolution of $\{L_E(t), x(v,t)\}_{t=0}^{\infty}$ is consistent with profit maximization by the final good producer.

The model is solved along the BGP, which assumes that all variables are growing at the same constant rate, denoted as $r(t) = r^*$. The relationships between the variables are as follows:

$$\frac{L_V(t)^{\beta} \cdot [D_V(t) \cdot L(t)]^{1-\beta}}{N(t) \cdot A^{\max}(t)} \to m; \tag{30}$$

$$\frac{\dot{D_H}(t)}{D_H(t)} = \frac{\dot{D_V}(t)}{D_V(t)}; \tag{31}$$

$$\frac{\dot{D_H}(t)}{D_H(t)} = \frac{1+\xi}{1-\beta-\theta} \cdot \frac{\dot{N}(t)}{N(t)}; \tag{32}$$

$$\frac{\dot{y}(t)}{y(t)} = \frac{\dot{c}(t)}{c(t)} = \frac{\dot{A^{\max}}(t)}{\int_0^{N(t)} A(v,t)\,dv} + \frac{A(v,t)}{\int_0^{N(t)} A(v,t)\,dv} \cdot \dot{N}(t); \tag{33}$$

Equation (30) states that on the BGP, the ratio of labor and data in vertical R&D sectors to the aggregate level of labor and the leading-edge productivity parameter tends to a certain value, represented as m in this paper.

Equation (31) indicates that the growth rate of per capita data provision in horizontal R&D sectors is the same as the growth rate of per capita data provision in vertical R&D sectors.

Equation (32) further explains the relationship between the growth rate of per capita data provision and the growth rate of varieties of intermediate goods. On the BGP, it is determined by the parameter values and is given by a specific ratio.

Equation (33) decomposes the sources of per capita output growth on the BGP. It reveals that the growth in per capita output can be attributed to two factors: the quality improvements in production achieved by vertical R&D sectors and the growth rate of varieties of intermediate goods in horizontal R&D sectors. These factors contribute to the overall increase in output per capita.

3.1. Growth rate in the decentralized economy

Proposition 1. *The economic growth rates of the decentralized economy on the BGP can be expressed as follows:*

$$m=\frac{n}{\mu\cdot\varphi}\cdot\frac{(1+\xi)\cdot(\sigma-\beta)}{(1-\beta-\theta)\cdot(1-\gamma)+(1+\xi)\cdot[(\beta-1)\cdot(1-\gamma)+\sigma]}>0; \tag{34}$$

$$\frac{\dot{D_H}(t)}{D_H(t)}=\frac{\dot{D_V}(t)}{D_V(t)}=(\mu\cdot\varphi\cdot m-n)\cdot\frac{1+\xi}{\theta+(1-\beta)\cdot\xi}<0; \tag{35}$$

$$\frac{\dot{N}(t)}{N(t)}=(\mu\cdot\varphi\cdot m-n)\cdot\frac{1-\theta-\beta}{\theta+(1-\beta)\cdot\xi}>0; \tag{36}$$

$$l_H^*(t)=\frac{1}{1+\frac{\beta\cdot\theta}{(1-\beta)\cdot(1-\theta)}+\frac{[\mu\cdot m\cdot(\gamma\cdot\varphi+1)+\rho-n]\cdot[\theta+(1-\beta)\cdot\xi]}{(\mu\cdot\varphi\cdot m-n)\cdot(\theta+\beta-1)\cdot(1-\theta)\cdot(1-\alpha)}}; \tag{37}$$

Equation (34) represents the expression for m on the BGP, and it can be easily proven that$(\mu\cdot\varphi\cdot m-n)<0$ is tenable. Equation (35) indicates that in the decentralized economy, the per capita data provision of both types of R&D sectors will experience a downward trend, effectively alleviating the problem of consumer privacy disclosure. However, the aggregate data provision can still grow in the long term [similar to Cong et al. (2021)].

In equation (36), it is necessary to assume that $\theta+\beta-1>0$. The first partial derivatives of m with respect to key variables are as follows:

$$\frac{\partial m}{\partial\beta}=\frac{(1+\xi)\cdot\sigma\cdot(-1-2\cdot\xi+\gamma\cdot\xi)+(1-\gamma)\cdot(\theta+\xi)}{\{(1-\gamma)\cdot[(\beta-1)\cdot\xi-\theta]+\sigma\cdot(1+\xi)\}^2}<0, \tag{38}$$

$$\frac{\partial m}{\partial\sigma}=\frac{(1+\xi)\cdot[(1-\gamma)\cdot(\beta\cdot\xi-\theta-\xi)+\beta\cdot(1+\xi)]}{\{(1-\gamma)\cdot[(\beta-1)\cdot\xi-\theta]+\sigma\cdot(1+\xi)\}^2}>0; \tag{39}$$

Proposition 2. *The innovation success rate of vertical R&D sectors is influenced by various factors. Specifically, it has a positive relationship with the parameter representing the disutility of data leakage and misuse σ, the natural population growth rate n, and a negative relationship with the contribution of labor in vertical innovation β, and the elasticity of intertemporal substitution of consumption γ.*

Most of the findings from Proposition 1 are in line with expectations. However, one notable result is the negative effect of the contribution of labor in the process of vertical innovation. This implies that a higher contribution rate of data to the success of innovation leads to a higher success rate. It demonstrates that data integration into vertical innovation processes can reduce uncertainty and improve the likelihood of success. There are two main reasons why data utilization can enhance the success rate of innovation. First, according to equation (16), wages increase with the expansion of intermediate goods varieties and improvements in productivity parameters. Consequently, labor costs continue to rise. However, equation (7) shows that the price of data is dependent on per capita data usage. If data remains unused, the cost of utilizing it would be relatively low. Therefore, from a cost perspective, incorporating data can significantly reduce innovation costs, thereby increasing the monopoly profits of intermediate producers. Second, equation (14) reveals that monopoly profits are contingent on the labor used in final goods production. Consequently, if vertical R&D sectors replace some labor with data, there will be an increase in labor allocated to final goods production, resulting in improved monopoly profits.

Hence, when data utilized in the innovation process offers greater revenue incentives, the success rate of innovation is substantially enhanced.

In contrast to the conclusion drawn by Howitt and Aghion (1998), data differ fundamentally from physical capital. While physical capital incurs production costs that affect the interest rate and consequently impact the monopoly profits of intermediate producers, data, as a by-product of consumption, only entails the usage cost associated with privacy concerns and lacks production costs. Although both factors can enhance the success rate of innovation when incorporated into the innovation process, their mechanisms are distinct.

The innovation success rate is positively influenced by the parameter representing the disutility of data leakage and misuse. This suggests that in equilibrium, consumers require a higher success rate of innovation to compensate for the increased disutility resulting from heightened privacy concerns. Furthermore, the success rate of innovation is positively associated with the natural population growth rate. This implies that with a larger population, there is an increase in both the amount of labor involved in the innovation process and the overall availability of data for innovation purposes. Conversely, the success rate of innovation is negatively correlated with the elasticity of intertemporal substitution of consumption. This indicates that if individuals prefer substituting current consumption with future consumption, the amount of data accessible to vertical R&D sectors will be diminished, leading to a lower success rate of innovation.

4. Social planner's economy on the balanced growth path

Subsequently, we proceed to derive the growth rate of key variables allocated under socially optimal allocations on the BGP. This serves as a benchmark for comparing the outcomes with those observed in the decentralized economy, given that the equilibrium in the decentralized economy is not socially optimal. In a social planner's economy, the objective of the social planner is to maximize the utility of representative consumers, subject to resource constraints. This implies that the aggregate net output, denoted as $\tilde{Y}(t)$, equals the aggregate total output minus the total cost of intermediate goods and can be expressed as

$$\max_{x(v,t)} \tilde{Y}(t) = \left\{ L_E(t)^{\alpha} \cdot \int_0^{N(t)} A(v,t)^{\alpha} \cdot x(v,t)^{1-\alpha}\, dv - \int_0^{N(t)} \psi \cdot x(v,t)\, dv \right\}; \tag{40}$$

The social planner optimizes the aggregate net output by selecting the optimal amount of intermediate goods input, denoted as $x^*(v,t)$, which can be calculated as

$$x^*(v,t) = \left(\frac{1-\alpha}{\psi}\right)^{\frac{1}{\alpha}} \cdot L_E(t) \cdot A(v,t); \tag{41}$$

Consequently, the aggregate net output is given by

$$\tilde{Y}(t) = \alpha \cdot \left(\frac{\psi}{1-\alpha}\right)^{1-\frac{1}{\alpha}} \cdot L_E(t) \cdot \int_0^{N(t)} A(v,t)\, dv; \tag{42}$$

As is widely recognized, the aggregate consumption, $C(t) = c(t) \cdot L(t)$, is equivalent to the aggregate net output. Thus, the average consumption can be represented as

$$c(t) = \alpha \cdot \Delta \cdot l_E(t) \cdot \int_0^{N(t)} A(v,t)\, dv, \text{ where } \Delta = \left(\frac{\psi}{1-\alpha}\right)^{1-\frac{1}{\alpha}}; \tag{43}$$

In comparison to the aggregate net output in the decentralized economy (15), the aggregate net output in the social planner's economy is consistently higher when labor and technology are held at the same level. This disparity can be attributed to the monopoly power prevailing in the decentralized economy.

The optimization problem faced by the social planner can be described as follows:

$$\max_{c(t),D_H(t),D_V(t)} \int_0^\infty e^{-(\rho-n)\cdot t}\left[\frac{c\,(t)^{1-\gamma}-1}{1-\gamma}-\chi_H\cdot D_H\,(t)^\sigma-\chi_V\cdot D_V\,(t)^\sigma\right]dt; \tag{44}$$

Subject to the following constraints:

$$\dot{N}(t)=\frac{\eta\cdot N\,(t)^\xi\cdot l_H\,(t)^{1-\theta}\cdot D_H\,(t)^\theta\cdot L\,(t)}{A^{\max}\,(t)}; \tag{45}$$

$$\dot{A}^{\max}\,(t)=\frac{\mu\cdot\varphi\cdot l_V\,(t)^\beta\cdot D_H\,(t)^{1-\beta}\cdot L\,(t)}{N\,(t)}; \tag{46}$$

$$c\,(t)=\alpha\cdot\Delta\cdot l_E\,(t)\cdot\int_0^{N(t)} A\,(v,t)\,dv; \tag{47}$$

$$l_H\,(t)+l_V\,(t)+l_E\,(t)=1; \tag{48}$$

Here, equation (45) represents the horizontal innovation possibility frontier, equation (46) represents the vertical innovation possibility frontier, equation (47) represents the resource constraint, and equation (48) signifies labor market equilibrium. To solve this problem, a current-value Hamiltonian equation can be defined as

$$\begin{aligned} G=&\frac{c\,(t)^{1-\gamma}-1}{1-\gamma}-\chi_H\cdot D_H\,(t)^\sigma-\chi_V\cdot D_V\,(t)^\sigma\\ &+\kappa\,(t)\cdot\left[\alpha\cdot\Delta\cdot l_E\,(t)\cdot\int_0^{N(t)}A\,(v,t)\,dv-c\,(t)\right]+\omega\,(t)\cdot\frac{\eta\cdot N\,(t)^\xi\cdot l_H\,(t)^{1-\theta}\cdot D_H\,(t)^\theta\cdot L\,(t)}{A^{\max}\,(t)}\\ &+\varepsilon\,(t)\cdot\frac{\mu\cdot\varphi\cdot l_V\,(t)^\beta\cdot D_H\,(t)^{1-\beta}\cdot L\,(t)}{N\,(t)}, \end{aligned} \tag{49}$$

Here, $\kappa(t)$, $\omega(t)$, and $\varepsilon(t)$ represent the shadow prices corresponding to constraints (49), (51), and (52), respectively. The first-order conditions are derived with respect to $c(t), D_H(t), D_V(t), l_H(t), l_V(t), A^{\max}(t), N(t)$.

4.1. Growth rate in the social planner's economy

In the analysis of the social planner's economy on the BGP, the paper examines the relationship between key variables. The following relationships are established:

$$\frac{l_V\,(t)^\beta\cdot D_V\,(t)^{1-\beta}\cdot L\,(t)}{N\,(t)\cdot A^{\max}\,(t)}\to M; \tag{50}$$

$$\frac{\dot{D_H}\,(t)}{D_H\,(t)}=\frac{\dot{D_V}\,(t)}{D_V\,(t)}; \tag{51}$$

$$\frac{\dot{N}\,(t)}{N\,(t)}=\frac{1-\beta-\theta}{\xi}\cdot\frac{\dot{D_H}\,(t)}{D_H\,(t)}; \tag{52}$$

$$\frac{\dot{c}\,(t)}{c\,(t)}=\frac{\sigma}{1-\gamma}\cdot\frac{\dot{D_H}\,(t)}{D_H\,(t)}; \tag{53}$$

https://doi.org/10.1017/S1365100523000445 Published online by Cambridge University Press

Equation (50) indicates that on the BGP, the success rate of the vertical innovation approaches a specific value denoted as M. Equations (51) and (52) demonstrate that on the BGP, the growth rate of per capita data provision is the same for both types of R&D sectors, and it relates the growth rate of per capita data provision to the growth rate of varieties of intermediate goods. Equation (53) reveals that on the BGP, the growth rate of per capita consumption is determined by the growth rate of per capita data provision.

Proposition 3. *The economic growth rates of the social planner's economy on the BGP can be expressed as follows:*

$$M = \frac{n}{\mu \cdot \varphi} \cdot \frac{\sigma \cdot \xi}{\sigma \cdot \xi - (1-\gamma) \cdot [(1-\beta) \cdot (\xi - 1) + \theta]} > 0; \tag{54}$$

$$\frac{\dot{N}(t)}{N(t)} = n \cdot \frac{(1-\gamma) \cdot (1-\beta-\theta)}{\sigma \cdot \xi - (1-\gamma) \cdot [(1-\beta) \cdot (\xi - 1) + \theta]} > 0; \tag{55}$$

$$\frac{\dot{D_H}(t)}{D_H(t)} = \frac{\dot{D_V}(t)}{D_V(t)} = n \cdot \frac{(1-\gamma) \cdot \xi}{\sigma \cdot \xi - (1-\gamma) \cdot [(1-\beta) \cdot (\xi - 1) + \theta]} < 0; \tag{56}$$

$$l^*_{H,s}(t) = \frac{1-\theta}{\frac{(1+\beta)\cdot\theta}{(1-\beta)} + \frac{(1-\sigma-\beta)\cdot n \cdot (1-\gamma) \cdot \xi + \rho \cdot \Lambda}{n \cdot (1-\gamma) \cdot (1-\theta-\beta)} - (\theta + \xi)}, \tag{57}$$

where $\Lambda = \sigma \cdot \xi - (1-\gamma) \cdot [(1-\beta) \cdot (\xi - 1) + \theta]$.

According to equation (59), it is evident that $\frac{\partial M}{\partial \beta} < 0$ holds, indicating that in the social planner's economy, a higher success rate of innovation is achieved with a greater contribution of data in the vertical innovation process. This is primarily because fewer labor resources are employed in the R&D sectors, allowing for more labor in the final goods production sector and a looser resource constraint, as indicated by equation (14). On the other hand, equation (56) demonstrates that in the social planner's economy, the per capita data provision in both types of R&D sectors continues to decline, indicating that the issue of consumer privacy disclosure does not deteriorate over time.

4.2. Misallocation in the decentralized economy

In this section, we examine the misallocation of resources in a decentralized economy. By using standard values from the existing literature, with $n = 0.02, \gamma = 2.5, \sigma = 3, \theta = 0.5$, and $\xi = 0.85$, as well as reasonable discretionary values, Figure 1 illustrates the difference between the success rate of vertical innovation in two cases, as per equations (35) and (54). The observed difference varies with the contribution of labor in the process of vertical innovation.

Similarly, utilizing $n = 0.02, \gamma = 2.5$ (standard values from existing literature), $\sigma = 3, \beta = 0.7, \xi = 0.85$ (reasonable discretionary values), Figure 2 displays the difference between the growth rate of varieties of intermediate goods in the two cases, according to equations (36) and (55). This observed difference varies with the contribution of data in the process of horizontal innovation.

As depicted in Figures 1 and 2, both the success rate of vertical innovation and the growth rate of varieties of intermediate goods in the social planner's economy surpass those in the decentralized economy. Intuitively, due to the inherent competition and lack of coordination between the two types of research and development (R&D) sectors, each sector independently maximizes its own profits in the decentralized economy, resulting in negative externalities between them. Conversely, in the social planner's economy, the R&D intensity of both sectors is determined by the maximization of total utility, effectively mitigating the negative externalities.

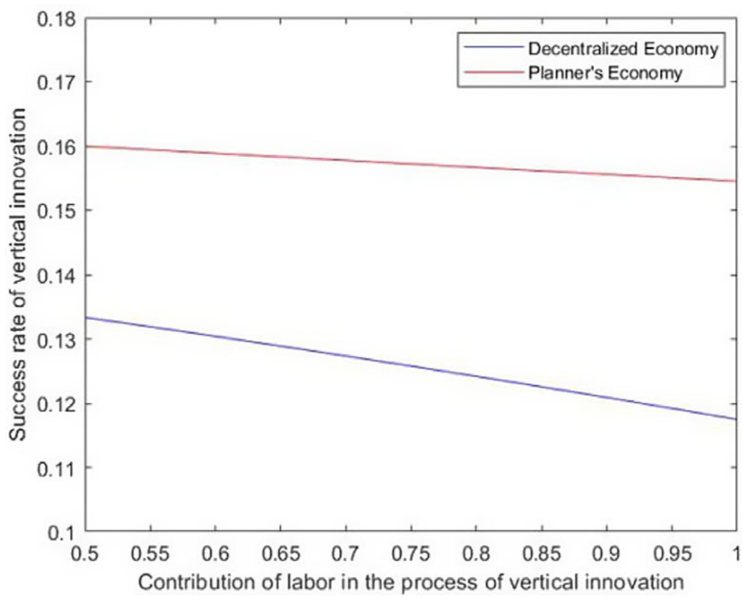

Figure 1. Success rate of the vertical innovation in both types of economies.

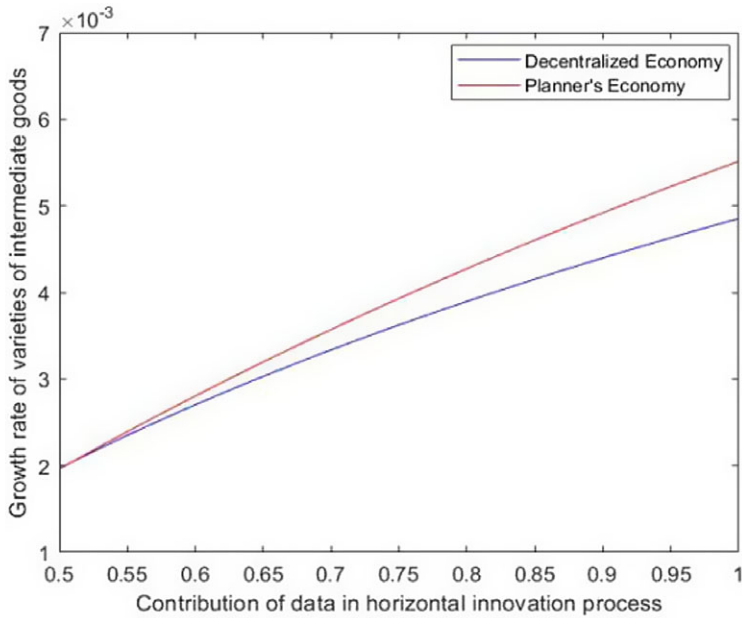

Figure 2. Growth rate of varieties of intermediate goods in both types of economies.

To further investigate the misallocation within the decentralized economy, this paper examines the differences in the ratio of labor employed in R&D sectors and the growth rate of data provision. By employing standard values ($n = 0.02, \gamma = 2.5, \rho = 0.03, \alpha = \frac{2}{3}, \beta = 0.7, \theta = 0.5, \xi = 0.85, \mu = 0.1$, and $\varphi = 10$) from the existing literature, Figure 3 presents the difference between the ratios of labor employed in horizontal R&D sectors in the two cases, as per equations (37) and (57). The observed difference varies with the disutility effect of data. Notably, since the ratio of labor

https://doi.org/10.1017/S1365100523000445 Published online by Cambridge University Press

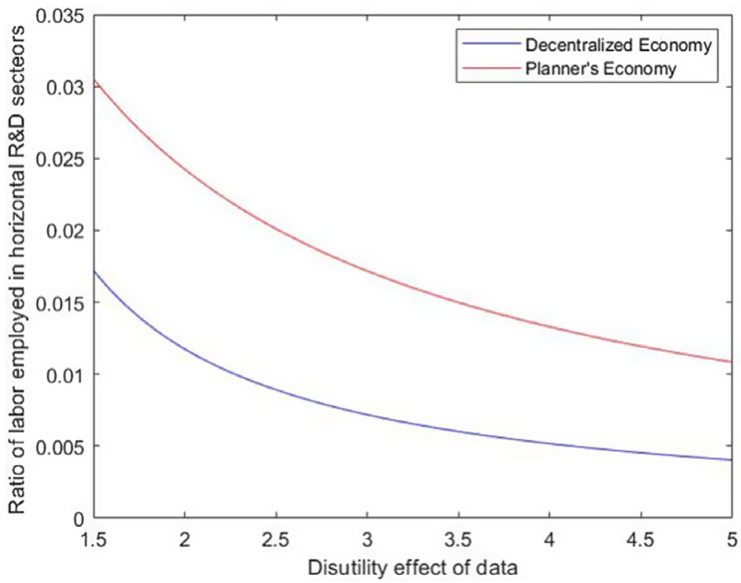

Figure 3. Ratio of labor employed in horizontal R&D sectors in both types of economies.

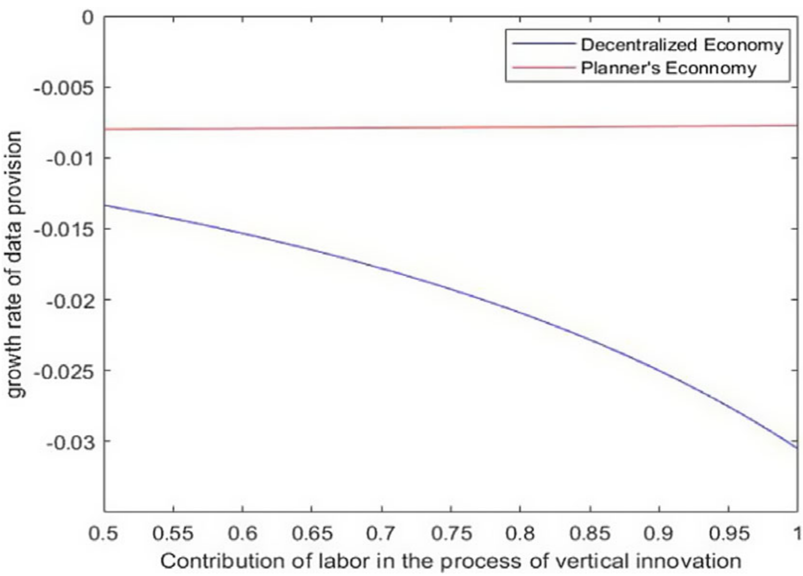

Figure 4. Growth ratio of data provision in R&D sectors in both types of economies.

employed in vertical R&D sectors is the same proportion in both economies on the BGP, our analysis focuses solely on the ratio of labor employed in horizontal R&D sectors.

With $n = 0.02, \gamma = 2.5, \rho = 0.03, \alpha = 0.67$ taking on standard values from the existing literature, as well as $\sigma = 3, \theta = 0.5, \xi = 0.85, \mu = 0.1, \varphi = 10$ as reasonable discretionary values, Figure 4 is presented to analyze the difference in the growth rate of data provision employed in R&D sectors between the two cases, as per equations (35) and (56). The observed difference varies with the contribution of labor in the vertical innovation.

https://doi.org/10.1017/S1365100523000445 Published online by Cambridge University Press

Notably, Figure 3 illustrates that on the BGP, the ratio of labor employed in horizontal R&D sectors in the decentralized economy is consistently lower than that in the social planner's economy, regardless of the disutility effect of data. This outcome arises due to the high prices set by monopoly intermediate producers for their intermediate goods, which prompts final goods producers to substitute labor for some intermediate goods. Consequently, there is an undersupply of labor employed in R&D sectors, aligning with the findings of Jones (1995).

Furthermore, Figure 4 demonstrates that on the BGP, the growth ratio of data provision in R&D sectors is consistently lower in the decentralized economy compared to the social planner's economy, regardless of the contribution of labor in the process of vertical innovation. This discrepancy primarily arises because, in the decentralized economy, R&D sectors need to purchase data using financial resources instead of receiving direct allocations for innovation. Moreover, according to equations (6) and (7), the price of data is positively influenced by data provision and per capita consumption, indicating that the price of data introduces distortions in data provision. Consequently, on the BGP, both types of R&D sectors employ less labor and data in the decentralized economy. There are two underlying reasons for this, which imply that the R&D intensity in the decentralized economy is lower compared to the planned economy.

5. Data ownership: innovators versus consumers

In the benchmark model, consumers possess ownership of their data, allowing them to enhance their utility by selling the data. However, per capita data usage gradually declines on the BGP due to privacy concerns among consumers. In this chapter, we examine the growth rate of the digital economy and data usage on the BGP when innovators assume ownership of data.

When innovators own the data, they no longer need to compensate consumers. To determine the amount of data usage, we introduce a data processing cost for innovators. Consequently, the profit maximization problem for the vertical R&D sector becomes:

$$\max_{D_V(t),l_V(t)} \left\{\Phi\left(t\right)\cdot V\left(t\right)-w\left(t\right)\cdot L_V\left(t\right)-\left[D_V\left(t\right)\cdot L\left(t\right)\right]^{\Gamma 1}\right\}; \tag{58}$$

Likewise, the profit maximization problem for the horizontal R&D sector becomes:

$$\max_{l_H(t),D_H(t)} \left\{E\left[\frac{A\left(v,t\right)}{A^{\max}\left(t\right)}\right]\cdot V\left(t\right)\cdot \dot{N}\left(t\right)-w\left(t\right)\cdot L_H\left(t\right)-\left[D_H\left(t\right)\cdot L\left(t\right)\right]^{\Gamma 2}\right\}; \tag{59}$$

Here, $\Gamma 1$ parameterizes the processing cost of data for the vertical sectors, which depends on the quantity of data usage. Similarly, $\Gamma 2$ parameterizes the processing cost of data for the horizontal sectors, also contingent on the quantity of data usage.

Consequently, equations (22) and (27) are modified as follows[9]:

$$\frac{(1-\beta)\cdot\mu\cdot l_V\left(t\right)^{\beta}\cdot D_V\left(t\right)^{1-\beta-\Gamma 1}\cdot\tilde{\pi}\left(t\right)\cdot L(t)^{1-\Gamma 1}\cdot\left[r\left(t\right)-n\right]}{\Gamma 1\cdot\left[r\left(t\right)+\Phi\left(t\right)-n\right]^2\cdot N\left(t\right)\cdot A^{\max}\left(t\right)}=1; \tag{60}$$

$$\eta\cdot\theta\cdot N\left(t\right)^{\xi}\cdot l_H\left(t\right)^{1-\theta}\cdot D_H\left(t\right)^{\theta-\Gamma 2}\cdot V\left(t\right)\cdot L(t)^{1-\Gamma 2}\cdot E\left[\frac{A\left(v,t\right)}{A^{\max}\left(t\right)}\right]=\Gamma 2\cdot A^{\max}\left(t\right); \tag{61}$$

Moreover, the consumers' budget constraint (2) becomes:

$$\dot{a}\left(t\right)=\left(r\left(t\right)-n\right)\cdot a\left(t\right)+w\left(t\right)-c\left(t\right); \tag{62}$$

Combining equations (21), (26), (28), (60), and (61), we derive the following results:

$$\frac{\dot{N}\left(t\right)}{N\left(t\right)}=\frac{n\cdot(\Gamma 2-\theta)\cdot(\Gamma 1-2)\cdot(1-\beta)}{1-\beta+(1+\xi)\cdot(\beta+\Gamma 1-1)}>0; \tag{63}$$

https://doi.org/10.1017/S1365100523000445 Published online by Cambridge University Press

$$\frac{\dot{A^{\max}}(t)}{A^{\max}(t)} = \frac{-n \cdot [\xi \cdot \Gamma 1 \cdot \Gamma 2 \cdot (2 - \Gamma 1) + \Gamma 1 (\Gamma 2 - \theta) - \Gamma 2 \cdot \xi \cdot (1 - \Gamma 1) \cdot (\beta + \Gamma 1 - 1)]}{1 - \beta + (1 + \xi) \cdot (\beta + \Gamma 1 - 1)} > 0; \tag{64}$$

$$\frac{\dot{D_H}(t)}{D_H(t)} = \frac{n \cdot \xi \cdot (1 - \beta) \cdot (2 - \Gamma 2)}{1 - \beta + (1 + \xi) \cdot (\beta + \Gamma 1 - 1)} > 0; \tag{65}$$

$$\frac{\dot{D_V}(t)}{D_V(t)} = \frac{n \cdot (2 - \Gamma 1) \cdot [\Gamma 2 \cdot (1 + \xi) - \theta]}{1 - \beta + (1 + \xi) \cdot (\beta + \Gamma 1 - 1)} > 0; \tag{66}$$

Consequently, if innovators have ownership of data, both of them will use more data to maximize their profits on the BGP, irrespective of consumers' privacy concerns. This finding deviates from the conclusion of the baseline model. Even if intermediate goods (digital knowledge) exhibit dynamic nonrivalry and the data processing cost is convex, the quantity of data usage continues to increase. Intuitively, as the marginal contribution of data to innovation diminishes, innovators need to utilize more data to ensure maximum profit. Thus, it becomes evident that only when data belongs to consumers, can privacy concerns be adequately addressed, preventing them from becoming increasingly severe in the long run.[10]

6. The role of human capital

In the previous discussion, we identified a significant issue in the development of the digital economy, namely the conflicting relationship between economic growth and the quantity of data usage. However, in the era of the digital economy, the rapid advancement of digital technology has had a positive impact on the dissemination and preservation of digital knowledge [Xie and Yang (2022)]. This, in turn, generates a notable positive externality effect on the accumulation of human capital, offering a potential solution to address both economic growth and privacy concerns simultaneously.

Drawing from the work of Lucas (1988) and Wu and Zhang (2022), the cumulative growth equation for human capital for each worker can be represented as follows:

$$\dot{h}(t) = h(t)^{\delta} \cdot N(t)^{\lambda}; \tag{67}$$

Here, $\delta \in (0, 1)$ represents the coefficient of workers' self-learning ability. It signifies that the agents with stronger self-learning abilities experience a faster cumulative growth rate of human capital, assuming they spend the same amount of time on learning. On the other hand, $\lambda \in (0, 1)$ denotes the coefficient of digital knowledge in promoting the accumulation of human capital. It represents the quality of local digital infrastructure and digital technology, indicating that better digital infrastructure and technology contribute to a faster cumulative growth rate of human capital under the condition of the same learning time. Clearly, the growth of human capital is influenced by both workers' self-learning ability and the digital knowledge generated by the horizontal R&D sectors.

In order to examine the role of human capital in R&D, we now consider a disaggregated model based on the benchmark model. The consumer's budget constraint (2) is modified as follows:

$$\dot{a}(t) = (r(t) - n) \cdot a(t) + w(t) \cdot h(t) + p_{D,H}(t) \cdot D_H(t) + p_{D,V}(t) \cdot D_V(t) - c(t); \tag{68}$$

Additionally, the production function (8) of the representative final goods producer becomes:

$$Y(t) = \left[u(t) \cdot h(t) \cdot L_E(t)\right]^{\alpha} \cdot \int_0^{N(t)} A(v,t)^{\alpha} \cdot x(v,t)^{1-\alpha}\, dv; \tag{69}$$

https://doi.org/10.1017/S1365100523000445 Published online by Cambridge University Press

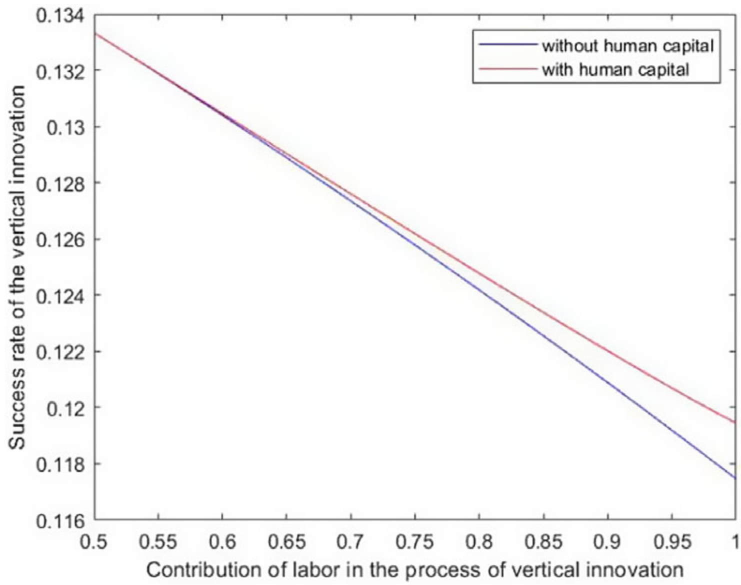

Figure 5. Success rate of the vertical innovation with (without) human capital.

The profit optimization problem (20) of the vertical R&D sectors is modified as

$$\max_{D_V(t),l_V(t)} \left\{\phi\,(\mathrm{t}) \cdot V\,(t) - h\,(t) \cdot w\,(t) \cdot L_V\,(t) - p_{D,V}\,(t) \cdot D_V\,(t) \cdot L\,(t)\right\}; \tag{70}$$

Similarly, the profit optimization problem (25) of the horizontal R&D sectors becomes:

$$\max_{l_H(t),D_H(t)} \left\{\begin{array}{c} E\left[\dfrac{A\,(v,t)}{A^{\max}\,(t)}\right] \cdot V\,(t) \cdot \dot{N}\,(t) - h\,(t) \cdot w\,(t) \cdot L_H\,(t) \\ -p_{D,H}\,(t) \cdot D_H\,(t) \cdot L\,(t) \end{array}\right\}; \tag{71}$$

By following the same optimization process as in the baseline model, we can derive the growth rates of various variables under BGP. The growth rate of the success rate of vertical innovation in the two cases can be obtained as

$$m' = \frac{n}{\mu \cdot \varphi} \cdot \frac{(1-\delta) \cdot (1+\xi) \cdot (\sigma - \beta) + \lambda \cdot (\beta + \theta - 1) \cdot (1 + \beta - \gamma - \sigma)}{Q+K}, \tag{72}$$

where $Q = (\beta + \theta - 1)[\lambda(2 - \sigma - 2\gamma) - (1-\delta)(1-\gamma)]$,
and $K = (1-\delta)(1+\xi)[(\beta - 1)(1-\gamma) + \sigma]$;

$$\frac{\dot{D_H}\,(t)}{D_H\,(t)} = \frac{\dot{D_V}\,(t)}{D_V\,(t)} = \frac{(\mu \cdot \varphi \cdot m - n) \cdot [\lambda \cdot (\beta + \theta - 1) - (1+\xi) \cdot (1-\delta)]}{(\lambda + \delta - 1) \cdot (\beta + \theta - 1) - (1+\xi) \cdot (1-\beta) \cdot (1-\delta)}; \tag{73}$$

$$\frac{\dot{N}\,(t)}{N\,(t)} = \frac{(\mu \cdot \varphi \cdot m - n) \cdot [(\beta + \theta - 1) \cdot (1-\delta)]}{(\lambda + \delta - 1) \cdot (\beta + \theta - 1) - (1+\xi) \cdot (1-\beta) \cdot (1-\delta)}; \tag{74}$$

Using the reasonable values ($n = 0.02, \gamma = 2.5, \sigma = 3, \theta = 0.5, \xi = 0.85, \lambda = 0.3, \delta = 0.8$) from existing literature, Figure 5 illustrates the difference between the success rates of vertical innovation in the two cases according to equations (34) and (72). The difference varies with the contribution of labor in the process of vertical innovation.

Clearly, when considering human capital, the success rate of vertical innovation is higher compared to the case without human capital, given the same contribution of labor in the innovation

https://doi.org/10.1017/S1365100523000445 Published online by Cambridge University Press

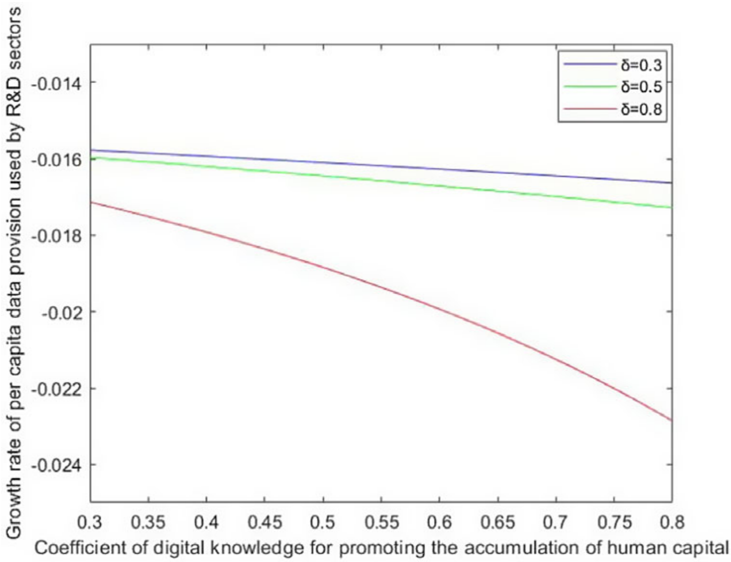

Figure 6. Relationship between human capital's parameters and the growth rate of per capita data provision.

process. Furthermore, even with the accumulation of human capital, the data used in the innovation process can still enhance the success rate of innovation. Unlike the conclusion drawn by Zeng (2003), we highlight that the accumulation of human capital relies on the positive externality of digital knowledge rather than physical capital.

Digital knowledge plays a crucial role in promoting human capital due to its vertical nonrivalry, which means that its utilization in the production process and data in the innovation process does not lead to their depletion. Thus, data can still improve the success rate of vertical innovation even when human capital is accumulated.

The first partial derivative of m with respect to λ (the coefficient of digital knowledge) is

$$\frac{\partial m}{\partial \lambda} = \frac{n}{\mu \cdot \varphi} \cdot \left\{ \begin{array}{c} \dfrac{-(\beta+\theta-1)^2(1+\beta-\gamma-\sigma)(1-\delta)(1-\gamma)}{(Q+K)^2} \\ +\dfrac{K-(\beta+\theta-1)(2-\sigma-2\gamma)(1-\delta)(1+\xi)(\sigma-\beta)}{(Q+K)^2} \end{array} \right\} > 0; \tag{75}$$

This derivative indicates that the success rate of vertical innovation is positively influenced by the coefficient of digital knowledge for promoting the accumulation of human capital. Therefore, a better digital infrastructure and digital technology will result in a higher success rate of vertical innovation.

Using the reasonable values ($n = 0.02, \gamma = 2.5, \sigma = 3, \theta = 0.5, \xi = 0.85, \lambda = 0.3, \delta = 0.8$) from existing literature, Figure 6 illustrates how parameters related to human capital affect the growth rate of per capita data provision according to equation (75).

As depicted in Figure 6, the growth rate of per capita data provision is negatively influenced by the coefficient of digital knowledge for promoting the accumulation of human capital. Essentially, the promotion of digital knowledge on human capital can be seen as another way of utilizing past data. This implies that the more effectively historical data are utilized, the higher the labor productivity, and the lower the current demand for data. Consequently, this approach effectively mitigates the issues of data leakage and misuse.

https://doi.org/10.1017/S1365100523000445 Published online by Cambridge University Press

Table 1. Parameters for studying transitional dynamics

Variable	Meaning	Value	Source
α	Contribution of labor in final good production	2/3	Antras (2004)
γ	Reciprocal of elasticity of intertemporal substitution	2.5	Vissing-Jørgensen (2002)
ρ	Subjective discount factor	0.025	Krupka and Stephens (2013)
ξ	Spillover effect of knowledge in horizontal innovation	0.85	Cong et al. (2021)
θ	Contribution of data in horizontal innovation	0.5	Cong et al. (2021)
β	Contribution of labor in vertical innovation	0.7	Discretionary
n	Population growth rate	0.02	Jones (1999)
η	Efficiency term in horizontal innovation	1	Jones (1995)
σ	Severity of consumers' privacy concern	1.8	Cong et al. (2021)
μ	Productivity of vertical innovation	10	Caballero and Jaffe (1993)
φ	Marginal impact of vertical innovation on the stock of public knowledge	0.1	Caballero and Jaffe (1993)

7. Transitional dynamics: numerical analysis

In the preceding sections, the analysis revealed that the development of the digital economy leads to a decline in data usage in two types of R&D sectors, indicating a gradual easing of concerns regarding data privacy as regulations and digital infrastructure become less restrictive. This section focuses on examining the transitional dynamics of the digital economy from the initial state to the BGP. Additionally, considering the limited provision of data for vertical and horizontal innovation, we explore the transitional dynamics of the digital economy in contrast to the benchmark model.

7.1. Methodology and calibration

To facilitate tractability, this paper employs a numerical method based on the social planner's economy, following a similar approach to Jones (2016). Instead of undertaking formal calibration to replicate the circumstances of a specific country, this analysis aims to demonstrate the basic transitional dynamics achievable within the theoretical framework.[11] By deriving the first-order conditions with respect to variables such as $c(t), D_H(t), D_V(t), l_H(t), l_V(t), A^{\max}(t), N(t)$ from equation (49) in Section 4, the dynamical systems can be expressed as a set of differential equations.[12] Since $l_H(t), l_V(t), A^{\max}(t), N(t)$ are all state variables, this paper sets their initial values to a sufficiently small ϵ to drive economic growth.[13]

Table 1 provides a summary of the parametrization choices made in this study. Based on estimates from Antras (2004), the elasticity of substitution falls within the range of approximately 0.6–0.8. Therefore, we set the contribution of labor in final goods production α to 2/3. Considering estimates from Vissing-Jørgensen (2002), which suggest that the elasticity of intertemporal substitution (EIS) ranges from 0.3 to 0.4, we select the reciprocal of EIS γ as 2.5. Drawing from estimates in Krupka and Stephens (2013), the subjective discount factor ρ is set at 0.025. Population growth is estimated at around 1% per year according to Jones and Tonetti (2020), and the growth rate of R&D labor in developed economies is approximately 4%. Hence, we choose the value n to be 0.02. Considering estimates from Caballero and Jaffe (1993), the productivity of vertical innovation μ is chosen as 10 and the marginal impact of vertical innovation on the stock of public knowledge φ as 0.1.

Because they could not be accurately estimated, the following parameter values are selected from Jones (1995) and Cong et al. (2021): the contribution of data in horizontal innovation θ as 0.5, the spillover effect of knowledge in horizontal innovation ξ as 0.85, the efficiency term in

https://doi.org/10.1017/S1365100523000445 Published online by Cambridge University Press

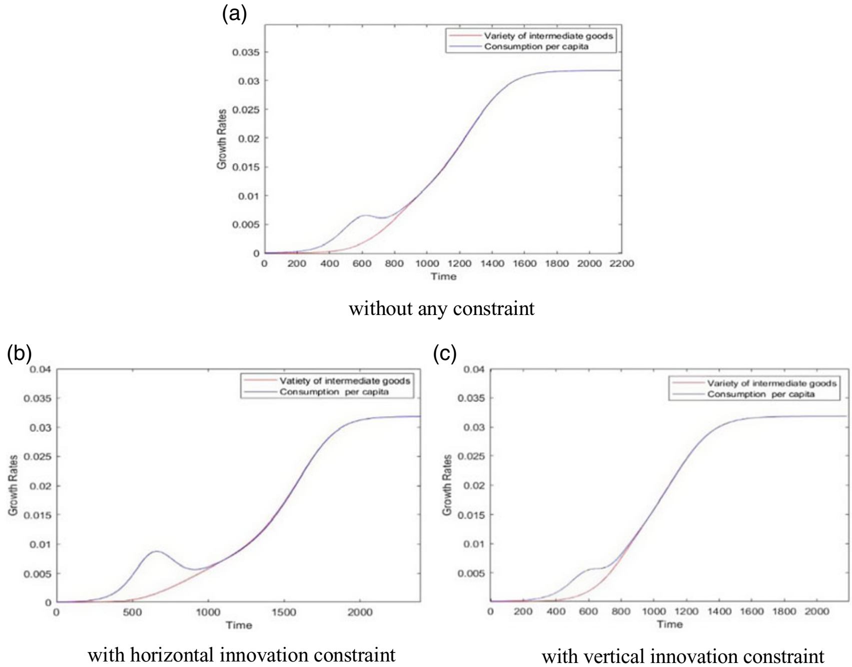

Figure 7. Growth rates of consumption and intermediate goods varieties. (a) Without any constraint. (b) With horizontal innovation constraint. (c) With vertical innovation constraint.

horizontal innovation η as 1, and the severity of consumers' privacy concern σ as 1.8. To meet the requirement $\theta + \beta - 1 > 0$, the contribution of labor in vertical innovation β is set at 0.7.

In the forthcoming figures, we examine the potential impact on the economy when the quantity of data available to either of the R&D sectors is limited.[14] To incorporate this constraint, we reformulate equations (3) and (4) as equations (76) and (77), respectively. To enforce the constraint on data provision for each R&D sector, we set the values of s_1 or s_2 as 0.01. This ensures that the provision of data used by the respective sectors adheres to the constraint. Conversely, if we do not intend to impose a constraint on the data provision used by the R&D sectors, we set the values of s_1 and s_2 as 100, rendering the constraint (3) or (4) ineffective.

$$\frac{\dot{D_H}(t)}{D_H(t)} \le \frac{\dot{c}(t)}{c(t)} + s_1; \tag{76}$$

$$\frac{\dot{D_V}(t)}{D_V(t)} \le \frac{\dot{c}(t)}{c(t)} + s_2; \tag{77}$$

Equation (76) states that the growth rate of $\frac{\dot{D_H}(t)}{D_H(t)}$ should not exceed the growth rate of $\frac{\dot{c}(t)}{c(t)} + s_1$, while equation (77) imposes a similar constraint for $\frac{\dot{D_V}(t)}{D_V(t)}$ compared to $\frac{\dot{c}(t)}{c(t)} + s_2$.

7.2. Results and discussions

In Panel (a) of Figure 7, the paper reveals that in the absence of constraints, different types of innovations play distinct roles during different periods of economic growth. In the early stage, vertical innovation assumes greater significance due to the limited availability of intermediate

https://doi.org/10.1017/S1365100523000445 Published online by Cambridge University Press

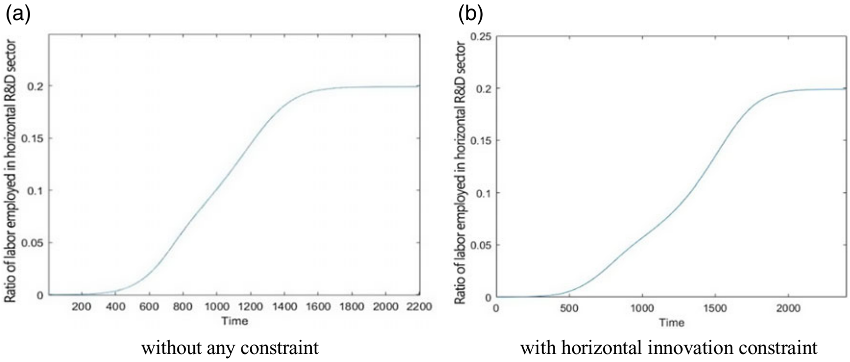

Figure 8. Rate of labor employed in horizontal R&D sector. (a) Without any constraint. (b) With horizontal innovation constraint.

goods varieties. However, as the number of intermediate goods varieties gradually increases, horizontal innovation becomes more crucial in the later stage. In comparison to the findings of Cong et al. (2021), by placing greater emphasis on vertical innovation, this economy achieves a higher economic growth rate along the transitional path in the long run.

Panel (b) of Figure 7, as per equation (6), illustrates the performance of the two types of innovations during different stages of economic growth when a constraint is imposed on horizontal innovation. The constraint implies insufficient data usage in the horizontal R&D sectors. It is evident that there is more pronounced fluctuation during the early period of economic development. Moreover, an economy constrained by horizontal innovation requires additional time to reach the BGP. Intuitively, when the data usage in horizontal R&D sectors is inadequate, the rate at which labor moves away from final goods production slows down. This can be observed in both Panels (a) and (b) of Figure 8, similar to the findings of Cong et al. (2021). The insufficient data provision and labor in horizontal R&D sectors result in a lack of driving force for horizontal innovation. Consequently, vertical innovation assumes a more prominent role in the later stage. However, as per capita consumption increases and the constraint on data usage becomes more significant, the speed of intermediate goods variety expansion accelerates. The impact of vertical innovation is subject to a certain spillover effect, leading to greater fluctuations.

In Panel (c) of Figure 7, it is observed that the economy with a constraint on vertical innovation reaches the BGP in approximately the same period as the economy without any constraints. Intuitively, due to the spillover effect of the expansion of intermediate product varieties, vertical innovation has minimal influence on the timing of reaching the BGP. Although the economy performs smoothly without significant fluctuations in this scenario, the growth rate of per capita consumption along the transitional path is lower compared to Panel (a) of Figure 7.

For further analysis, Figure 9 is examined. Both images demonstrate that the growth rate of the leading-edge productivity parameter initially increases to a maximum and then stabilizes. This suggests that when there are relatively few varieties of intermediate goods, vertical innovation is more easily realized. Additionally, Figure 10 shows an increasing number of laborers swiftly transitioning from final goods production to vertical R&D sectors.

However, as the scale of intermediate goods gradually expands, vertical innovation becomes more complex and plays a less significant role. Consequently, the growth rate of the leading-edge productivity parameter gradually slows down, with some laborers in vertical R&D sectors transitioning to other sectors. Furthermore, an economy with limited data usage requires additional time to reach the maximum value of the growth rate of the leading-edge productivity parameter, influenced by the stronger spillover effect. This is why such an economy experiences smoother progress along the traditional path with minimal fluctuation.

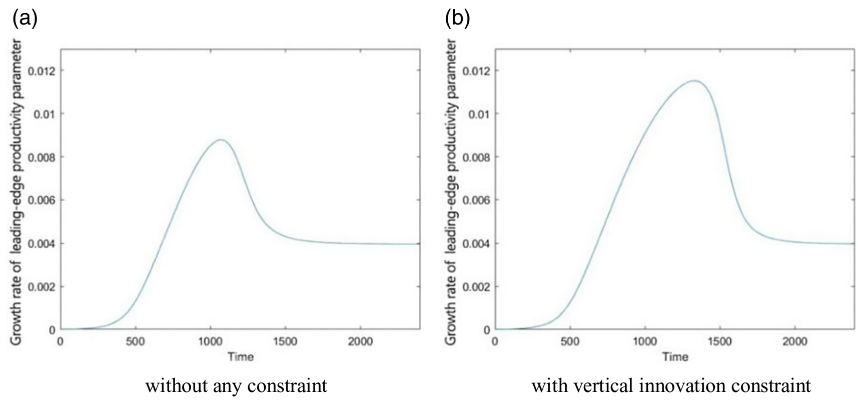

Figure 9. Growth rate of leading-edge productivity parameter. (a) Without any constraint. (b) With vertical innovation constraint.

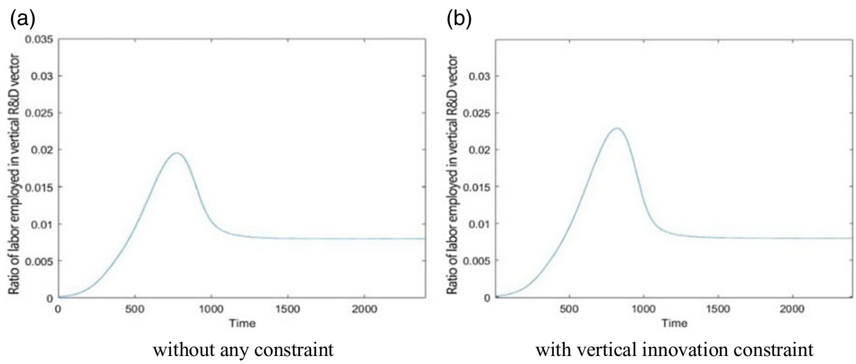

Figure 10. Rate of labor employed in vertical R&D sector. (a) Without any constraint. (b) With vertical innovation constraint.

In summary, when the data provision in horizontal R&D sectors is limited, the economy exhibits more severe fluctuations along the transitional path, resulting in increased uncertainty during the development process and requiring additional time to reach the BGP. If the data usage in vertical R&D sectors is constrained, the economy experiences a lower growth rate without significant fluctuations, indicating that the welfare level of the local population is relatively lower in the early stages of development.

8. Conclusion

Based on the conclusions drawn from this study, several policy recommendations are proposed. First, to expedite the transformation of the economy from being driven by factors to being driven by innovation, it is crucial to fully leverage the existing digital infrastructure and relax regulations pertaining to privacy issues. This will encourage the application of data in scientific and technological innovation, ultimately increasing the success rate of innovation. Second, taking advantage of the relatively advanced digital infrastructure, it is important to harness the spillover effect of digital knowledge. With continuous upgrades in digital technology and improved digital infrastructure, digital knowledge can facilitate economic growth while alleviating privacy concerns, by the accumulation of human capital. Lastly, in order to mitigate economic fluctuations and maintain higher growth rates in the long-term development process, it is essential to further enhance digital infrastructure, refine relevant laws, and establish appropriate national privacy supervision policies. These measures will ensure the optimal utilization of both vertical and horizontal R&D sectors.

Notes

1 The data discussed in this paper specifically pertain to economic activities that raise privacy concerns. Jones and Tonetti (2020) mentioned that every agent will generate data when engaging in the consumption activities. Although small amounts of data may seem like useless information, professionals can generate enormous value by processing large amounts of data. Besides, data are nonrival, which means that existing data can be used by any number of firms or people simultaneously without being diminished.
2 Aghion and Howitt (1992) mentioned that creative destruction, as a type of vertical innovation, constitutes the source of economic growth. Commonly, creative destruction and vertical innovation are interchangeable terms in the literature. Jones and Tonetti (2020) and Cong et al. (2021) discovered that the more incumbent firms sell their own data, the greater the likelihood of them being replaced by potential entrants.
3 It is intuitive that when firms or R&D sectors possess sufficient data, they can enhance product quality by modifying supplier management processes, establishing databases, and internally analyzing accumulated data.
4 To simplify calculation, this paper assumes that the marginal disutility of selling data to two different R&D sectors is the same. In order to distinguish them, the different weights are given to sell data to different sectors. Considering that we could distinguish clearly between the trade-offs of data in two kinds of innovation, with avoid their interplay, the disutility equation adopts a linearly separable form.
5 Unlike Jones (1995), intermediate goods not only have spillover effect on horizontal innovation but also have an inhibition effect vertical innovation.
6 According to Cong et al. (2021), if data processing cost is considered, the analysis of data transactions will be disrupted. Therefore, we will ignore this cost in the benchmark model.
7 More detail can be acquired from the appendix of Howitt's paper (1999).
9 Obviously, $\Gamma 1$ and $\Gamma 2$ should be greater than 2. On one hand, $N(t)$ and $A^{\max}(t)$ need to ensure positive growth rates on the BGP, and on the other hand, the sufficient convexity of processing cost is quite reasonable.
10 Cong et al. (2021) mentioned that data create knowledge spillovers to future periods by creating new intermediate goods, which is namely dynamic nonrivalry.
11 Due to the inability to precisely calibrate some parameters, this article will use the parameters from previous articles for numerical simulation. On the one hand, it can help us better understand the evolution process of this economy, and on the other hand, it can be compared with the results of previous articles.
12 Using these differential equations in Section 4, we could obtain the dynamical system of all variables, which could help us understand the transition dynamics of each variable from initial state to steady state.
13 Here, ϵ takes the value of 0.0001. Because it is small enough, we could approximate that the growth rate of digital economy starts at zero.
14 Since this is a numerical simulation, the time period does not have realistic meaning, such as one year or one month. In reality, data have been a key factor in business (or product) innovation for decades, but it has not been fully utilized yet. There are grounds to believe that the digital economy in a period of transition from an initial state to a steady state. Therefore, we could analyze the impact of data restrictions on economic growth by comparing the time required for the economy to enter a steady state under different circumstances.

References

Acemoglu, D., A. Makhdoumi, A. Malekian and A. Ozdaglar. (2019) Too Much Data: Prices and Inefficiencies in Data Markets. NBER Working Paper No. w26296.

Admati, A. R. and P. Pfleiderer. (1990) Direct and indirect sale of information. *Econometrica* 58(4), 901–928.

Aghion, P. and P. W. Howitt. (1992) A model of growth through creative destruction. *Econometrica* 60(2), 323–351.

Akçura, M. T. and K. Srinivasan. (2005) Research note: Customer intimacy and cross-selling strategy. *Management Science* 51(6), 1007–1012.

Antras, P. (2004) Is the US aggregate production function Cobb-Douglas? New estimates of the elasticity of substitution. *The BE Journal of Macroeconomics* 4(1), 1–36.

Caballero, R. J. and A. B. Jaffe. (1993) How high are the giants' shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. *NBER Macroeconomics Annual* 8, 15–74.

Casadesus-Masanell, R. and A. Hervas-Drane. (2015) Competing with privacy. *Management Science* 61(1), 229–246.

Choi, J., D. Jeon and B. Kim. (2019) Privacy and personal data collection with information externalities. *Journal of Public Economics* 173, 113–124.
Cong, L. W., D. Xie and L. Zhang. (2021) Knowledge accumulation, privacy, and growth in a data economy. *Management Science* 67(10), 6480–6492.
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.
Fainmesser, I. P., A. Galeotti and R. Momot. (2022) Digital privacy. *Management Science* 69(6), 3157–3173.
Farboodi, M. and L. Veldkamp. (2021) A Growth Model of the Data Economy. NBER Working Paper w28427.
Hirshleifer, J. (1971) The private and social value of information and the reward to inventive activity. *American Economic Review* 61(4), 561–574.
Howitt, P. (1999) Steady endogenous growth with population and R&D inputs growing. *Journal of Political Economy* 107(4), 715–730.
Howitt, P. W. and P. Aghion. (1998) Capital accumulation and innovation as complementary factors in long-run growth. *Journal of Economic Growth* 3(2), 111–130.
Ichihashi, S. (2020) Online privacy and information disclosure by consumers. *American Economic Review* 110(2), 569–595.
Ichihashi, S. (2021a) Competing data intermediaries. *The RAND Journal of Economics* 52(3), 515–537.
Ichihashi, S. (2021b) The economics of data externalities. *Journal of Economic Theory* 196, 105316.
Jones, C. I. (1995) R&D-based models of economic growth. *Journal of Political Economy* 103(4), 759–784.
Jones, C. I. (1999) Growth: With or without scale effects? *American economic review* 89(2), 139–144.
Jones, C. I. (2016) Life and growth. *Journal of Political Economy* 124(2), 539–578.
Jones, C. I. and C. Tonetti. (2020) Nonrivalry and the economics of data. *American Economic Review* 110(9), 2819–2858.
Krupka, E. L. and M. Stephens Jr.. (2013) The stability of measured time preferences. *Journal of Economic Behavior & Organization* 85, 11–19.
Lucas, R. E. (1988) On the mechanics of economic development. *Journal of Monetary Economics* 22(1), 3–42.
Murphy, R. S. (1996) Property rights in personal information: An economic defense of privacy. *Georgetown Law Journal* 84(7), 2381–2417.
Romer, P. M. (1990) Endogenous technological change. *Journal of Political Economy* 98(5, Part 2), S71–S102.
Sun, T., Z. Yuan, C. Li, K. Zhang and J. Xu. (2022) The value of personal data in internet commerce: A high-stake field experiment on data regulation policy. *Management Science* Forthcoming.
Vissing-Jørgensen, A. (2002) Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy* 110(4), 825–853.
Wu, M. and L. Zhang. (2022). Endogenous Growth and Human Capital Accumulation in a Data Economy. *Available at SSRN 4248918*.
Xie, D. and B. Yang. (2022). Endogenous Information Carrier Technology. *Available at SSRN 4114985*.
Zeng, J. (2003) Reexamining the interaction between innovation and capital accumulation. *Journal of Macroeconomics* 25(4), 541–560.

Cite this article: Liao L, Qin X, Li X and Gong L (2024). "Creative destruction, human capital accumulation, and growth in a digital economy." *Macroeconomic Dynamics* **28**, 1206–1230. https://doi.org/10.1017/S1365100523000445