Unveiling the Power of Basic Research: Innovation and Economic Resilience In Production Network

JIANDA BI[*], XIUPING HUA[†], YONG WANG[‡], HAISHENG YANG[§]

The theoretical literature highlights the critical role of resource allocation efficiency in driving the economic impact of innovation and industrial policies. However, few empirical studies have examined how basic research investment should be allocated across different disciplines. This study addresses this gap by investigating the effects of basic research investment on economic productivity and resilience. Specifically, we extend the framework of Acemoglu and Azar (2020) to allow for heterogeneous industry-level total factor efficiency change and investigation on labor market shares. Our findings reveal that (i) basic research investment significantly enhances the efficiency of various industries in China; (ii) industries with stronger interconnections get greater economic benefits and exhibit higher employment shares; (iii) counterfactual analysis suggests that basic research investment in high-tech sectors can substantially improve overall productivity and economic resilience; and (iv) improving the efficiency of complementary industries can yield economic outcomes comparable to directly targeting constrained sectors. These results underscore the strategic importance of basic research investment in fostering sustainable economic growth and resilience.
JEL: C67,E10,E23,L23,O41
Keywords: Basic Research, Machine Learning, Economic Resilience

The economic impact of basic research requires careful investigation due to its critical role in driving technological progress and innovation. While profit-driven R&D often focuses on creating patentable techniques and products (Romer, 1990; Grossman and Helpman, 1991; Aghion and Howitt, 1992), basic research enhances a society's understanding of natural phenomena and regularities (Akcigit, Hanley and Serrano-Velarde, 2021; Kantor and Whalley, 2014). Unlike applied research, which benefits from excludability through patents, the non-excludable nature of basic research limits its profitability, making public funding essential (Marchiori and Minelli, 2023).

Increasing public funding for basic research may incur short-term economic costs by diverting resources away from more immediately productive sectors (Prettner and Werner, 2016). Additionally, it could crowd out private sector R&D efforts, while firms exposed to uncertainties—such as reduced profitability from lower technological bar-

[*] Bi: Business School, Zhejiang Wanli University, Jianda.Bi@nottingham.edu.cn.
[†] Hua: Nottingham University Business School China, University of Nottingham Ningbo China, Xiuping.Hua@nottingham.edu.cn.
[‡] Wang: Institute of New Structural Economics, Peking University, yongwang@nsd.pku.edu.cn.
[§] Yang: Lingnan College, Sun Yat-Sen University, yhaish@mail.sysu.edu.cn.

riers and heightened market competition—may face exacerbated challenges (Bloom, Schankerman and Van Reenen, 2013). Nevertheless, empirical studies consistently demonstrate a positive link between public investment in basic research and long-term economic prosperity (Mansfield, 1980; Toole, 2012). This highlights the importance of enhancing the resource allocation efficiency of basic research, ensuring that the long-term benefits of such investments outweigh their short-term tradeoffs. Despite its importance, limited empirical work has quantified the effects of basic research on efficiency gains at a sectoral level, and has explored its impact on economic resilience from the perspective of interconnected production networks.

Our paper aims to address this question by focusing on how investments in basic research accelerate technological innovation for different sectors and strengthen the resilience of upstream and downstream industries. One challenge is to measure the effect of basic research investment (Mansfield, 1981; Link, 1981). Due to the pervasive externalities of R&D activities, knowledge in one discipline often spillover over multiple industries (Mansfield, 1980; Griliches, 1986; Myers and Lanahan, 2022). For example, advancements in the automation sector can simultaneously drive progress in the manufacturing of synthetic materials and electronics. Such spillovers can be found among patent citations (Jaffe, 1986; Jaffe, Trajtenberg and Henderson, 1993) and inventors (Bloom, Schankerman and Van Reenen, 2013). A common way in the previous studies is to match patents with disciplines and firms to estimate industry-level basic research inputs and outputs. e.g., (Azoulay et al., 2019). However, this approach typically assumes a specific functional form between research investments and innovation outputs. Traditional econometric models with such assumptions may suffer from specification errors, leading to biased estimates.

To overcome this challenge, we first employ a random forest model to precisely measure the impact of basic research investments on technological efficiency across various industries (Gelman et al., 2023; Van Binsbergen, Han and Lopez-Lira, 2023). Compared to traditional econometric approaches, the random forest model provides a robust framework for estimating nonlinear relationships in economic activities, such as estimating the effect of basic research investment. It effectively mitigates challenges like multicollinearity and the loss of degrees of freedom associated with incorporating many variables, yielding consistent and reliable estimates (So, 2013; Gu, Kelly and Xiu, 2020). We thus rely on this model to quantify the effect of basic research on the sectoral innovation outputs.

We then apply our model estimates to a dataset from China, a country with rapidly increasing investment in R&D. In 2022, China's total R&D expenditure surpassed 3 trillion RMB, with an R&D intensity[1] of 2.55%. By 2023, R&D expenditure had risen to 3.33 trillion RMB, an 8.1% year-on-year increase.[2] However, the proportion of basic research within China's R&D expenditure remains relatively low,[3] significantly lagging behind the levels observed in developed countries across Europe and North America.

[1]R&D expenditure as a share of GDP
[2]`http://theory.people.com.cn/n1/2023/0303/c40531-32635510.html`
[3]R&D comprises two components: "basic research" (R) and "applied development" (D).

The data on basic research is sourced from the National Science Foundation of China (NSFC), a key institution for public funding of basic research. Government support for this area has steadily increased over the years (Zhou and Zhao, 2019). Recognizing the critical role of basic research, China's 14th Five-Year Plan raised the share of basic research within total R&D investment to 8%. Using NSFC data, we constructed a discipline-year funding dataset, which we combined with industry-level patent data and macroeconomic indicators such as GDP and employment shares to create a comprehensive panel dataset covering the period from 2007 to 2017. Then, we employed this dataset to estimate the impact of basic research investments on industry-level technological efficiency. The results indicate that disciplines such as Automation, Geophysics, and space physics, and Geophysics have the highest contributions to driving technological efficiency across industries, highlighting their pivotal role in translating foundational research into productivity gains.

Furthermore, we do several counterfactual experiments to examine how industry-level productivity gains affect GDP, employment shares, and economic resilience in China. Building on Acemoglu and Azar (2020), we extend the model to incorporate heterogeneous technology improvements across sectors and account for labor market clearing within a general equilibrium model. This model shows that technological efficiency improvements in one sector propagate through the production network and amplify their impact economy-wide, which is consistent with the prior studies (Hertzel et al., 2008).

Using estimates from our random forest model, we analyze the effects of exogenous technology shocks on China's GDP and employment shares. We find that improvements in technological efficiency, particularly in industries like "electronic components", can increase economic resilience, and mitigate the impact of price shocks in agriculture, mining, and high-tech sectors. To examine economic resilience, we follow Dew-Becker (2023) and construct two measures: left-tail centrality, which measures GDP sensitivity to technological shocks in specific industries, and local elasticity, which captures how changes in a sector's GDP influence the broader economy. These variables enable us to quantify the impact of major industry productivity shifts on national GDP, providing insights into economic resilience. Finally, we examine the complementary effects among industries. Specifically, we explore how technological improvements in sectors like computers and power systems can offset the negative impacts of restrictions on electronic components. Our results show that peripheral industries such as ceramics or construction materials require significantly larger efficiency gains to yield comparable economic effects, reflecting their limited roles within the input-output network.

Our paper is linked to a number of different literatures. First, it is most closely related to the research on industrial linkages and spillover effects within production networks (e.g., Acemoglu et al. (2012); Baqaee (2018); Acemoglu and Azar (2020); Elliott, Golub and Leduc (2022)). This body of work highlights how production networks facilitate the diffusion of innovation and exogenous shocks through the economy (Atalay et al., 2011). For instance, Acemoglu, Akcigit and Kerr (2016) differentiate between the direct effects of exogenous shocks on the immediately impacted sectors and the indirect effects that propagate through the network. Similarly, Carvalho and Voigtländer (2014) investigate the dynamic evolution of production networks, demonstrating that a network's initial configuration significantly influences its development. Building on this literature, we extend the endogenous production network model of Acemoglu and Azar (2020) by incorporating labor market clearing constraints and developing it into a general equilibrium framework. Our contribution lies in introducing productivity shocks with sectoral heterogeneity, estimated through the random forest model in our empirical analysis. These heterogeneous shocks allow for a detailed examination of how technological progress drives the formation of new industrial linkages, stimulates GDP growth, and reshapes employment structures.

Second, our paper is related to the growing literature on labor market dynamics and technological change (Bárány and Siegel, 2018; Bergholt, Furlanetto and Maffei-Faccioli, 2022; Leduc and Liu, 2024; Alvarez-Cuadrado and Poschke, 2011). Early studies primarily examined the labor-displacing effects of technological progress, showing that advancements in technology reduce the relative prices of investment goods compared to consumer goods, resulting in crowding-out effects that lower labor's share of income (Frey and Osborne, 2017; Karabarbounis and Neiman, 2014; Fadinger, Ghiglino and Teteryatnikova, 2022). For instance, Hémous and Olsen (2022) find that automation and product-level innovations increase technology premiums, further diminishing labor's share in production. However, technological progress also enhances the division of labor, fosters the creation of new industries, and mitigates underemployment (Acemoglu and Restrepo, 2018, 2022, 2019, 2020). High-tech advancements—such as 5G, big data, cloud computing, and the Internet of Things—have heightened the demand for skilled professionals in fields like computing, data analysis, and intelligent algorithms, thereby reshaping labor demand structures. While these technologies often displace workers in routine manual roles, they simultaneously increase the demand for non-routine, interactive, and specialized skills (Autor, Levy and Murnane, 2003). Despite this, relatively few studies examine the effects of basic research on employment. Existing literature tends to focus on macroeconomic outcomes or single-sector analyses to assess employment impacts (Herrendorf, Rogerson and Valentinyi, 2013), with limited exploration of how basic research influences labor allocation across sectors within a general equilibrium framework. This study addresses this gap by analyzing how technological progress in high-tech sectors redistributes employment shares across other sectors, offering valuable insights for designing targeted employment policies and optimizing labor market structures.

Our paper also contributes to the literature on measuring economic resilience, defined as an economy's capacity to adapt its structure and growth patterns to effectively absorb internal and external shocks while ensuring sustainable development (Martin, 2012; Bazzi and Blattman, 2014). Recent research underscores the pivotal role of high-tech sectors in enhancing economic resilience. For example, Gu, Yang and Huo (2021) demonstrates that adopting IT systems improves information processing capabilities along supply chains, thereby strengthening supply chain resilience. However, much of the existing literature focuses on micro-level data, often relying on theoretical frameworks rather than empirical evaluations of resilience. This study extends the concept of "left-tail centrality" introduced by Dew-Becker (2023) to empirically measure economic resilience. Specifically, it examines how productivity improvements in high-tech sectors bolster resilience against sector-specific price shocks. By employing Dew-Becker (2023)'s methodology, this study assesses economic resilience by comparing an economy's capacity to respond to exogenous price shocks before and after changes in input-output linkages.

The rest of the paper is organized as follows. Section I presents the model. Section II describes our data and quantitative framework. Section III presents our quantified parameter estimates and provides validation tests. Section IV examines the impact of counterfactual policy experiments on GDP and employment. Section V identifies economic resilience and explores the complementary effect among sectors. The last section concludes, while Appendix A contains additional quantitative results.

I. Model

In this section, we present our theoretical framework, which extends the model proposed by Acemoglu and Azar (2020) to analyze the effects of exogenous technological changes.

A. Production and Consumer Preferences

Assume there are n industries in the economy, with each industry i having a representative firm. The production function Y_i of each firm can be expressed as a C-D production function:

$$\begin{aligned} Y_i &= F_i(S_i, A_i(S_i), L_i, X_i) \\ &= \frac{1}{(1-\sum_{j\in S_i}\alpha_{ij})^{1-\sum_{j\in S_i}\alpha_{ij}}\prod_{j\in S_i}\alpha_{ij}^{\alpha_{ij}}} A_i(S_i) L_i^{1-\sum_{j\in S_i}\alpha_{ij}} \prod_{j\in S_i} X_{ij}^{\alpha_{ij}} \end{aligned} \tag{1}$$

where $S_i \subseteq \{1,2,...,n\}$ represents the set of intermediate input suppliers for firm i, $A_i(S_i)$ is the productivity of firm i, L_i is labor input, X_{ij} is the input of intermediate goods from supplier j, and α_{ij} denotes the output elasticity of intermediate input X_{ij}. The firm's productivity function is given by:

$$A_i(S_i) = \prod_{j\in S_i} B_{i0} B_{ij} \tag{2}$$

where B_{i0} is the baseline productivity of firm i, and B_{ij} represents the impact of intermediate supplier j on the productivity of firm i. The firm's cost minimization decision can be expressed as:

$$\min_{L_i, X_{ij}} K_i\left(S_i, A_i\left(S_i\right), P\right) = \min \frac{1}{Y_i}\left(L_i + \sum_{j \in S_i} P_j X_{ij}\right) \tag{3}$$
$$s.t \quad F_i\left(S_i, A_i\left(S_i\right), L_i, X_i\right) = Y_i$$

Our economy is in continuous time and admits a representative household with the following utility maximization function:

$$\max u(C_1, C_2, ..., C_n) = \max \prod_{i=0}^{n} \beta_i^{-\beta_i} \prod_{i=0}^{n} C_i^{\beta_i} \tag{4}$$
$$s.t \quad \sum_{i=1}^{n} P_i C_i \leq 1$$

where C_i represents the consumption of the goods produced by sector i. Each household provides one unit of labor and receives a wage W, which is standardized to 1.

B. Equilibrium Prices

With market clearing conditions,

$$C_i^* + \sum_{j=1}^{n} X_{ji}^* = Y_i^*, \quad \forall i \in \{1, 2, ..., n\} \tag{5}$$
$$s.t \quad \sum_{i=1}^{n} L_i^* = 1,$$

the production of firm i equals the sum of intermediate inputs and consumption, and the total labor input across all firms is normalized to 1. Solving for the utility maximization of the representative household and the cost minimization problem of firms gives us the equilibrium values P^*, S^*, C^*, L^*, X^*, and Y^*, which satisfy the following conditions:

$$P_i^* = K_i(S_i^*, A_i(S_i^*), P^*). \tag{6}$$

At equilibrium, the cost function of firm i is:

$$K_i = \frac{\prod_{j \in S_i} P_j^{\alpha_{ij}}}{A_i(S_i)} \tag{7}$$

The cost depends on the choice of intermediate suppliers S_i, and the firm faces a trade-off between higher productivity $A_i(S_i)$ and higher prices P_i when choosing suppliers.

Combining this with Equation (6), we obtain the equilibrium price:

$$P_i^* = \frac{\prod_{j \in S_i} P_j^{*\,\alpha_{ij}}}{A_i(S_i)} \tag{8}$$

Take the logarithm of both sides, we have

$$p_i^* = \sum_{j \in S_i} \left(\alpha_{ij} p_j^* \right) - a_i(S^*) \tag{9}$$

PROPOSITION 1: *Let the $n \times n$ matrix $\alpha(S) \in R^{n \times n}$ be defined as:*

$$\alpha_{ij}(S) = \begin{cases} \alpha_{ij} & j \in S_i \\ 0 & otherwise \end{cases}$$

and the sectoral price $\boldsymbol{p} = (p_1, p_2, ..., p_n)^T$, equilibrium price $\boldsymbol{p}^*$ can then be written as:

$$\boldsymbol{p}^* = -(\boldsymbol{I} - \alpha(S^*))^{-1} a(S^*)$$

where $\boldsymbol{I}$ is the identity matrix and $a(S^*) = (a_1(S^*), a_2(S^*), ..., a_n(S^*))$.

II. Estimation and Quantitative Analysis

To examine the impact of basic research as outlined in the introduction, we begin by estimating the significance of each discipline using a random forest (RF) model. This section outlines the dataset and estimation methodology, and the subsequent sections present the results and policy counterfactual analyses.

A. Data

We utilize multiple data sources for this study, including the National Natural Science Foundation of China (NSFC),[4] the INCOPAT patent database, the EPS database, and statistics from the National Bureau of Statistics of China. The NSFC is the backbone of our analysis, providing data on annual basic research investments in China. Specifically, the amount of funding allocated to each research project is used as a proxy for basic research investment across various disciplines from 2007 to 2017.

The INCOPAT patent database offers detailed patent information, such as applicant names, IPC classification codes, application years, patent citations, and classifications by national economic sector. Changes in patent numbers across industries are used as a measure of technological efficiency within those sectors. Macroeconomic indicators and input-output data, including intermediate inputs, value-added, and the Consumer Price Index (CPI), are obtained from the National Bureau of Statistics of China.

[4] `https://www.nsfc.gov.cn/publish/portal0`

To account for industry-specific factors influencing output, we also include variables related to an industry's factor endowments, which shape its structure and productivity. Specifically, 26 control variables spanning six dimensions—natural conditions, human resources, physical capital, technology, economic structure, and institutional factors—are included. Data for these control variables is drawn from the EPS database, which aggregates information from sources such as the Chinese Environmental dataset, the Chinese Agricultural and Forestry dataset, the Chinese Land Resources dataset, and the Chinese Macro dataset. Detailed descriptions of these variables and their calculations are provided in A1.

We construct the sample through three steps. First, we aggregate NSFC funding data by discipline and year. Nest, we compile patent data by industry and year, assigning each patent to its corresponding industry using national economic classification codes, with a focus on invention patents.[5] Finally, we match the industry-level data with input-output tables using the same "discipline-year". This process yields a panel dataset from 2007 to 2017, including discipline-level funding amounts, granted patent numbers, macroeconomic indicators, and changes in input-output shares across industries.

B. Comparative Analysis

Based on the model, we analyze how a productivity shift in a specific sector impacts the equilibrium of the existing production network. Specifically, we investigate the effects of a positive productivity shock. Let the initial productivity of industry i be b_{i0}. The change in productivity is represented as follows:

$$b_{i0}^{new} = b_{i0} + \varepsilon$$

where $\varepsilon > 0$ denotes a positive technological shock, indicating a relative increase in the productivity of industry i while the productivity of other sectors remains unchanged. Such a shock could stem from targeted industrial policies or breakthrough interventions. To analyze the impact, we simulate the production network. Following Acemoglu and Azar (2020), the change of the updated input-output matrix, $\tilde{\alpha}_{ij}$, are defined as follows:

$$\tilde{\alpha}_{ij} = \begin{cases} \alpha_{ij} & j \in S_i \\ 0.95 \times \left(1 - \sum_{j' \in S_i} \alpha_{ij'}\right) \frac{\sum_{i', j \in S_{i'}} \alpha_{i'j}}{\sum_{i', j' \in S_{i'}} \alpha_{i'j'}} & j \notin S_i \end{cases}$$

where S_i represents the set of industries originally linked to industry i in the production network. If industries i and j were previously connected ($j \in S_i$), the input coefficient α_{ij} remains unchanged. Conversely, if industries i and j were not initially connected ($j \notin S_i$), the updated coefficient $\tilde{\alpha}_{ij}$ is assigned a value proportional to the relative "importance" of industry j. This importance is calculated as the ratio of the column sum of $\alpha_{i'j}$ (indicating the total input contribution of j across all industries) to the sum of all

[5]The INCOPAT database assigns a three-digit national economic classification code to each patent. For any missing data, we supplemented the information using the Reference Table for International Patent Classification and National Economic Industry Classification.

elements in the input-output matrix. This adjustment ensures that the value is less than 1 and thus reflects the relative role of industry j in the production network. Note that the proportionality factor is set to 0.95. This parameter is validated using the input-output table in 2017, where the simulated GDP for 2018 and 2020 aligns most closely with the actual value. While alternative parameters have been tested, we find that the choice of 0.95 consistently minimized the deviation between simulated and observed GDP.

C. *Parameter Estimation and Shock Specifications*

In this study, we need to estimate parameters for the random forest model (RF) and the technology efficiency in the production network.

We proceed to estimate by RF parameters. Technological advancement within an industry is influenced by numerous interconnected factors, making linear models less effective for capturing the complex effects of basic research. To address this limitation, we follow Van Binsbergen, Han and Lopez-Lira (2023) and utilize a nonlinear machine learning model, the RF model, for our analysis.

The RF model is a nonlinear, nonparametric ensemble learning method built on decision trees. It operates by aggregating weak predictors through averaging multiple decision trees, selecting optimal variables and thresholds to minimize mean squared error. This approach delivers asymptotically unbiased predictions (Breiman, 2001). Due to its robustness and flexibility, the RF model is particularly well-suited for complex, heterogeneous datasets. It accurately estimates individual treatment effects by effectively handling both discrete and continuous variables while imposing minimal assumptions about autocorrelation. This method thus enables a more precise evaluation of the impact of R&D investments on technological progress. The baseline model is specified as follows:

$$\Delta b_{i,t+1} = RF[Fundamentals_{i,t}, Macro_t, b_{i,t}] \tag{10}$$

where $\Delta b_{i,t+1}$ is the technology change for industry i in year $t+1$, calculated using the changes in patent numbers. $Fundamentals_{i,t}$ includes innovation-related factors such as R&D inputs, patent counts, and technological position for industry i in year t. $Macro_t$ and $b_{i,t}$ capture macroeconomic variables and patent numbers at year t, respectively. Model hyperparameters are optimized using cross-validation, as detailed in Table 1. The Specific definitions of these variables are in Table A1.

Second, we estimate the parameters of the production network using a Cobb-Douglas (C-D) form for both the production and utility functions. The share of intermediate inputs supplied from sector j to sector i is denoted as α_{ij}, while the proportion allocated to consumption is represented by β_i. Following Acemoglu and Azar (2020), we assume an equilibrium state in the production network to determine the initial productivity b_{i0} and the intermediate input productivity b_{ij}. Specifically, b_{i0} is assumed to have a truncated normal distribution with a mean of μ and a standard deviation of 1, while b_{ij} is assumed to have a mean of μ/n and a standard deviation of $1/n$. These parameters are calibrated

TABLE 1—HYPERPARAMETERS FOR THE RANDOM FOREST MODEL

Hyperparameter	Value
Number of Trees	1,500
Maximum Depth	20
Sample Fraction	33%
Minimum Node Size	7

Note: Parameters were chosen through 5-fold cross-validation. The number of trees represents the number of decision trees in the random forest model, and the maximum depth indicates the maximum allowable splits for each tree. Sample fraction refers to the proportion of observations used to train each tree, and the minimum node size determines when splitting stops.

to ensure that $\mu = \min_{b_{ij}} |GDP_{\text{real}} - GDP|$, thereby aligning the simulated GDP with the real GDP.

Table 2 presents the mean productivity value (μ) for the years 2017, 2018, and 2020, along with a comparison between actual GDP and the simulated GDP. The results indicate minimal differences between the simulated and actual GDP values, suggesting that the selected parameters are both appropriate and reliable.

TABLE 2—AVERAGE PRODUCTIVITY AND COMPARISON OF ACTUAL AND SIMULATED GDP

Year	μ	Actual GDP	Simulated GDP
2017	3.3262	823,215.71	823,215.71
2018	3.4056	922,057.17	922,057.28
2020	3.5337	1,016,422	1,016,422

Note: GDP is measured in billion CNY.

III. Results

In this section, we present our estimation results and evaluate the fit of our model in the data.

A. *Parameter Estimates*

We use changes in the number of patents as a proxy for sectoral technological efficiency, which serves as the dependent variable in our model. The explanatory variables include 30 different factors, such as NSFC funding amounts by discipline, industry factor endowments (26 variables), R&D intensity, total number of employees, and average firm age.

Figure 1 presents the estimation results of equation (10). The ranking reflects the contribution of each factor to the reduction in the residual sum of squares in the random forest model.[6] A larger contribution corresponds to a higher rank and greater weight. Among the disciplines, Automation, Geophysics and space physics, and Geophysics exhibit the highest weights in predicting technological progress across industries, with standardized weights of approximately 4.88%, 3.73%, and 2.56%, respectively.

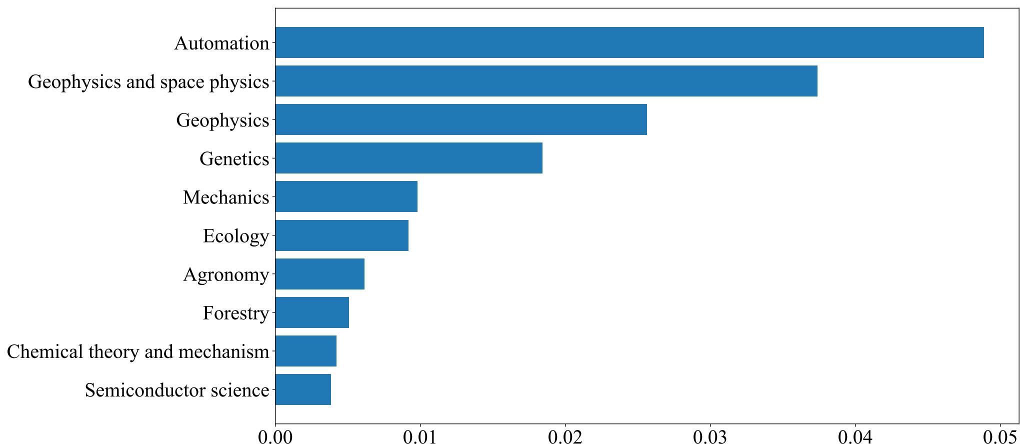


FIGURE 1. IMPORTANCE OF DIFFERENT DISCIPLINES

B. Goodness of Fit

To ensure the objectivity of our model's parameters, we follow the approach of Chen et al. (2022) and further validate the results using a Gradient Boosting Decision Tree (GBDT) model. Unlike random forest models, GBDT enhances predictive accuracy by iteratively refining prior predictions through a boosting mechanism. By introducing randomness in each iteration and using random samples, this method reduces the correlation between estimates across iterations, effectively minimizing variance (Friedman, 2002). Using the same variables, we re-estimated the baseline model and derived the weights for each discipline, as illustrated in Figure 2.

The new top ten disciplines show significant overlap with the previous list, with nine disciplines appearing in both rankings. The consistent presence of disciplines such as Automation, Geophysics and space physics, and Geophysics underscores their enduring foundational significance across diverse scientific and industrial domains. However, Chemical metrology are newly included, replacing Chemical theory and mechanism. We therefore compare the goodness-of-fit of various nonlinear machine learning algorithms within the sample. The results, illustrated in Figure 3, demonstrate that, based on MSE and R^2 as evaluation metrics, the random forest algorithm achieves the highest goodness-of-fit and the lowest error.

[6]Only disciplines with the highest weights are displayed.

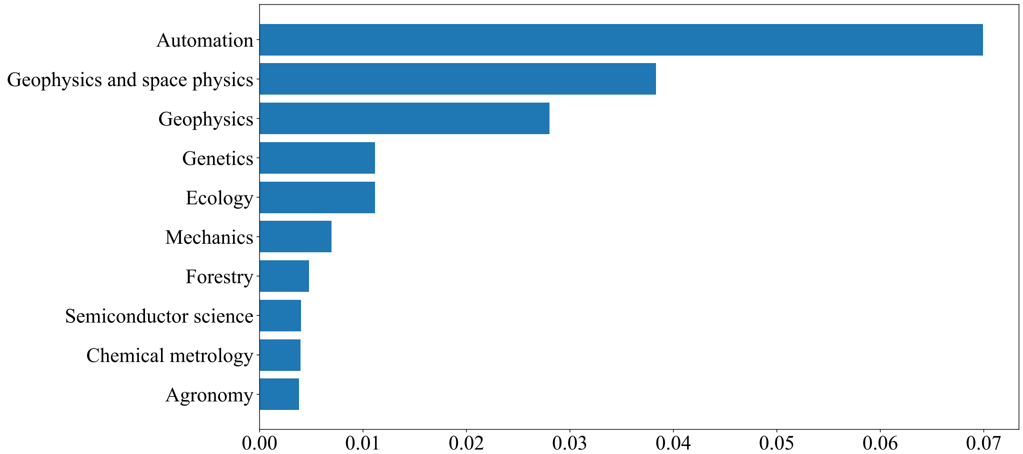


FIGURE 2. IMPORTANCE OF DIFFERENT DISCIPLINES USING GBDT

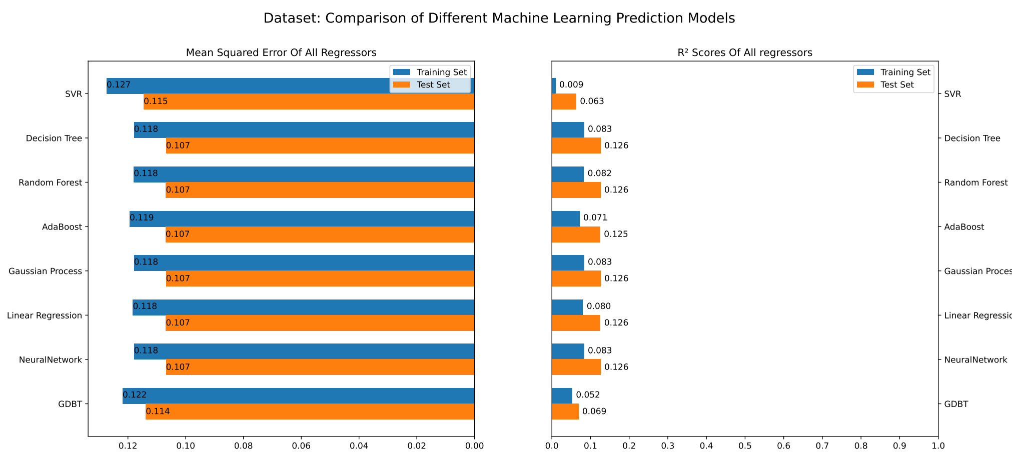


FIGURE 3. COMPARISON OF DIFFERENT MACHINE LEARNING MODELS

IV. Counterfactual Analysis

So far, we have identified the weights measuring the contributions of basic research investments to technological progress. In this section, we examine the role of high-tech sectors in shaping the overall economy through a production network framework. We conduct a counterfactual analysis to assess changes in GDP resulting from an exogenous technological shock. High-tech sectors are selected for their critical role as drivers of innovation and their frequent access to government support. For example, high-tech firms often receive subsidies in China, including direct financial assistance through national initiatives such as the Spark Program, Torch Program, and 863 Program.

Following the classification standards of the National Bureau of Statistics (GB/T 4754), we categorize high-tech manufacturing industries into eight groups: general equipment, specialized equipment, transportation equipment, electrical machinery, computers and other electronic equipment, communication equipment, instruments and meters, and cultural and office machinery. By matching these categories with sectors in the input-output tables, we identify 28 sectors representing the high-tech industries in the national input-output framework.[7] Unless otherwise specified, the term "high-tech" in the subsequent analysis refers to these 28 sectors.

A. *Impact on GDP*

To proceed, we assume the evolution of the technological shock takes the following form for simplicity:

$$b_{i0}^{new} = (1+\delta)b_{i0} \tag{11}$$

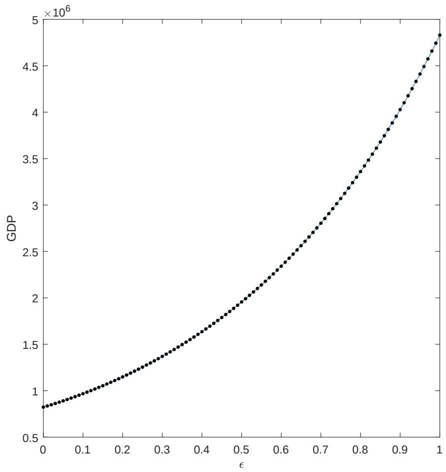
where $\delta > 0$ represents the technology shock. Using this evolution function, we evaluate the impact of this shock on the overall production network.

Given the interdependencies among production sectors, technological progress can affect overall GDP through input-output network relationships. To analyze this, we use the national input-output tables for 2017, 2018, and 2020 to assess how a positive technological shock in high-tech industries influences GDP growth. Specifically, we calculate the change of sector *i*'s productivity and its influence on the other sectors in the network, assuming that the productivity of other sectors remains unchanged. We adjust the shock coefficient δ from 0 to 1 in increments of 0.01, calculating the corresponding changes in GDP. The relationship between the shock coefficient and the resulting GDP growth is depicted in Figure 4.

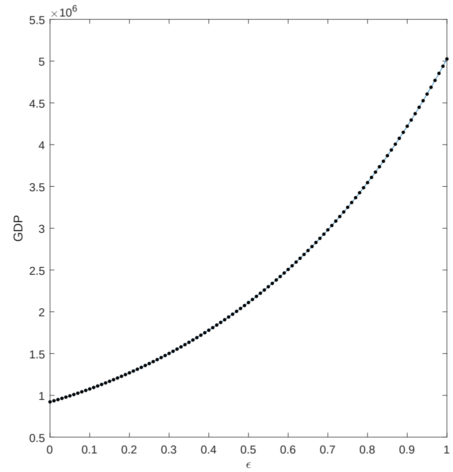
Figure 4 illustrates the impact of technological shocks on GDP changes, revealing a nonlinear relationship. Table 3 provides the GDP growth rates corresponding to different values of δ. As δ increases, GDP growth rates also rise, underscoring the growing significance of high-tech industries in driving economic expansion. This trend highlights the critical role of technological advancements in fostering innovation and supporting sustainable economic development.

The observed nonlinear relationship emphasizes the substantial influence of technological innovation on economic growth, suggesting that high-tech industries will continue to be major contributors to GDP growth in the future. To fully realize this potential, policymakers and industry leaders may actively track advancements in high-tech sectors and use our framework to monitor the effect of technological innovation.

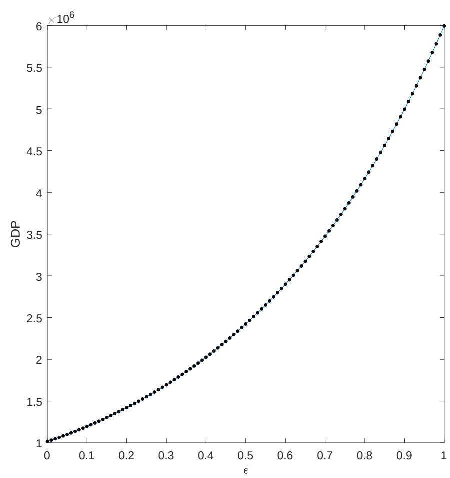
[7] see A3 for the matches



(a) 2017



(b) 2018



(c) 2020

FIGURE 4. NONLINEAR IMPACT OF TECHNOLOGICAL SHOCKS ON GDP

Note: GDP is measured in a hundred million CNY.

B. Impact on Employment Share

Technological advancements play a pivotal role in reshaping labor allocation across sectors. Our theoretical model assumes that a representative household supplies a fixed unit of labor to the economy. Consequently, the equilibrium employment levels in each

TABLE 3—GDP GROWTH RATES FOR DIFFERENT δ (%)

δ	2017	2018	2020
0.010	0.015	0.014	0.015
0.020	0.031	0.029	0.031
0.030	0.046	0.044	0.046
0.040	0.062	0.060	0.063
0.050	0.079	0.075	0.079
0.060	0.095	0.091	0.095
0.070	0.112	0.107	0.112
0.080	0.128	0.123	0.129
0.090	0.145	0.139	0.146
0.100	0.162	0.155	0.163

sector reflect the proportion of total labor allocated to that sector. To assess the impact of technological shocks on labor distribution, we analyze changes in employment shares across sectors rather than absolute employment levels, as the initial shares and sizes of sectors vary. Figure 5 depicts the percentage changes in employment shares for high-tech sectors and newly connected industries following a positive productivity shock, and Table 4 provides a detailed comparison of employment shares and quantities across high-tech, newly connected, and other sectors.

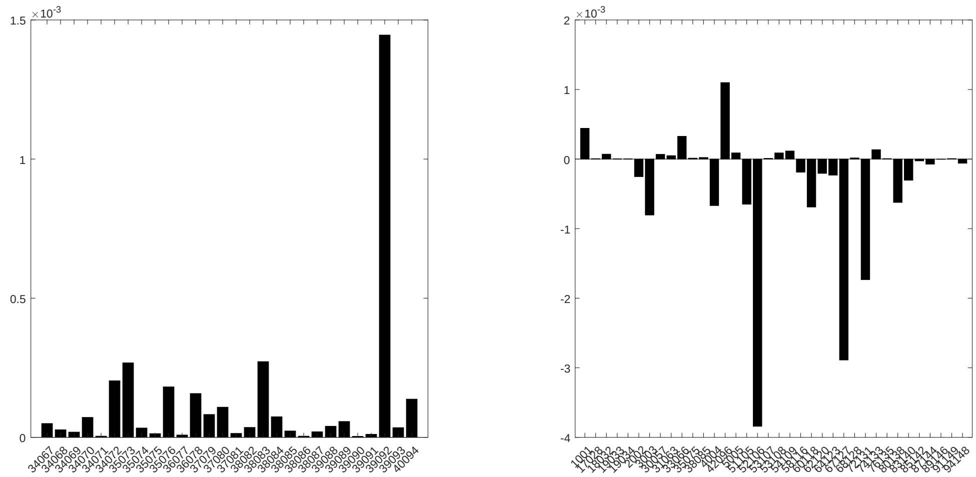
Over three years, the aggregate labor share of high-tech sectors increased by 8.60%, 8.93%, and 9.07%, respectively. Notable employment growth was observed in industries such as electronic components (39,092), mining (35,073), general-purpose equipment (34,072), automotive parts (36,078), power transmission (38,083), and shipbuilding (37,080). Conversely, newly connected sectors experienced a decline in labor share, reflecting a substitution effect that shifted labor away from these industries. This trend underscores the profound impact of technological progress on the employment structure and highlights the need for policymakers and business leaders to adopt proactive strategies to manage labor market transitions effectively.

TABLE 4—CHANGES IN EMPLOYMENT SHARE AND EMPLOYMENT LEVELS

Year	High-Tech Sectors		New Sectors		Other Sectors	
	Δ EMPS	Emp	Δ EMPS	Emp	Δ EMPS	Emp
2017	8.60	286.65	-1.98	-82.84	1.73	56.73
2018	8.93	245.06	-1.64	-77.63	2.02	56.25
2020	9.07	218.79	-2.24	-95.09	2.43	73.21

Note: Δ EMPS means Employment Share Change (%); *Emp* denotes to Employment (10,000 persons); Other Sectors refers to sectors excluding the high-tech core sectors and those newly linked to them.

Table 5 presents the new linkages established in the production network after a produc-



(a) High-tech Sectors (b) Newly Connected Sectors

FIGURE 5. PERCENTAGE CHANGES IN EMPLOYMENT SHARES

Note: The figure is drawn using data from 2017. See Table A2 for the results of 2018 and 2022.

tivity shock to high-tech sectors. Industries such as general-purpose equipment, specialized equipment, transportation equipment, electrical machinery, computers, electronic devices, and communication equipment exhibit an increase in their outdegree. Simultaneously, primary and service sectors, including agriculture, forestry, fisheries, postal services, road freight transport, and auxiliary transport services, form new connections with these high-tech core industries by utilizing their products or services. These newly established linkages expand the role of technology in supporting the broader economy.

The growth of these industrial connections not only fosters greater collaboration and interdependence across sectors but also creates opportunities for technological innovation and industrial upgrading. Policymakers can capitalize on this by implementing strategies to strengthen intersectoral linkages, thereby enhancing the complementarity of industrial and supply chains, reducing dependence on external technologies, and accelerating the pace of innovation and industrial development. Such measures will contribute to sustained economic growth and industrial advancement.

V. Economic Resilience

In this section, we examine the impact of technological advancements in high-tech sectors on the economy's resilience to price shocks across three sectors: high-tech sectors, primary sectors, and mining sectors. Strengthening the resilience of these sectors is essential for achieving sustainable development. For instance, a country heavily relying on agricultural imports faces significant vulnerabilities due to price fluctuations caused by extreme weather events, export restrictions, and geopolitical conflicts. Using the 2017

TABLE 5—NEW SECTORS AFTER SHOCK

Year	Pre-Shock Outdegree	Post-Shock Outdegree	New Linkages
2020	3221	3281	**General Equipment** → Road Freight Transport and Auxiliary Activities; Postal Services; Broadcasting and Satellite Transmission Services; Monetary Finance and Other Financial Services; Capital Market Services; Residential Services; Social Work; Broadcasting, Television, Film and Audio-Visual Recording Production; Social Security **Specialized Equipment** → Agricultural Products; Forestry Products; Agricultural, Forestry, Livestock, and Fishery Services; Feed Products; Seasonings and Fermented Products; Textile Products; Glass and Glass Products; Gas Production and Supply; Retail; Railway Passenger Transport; Information Technology Services; Professional Technical Services; Cultural Arts; Sports; Entertainment; Social Security **Transportation Equipment** → Forestry Products; Fishery Products; Construction Decoration, Renovation, and Other Construction Services; Railway Passenger Transport; Railway Freight Transport and Auxiliary Activities; Telecommunications; Broadcasting and Satellite Transmission Services; Monetary Finance and Other Financial Services; Ecological Protection and Environmental Governance; Public Facilities and Land Management; Education; Social Work; News and Publishing; Social Security; Public Administration and Social Organizations **Electrical Machinery and Equipment** → Wholesale; Postal Services; Catering; Information Technology Services; Water Conservancy Management; Social Security; Public Administration and Social Organizations **Computers and Other Electronic Equipment** → Agricultural Products; Postal Services; Business Services **Communication Equipment** → Social Security

national input-output table in China, we analyze the GDP recovery rate following such shocks to evaluate the long-term benefits of basic research.

Following Dew-Becker (2023), we define economic resilience as follows:

$$\gamma_i^L \equiv \lim_{\Delta z_i \to -\infty} \frac{\Delta GDP}{\Delta z_i} \tag{12}$$

where Δz_i and ΔGDP denote the deviations of sector i's productivity and overall GDP, respectively, from their steady-state levels. A negative limit means a sufficiently large negative productivity shock that disrupts the entire production network without being absorbed. A higher γ_i^L value indicates that a negative productivity shock in sector i generates greater systemic risk. Therefore, γ_i^L can be interpreted as the magnitude of a sector's influence on the overall economy or the sensitivity of total output to productivity shocks in that sector. When γ_i^L values are high across sectors, it suggests that negative shocks in individual sectors are more likely to cause substantial GDP declines, indicating weaker economic resilience.

A. *Resilience to Price Shocks*

We then utilize the economic resilience, γ_i^L, to assess the economy's capacity to respond to a supply-side price shock. Specifically, we analyze changes in overall GDP resulting from price increases in the primary sector, mining sector, and high-tech sector. Additionally, we simulate the effects of technological shocks in the high-tech sector at

1%, 2%, and 3%, with $\delta = 0$ representing the baseline scenario without any technological shock.

The horizontal axis in Figure 4 indicates the percentage increase in product prices across various sectors, while the vertical axis represents the corresponding GDP change. Figure 4 demonstrates that price increases for intermediate goods negatively impact GDP. However, as technological advancements occur within the high-tech sector, the adverse effects of these price increases are progressively mitigated. These findings suggest that fostering innovation and technological progress in the high-tech sector can reduce the negative impact of price shocks, enhance GDP stability, and improve overall economic resilience.

Furthermore, we investigate the heterogeneous effects within high-tech sectors. In the unreported figures, our analysis reveals that individual sectors exhibit relatively minor variations following positive productivity shocks. This finding suggests that technological advancements in these sectors have a limited impact on enhancing GDP's resilience to price increases in intermediate goods.

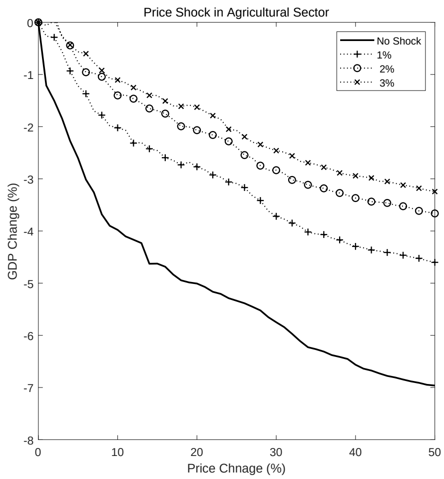
However, the "electronic components" sector stands out as an exception, making a significant contribution to GDP stability compared to other core sectors of the high-tech economy. As illustrated in Figure 7, the upward shift of the curve with increasing shock intensity highlights that productivity growth in the electronic components sector substantially enhances GDP's ability to withstand price shocks originating from the agricultural, mineral, and high-tech sectors, thereby fostering stable economic growth.

This notable effect is likely due to the critical role of the electronic components sector within the high-tech industrial chain. Electronic components are essential to modern information technology and communication systems, underpinning devices such as computers, telecommunications equipment, and even medical and military hardware. The sector's applications span diverse industries, including electronics, communications, automotive, healthcare, and aerospace, thereby influencing both upstream and downstream segments of the value chain. Our findings thus show the importance of the electronic components industry for policymakers, offering valuable insights for anticipating future trends and formulating strategic responses. Therefore, prioritizing the development of this sector is crucial for enhancing economic stability and fostering sustainable growth.

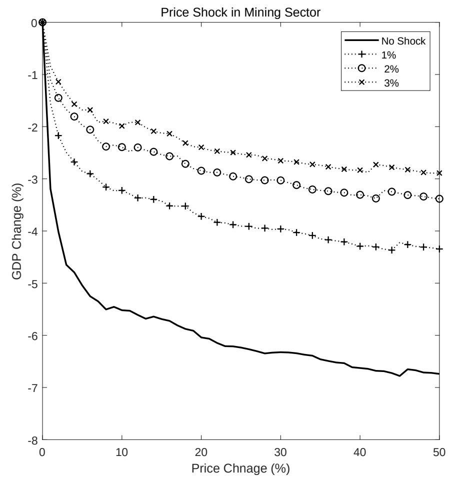
B. Simulation of Complementary Effect

Next, we examine the complementary effect in the production network, focusing on the interconnected sectors of the electronic components sector. Specifically, we investigate whether improving the technological efficiency of sectors linked to electronic components has a comparable effect on GDP growth as directly enhancing the technological efficiency of the electronic components sector itself. As a cornerstone of modern infrastructure, electronic components are essential to the functionality of advanced technological systems. By improving the technological efficiency of these linked sectors, we simulate potential outcomes for enhancing economic resilience.

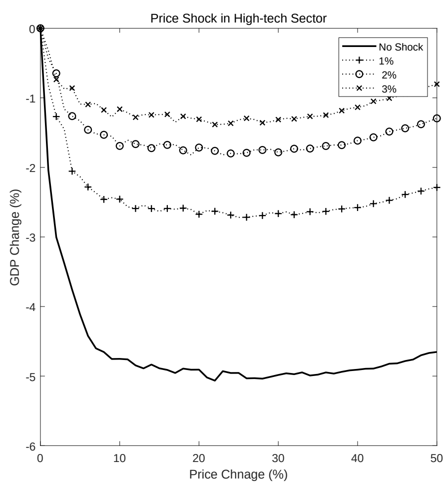
Using the 2017 input-output table, our analysis reveals that a 1% increase in the technological efficiency of the electronic components sector ($\delta = 1\%$) results in a 0.011%



(a) Agriculture Sector



(b) Minning Sector

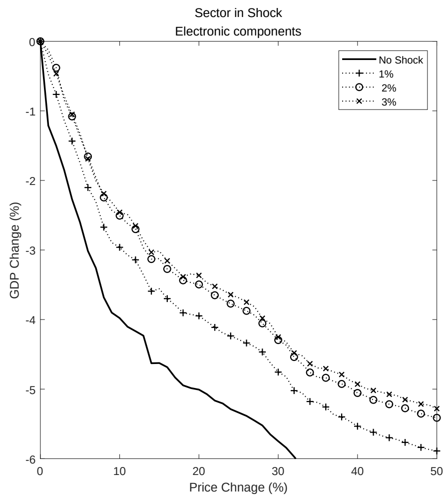


(c) High-tech Sector

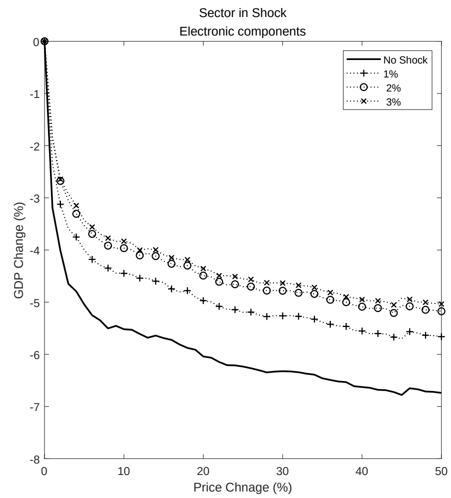
FIGURE 6. ECONOMIC RESILIENCE AND PRICE SHOCKS

increase in GDP. We also evaluate the impact of technological progress across various industries by applying different technological improvement parameters (δ) ranging from 0% to 100%. The detailed results are provided in Table A5.

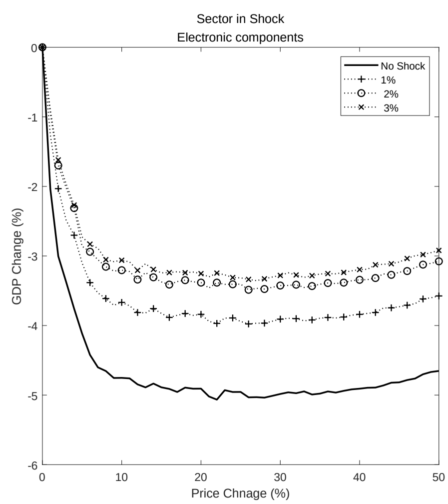
Our findings indicate that industries more distantly positioned in the production network, such as ceramics ($\delta = 46.1\%$), construction materials (e.g., bricks, tiles, and stone, $\delta = 60.5\%$), and shipbuilding ($\delta = 39.4\%$), require significantly higher levels of tech-



(a) Agriculture Sector



(b) Minning Sector



(c) High-tech Sector

FIGURE 7. PRICE SHOCKS TO THE ELECTRONIC COMPONENTS SECTOR

nological improvement to achieve the same GDP impact as the electronic components sector. In contrast, industries closely linked to electronic components, such as the computer industry ($\delta = 3.4\%$), power production and supply ($\delta = 0.6\%$), and automotive parts ($\delta = 1\%$), achieve similar GDP gains with much smaller technological shocks. Meanwhile, sectors with no direct connection to electronic components, such as construction, social work, and housing construction, exhibit negligible GDP growth even under a 100% technological improvement.

These results highlight the complementary effect among different sectors, suggesting that targeted investments in complementary sectors can unlock significant economic benefits. Policymakers may prioritize overcoming technological bottlenecks in these complementary industries by increasing investments in basic research.

It is noteworthy that sectors like computers, power production, and automotive parts belong to China's strategic emerging industries, which encompass advanced manufacturing and modern services. These industries often occupy central positions in production networks, with strong upstream and downstream linkages. Table A4 shows that these strategic emerging industries are actually complementary sectors to the electronic components sector, with lower δ values, underscoring their efficiency in driving GDP growth.

Our analysis thus reaffirms the importance of strategic emerging industries as critical drivers of economic development. Since the launch of the "Strategic Emerging Industries Development Plan" in 2010, the Chinese government has consistently emphasized its development. This study provides additional evidence of their central role in shaping China's economic future and highlights the importance of continued government support for these sectors to foster sustainable growth and innovation.

VI. Conclusion

In conclusion, this study highlights the critical role of basic research investment in driving technological progress in high-tech sectors, particularly the electronic components industry, to foster economic resilience, aggregate output, and employment shares. By utilizing an endogenous production network model and analyzing data from the National Natural Science Foundation and patent counts, this study clarifies how investments in basic research stimulate innovation, enhance industrial linkages, and bolster GDP and employment shares. These findings provide insights into the transformative potential of technologies in transitioning China's economy from "super-large" to "super-strong."

The study reveals that sectors with higher technological interconnectivity reap greater economic benefits, underscoring the importance of disciplines such as Automation, Geophysics and space physics, and Ecology in driving innovation. Key industries like "automobile components" and "electronic components" emerge as pivotal contributors to enhancing China's economic resilience. Furthermore, technological advancements in high-tech sectors catalyze growth across upstream and downstream industries, strengthening the overall value chain.

For policymakers, this research emphasizes the importance of increasing investments in basic research with significant technological impact, addressing bottleneck sectors strategically, and prioritizing high-quality innovation. By targeting specific areas of basic research, the government can enhance the efficiency of translating research investments into practical innovations, thereby driving economic progress and resilience.

REFERENCES

Acemoglu, Daron, and Pablo D Azar. 2020. "Endogenous production networks." *Econometrica*, 88(1): 33–82.

Acemoglu, Daron, and Pascual Restrepo. 2018. "The race between man and machine: Implications of technology for growth, factor shares, and employment." *American Economic Review*, 108(6): 1488–1542.

Acemoglu, Daron, and Pascual Restrepo. 2019. "Automation and new tasks: How technology displaces and reinstates labor." *Journal of Economic Perspectives*, 33(2): 3–30.

Acemoglu, Daron, and Pascual Restrepo. 2020. "Robots and jobs: Evidence from US labor markets." *Journal of Political Economy*, 128(6): 2188–2244.

Acemoglu, Daron, and Pascual Restrepo. 2022. "Tasks, automation, and the rise in US wage inequality." *Econometrica*, 90(5): 1973–2016.

Acemoglu, Daron, Ufuk Akcigit, and William Kerr. 2016. "Networks and the macroeconomy: An empirical exploration." *NBER Macroeconomics Annual*, 30(1): 273–335.

Acemoglu, Daron, Vasco M Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. 2012. "The network origins of aggregate fluctuations." *Econometrica*, 80(5): 1977–2016.

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica*, 60(2): 323–351.

Akcigit, Ufuk, Douglas Hanley, and Nicolas Serrano-Velarde. 2021. "Back to basics: Basic research spillovers, innovation policy, and growth." *The Review of Economic Studies*, 88(1): 1–43.

Alvarez-Cuadrado, Francisco, and Markus Poschke. 2011. "Structural change out of agriculture: Labor push versus labor pull." *American Economic Journal: Macroeconomics*, 3(3): 127–158.

Atalay, Enghin, Ali Hortacsu, James Roberts, and Chad Syverson. 2011. "Network structure of production." *Proceedings of the National Academy of Sciences*, 108(13): 5199–5202.

Autor, David H, Frank Levy, and Richard J Murnane. 2003. "The skill content of recent technological change: An empirical exploration." *The Quarterly Journal of Economics*, 118(4): 1279–1333.

Azoulay, Pierre, Joshua S Graff Zivin, Danielle Li, and Bhaven N Sampat. 2019. "Public R&D investments and private-sector patenting: evidence from NIH funding rules." *The Review of Economic Studies*, 86(1): 117–152.

Baqaee, David Rezza. 2018. "Cascading failures in production networks." *Econometrica*, 86(5): 1819–1838.

Bárány, Zsófia L, and Christian Siegel. 2018. "Job polarization and structural change." *American Economic Journal: Macroeconomics*, 10(1): 57–89.

Bazzi, Samuel, and Christopher Blattman. 2014. "Economic shocks and conflict: Evidence from commodity prices." *American Economic Journal: Macroeconomics*, 6(4): 1–38.

Bergholt, Drago, Francesco Furlanetto, and Nicolò Maffei-Faccioli. 2022. "The decline of the labor share: new empirical evidence." *American Economic Journal: Macroeconomics*, 14(3): 163–198.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen. 2013. "Identifying technology spillovers and product market rivalry." *Econometrica*, 81(4): 1347–1393.

Breiman, Leo. 2001. "Random forests." *Machine Learning*, 45: 5–32.

Carvalho, Vasco M, and Nico Voigtländer. 2014. "Input diffusion and the evolution of production networks." National Bureau of Economic Research.

Chen, Xi, Yang Ha Cho, Yiwei Dou, and Baruch Lev. 2022. "Predicting future earnings changes using machine learning and detailed financial data." *Journal of Accounting Research*, 60(2): 467–515.

Dew-Becker, Ian. 2023. "Tail risk in production networks." *Econometrica*, 91(6): 2089–2123.

Elliott, Matthew, Benjamin Golub, and Matthew V Leduc. 2022. "Supply network formation and fragility." *American Economic Review*, 112(8): 2701–2747.

Fadinger, Harald, Christian Ghiglino, and Mariya Teteryatnikova. 2022. "Income differences, productivity, and input-output networks." *American Economic Journal: Macroeconomics*, 14(2): 367–415.

Frey, Carl Benedikt, and Michael A Osborne. 2017. "The future of employment: How susceptible are jobs to computerisation?" *Technological Forecasting and Social Change*, 114: 254–280.

Friedman, Jerome H. 2002. "Stochastic gradient boosting." *Computational Statistics & Data Analysis*, 38(4): 367–378.

Gelman, Michael, Yuriy Gorodnichenko, Shachar Kariv, Dmitri Koustas, Matthew D Shapiro, Dan Silverman, and Steven Tadelis. 2023. "The response of consumer spending to changes in gasoline prices." *American Economic Journal: Macroeconomics*, 15(2): 129–160.

Griliches, Zvi. 1986. "Productivity, R&D, and Basic Research at the Firm Level in the 1970's." *American Economic Review*, 76(1).

Grossman, Gene M., and Elhanan Helpman. 1991. "Quality Ladders in the Theory of Growth." *The Review of Economic Studies*, 58(1): 43–61.

Gu, Minhao, Lu Yang, and Baofeng Huo. 2021. "The impact of information technology usage on supply chain resilience and performance: An ambidexterous view." *International Journal of Production Economics*, 232: 107956.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu. 2020. "Empirical asset pricing via machine learning." *The Review of Financial Studies*, 33(5): 2223–2273.

Hémous, David, and Morten Olsen. 2022. "The rise of the machines: Automation, horizontal innovation, and income inequality." *American Economic Journal: Macroeconomics*, 14(1): 179–223.

Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi. 2013. "Two perspectives on preferences and structural transformation." *American Economic Review*, 103(7): 2752–2789.

Hertzel, Michael G, Zhi Li, Micah S Officer, and Kimberly J Rodgers. 2008. "Inter-firm linkages and the wealth effects of financial distress along the supply chain." *Journal of Financial Economics*, 87(2): 374–387.

Jaffe, Adam B. 1986. "Technological Opportunity and Spillovers of R & D: Evidence from Firms' Patents, Profits, and Market Value." *The American Economic Review*, 984–1001.

Jaffe, Adam B, Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic localization of knowledge spillovers as evidenced by patent citations." *the Quarterly Journal of Economics*, 108(3): 577–598.

Kantor, Shawn, and Alexander Whalley. 2014. "Knowledge spillovers from research universities: evidence from endowment value shocks." *Review of Economics and Statistics*, 96(1): 171–188.

Karabarbounis, Loukas, and Brent Neiman. 2014. "The global decline of the labor share." *The Quarterly Journal of Economics*, 129(1): 61–103.

Leduc, Sylvain, and Zheng Liu. 2024. "Automation, bargaining power, and labor market fluctuations." *American Economic Journal: Macroeconomics*, 16(4): 311–349.

Link, Albert N. 1981. "Basic research and productivity increase in manufacturing: additional evidence." *American Economic Review*, 71(5): 1111–1112.

Mansfield, Edwin. 1980. "Basic research and productivity increase in manufacturing." *The American Economic Review*, 70(5): 863–873.

Mansfield, Edwin. 1981. "Composition of R and D expenditures: Relationship to size of firm, concentration, and innovative output." *The Review of Economics and Statistics*, 610–615.

Marchiori, Carmen, and Enrico Minelli. 2023. "Talent, basic research and growth." *Journal of Economic Theory*, 213: 105721.

Martin, Ron. 2012. "Regional economic resilience, hysteresis and recessionary shocks." *Journal of Economic Geography*, 12(1): 1–32.

Myers, Kyle R, and Lauren Lanahan. 2022. "Estimating spillovers from publicly funded R&D: Evidence from the US Department of Energy." *American Economic Review*, 112(7): 2393–2423.

Prettner, Klaus, and Katharina Werner. 2016. "Why it pays off to pay us well: The impact of basic research on economic growth and welfare." *Research Policy*, 45(5): 1075–1090.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy*, 98(5): 71 – 102.

So, Eric C. 2013. "A new approach to predicting analyst forecast errors: Do investors overweight analyst forecasts?" *Journal of Financial Economics*, 108(3): 615–640.

Toole, Andrew A. 2012. "The impact of public basic research on industrial innovation: Evidence from the pharmaceutical industry." *Research Policy*, 41(1): 1–12.

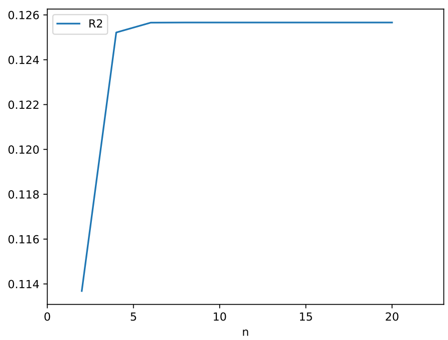
Van Binsbergen, Jules H, Xiao Han, and Alejandro Lopez-Lira. 2023. "Man versus machine learning: The term structure of earnings expectations and conditional biases." *The Review of Financial Studies*, 36(6): 2361–2396.

Zhou, Zhonghe, and Weijie Zhao. 2019. "Funding system reform for excellence in science: an interview with Jinghai Li, President of NSFC." *National Science Review*, 6(1): 177–181.

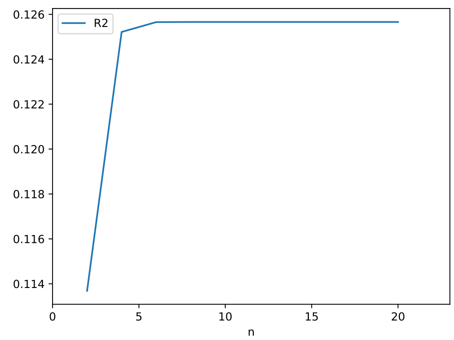
Appendix A: Additional Results

TABLE A1—VARIABLE DEFINITION

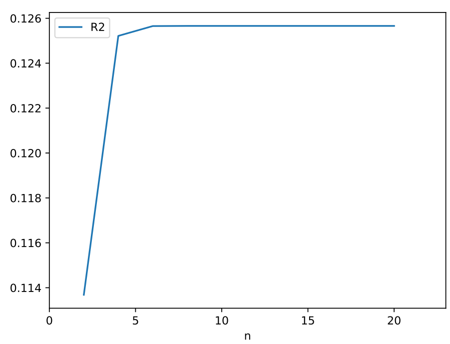
Endowment	**Primary indicators**	**Secondary indicators**
Natural factors	Natural conditions	Total water resources, precipitation
	Land resources	Forest area, total crop planting area
	Mineral resources	Annual mineral output
Human resources factors	Education structure	Percentage of employed people with no schooling and primary school education, percentage of employed people with junior high school and high school education, percentage of employed people with technical college education and university education
	Labor input	Number of employed people in urban units at the end of the year, average wage of employed people in urban units
Material capital	Level of material capital	Capital stock, urban industrial land
	Road mileage	Railway operating mileage, highway mileage and inland waterway mileage
Technology factors	Technology input and output	Full-time equivalent of research and development personnel, high-tech industry technology introduction expenditure, R&D institution project investment funds, R&D investment intensity, number of domestic patent applications
Structural factors	Economic structure	Urban-rural income gap (Urban residents' income divided by rural residents' income) Urbanization rate (The ratio of urban population to total population)
Institutional factors	Openness to the outside world	Total import and export as a percentage of GDP
	Government intervention	Local fiscal expenditure as a percentage of GDP
	Degree of economic privatization	Number of self-employed employees in private enterprises and employees in other units divided by total employees
	Degree of industry monopoly	Marketization index



(a) Number of trees



(b) Maximum depth



(c) Minimum node size

FIGURE A1. CROSS-VALIDATION RESULTS OF HYPERPARAMETERS

Note: This figure illustrates the evolution of various hyperparameters during cross-validation. Panel (a) shows that the predictive performance of the random forest model improves as the number of trees increases, eventually stabilizing. However, the depth of each decision tree determines the overall complexity of the model. While more complex models are generally more prone to overfitting, the inherent randomness of the random forest algorithm effectively mitigates this risk under most circumstances. Panel (b) demonstrates a positive correlation between model performance and model complexity, with performance peaking when the tree depth reaches seven. The sample fraction, which represents the proportion of original samples selected for training, is determined by random sampling without replacement. Finally, panel (c) depicts the relationship between the minimum node size and the predictive performance of the random forest. A comprehensive comparison reveals that the predictive performance of the model initially increases with a higher sample proportion but eventually plateaus or declines.

TABLE A2—COMPARISON OF THE INPUT-OUTPUT TABLES BETWEEN 149 AND 153 DEPARTMENTS

149 Department	153 Department
Other general equipment subordinate departments	Ovens, fans, packaging and other equipment
Other special equipment subordinate departments	Medical instruments and equipment
Housing construction	Residential housing construction Stadiums and other housing construction
Civil engineering construction	Railway, road, tunnel and bridge engineering construction Other civil engineering construction

TABLE A3—THE CORRESPONDENCE TABLE FOR HIGH-TECH SECTOR AND INPUT-OUTPUT TABLE

Industry classification of the National Bureau of Statistics	Input-output table industry classification	Code
General equipment	Boilers and prime movers	34067
	Metal processing machinery	34068
	Material handling equipment	34069
	Pumps, valves, compressors and similar machinery	34070
	Other general equipment	34072
Special equipment	Mining, metallurgy, construction special equipment	35073
	Chemical, wood, non-metal processing special equipment	35074
	Agriculture, forestry, animal husbandry, fishery special machinery	35075
	Other special equipment	35076
Transportation equipment	Complete vehicles	36077
	Auto parts and accessories	36078
	Railway transportation and urban rail transit equipment	37079
	Ships and related equipment	37080
	Other transportation equipment	37081
Electrical machinery and equipment	Motors	38082
	Power transmission and distribution and control equipment	38083
	Wires, cables, optical cables and electrical equipment	38084
	Batteries	38085
	Household appliances	38086
	Other electrical machinery and	38087
Computers and other electronic equipment	equipment Computers	39088
	Radio and television equipment, radar and supporting equipment	39090
	Audiovisual equipment	39091
	Electronic components	39092
	Other electronic equipment	39093
Communication equipment	Communication equipment	39089
Instruments	Instruments	40094
Cultural and office machinery	Cultural and office machinery	34071

TABLE A4—NEW SECTORS AFTER SHOCK

Year	Pre-Shock Outdegree	Post-Shock Outdegree	New Linkages
2017	3222	3287	**General Equipment** → Catering; Internet and Related Services; Insurance; Water Conservancy Management; Radio, Television, Film and Video Recording Production; Social Security; **Special Equipment** → Forest Products; Fishery Products; Agricultural, Forestry, Animal Husbandry and Fishery Service Products; Wool Textiles and Dyeing and Finishing Products; Glass and Glass Products; Metal Products; Agricultural, Forestry, Animal Husbandry and Fishery Special Machinery; Waste Resources and Waste Materials Recycling Products; Retail; Railway Passenger Transport; Urban Public Transport and Highway Passenger Transport; Catering; Business Services; Professional and Technical Services; Resident Services; Social Work; Sports; Social Security; **Transportation equipment** → agricultural products; forest products; animal husbandry products; textiles, clothing and accessories; leather, fur, feathers and their products; shoes; iron and ferroalloy products; batteries; wholesale; retail; railway passenger transport; railway freight transport and transportation auxiliary activities; Internet and related services; capital market services; business services; resident services; education; sports; social security; public administration and social organizations; **Electrical machinery and equipment** → wholesale; multimodal transport and transport agents; Internet and related services; capital market services; business services; water conservancy management; social security; public administration and social organizations; **Computers and other electronic equipment** → agricultural products; postal services; business services; social security; **Communication equipment** → forest products; fishery products **General equipment** → wholesale; handling and warehousing; postal services; telecommunications; Internet and related services; capital market services; insurance; ecological protection and environmental governance; social security
2018	3221	3290	**Special equipment** → forest products; animal husbandry products; tobacco products; rubber products; refractory products; other general equipment; waste resources and waste materials recycling products; wholesale; railway passenger transportation; postal service; monetary finance and other financial services; insurance; real estate; ecological protection and environmental governance; education; radio, television, film and video recording production; sports; social security; public administration and social organizations **Transportation equipment** → agricultural products; forest products; knitting or crochet and its products; textile products; steel; chemical, wood, non-metal processing special equipment; power transmission and distribution and control equipment; building decoration, decoration and other construction services; wholesale; railway freight transportation and transportation auxiliary activities; postal service; telecommunications; monetary finance and other financial services; insurance; real estate; ecological protection and environmental governance; resident services; social work; radio, television, film and video recording production; entertainment; social security **Electrical machinery and equipment** → wholesale; multimodal transport and transport agency; postal service; Internet and related services; software services; capital market services; water conservancy management; social security **Computers and other electronic equipment** → agricultural products; postal service; business services **Communication equipment** → social security **Instruments and meters** → Capital market services

Table A5—: ε for Complementary Industry

Industry	ε
Accommodation	0.038
Agricultural products	0.003
Agriculture, forestry, animal husbandry and fishery service products	0.073
Air cargo transport and transport support activities	0.055
Air passenger transportation	0.059
Alcohol and wine	0.03
Animal husbandry products	0.01
Aquatic products	0.046
Arts and crafts	0.14
Audiovisual equipment	0.119
Auto parts and accessories	0.011
Basic chemical raw materials	0.018
Batteries	0.06
Beverages	0.03
Boilers and prime movers	0.135
Bricks, tiles, stones and other building materials	0.605
Broadcasting, television and satellite transmission services	0.374
Broadcasting, television, film and video recording production	0.13
Building decoration, renovation and other construction services	0.062
Building installation	0
Business services	0.006
Capital market services	0.045
Catering	0.011
Cement, lime and gypsum	0.351
Ceramic products	0.461
Chemical fiber products	0.07
Chemical, wood, non-metal processing equipment	0.253
Civil engineering	0
Coal mining and washing products	0.014
Coal processing	0.071
Communication equipment	0.026
Complete vehicles	0.014
Computers	0.034
Condiments, fermented products	0.077
Convenient food	0.063
Cotton, chemical fiber textiles and printing and dyeing products	0.015
Cultural and office machinery	0.283
Cultural, educational, sports and entertainment products	0.072

Continued on next page

Table A5 – continued from previous page

Industry	ε
Culture and art	0.349
Daily chemical products	0.058
Dairy products	0.125
Ecological protection and environmental management	0.314
Education	0.022
Electricity, heat production and supply	0.006
Electronic components	0.011
Entertainment	0.03
Fertilizer	0.034
Fishing products	0.022
Forest products	0.094
Furniture	0.128
Gas production and supply	0.033
Glass and glass products	0.103
Grain mill products	0.027
Graphite and other non-metallic mineral products	0.135
Gypsum, cement products and similar products	0.54
Health	0.02
Hemp, silk textiles and processed products	0.155
Household appliances	0.058
Housing	0
Information technology services	0.035
Instruments	0.07
Insurance	0.019
Internet and related services	0.046
Iron and ferroalloy products	0.087
Knitting or crocheting and its products	0.151
Leasing	0.075
Leather, fur, feathers and their products	0.037
Loading and unloading and warehousing	0.083
Machinery for agriculture, forestry, animal husbandry and fishery	0.394
Material handling equipment	0.216
Metal processing machinery	0.245
Metal products	0.017
Metal products, machinery and equipment repair services	0.382
Mining support activities and other mining products	0.025
Mining, metallurgy, construction equipment	0.037
Monetary and other financial services	0.003
Motors	0.123
Multimodal transport and transport agents	0.064
News and publishing	0.143

Continued on next page

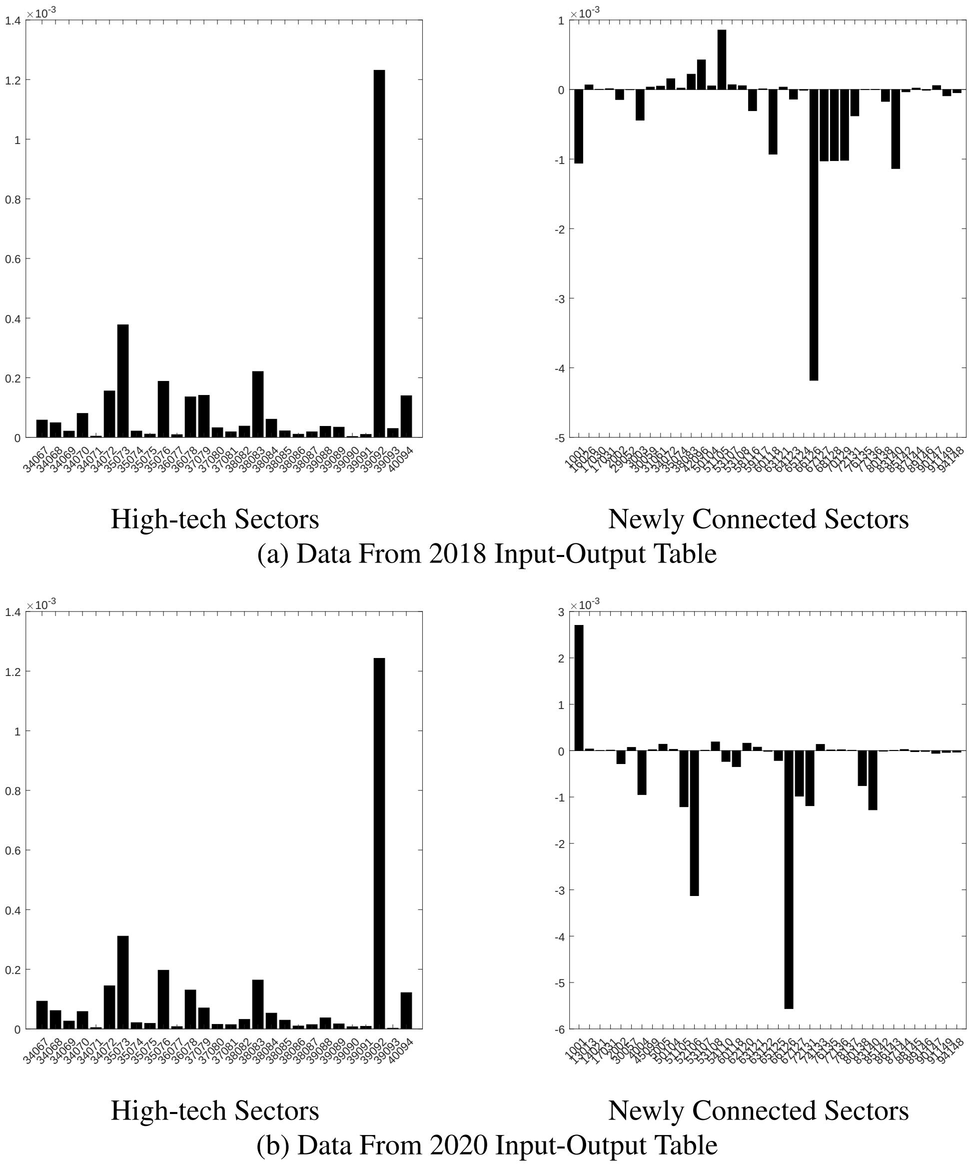
Table A5 – continued from previous page

Industry	ε
Non-ferrous metal mining and dressing products	0.044
Non-ferrous metal rolling products	0.049
Non-ferrous metals and their alloys	0.013
Non-metallic mineral mining and dressing products	0.153
Other electrical machinery and equipment	0.248
Other electronic equipment	0.104
Other foods	0.032
Other general equipment	0.062
Other manufacturing products	0.094
Other services	0.031
Other special equipment	0.059
Other transportation equipment	0.134
Paints, inks, pigments and similar products	0.115
Paper and paper products	0.019
Pesticides	0.074
Pharmaceutical products	0.012
Pipeline transportation	0.154
Plastic products	0.02
Postal	0.042
Power transmission and distribution and control equipment	0.024
Printing and recording media reproduction	0.028
Processed feed products	0.055
Processed vegetable oil products	0.018
Products of ferrous metal ore mining	0.041
Products of oil and gas extraction	0.032
Products of refined petroleum and nuclear fuel processing	0.009
Professional technical services	0.093
Public administration and social organizations	0.122
Public facilities and land management	0.169
Pumps, valves, compressors and similar machinery	0.183
Radio and television equipment, radar and supporting equipment	0
Railway freight transport and transport support activities	0.091
Railway passenger transportation	0.058
Railway transportation and urban rail transportation equipment	0.26
Real estate	0.003
Refined tea	0.138
Refractory products	0.154
Research and experimental development	0
Residential services	0.017
Retail	0.007
Road freight transport and transport support activities	0.009

Continued on next page

Table A5 – continued from previous page

Industry	ε
Rubber products	0.047
Ships and related equipment	0.394
Shoes	0.061
Slaughter and meat processing	0.01
Social security	0
Social work	0
Software services	0.293
Special chemical products and explosives, pyrotechnics, and fireworks products	0.022
Sports	0.462
Steel	0.212
Steel rolling products	0.032
Sugar and sugar products	0.105
Synthetic materials	0.029
Technology promotion and application services	0.058
Telecommunications	0.019
Textile products	0.157
Textiles, clothing and accessories	0.02
Tobacco products	0.04
Urban public transport and road passenger transport	0.033
Vegetables, fruits, nuts and other agricultural and sideline food products	0.024
Waste resources and waste materials recycling products	0.042
Water freight transport and transport support activities	0.072
Water management	0.205
Water passenger transport	0.383
Water production and supply	0.106
Wholesale	0.013
Wires, cables, optical cables and electrical equipment	0.067
Wood processing and wood, bamboo, rattan, palm and grass products	0.037
Wool textiles and dyeing and finishing products	0.111



High-tech Sectors Newly Connected Sectors

(a) Data From 2018 Input-Output Table

High-tech Sectors Newly Connected Sectors

(b) Data From 2020 Input-Output Table

FIGURE A2. PERCENTAGE CHANGES IN EMPLOYMENT SHARES

Note: The codes and names of the newly added departments in the table of 149 department production and output correspond to the following: 1001-agricultural products; 2002-forest products; 3003-livestock products; 4004-fishery products; 5005-agricultural, forestry, animal husbandry and fishery service products; 17028-wool textiles and dyeing and finishing products; 18032-textile clothing and accessories; 19033-leather, fur, feathers and their products; 19034-shoes; 30057-glass and glass products; 31063-iron and ferroalloy products; 33066-metal products; 35075 - Machinery for agriculture, forestry, animal husbandry and fishery; 38085 - Batteries; 42096 - Recycled products of waste resources and waste materials; 51105 - Wholesale; 52106 - Retail; 53107 - Railway passenger transport; 53108 - Railway freight transport and transport support activities; 54109 - Urban public transport and road passenger transport; 58116 - Multimodal transport and transport agents; 60118 - Postal service; 62120 - Catering; 64123 - Internet and related services; 67127 - Capital market services; 68 128—Insurance; 72131—Business services; 74133—Professional and technical services; 76135—Water conservancy management; 80138—Residential services; 83140—Education; 85142—Social work; 87144—Radio, television, film and video recording production; 89146—Sports; 94148—Social security; 91149—Public administration and social organizations; 16026—Tobacco products; 17030—Knitting or crocheting and its products; 29052—Rubber products; 30059—Refractory materials Products; 31061-Steel; 34072-Other general equipment; 35074-Special equipment for chemical, wood and non-metal processing; 38083-Power transmission and distribution and control equipment; 50104-Architectural decoration, renovation and other construction services; 59117-Loading and unloading and warehousing; 63122-Radio, television and satellite transmission services; 65125-Information technology services; 77136-Ecological protection and environmental management; 78137-Public facilities and land management; 86143-News and publishing; 88145-Culture and art.