Federal Reserve Bank of Cleveland Working Paper Series

Rising Skill Supply, Technological Changes, and Innovation: A Quantitative Exploration of China

Shijun Gu and Chengcheng Jia

Working Paper No. 25-15

June 2025

Suggested citation: Gu, Shijun and Chengcheng Jia. 2025. "Rising Skill Supply, Technological Changes, and Innovation: A Quantitative Exploration of China." Working Paper No. 25-15. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202515.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

Rising Skill Supply, Technological Changes, and Innovation: A Quantitative Exploration of China*

Shijun Gu† Chengcheng Jia‡

June 10, 2025

Abstract

Can the expansion of higher education lead to firm productivity growth? In this paper, we examine how China's college expansion program contributes to the rapid growth of firms' R&D expenditure and productivity. In our model, heterogeneous firms make endogenous R&D decisions, requiring them to allocate skilled workers between production and R&D. We structurally estimate the model using firm-level data on the level and distribution of R&D, as well as macro-level data on skill prices and sectoral allocation. Quantitative analysis reveals that between 2004 and 2018, the combination of the R&D-sector-biased technology shock, the skill-biased technology shock, and the skilled-labor supply shock leads to a 12 percent increase in total factor productivity (TFP), of which one-fifth is explained by the rising supply of skilled labor. Counterfactual analysis shows that a further increase in the share of skilled labor has the potential to increase TFP by an additional 2 percent, but the marginal effect diminishes due to the rising wages of unskilled labor.

Keywords: R&D, TFP, Skilled Labor, College Expansion, Chinese Economy
JEL classification: O31, O32, J24

*We are grateful to Kaiji Chen, Dirk Krueger, Ayse Imrohoroglu, Helu Jiang, Vincenzo Quadrini, Kjetil Storesletten, and Le Xu, as well as seminar and conference participants at the Federal Reserve Bank of Cleveland, 2022 Asian Meeting of the Econometric Society, and 3rd Forum of Chinese Macroeconomic Scholars for helpful comments and conversations. The views stated herein are those of the authors and are not necessarily those of the Federal Reserve Bank of Cleveland or the Board of Governors of the Federal Reserve System.

†School of Economics, Central University of Finance and Economics, China. Email: shijungu@cufe.edu.cn
‡Federal Reserve Bank of Cleveland, US. Email: Chengcheng.Jia@clev.frb.org

1 Introduction

During the past two decades, China has been striving to transition from a capital-investment-driven economy to an innovation-driven economy to sustain economic growth. In the industrial sector, the aggregate R&D-expenditures-to-sales ratio increased from 0.6 percent in 2004 to 1.2 percent in 2018. As emphasized by Romer (2000) and Bloom, Van Reenen, and Williams (2019), innovation activity needs workers who have the skills to carry out research; thus, increasing the quantity of innovative activity requires increasing the supply of skilled workers, which may be achieved by policy reforms in the education system. In this paper, using China's higher-education expansion policy as an example, we explore to what extent increasing the supply of skilled labor can foster firm innovation and promote total factor productivity (TFP) growth.

In our model, heterogeneous firms make endogenous R&D choices to improve their productivity. One key novelty of the model is that firms decide how to allocate skilled workers between the production sector and the R&D sector. Wages of skilled and unskilled labor are determined by the equilibrium conditions of the labor market. An increase in the share of skilled workers reduces the wage premium, all else equal, and lowers the cost of R&D. We calibrate three aggregate shocks: a labor market shock that captures the increase in the share of skilled workers caused by the college expansion program, a skill-biased technology shock that increases the labor productivity of all skilled workers, and a sector-biased technology shock that increases productivity in the R&D sector. Together, these three factors lead to an 11.8 percent increase in TFP between 2004 and 2018, of which 2.2 percent is explained by the college expansion program. We predict that the share of skilled workers will continue to rise, but the effect on TFP is diminishing. TFP can potentially further increase by roughly 2 percent before it reaches a plateau.

We start by empirically documenting some salient trends in Chinese firms' R&D expenditure and China's labor market. We first show a dramatic increase in both aggregate R&D expenditures and the R&D-to-sales ratio in China's industrial sector since 2004. We argue that the concurrent massive expansion of higher education, which significantly increased the supply of skilled labor, has been a key contributing factor. We also find that an increasing share of skilled workers was allocated to the R&D sector, and at the same time the research productivity in the R&D sector (measured by patents granted per R&D worker) increased significantly. This suggests that a technology shock that is biased toward the R&D sector may also contribute to the expansion of the R&D sector. Additionally, we observe that the college wage premium has remained largely stable despite a substantial rise in the share of college-educated workers, indicating the existence of a skill-biased technological change.

These macro-level trends suggest that the surge in firms' R&D expenditure in China may be driven by three shocks: a labor supply shock, an R&D sector-biased technology shock, and a skill-biased technology shock.

In the empirical section, we also provide suggestive evidence of the determinants of firms' R&D decisions and the effects of R&D on productivity growth using firm-level data, which contain information on firms' production variables, balance sheets, and R&D expenditures. We find that the fraction of firms engaging in R&D activities (referred to as "R&D-active firms" henceforth) increases with productivity levels and asset levels. These empirical findings imply that financial frictions play a significant role in firms' R&D decisions, highlighting the need to incorporate financial constraints into our model.

We then develop a model of heterogeneous firms with endogenous R&D decisions under financial frictions. In the economy, workers are heterogeneous in their labor productivity and education levels. We classify workers with a bachelor's degree as skilled labor. There is a continuum of firms that are heterogeneous in their productivity and net worth. All firms produce homogeneous manufactured goods using a decreasing-returns-to-scale production technology that combines capital, skilled labor, and unskilled labor, where capital demand is subject to a collateral constraint.

In each period, firms decides whether to perform R&D and how many *ideas* (the output of R&D activity) to produce. The cost of R&D consists of a fixed cost incurred every period and a variable cost, which includes wage expenses for skilled workers and expenditures on intermediate inputs. If a firm chooses not to perform R&D, its productivity evolves according to an AR(1) process. If a firm invests in R&D, its productivity in the next period improves on top of the AR(1) process, with the magnitude of the improvement depending on the number of R&D *ideas* it produces. The endogenous R&D decisions create a trade-off between current-period profits and future productivity gains.

We calibrate model parameters by matching the stationary equilibrium of our model with the Chinese economy in the early 2000s. We employ a standard two-step calibration strategy. First, a subset of parameters is selected outside the model. Second, the remaining parameters, reflecting unique characteristics of the Chinese economy, are internally calibrated to match the aggregate and distributional patterns in China's Annual Survey of Industries (ASI) and the China Economic Census.[1] Under our calibration, the R&D production function exhibits a higher degree of decreasing returns to scale compared to the manufactured goods production function, and the two input factors are gross substitutes. Doubling the amount of R&D ideas leads to a 6 percent increase in productivity in the following period. To test our model's performance, we also compare the model predictions and the data counterparts

[1]China Economic Census Yearbooks are available at https://www.stats.gov.cn/sj/pcsj/.

for the moments that we do not explicitly target in the calibration.

We feed three aggregate shocks into the estimated model to capture all relevant factors driving the rapid growth in firms' R&D investment between 2004 and 2018. The first shock is an increase in the share of skilled labor from 4 percent to 14 percent due to the college expansion program. The second shock is the skill-biased technological shock that is constructed to match the change in the college wage premium, The third shock is the R&D-sector-biased technological shock that is calibrated to match the change in the R&D expenditures-to-output ratio in the industrial sector.

The combined effect of the three shocks leads to a 154 percent increase in R&D ideas per R&D worker. In the data, patents per R&D worker increased by 144 percent, aligning with our model's prediction. Our model predicts that the combined effect of all three shocks leads to a 12 percent increase in TFP. The influx of skilled workers reduces the labor cost of R&D for a given number of *ideas* produced. Both the sector-biased and the skill-biased technological changes generate more *ideas* for a given amount of R&D inputs.

The model implies that the R&D sector expands the employment of skilled labor to a larger extent than the production sector in response to aggregate shocks, which is consistent with the data. There are two main reasons for this result. First, an R&D-sector-biased technological shock disproportionately favors the production of R&D *ideas*, thereby increasing the demand for skilled labor in the R&D sector relative to the production sector. Second, because the substitutability between skilled labor and other inputs is higher in the R&D sector than in the production sector, the positive skill-biased technology shock leads to a higher increase in the demand for skilled labor in the R&D sector than in the production sector.

To isolate the effect of the rising skilled labor supply, we simulate a counterfactual economy in which the share of skilled labor is fixed at its 2004 level. In this case, the model implies that the skilled wage rate would be 158 percent higher than in the benchmark model, implying a significantly higher cost of R&D. As a consequence, the total number of *ideas* produced by the R&D sector would decrease by 24 percent, but the aggregate R&D expenditure is reduced by only 7.2 percent. The smaller reduction in R&D expenditure relative to R&D output implies a decline in R&D productivity. This is because, when firms face a higher wage rate for skilled labor, they substitute away from skilled workers to intermediate goods inputs when producing R&D *ideas*. As firms use fewer skilled workers in R&D, they also get less benefit from the skill-biased technology shock that largely improves the productivity of R&D workers in the benchmark economy. Our simulation suggests that 2.2 percent of the model-implied TFP improvement would be lost without the college expansion program.

China's college expansion program has primarily increased the college enrollment rate among the population under 40 years old. As younger cohorts gradually replace older ones in the labor market, the share of skilled workers will continue to rise in the coming decades. However, our model predicts a diminishing marginal effect on aggregate TFP in the long run. We find that if the share of skilled labor reaches 50 percent, TFP may only increase by additional 2 percent. This is because the equilibrium effect of a rising wage for unskilled workers, which reduces firms' after-tax profit, making firms less willing to invest in R&D.

Related Literature

Our paper connects three strands of the literature. First, we contribute to the growing literature that studies the drivers and consequences of the recent surge of R&D expenditure in China. Most of the previous research focuses on fiscal policies that promote firms' R&D, including institutional factors (Hu and Jefferson (2009)) and tax incentives, (Li (2012), Dang and Motohashi (2015), Jia and Ma (2017), and Dai and Wang (2019)). Chen et al. (2021) also point out that the increase in reported R&D could be partly driven by relabeling other expenses as R&D. Ma (2024) is the first to evaluate the effects of the increase in the share of skilled labor on firms' innovation choices in a quantitative model. Ma (2024) emphasizes how the effects of a labor supply shock on decisions to innovate interact with trade and industry structure, whereas we emphasize how the effects of a labor supply shock also depend on input substitutability and technological change.

Second, our paper is related to the literature on the connection between human capital and innovation. The previous literature has provided empirical evidence (e.g., Aghion et al. (2009), Toivanen and Väänänen (2016), and Aghion et al. (2017)) and theoretical explanations (e.g., Akcigit, Pearce, and Prato (2025) and Bloom, Van Reenen, and Williams (2019)) for the relationship between human capital, firm R&D, and productivity growth. In the context of China, Che and Zhang (2018) show empirical evidence that the policy-induced increase in college graduates in China leads to more productivity growth in human-capital intensive industries.

Third, our paper contributes to the large literature on the barriers to productivity growth in developing economies. A large body of research on economic growth has pointed out that differences in aggregate TFP between rich and poor countries could be largely due to resource misallocation, and resource reallocation has been a key driver of growth in many developing economies, including the Chinese economy (e.g. Hsieh and Klenow (2009), Song, Storesletten, and Zilibotti (2011), Brandt, Tombe, and Zhu (2013), and Midrigan and Xu (2014)). Our paper is more related to the recent efforts in this field that incorporate firms' endogenous R&D decisions in the presence of idiosyncratic distortions in China, for example

König et al. (2022) and Buera and Fattal Jaef (2018). Relative to König et al. (2022) and Buera and Fattal Jaef (2018), we consider the limited supply of skilled labor capable of conducting R&D as a barrier to innovation, and thus to TFP growth, and quantify the effects of alleviating such barriers, particularly in the presence of technological changes.

2 Motivating Facts and Descriptive Evidence

In this section, we first examine the changes in the share of skilled labor and the wage premium since China started the college expansion program in 1999 (Section 2.1). In Section 2.2, we demonstrate that the increasing share of skilled labor coincides with a substantial surge in R&D investment in the industrial sector. Furthermore, we document both an expanding allocation of skilled labor to the R&D sector and rising research productivity among R&D workers. Finally, in Section 2.3, we present firm-level evidence concerning the determinants of firms' R&D decisions and their effects on productivity growth.

2.1 The Change in the Skill Composition and Skill Prices

In China, university education is highly controlled by the state. The Ministry of Education manages the admission process and assigns quotas of admitted students in each year. In 1999, the central government initiated the "college expansion program." As a result, the number of college graduates increased from slightly above 1 million to near 4 million between 2004 and 2018.[2] (Figure 1, Panel A) In the same panel, the college graduate share (the share of college graduates in the population aged 22) increased from approximately 5 percent to 19 percent between 2004 and 2018.

In addition to enrolling more college students, China has implemented other higher education reforms, such as promoting world-class research universities, increasing research funding, encouraging international collaboration, and adopting a more flexible college entrance exam system to improve educational quality and accessibility.

Does the rapid influx of skilled labor push down the college wage premium? We study this question by performing a Mincer regression using micro-level data from the Chinese Household Income Project(CHIP). CHIP data provide rich self-reported information on respondents' labor market activities, including annual salary, working hours, industry, etc.; educational attainment; and other characteristics, including gender, age, and household composition. We use the 2002, 2007, 2013, and 2018 waves of CHIP surveys to study the

[2] We only account for 4-year college programs and do not account for 2- or 3-year vocational college programs. Data source: China Statistical Yearbooks, available at https://www.stats.gov.cn/sj/ndsj/.

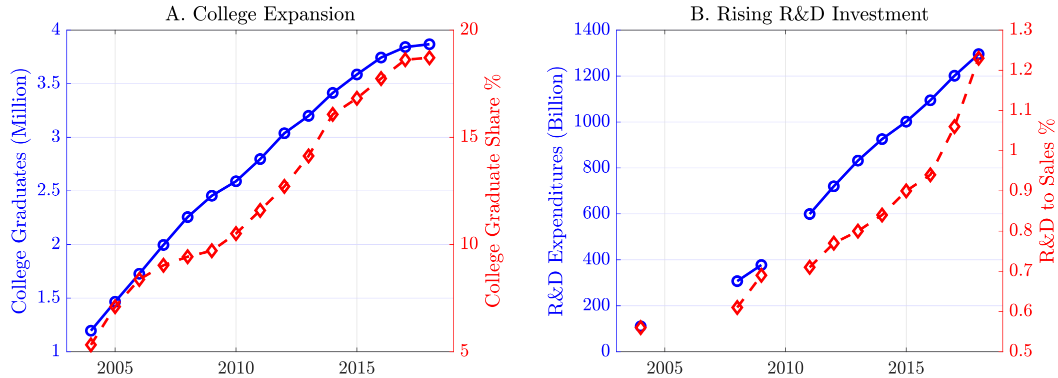

Figure 1: College Expansion and Rising R&D Investment in China

Note: College graduates refers to the number of four-year college degree recipients in the respective years, measured in millions. College graduate share refers to the ratio of college graduates to the total population aged 22. In our model calibration, we use the urban population instead of the total population, as we focus on productivity and R&D investments in the industrial sector. R&D expenditures refers to the total R&D spending of all industrial enterprises above a designated size, measured in billions of RMB. R&D to sales refers to the ratio of R&D expenditures to sales revenue for all industrial enterprises above a designated size. Data sources include the China Statistical Yearbooks and the China Economic Census Yearbooks.

change in the wage premium for college graduates. One unique feature of CHIP data is that the 2007, 2013 and 2018 waves of the CHIP survey also contain self-reported information on respondents' college entrance exam (Gaokao) scores, which can be regarded as a measure of cognitive skills and a proxy for labor productivity.[3]

To estimate the college wage premium in each of the four waves, we need to control for other factors that may affect individuals' wage rate. Following the literature, we run the following regression:

$$ln(wage_i) = \beta_0 + \gamma_1 Col_i + \gamma_2 Cog_i + \beta_1 YE_i + \beta_2 YE_i^2 + \beta_3 X_i + \epsilon_i \tag{1}$$

where Col_i is a dummy variable, indicating if the person has graduated from a four-year college. Cog_i denotes cognitive capability (normalized Gaokao score). YE_i is the number of years of work experience. X_i controls for other labor characteristics, including gender and province of residence.

After controlling for individual characteristics, including the college entrance exam score, the wage premium slightly decreased from 1.587 ($e^{0.462}$) in 2007 to 1.474 ($e^{0.388}$) in 2018 (Table 1). Considering the large increase in the share of the labor force that has a college degree, the decrease in the wage premium is rather moderate.

[3]A summary of statistics and our sample selection procedure are provided in Appendix A.

Table 1: Estimating the College Wage Premium in China

	2002	2007		2013		2018	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
4-year college	0.446	0.531	0.462	0.343	0.296	0.433	0.388
	(0.020)	(0.028)	(0.030)	(0.023)	(0.025)	(0.019)	(0.022)
Gaokao z-score			0.092		0.059		0.050
			(0.015)		(0.013)		(0.011)
experience	0.021	0.038	0.039	0.039	0.040	0.040	0.042
	(0.003)	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)
experience^2	-0.000	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No. of observations	7193	2158	2158	3143	3143	4839	4839
R-squared	0.132	0.188	0.201	0.133	0.139	0.136	0.140

Notes: 4-year college is a dummy variable indicating the completion of four years of college. Gaokao z-score is a proxy for individual ability. (See Appendix A for the normalization method for Gaokao score.) Experience is the number of years of work experience. We also controlled for individual characteristics, including gender and province of residence.

2.2 R&D Expenditure, Employment, and Productivity

As shown in Panel B in Figure 1, along with the rising share of workers with college degrees there is a surge in aggregate R&D expenditure. Total R&D expenditure increased approximately eight times between 2004 and 2018, while the R&D-to-sales ratio rose from around 0.6 percent to 1.2 percent.

In this paper, we conjecture that the increasing supply of college-educated labor can promote firm innovation in China, as skilled workers are essential inputs for R&D tasks. Data on the use of R&D workers and research output in Chinese industrial enterprises from 2004 to 2018 provide further evidence for our hypothesis.

The first two rows of Table 2 show that the number of workers engaged in R&D tasks in Chinese industrial enterprises increased 5.5-fold from 2004 to 2018, and the share of R&D personnel in aggregate employment rose from 0.8 percent to 3.6 percent.[4] Meanwhile, as shown in the third and fourth rows, the share of wage expenditures for R&D workers in total R&D spending increased by nearly 12 percentage points,[5] and an increasingly larger

[4]The number of R&D workers is measured by full-time equivalents (*R&D quan shi dang liang*), which accounts for both full-time and part-time employees involved in R&D. The number of part-time R&D workers is converted to full-time equivalents using their average working hours on R&D activity.

[5]A widely known issue in Chinese plant-level data is that the National Bureau of Statistics (NBS) understates actual labor cost. For example, the aggregate share of wage cost is roughly 30 percent, which is significantly lower than the aggregate labor share in the manufacturing sector in the Chinese national account. This inconsistency is also documented in Bai, Hsieh, and Qian (2006) and Hsieh and Klenow (2009) A possible explanation for this inconsistency is that the NBS only reports wage income but does not provide information on non-wage benefits, including bonuses, pensions, insurance, and housing allowance. As a result, when calculating the cost share of labor in R&D activities, we assume that non-wage benefits are a fixed portion of total labor compensation, and we adjust the raw data (22.5 percent in the 2004 China Economic Census Yearbook) by the same factor (1.6), following Bai, Hsieh, and Qian (2006) and Hsieh and

proportion of college-educated workers were assigned to R&D tasks. These data not only reflect a significant rise in the demand for skilled labor capable of conducting R&D tasks but also highlight the growing importance of skilled labor in the process of firm innovation.

Table 2: R&D Workers: Employment, Cost, and Productivity

	2004	2008	2013	2018
Skill share and allocation				
R&D workers ($\times$1,000)	542	1,230	2,494	2,981
Share of R&D worker in labor force %	0.82	1.39	2.55	3.57
R&D workers cost share %	36.0	32.5	43.7	47.9
Share of R&D worker in skilled labor force %	21.2	20.3	28.7	31.2
Patent production				
Patent applications ($\times$1,000)	20	59	205	371
Patent applications per R&D worker	0.04	0.05	0.08	0.13
Patents granted per R&D worker	0.014	0.016	0.021	0.034
Patent application approval rate %	37.9	32.3	25.2	28.0

Note: R&D workers refers to the number of employees engaged in R&D tasks in industrial enterprises in the respective years, multiplied by their average hours spent on R&D tasks. Share of R&D workers in labor force refers to the number of R&D workers divided by the total number of employees in industrial enterprises. R&D workers cost share refers to the ratio of compensation for R&D workers to total R&D expenditures. Share of R&D workers in skilled labor force refers to the ratio of the number of R&D workers to the number of college-educated workers. Patent applications refers to the total number of invention patent applications filed by industrial enterprises in the respective years. Patent applications per R&D worker refers to the ratio of the total number of invention patent applications to the number of R&D workers. Patents granted per R&D worker refers to the ratio of the total number of invention patents granted to the number of R&D workers. Patent application approval rate refers to the ratio of the total number of invention patents granted to the total number of invention patent applications. Data sources include the China Statistical Yearbook and the China Economic Census.

The lower panel of Table 2 shows China's R&D output, which is measured by patent production, and R&D productivity, which is measured by patents per R&D worker. To control for the quality of patents, patent data in Table 2 only include invention patents and exclude utility model patents and design patents. From 2004 to 2018, the number of patent applications filed by Chinese industrial enterprises increased more than 18.5-fold, and the number of patent applications per R&D worker increased by 255 percent. If we only consider patents that are eventually granted, R&D workers' productivity, measured by patents granted per R&D worker, increased by 144 percent. These changes reflect that R&D workers have become more productive.[6]

Klenow (2009).

[6]One might argue that the significant rise in patents per R&D worker in China could have stemmed from the government's stricter enforcement of patent laws rather than an improvement in research productivity. This argument is more plausible in the early 2000s, when the Chinese government took significant actions

2.3 Firm-Level Evidence on R&D

In this section, we use firm-level data to investigate the determinants and the effects of R&D expenditure. We obtain firm-level data from the Annual Survey of Industries (ASI) conducted by China's National Bureau of Statistics, which is a census of all state-owned firms and private firms in the manufacturing industry with annual revenue above 5 million RMB.[7]

To investigate how firms' R&D decisions are affected by other factors, we do the following regression analysis using a linear probability model. The dependent variable is a dummy variable whose value is 1 for R&D firms, and 0 for non-R&D firms. The main independent variable is the firm's TFP in the current year,[8] Columns (1) - (3) in Table 3 show that the fraction of R&D-active firms robustly increases with firm TFP, after controlling for the amount of equity in the firm and the firm's ownership type.[9] In addition, firms with higher net worth are more likely to invest in R&D. State-owned enterprises (SOEs) make more R&D investments compared to private firms.

To further quantify the effects of R&D expenditure on productivity growth, we estimate the following TFP process:

$$\log(TFP)_{i,t+1} = \rho \log(TFP)_{i,t} + \gamma\ X_{i,t} + \epsilon_{i,t+1} \tag{2}$$

where $X_{i,t}$ measures R&D activities. We first estimate whether making positive R&D expenditures increases productivity in the next period. To do so, we let $X_{i,t}$ be the dummy variable taking the value 1 for R&D firms, and 0 for non-R&D firms. Then, for all R&D firms, we test whether making greater R&D investments leads a higher growth rate in the

against patent infringement, in order to meet the the requirements for joining the WTO. However, we find that most of the increase in patents per R&D worker occurred between 2013 and 2018, a period when China's patent protection system had already been well established. This suggests that stricter enforcement of patent laws is not the primary driver of the increase in patents per R&D worker.

[7] We do not use ASI after 2008 because value-added firm output is no longer reported after 2008. Information on R&D is only available in the years 2001 and 2005 - 2007.

[8] We estimate firm TFP using the Solow residual method, which requires a specification of the production function. We assume a decreasing-return-to-scale production function, taking the form of $Y_t = Z_t K_t^{\alpha} L_t^{\theta}$, where Y_t, K_t, and L_t are value-added output, capital, and labor of a firm, correspondingly. To account for the heterogeneity in labor inputs, we use total wage bills to approximate L_t. We chose $\alpha = 0.40$ and $\theta = 0.45$ consistent with Bai, Hsieh, and Qian (2006). We also include control variables of the firm's equity value and ownership type. We also regress TFP (in logs) on start year, province, and 2-digit industry to control for variation caused by these factors.

[9] The dummy variable SOE takes the value 1 if a firm is a state-owned enterprise according to its registration type, or the state has an absolute or relative controlling share in the firm. The two definitions of state ownership (registration type and controlling share) do not always overlap since China started the reform on SOEs in 1998, and since then, a growing fraction of registered SOEs have corporatized. See Hsieh and Song (2015) and Gu and Jia (2022) for more information.

Table 3: Characteristics of Firm R&D Decisions

	(1) R&D$_d$	(2) R&D$_d$	(3) R&D$_d$	(4) log(TFP$_{t+1}$)	(5) log(TFP$_{t+1}$)
log(TFP$_t$)	0.033	0.021	0.028	0.681	0.698
	(0.004)	(0.003)	(0.003)	(0.018)	(0.011)
log(equity)		0.071	0.066		
		(0.005)	(0.005)		
SOE			0.121		
			(0.014)		
R&D$_d$				0.050	
				(0.014)	
log(R&D)					0.032
					(0.004)
R-squared	0.149	0.250	0.259	0.452	0.530

Notes: The dependent variable, $R\&D_d$ is a dummy variable whose value is 1 for R&D-active firms, and 0 for R&D-inactive firms. $\log(TFP)$ is the logarithm of firm TFP. Equity is calculated as the difference between total asset and total debt. SOE is a dummy variable with a value 1 for state-owned enterprises. All regressions include industry, start year, and province fixed effects. Regressions are weighted by the size of total wage bills. Standard errors are clustered at the industry level and are reported in parenthesis. We use a balanced panel between 2001 and 2002.

next period. To this end, we restrict the sample to be all R&D firms, and $X_{i,t} = \log(R\&D)_{i,t}$ is the logarithm of the amount of R&D expenditure.

The last two columns in Table 3 show that the auto-correlation coefficient of the productivity process is roughly 0.7 at an annual rate. Conditional on the current TFP level, making positive R&D investments increases TFP by 5 percent on average in the next period. In addition, conditional on making positive R&D investments, doubling R&D increases TFP in the next period by 3 percent. This number is at the lower end in the range of estimates of the private returns to R&D. Most empirical studies on this topic are based on micro data in developed countries. Hall, Mairesse, and Mohnen (2010) review this literature and find this R&D elasticity parameter to have a broad range between 2 percent and 25 percent. Chen et al. (2021) use a different source of data from the Chinese State Administration of Tax and find the R&D elasticity parameter to be 9 percent from 2006 to 2011.

Summary. In this section, we first present the institutional background of China's college expansion program and provide suggestive evidence linking the increased supply of skilled labor to rising R&D expenditures. Our analysis reveals that the dramatic growth in the share of college-educated workers resulted in only a modest decline in the college wage premium, indicating the presence of skill-biased technological change. Moreover, we document significant increases in both the quantity and the productivity of R&D workers between 2004 and 2018, suggesting the existence of a technological shock that exclusively enhances productivity in the R&D sector. Using firm-level data, we further demonstrate that a firm's likelihood of investing in R&D rises with its productivity and net worth, and that such investments significantly enhance subsequent productivity.

These empirical patterns motivate our theoretical framework, where skilled labor is a crucial input in R&D activities, and firms optimally make R&D investment decisions and allocate skilled labor between the production and R&D sectors. The evidence suggests three potential drivers of R&D expenditure growth: (1) a labor supply shock that increases the availability of skilled labor, (2) an R&D-sector-biased technology shock that enhances research productivity, and (3) a skill-biased technological change that increases the productivity of all skilled labor. The rest of the paper aims to quantify the effects of these shocks on R&D expenditures and TFP in the Chinese economy, with a particular focus on isolating the contribution of the labor supply shock.

3 Model

We develop a heterogeneous-firms model with endogenous R&D decisions. All firms produce a homogeneous manufactured good for sale, and may choose to invest in R&D to improve their firm-specific productivity in the next period. Firms' capital demand may be constrained due to financial frictions, which also affect their decisions on R&D expenditure. We outline the model assumptions in Section 3.1, formulate the recursive problem in Section 3.2, and define the stationary equilibrium in Section 3.3.

3.1 Environment

Time is discrete and the horizon is infinite with $t \in \{1, 2, ..., \infty\}$.[10] There is no aggregate uncertainty.

Demographics. There is a continuum of firms of measure one. The total labor supply is also normalized to one. N_t^s denotes the share of skilled labor, corresponding to the share of workers with a bachelor-or-above degree in the data.[11] The remaining workers, N_t^u, make up the unskilled labor force.

[10] The time subscript indicates that the time variation in the corresponding variable can be treated as an aggregate shock later in Section 5. Otherwise, we omit the time index.

[11] Those who have completed vocational or part-time college degrees are not treated as skilled workers in our model. We set a relatively high bar for skilled workers because in our model, all skilled workers are capable of doing R&D, and R&D requires more advanced skills than part-time or vocational college programs typically provide.

Manufactured good production. All firms produce a homogeneous manufactured good and operate under a decreasing-returns-to-scale technology. Firms are heterogeneous in productivity, and demand skilled labor, unskilled, labor, and capital in production.

The production function of the manufactured good is given by

$$y = F(k, n_{s,y}, n_u) = Azk^{\alpha}\left[\gamma(S_t n_{s,y})^{\sigma} + (1-\gamma)n_u^{\sigma}\right]^{\frac{\theta}{\sigma}}, \tag{3}$$

where A is the aggregate productivity level, and z is the idiosyncratic productivity level. $n_{s,y}$ and n_u denote the efficient units of skilled and unskilled labor used in production, respectively. S_t is a skill biased technology shock that will be calibrated in Section 5. α is capital share, θ is labor share, γ is skilled labor share in producing the manufactured good, and σ controls for the elasticity of substitution between skilled and unskilled labor.

Financial market. In our model, we assume that firms borrow intra-period loans to finance their capital demands. The cost of capital is the sum of the capital rental rate (r) and the depreciation rate of capital (δ). We assume the financial market is imperfect, and firms are subject to a collateral constraint. A firm's choice of capital may be constrained, and the borrowing limit depends on the firm's net worth, a. Specifically, the maximum loan that a firm can take out is given by

$$k \leq \lambda a, \tag{4}$$

where λ is a parameter reflecting the tightness of the financial constraint.

R&D. In each period, firms may pay a fixed cost to perform R&D activities (become R&D-active firms). We assume that the amount of the fixed cost increases with the firm's current productivity, given by fz^{η}, where f and η govern the scale and shape of the fixed cost.

Conditional on becoming an R&D-active firm, the firm then chooses the amount of R&D activities it performs. We call the output of R&D activities "ideas," denoted by x. Ideas are firm-specific, meaning they can only improve firm-specific productivity in the next period and cannot be traded. Firms combine two factors in producing R&D ideas: skilled labor, $n_{s,x}$, and the manufactured good as intermediate inputs, i.

The production function of R&D is given by

$$x = J(n_{s,x}, i) = B_t\left[\psi\left(S_t n_{s,x}\right)^{\kappa} + (1-\psi)i^{\kappa}\right]^{\frac{\mu}{\kappa}} \tag{5}$$

where B_t is the R&D-sector-biased technology shock and is time varying. There is no idiosyncratic technology variation in R&D activity. ψ is the share of skilled labor in the R&D sector, κ controls for the elasticity of substitution between skilled labor and intermediate input, and $\mu < 1$ implies that the production of the R&D ideas exhibits decreasing returns to scale. The skill-biased technology shock, S_t, also affects the productivity of the skilled labor in the R&D sector.

Productivity process. If a firm chooses not to perform R&D (an R&D-inactive firm), its idiosyncratic productivity z follows a standard AR(1) process in logs

$$\log(z') = \rho \log(z) + \epsilon',\ \epsilon' \sim \mathrm{N}(0, \sigma_\epsilon), \tag{6}$$

where ρ governs the persistence of the productivity process and σ_ϵ is the standard deviation of the *i.i.d.* shock, ϵ.

If a firm chooses to perform R&D (an R&D-active firm), its next-period productivity has an endogenous increase on top of the stochastic AR(1) process. The size of the increase depends on the amount of R&D ideas the firm makes in the current period. Specifically, the productivity process on an R&D-active firm follows:

$$\log(z') = \rho \log(z) + \phi \log(1 + x) + \epsilon',\ \epsilon' \sim \mathrm{N}(0, \sigma_\epsilon), \tag{7}$$

where ϕ controls the elasticity of productivity gains with respect to the amount of R&D ideas.

Households. We assume that workers are heterogeneous in their labor productivity, h, which follows a Gamma distribution $\Gamma(\zeta_1, \zeta_2)$. The two shape parameters, ζ_1 and ζ_2, are time invariant and will be calibrated externally (see Section 4.1).

In addition, a worker is either a skilled worker or an unskilled worker, $e \in \{s, u\}$, determined by his education attainment. We do not allow households to choose educational attainment in our model; rather, their educational attainments depend solely on their labor productivity. There is a productivity cutoff, χ_t, above which all individuals receive a college degree and become skilled workers. The change in the productivity cutoff reflects the policy shock of the college expansion program, which we discuss in Section 5.1.

We keep the household problem as simple as possible, as the focus of our analysis is on the firm side. Households maximize their utility function $\log(c)$ subject to the budget constraint:

$$c \leq w_e h + D + TR, \tag{8}$$

where D represents aggregate firm dividends and TR denotes government transfers.

Government policies. The government taxes all profit-making firms at a uniform rate τ_c, and provides proportional subsidies to all R&D-active firms for their R&D expenses at the rate of τ_x.[12]

3.2 A Dynamic Problem of R&D Choices

Firms are heterogeneous in productivity and net worth. At the beginning of each period, after observing their net worth and productivity shocks, firms decide whether to become an R&D-active or an R&D-inactive firm. If a firm decides to be an R&D-inactive firm, it faces a standard intra-period profit-maximization problem and a inter-period dividend-saving decision. If a firm decides to be an R&D-active firm, it pays a fixed cost, and in addition to the profit-maximization and dividend-saving decisions, it chooses the number of R&D ideas to produce. For a given number of R&D ideas, the firm optimally chooses intermediate goods input and skilled labor to minimize the cost of R&D.

Profit maximization. Conditional on the current period's productivity and net worth, all firms maximize their operating profit[13]

$$\pi(a,z) = \max_{k,n_{s,y},n_u} \{F(k,n_{s,y},n_u) - (r+\delta)k - w_s n_{s,y} - w_u n_u\} \tag{9}$$

subject to

$$0 \le k \le \lambda a, n_{s,y} \ge 0, n_u \ge 0 \tag{10}$$

where w_s is the wage rate for skilled labor and w_u is the wage rate for unskilled labor.

R&D cost minimization. To produce a given x units of R&D ideas, the R&D-active firm chooses the optimal amount of intermediate goods input, i, and skilled labor, $n_{s,x}$, to minimize the cost of R&D. The R&D cost-minimization problem can be expressed as:

$$\Omega(x) = \min_{i,n_{s,x}} \{i + w_s n_{s,x}\} \tag{11}$$

subject to

$$x \le J(n_{s,x}, i), i \ge 0, n_{s,x} \ge 0 \tag{12}$$

[12]In reality, China has more complex tax incentives for R&D investments, as extensively explored by Chen et al. (2021)). We abstract from this complexity to focus on the other drivers of firms' R&D decisions.

[13]The operating profit defined here does not include the fixed and variable cost of R&D.

where $\Omega(x)$ is the minimum cost of producing x units of R&D ideas.

Dynamic problem. A firm chooses whether or not to be an R&D-active firm by comparing $v^E(a,z)$ and $v^R(a,z)$, where $v^E(a,z)$ and $v^R(a,z)$ denote the value of an R&D-inactive and an R&D-active firm, respectively. The value of a firm conditional on its net worth and productivity is given by

$$v(a,z) = \max\left\{v^E(a,z), v^R(a,z)\right\}. \tag{13}$$

An R&D-inactive firm maximizes the current period's operating profit and then decides the amount of net worth to carry over to the next period. Its dynamic problem can be expressed as:

$$v^E(a,z) = \max_{c,a'}\left\{u(c) + \beta \mathbb{E}_{z'}\left[v(a',z')\right]\right\} \tag{14}$$

subject to

$$c + a' + \tau_c \pi(a,z) = (1+r)a + \pi(a,z) \tag{15}$$

$$c \geq 0, a' \geq 0 \tag{16}$$

The evolution of productivity z follows Equation (6).

R&D-active firms face a trade-off where increasing R&D expenditure reduces profit in the current period, but increases productivity in the next period. The dynamic problem of an R&D-active firm can be expressed as:

$$v^R(a,z) = \max_{c,a',x}\left\{u(c) + \beta \mathbb{E}_{z'}\left[v(a',z')\right]\right\} \tag{17}$$

subject to

$$T(a,z,x) = \pi(a,z) - f z^{\eta} - (1-\tau_x)\Omega(x) \tag{18}$$

$$c + a' + \tau_c \max\{T(a,z,x), 0\} = (1+r)a + T(a,z,x) \tag{19}$$

$$c \geq 0, a' \geq 0, x \geq 0 \tag{20}$$

where $T(a,z,x)$ is the taxable corporate income, which equals the operating profit less the fixed and variable R&D costs. Productivity z evolves according to Equation (7).

3.3 Definition of Equilibrium

We study a competitive equilibrium for a small open economy, in which the world's interest rate is r. Firms are indexed by individual states $\mathbf{s} = \{a,z\}$. The recursive competitive equilibrium is defined by a set of value functions $\{v(\mathbf{s}), v^E(\mathbf{s}), v^R(\mathbf{s})\}$, the allocation of R&D-active firms $X^R = \{c(\mathbf{s}), a'(\mathbf{s}), k(\mathbf{s}), n_u(\mathbf{s}), n_{s,y}(\mathbf{s}), i(\mathbf{s}), n_{s,x}(\mathbf{s})\}$, the allocation of R&D-inactive

firms $X^E = \{c(\mathbf{s}), a'(\mathbf{s}), k(\mathbf{s}), n_u(\mathbf{s}), n_{s,y}(\mathbf{s})\}$, government policies $\{\tau_c, \tau_x\}$, prices $\{r, w_s, w_u\}$, and the joint cumulative distribution of firms $\Psi = \{\Psi^E(\mathbf{s}), \Psi^R(\mathbf{s}))\}$ such that

1. The value and policy functions solve the maximization problem described in Section 3.2, given factor prices and government policies.

2. The skilled labor market clears

$$\int_{\chi}^{\infty} \Gamma(h;\zeta_1,\zeta_2)dh = \int n_{s,y}(\mathbf{s})d\Psi^R(\mathbf{s}) + \int n_{s,y}(\mathbf{s})d\Psi^E(\mathbf{s}) + \int n_{s,x}(\mathbf{s})d\Psi^R(\mathbf{s}) \tag{21}$$

3. The unskilled labor market clears

$$\int_{0}^{\chi} \Gamma(h;\zeta_1,\zeta_2)dh = \int n_u(\mathbf{s})d\Psi^R(\mathbf{s}) + \int n_u(\mathbf{s})d\Psi^E(\mathbf{s}) \tag{22}$$

4. The amount of government transfers to all households is given by

$$TR = \tau_c \left(\int \pi(\mathbf{s})\Psi^R(\mathbf{s} + \int \pi(\mathbf{s})\Psi^E(\mathbf{s}) \right) - \tau_x \left(\Omega(x(\mathbf{s}) \right) \tag{23}$$

5. The distribution Ψ is a fixed point where its transition is consistent with the policy functions and the law of motion for Ψ:

$$\Psi = \Phi(\Psi) \tag{24}$$

   where Φ is a one-period-ahead transition operator such that $\Psi' = \Phi(\Psi)$.

4 Calibration

The initial steady state of the model is calibrated to the Chinese economy in 2001-2004.[14] We first explain the parameters that are calibrated externally (Section 4.1). The rest of the parameters, which represent the unique features of the Chinese economy, are estimated internally within the model (Section 4.2). Section 4.3 evaluates the performance of our calibration by comparing a set of non-targeted moments generated by our model with their empirical counterparts.

[14] Since the college expansion program was initiated in 1999 and its implementation was a gradual process, the share of skilled labor was stable in the early 2000s. We do not specify a single year as our calibration target, because some data are not released every year.

4.1 Externally Calibrated Parameters

The model is calibrated at an annual frequency. Table 4 lists all the parameters that we calibrated externally. The values we assign to these parameters are either standard in the literature or can be directly measured from the data.

The factor-neutral aggregate TFP level, A, is normalized to one. The degree of decreasing returns to scale is set to be $\alpha + \theta = 0.85$, which is standard in the literature. We set the depreciation rate of capital $\delta = 0.10$, and the capital share parameter $\alpha = 0.40$, consistent with estimates of the Chinese economy in the early 2000s (Bai, Hsieh, and Qian (2006)). The elasticity of substitution between skilled and unskilled labor, σ, is set to 0.30, following Li (2010). The capital rental rate is set to 6 percent.

Table 4: Externally Calibrated Parameters

Parameter		Value	Source
Production			
Total factor productivity	A	1.00	Normalization
Span-of-control (production)	$\alpha+\theta$	0.85	Standard
Share of capital	α	0.40	Bai, Hsieh, and Qian (2006)
Capital depreciation rate	δ	0.10	Bai, Hsieh, and Qian (2006)
Elas. of substitution btw. skills	σ	0.30	Li (2010)
Capital rental rate	r	0.06	Standard
Skill endowment			
Measure of workers	N	1.00	Normalization
Share of skills	N_t^s	0.04	China Population Census
Distribution (scale parameter)	ζ_1	9.15	CFPS 2010-18
Distribution (rate parameter)	ζ_2	0.11	CFPS 2010-18
Government policy			
Corporate income tax rate	τ_c	0.33	State Taxation Administration
R&D subsidy rate	τ_x	0.06	2000 R&D Census

The total supply of workers is normalized to one. The share of skilled labor is set to 0.04, matching the share of the working-age urban population that have college-or-above degrees in 2004.[15] In addition, we use math test scores in the labor force from the China Family Panel Studies (CFPS) to proxy the Γ distribution of labor productivity.[16] In the initial

[15]The Chinese National Bureau of Statistics carries out a population census every ten years, and the most recent ones were carried in 2000, 2010, and 2020. We calculate the share of college graduates in the working-age population (ages 25 - 54) in 2004, utilizing the information on educational attainments across age groups in the 2000 Population Census. (See Appendix B for a detailed breakdown of educational attainments across age groups.) Ideally, we should use data on the educational attainment of workers in the industrial sector, since our firm-level data are taken from the Annual Survey of Industries (ASI). However, the ASI did not release this information except in 2004, and thus we cannot calculate the change in the share of college graduates from the ASI. We thereby use data on urban workers from the Population Census as a proxy. We have compared the two sets of data and confirmed that the educational attainment of all urban workers in the Population Census is similar to that in the industrial sector in the ASI.

[16]Description and data of the CFPS are available at https://www.isss.pku.edu.cn/cfps/index.htm

steady state, labor productivity at the cutoff for skilled labor is 63 percent higher than the average labor productivity. (Figure 5 in Appendix C shows the calibrated labor productivity distribution and the cutoff for college admission.)

The corporate income tax rate is 33 percent, mirroring the tax rate for domestic firms in the early 2000s. The R&D subsidy rate is 6 percent, which is calculated as the aggregate tax deduction for R&D investments, divided by aggregate R&D expenditure.[17]

4.2 Internally Calibrated Parameters

The remaining 12 parameters are internally estimated within the model and are summarized in Table 5. The targeted moments, together with their empirical and model-predicted values, are displayed in Table 6. These parameters are calibrated to replicate the unique aggregate and distributional features of Chinese manufacturing firms, especially with regard to patterns of R&D investments.

Table 5: Internally Calibrated Parameters

Parameter		Value
Preference		
Discount factor	β	0.855
Manufactured good production		
Borrowing constraint	λ	2.839
Skills share in production	γ	0.195
R&D good production		
R&D sector-biased technology	B	0.906
Span-of-control (R&D)	μ	0.720
Skills share in R&D	ψ	0.554
Elas. of substitution	κ	0.693
R&D fixed cost		
Fixed cost, level ($\times$100)	f	0.044
Fixed cost, curvature	η	4.523
Productivity process		
Persistence of productivity	ρ_z	0.699
SD of productivity shocks	σ_ϵ	0.451
Scale effect of R&D	ϕ	0.063

We choose the 12 parameters to match 12 data moments using simulated method of moments (SMM).[18] Even though every targeted moment is determined simultaneously by

[17]Source: China Economic Census in 2004.

[18]Specifically, we minimize the weighted sum of squared percentage differences between the data and model moments. The vector of parameters, Θ, is chosen to minimize the minimum-distance-estimator criterion function

$$f(\Theta) = [\mathbf{m}_{data} - \mathbf{m}_{model}(\Theta)]'\mathbf{W}[\mathbf{m}_{data} - \mathbf{m}_{model}(\Theta)],$$

all parameters, in what follows, we discuss each of the moments in relation to the parameter for which the moment yields the most identification power.

Preferences. We choose the discount factor $\beta = 0.855$ to match the aggregate assets-to-output ratio. A higher β means firms are more patient and thus maintain a higher level of assets.

Manufactured good production. We choose the borrowing constraint $\lambda = 2.839$ to target the aggregate firms' debt-to-capital ratio (leverage ratio) of 0.563, following the standard practice in the literature (Karabarbounis and Macnamara (2021) and Ottonello and Winberry (2024)). γ is calibrated to match the wage premium of skilled workers in 2007.[19]

R&D fixed cost. The two parameters that govern the fixed cost of R&D are calibrated to target two moments in the firm-level data: the share of R&D-active firms (6.2 percent in both the model and the data[20]), and how the share of R&D-active firms depends on their current productivity levels. The latter is obtained by regressing the dummy variable of being an R&D-active firm on the firm's productivity level in logs.

R&D ideas production. The effect of R&D investments on next-period productivity is jointly governed by four parameters in the R&D function (the R&D productivity level, B, the span-of-control parameter, μ, the share of skilled labor, ψ, and the elasticity of substitution parameter, κ) and the effect of R&D on next-period productivity, ϕ.

We estimate $\psi = 0.554$ to target the 0.82 percent share of R&D workers in the labor force. The parameter κ has no steady-state implications. Instead, it determines how R&D-active firms substitute between input factors (skilled labor vs. intermediate goods) after the price or productivity of one factor changes relative to the other. We set $\kappa = 0.69$ to match the increase in the share of R&D workers in the labor force in the data (from 0.82 percent in 2004 to 3.57 percent in 2018) after all aggregate shocks are fed into the model.[21]

where $\mathbf{m}_{data}$ and $\mathbf{m}_{model}(\Theta)$ are the vectors of moments in the data and the model, and $\mathbf{W}_{ii} = \text{diag}(\omega_i)$ is a diagonal weighting matrix, where i indexes the i_{th} moment. We place additional weight, ω_i, on the data we view as more important to match and normalize $\sum_{t=1}^{12} \omega_i = 1$.

[19]See Section 2.1 for the estimation of the wage premium of skilled workers. Ideally, we should match the wage premium in 2002, but the 2002 CHIP does not include the Gaokao score, which we use to approximate individual ability. Also notice that without controlling for individual ability, columns (1) and (3) in Table 1 suggest that the wage premium does not differ much from 2002 to 2007.

[20]Source: the 2004 China Economic Census Yearbook. Among 276,474 above-the-scale industrial firms, 17,075 of them report positive R&D investments.

[21]See Section 5.1 for the explanation of how aggregate shocks are calibrated.

The R&D productivity level (B), the span-of-control parameter (μ), and the effect of R&D expenditure on next-period productivity (ϕ) are calibrated jointly. A larger B yields a greater number of R&D ideas for a given amount of inputs. A larger ϕ generates a larger productivity increase for a given number of R&D ideas. A firm is willing to invest more on R&D if either B or ϕ is larger. Therefore, various combinations of the two parameters could yield the same aggregate R&D-expenditure-to-output ratio. We calibrate these two parameters together by targeting the 1.9 percent aggregate R&D intensity in the data and the effect of R&D investments on next-period productivity, which is estimated to be 0.032 in Section 2.3. ϕ is estimated to be 0.063, indicating that doubling the number of R&D ideas would result in an average 6.3 percent increase in firm productivity in the next period.

Both a greater B and a greater μ yield a higher aggregate R&D intensity. However, for a given aggregate R&D intensity, a larger μ shifts the distribution of R&D expenditure to more productive firms, which is equivalent to a lower median R&D expenditure-to-output ratio. This is because a larger μ means that the R&D production function displays less decreasing returns to scale. Therefore, we use the combination of the aggregate and the median R&D expenditure-to-output ratio to identify both B and μ.[22]

Productivity process. ρ_z and σ_ϵ govern the the exogenous component of the productivity process. We estimate ρ_z by running a regression of a firm's next-period productivity on its current productivity using the 2001-02 ASI data. We estimate σ_ϵ by measuring the standard deviation of productivity growth in the 2001-02 ASI data.

4.3 Model's Performance

To evaluate our model's performance, we compare the non-targeted moments between our model predictions and the data counterparts, and report them in the lower panel of Table 6.

When calibrating the parameters that govern the fixed cost R&D production, we target how the decision to become an R&D-active firm relates to a firm's productivity. To evaluate our model's performance, we check if our calibrated model can reproduce the relationship between the decision to become an R&D-active and a firm's net worth. To this end, we simulate our calibrated model and regress the dummy variable of being an R&D-active firm on the firm's productivity level (in log) and asset level (in log). The estimated regression coefficient in our model (0.074) is closely aligned with that in the data (0.071). This implies that our calibrated model can capture how financial frictions shape firms' R&D decisions in

[22] Cao et al. (2024) estimate that the skill intensity is 0.22 for incremental innovation and 0.39 for radical innovation in China. In our estimation, the parameters μ and ψ together imply that the elasticity of R&D ideas to the skilled-labor input is 0.29, which is in line with the estimation in Cao et al. (2024).

Table 6: Targeted and Non-targeted Moments: Model vs. Data

Moment	Model	Data
Targeted		
Aggregate assets to output ratio	1.913	1.923
Aggregate leverage ratio	0.560	0.563
College wage premium	1.587	1.587
Aggregate R&D to output ratio %	1.914	1.914
Median R&D to output ratio %	2.084	2.062
Share of R&D workers in labor force %	0.818	0.818
Δ Share of R&D workers (18 vs. 04) %	2.758	2.740
Share of R&D firms	0.062	0.062
Regress R&D dummy on $\log(z)$	0.034	0.034
Regress $\log(z')$ on $\log(z)$	0.700	0.698
SD of productivity growth	0.500	0.499
Regress $\log(z')$ on $\log(i + w_s n_{s,x})$	0.032	0.032
Non-targeted		
Reg. R&D decision on log asset	0.074	0.071
R&D to output 95th percentile	0.248	0.265
R&D to output 90th percentile	0.155	0.163
R&D concentration Top 1%	0.575	0.566
R&D concentration Top 5%	0.820	0.776
R&D concentration Top 10%	0.909	0.863
R&D workers cost share	0.379	0.360

China, as we have shown empirically in Section 2.3.

The share of R&D workers in the labor force is one of our calibration targets. We now check to see whether the labor cost of R&D workers as a share of total R&D expenditure is similar between the model and the data. The last row of Table 6 shows that, in the data, 36.0 percent of R&D expenditure is allocated to R&D workers' compensation, while the model counterpart is 37.9 percent. The reason that our model is able to simultaneously match the labor cost share of R&D expenditure and the population share of R&D workers is due to the assumption on the Γ distribution of labor productivity. Skilled workers have higher average productivity and, consequently, receive higher labor income.

When calibrating the parameters of the R&D production function, we use the aggregate and the median R&D expenditure-to-output ratio. We now check if our calibrated model can reproduce the entire distribution of R&D expenditure in the data. To this end, we first compare the most R&D intensive firms in our model and in the data. Specifically, we calculate the R&D-to-output ratio for firms at the top 5 percent and 10 percent of the R&D expenditure-to-output distribution. Table 6 shows that our model predictions are consistent with the data.

Next, we examine whether the concentration of R&D expenditure aligns with the data, which is not a target in the calibration. Our calibrated model reproduces these empirical features well. For example, among all R&D-active firms, the top 1 percent firms that have the largest R&D investments account for 56.6 percent of aggregate R&D investments in the data and 57.5 percent in the model.

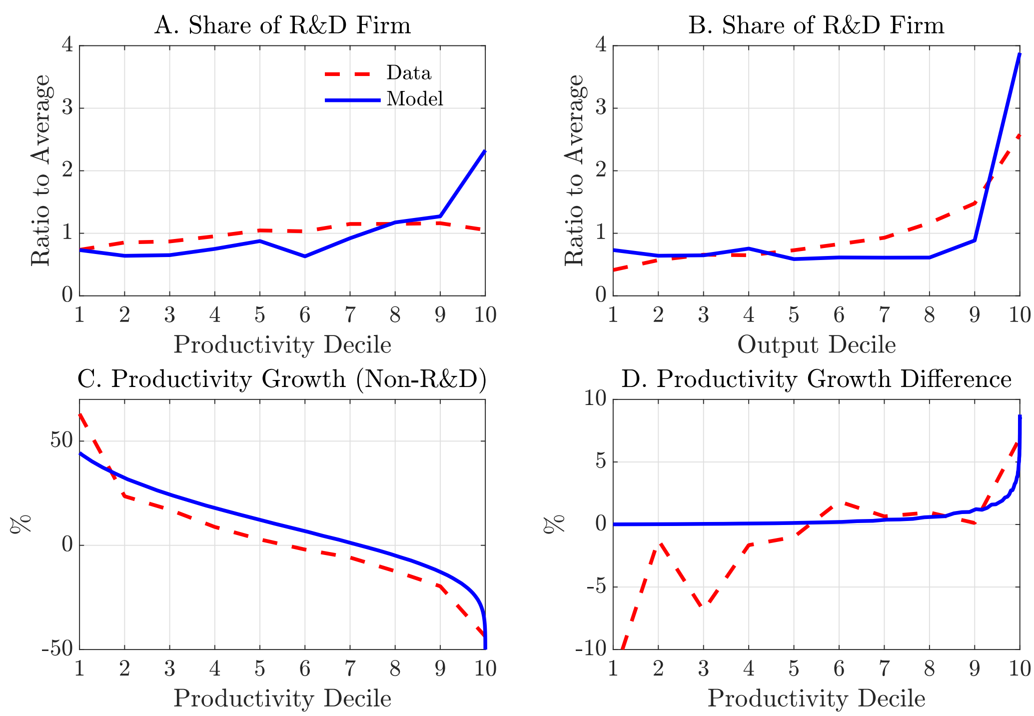

Figure 2: Model Performance

When calibrating the fixed cost of R&D, we target the share of R&D-active firms in all firms and the estimated linear relationship between being R&D active and the productivity level. To examine our model's performance, we check the share of R&D-active firms across all productivity and output levels. Figure 2 shows that while the data show that more productive firms are more likely to invest in R&D, this positive relationship is very moderate (Panel A). In contrast, the positive correlation between the share of R&D-active firms and output levels are more pronounced, especially at the top end of the output distribution (Panel B). Our model can reproduce both of these features. The reason is that in our model, both productivity levels and asset levels affect firms' R&D decisions. A firm at the top of the productivity distribution may choose not to invest in R&D if it does not have enough net worth. The reason is that a firm's future capital demand is more likely to be constrained if it has a low asset level, and therefore the expected profitability of an improvement in productivity is limited. On the other hand, the output size is determined by both productivity and net worth. Therefore, our model reproduces the empirical pattern of a more significant relationship between the share of R&D-active firms and output sizes, and a less significant relationship between the share of R&D-active firms and productivity levels.

Lastly, we check if our calibration of the productivity process can reproduce the productivity growth patterns in the data. First, we plot the empirical and the model-predicted productivity growth for R&D-inactive firms across the productivity distribution (Figure 2 Panel C). Our model predictions closely match the data pattern that more productive firms have lower productivity growth, implying a large discount factor (ρ) in the productivity process. Second, we plot the difference in growth rate between R&D-active firms and R&D-inactive firms across the productivity distribution. Our model predicts that at the middle and lower end of the productivity distribution, holding current productivity levels fixed, R&D-active and R&D-inactive firms have a similar productivity growth rate. In contrast, at the upper end of the productivity distribution, R&D-active firms have a much faster productivity growth than R&D-inactive firms. This is because under the calibrated productivity process, more productive firms have a faster depreciation of productivity, and therefore they invest more in R&D to improve their future productivity.

5 Quantitative Analysis

We now turn to quantifying the drivers and consequences of the surge in firm R&D expenditure, with a focus on the role of the college expansion program. We make a steady-state comparison between the Chinese economy in 2004 and 2018. We first construct the aggregate shocks that drive the increase in R&D intensity during this period (Section 5.1). Next, we analyze the aggregate and the distribution effects of the combination of these aggregate shocks (Section 5.2). In particular, we isolate the impact of the skilled labor supply shock by simulating a counterfactual economy without the college expansion policy. Lastly, we predict the long-run effect of the college expansion program by simulating counterfactual economies with a continued increase in the share of skilled labor (Section 5.3).

5.1 Constructing Aggregate Shocks

We model three aggregate shocks that drive the increase in R&D intensity from 2004 to 2018. The first aggregate shock is the exogenous increase in the share of skilled labor, which is calibrated externally using Chinese Population Census data. As reported by the 2020 Chinese Population Census, the share of skilled labor—measured as the fraction of urban workers aged 23 to 55 holding a bachelor's degree or higher—increased from 4 percent in 2004 to 14 percent in 2018. Under our assumption of heterogeneous labor productivity, the college expansion program brings in more individuals with lower labor productivity into colleges. As a result, the productivity cutoff for skilled labor reduces from 63 percent higher than the average productivity to 35 percent higher than the average. (See Figure 5 in Appendix C.)

As the share of skilled labor increases dramatically, *ceteris paribus*, the wage premium for skilled labor should decline sharply. However, the data only show a slight decline from 1.56 to 1.47. To explain this difference, we introduce a skill-biased technological shock, S_t that increases productivity for skilled labor in both the production and the R&D sectors. In our calibration, S_t increases from 1.00 (normalization) to 2.09 (5.4 percent increase annually) from 2004 to 2018.

Table 7: Calibrating the Aggregate Shocks

Parameter		2004	2018	Target	Model	Data
Share of skilled workers	N_t^s	0.04	0.14	Share of college-educated labor	0.14	0.14
Skill-biased tech. change	S_t	1.00	2.09	College wage premium	1.47	1.47
Sector-biased tech. change	B_t	0.91	2.96	Aggregate R&D to output ratio %	4.22	4.22

Non-Target	Model	Data
Δ patents per R&D worker %	154	144

We construct the third shock to capture all residual factors driving the relative expansion of the R&D sector compared to the production sector, and refer to it as the R&D-sector-specific technology shock, B_t. We calibrate this shock to match the observed change in the aggregate R&D-to-output ratio, which rises from 1.91 to 4.22 percent, implying an annual growth rate of 8.8 percent in B_t. This shock can be interpreted as an increase in R&D productivity, reflecting Chinese firms' growing ability to advance technological frontiers, as well as an improvement in the administrative and legal environment for intellectual property protection that enhances firms' capacity to secure returns on their R&D investments.

We validate our construction of the aggregate shocks by comparing our model prediction on the number of *ideas* per researcher with the data counterpart of the number of patents granted per R&D worker. Our calibrated shocks generate a 154 percent increase in number of *ideas* produced per R&D worker, and the data show a 144 percent increase.[23]

[23] In the China Economic Census, the patents granted per R&D worker is 0.014 in 2004 and 0.034 in 2018. We document this statistic in Table 2. Cao et al. (2024) report that the number of eventually granted patents per researcher in China increased from 0.14 in 2004 to 0.27 in 2014, using a different data source. This finding is consistent with our calculations, although data beyond 2014 are unavailable.

5.2 Quantifying the Combined Effects of Aggregate Shocks

In this section, we examine the aggregate and distributional effects of the aggregate shocks. In our analysis, we adjust the equilibrium wages of skilled and unskilled workers to ensure that both labor markets clear. Our main goal is to quantify how much the combination of these aggregate shocks contributes to the surge in R&D expenditure and how much the surge in R&D expenditure contributes to the increase in aggregate TFP.

Table 8: Quantifying the Impact of Aggregate Shocks

	w/o Tech. Shocks		w/ Tech. Shocks	
	w/o ΔN_t^s	w/ ΔN_t^s	w/o ΔN_t^s	w/ ΔN_t^s
	(1)	(2)	(3)	(4)
A. R&D input (%)				
Aggregate R&D to output ratio	1.9	2.4	4.0	4.2
Median R&D to output ratio	2.1	2.4	3.7	4.0
Fraction of R&D firms	6.2	8.1	16.4	17.6
R&D concentration Top 1%	57.4	59.5	68.7	70.4
R&D concentration Top 5%	81.9	82.8	87.1	87.8
R&D concentration Top 10%	90.8	91.2	93.2	93.6
Share of R&D workers in skilled labor force	19.0	29.4	16.3	24.9
B. Productivity (%Δ from bench.)				
Total factor productivity	-	1.5	9.6	11.8
Average productivity	-	0.1	1.1	1.4
TFP Losses	-	-0.4	-2.3	-2.9
C. Financial metrics (%Δ from bench.)				
Manufacturing output	-	-1.3	23.0	27.4
After-tax profit	-	-2.3	10.5	9.5
Net worth	-	-1.6	14.2	13.7
D. Skill price (%Δ from bench.)				
College wage	-	-47.5	141.7	27.1
Non-college wage	-	13.0	18.5	36.8

Note: In Panel A, we report the absolute values in each case. All statistics shown in panels B-D display the changes from the benchmark steady state. Case (1) displays the moments in the initial steady state. In case (2), we feed in the skilled labor supply shock but fix B_t at the level in the initial steady state. In case (3), we feed in both the skill-biased technology shock and the sector-biased technology shock, but fix the skilled labor supply at the initial steady state level. In case (4), we feed in all shocks described in Section 5.1. In all experiments, we adjust wages to clear the labor markets.

Table 8 presents the effects of the aggregate shocks on R&D input, productivity, financial metrics, and wages for skilled and unskilled labor. The difference between columns (1) and (4) shows the combined effect of all three aggregate shocks.

R&D input. The sector-specific technology shock, B_t is calibrated to hit the increase in the aggregate R&D-expenditure-to-output ratio in the data. The model-predicted median

R&D expenditure-to-output ratio increased by a similar size. The fraction of R&D-active firms increases from 6.2 percent in the initial steady state to 17.6 percent in the final steady state, whereas in the data, the fraction of R&D firms increases to 28 percent.

The key reason why our model underpredicts the increase in the share of R&D firms is that we do not model any fiscal or industrial policies that were implemented by the Chinese government to stimulate firms' R&D investment, such as tax deductions and exemptions. Studies show that firms may have over-reported their R&D expenditure to get benefits from these policies (Chen et al. (2021)). For this reason, both the officially released R&D intensity and the share of R&D firms might have suffered from over-reporting, but over-reporting might inflate the share of R&D firms more than the R&D intensity. This is because R&D investment is heavily concentrated in large firms and large private enterprises, which are subject to more stringent tax regulations. In contrast, it is easier for small and medium-sized enterprises (SMEs) to inflate their R&D expenditures in order to qualify as "high-tech enterprises" and gain tax benefits.

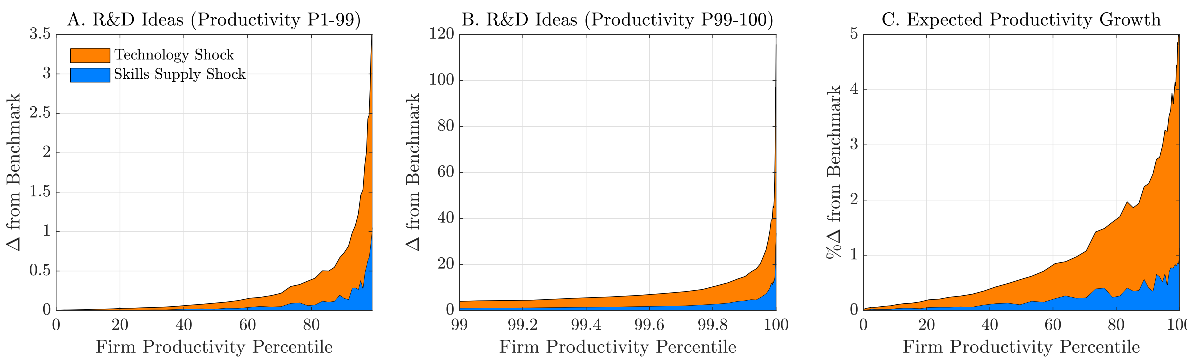

Figure 3: Decomposing the Impacts of Aggregate Shocks

The combined effect of the aggregate shocks makes R&D expenditure more concentrated at the top, especially at the top 1 percent. The share of R&D investment from the top 1 percent of the largest innovators increases by 13 percentage points (from 57 percent to 70 percent) (Table 8). In addition, we plot the change in R&D output from the initial steady state across the productivity distribution. As shown in Figures 3.A and 3.B, the number of *ideas* generated by firms at the top of the productivity distribution, especially at the top 1 percent, increases much more than for the rest of the firms.

Consistent with the empirical evidence in Table 2, the model predicts a 5.9-percentage-point increase in the share of R&D workers within the skilled labor force (compared to 10.0 percentage points in the data). This suggests that the R&D sector expands employment of skilled labor to a larger extent than the production sector does in response to aggregate shocks. There are two main reasons for this pattern. First, R&D-sector-biased technological shocks disproportionately favor the production of R&D *ideas*, thereby increasing the relative demand for skilled labor to perform R&D tasks. Second, skilled labor exhibits higher substitutability with other inputs in the R&D sector than in the production sector. Consequently, when skill-biased technological change occurs, the R&D sector experiences a more pronounced increase in skilled labor demand than the production sector does.

Productivity growth. As firms invest more in R&D following the aggregate shocks, productivity improves. The combination of the three aggregate shocks results in an 11.8 percent increase in TFP, much greater than the increase in average productivity, which only increases by 1.4 percent. The reason is that after the aggregate shocks, firms that are already at the top of the productivity distribution invest more in R&D than the rest, and thereby become more productive. As illustrated in Figure 3.C, the expected productivity growth of the most productive firms (top 1 percent) increases by approximately 5 percentage points, while firms up to the third quartile increase by less than 1 percentage point. As firms at the top of the productivity distribution use a larger amount of production factors (labor and capital) than the rest of the firms do, the productivity growth of highly productive firms has a disproportionately large impact on aggregate TFP. Therefore, the increase in TFP greatly exceeds the increase in average productivity.[24]

Allocative efficiency. The aggregate shocks improve allocative efficiency. The TFP losses due to the presence of financial frictions decrease by 2.9 percent.[25] This seems counter-intuitive at first glance, because as firms' productivity increases, their optimal capital demand should increase, implying a tighter constraint and thereby greater capital misallocation.

However, in response to the aggregate shocks, firms accumulate a higher level of net worth for two reasons. First, firms' after-tax profits increased by 9.5 percent, allowing them to allocate more resources toward accumulating net worth. Second, firms also expect to invest more in R&D in the future, and are therefore willing to save more today. Firms' willingness to accumulate assets and the ability to self finance alleviate the financial constraints, leading to an improvement in allocative efficiency.

[24] Average productivity is calculated as $\int z d\mu$, where μ denotes the distribution of all firms. The (measured) TFP is calculated as $Y/\left(AK^{\alpha}\left[\gamma(N_{s,y})^{\sigma}+(1-\gamma)N_u^{\sigma}\right]^{\frac{\theta}{\sigma}}\right)$, where the upper-case letters denote the aggregate variables.

[25] To quantify TFP losses, we apply the TFP accounting framework developed in Karabarbounis and Macnamara (2021). We first calculate an efficiency level of TFP, which can be computed as $\text{TFP}_e = (\int z^{\frac{1}{1-\alpha-\theta}} d\mu)^{1-\alpha-\theta}$. The TFP loss is defined as the percentage difference between the efficient and measured level of TFP: TFP losses $= (\text{TFP}_e/\text{TFP} - 1) \times 100$.

Wages. The labor supply shock alone increases the equilibrium wage for skilled workers and decreases the equilibrium wage for non-skilled workers. However, both the skilled-biased technology shock and the sector-biased technology shock increase the demand for skilled labor. The combined effect of all three shocks leads to a smaller wage increase for skilled workers than for unskilled workers, narrowing the wage gap between the two groups.

5.3 The Effect of the Rising Share of Skilled Labor

In this section, we first study the actual effect of the college expansion program. To isolate the labor supply shock, we compare the final steady state studied in Section 5.2 with an economy that is hit by the same skilled-biased technology shock and the sector-biased technology shock as described in Section 5.1, but the share of skilled labor is kept at the initial steady-state level. Next, we analyze a hypothetical scenario in which the share of skilled labor continues to increase in the future.

5.3.1 Isolating the effect of the labor supply shock

We first study the effect of the labor supply shock conditional on the two technology shocks, which is shown as the difference between columns (3) and (4) in Table 8. The most direct effect of the labor supply shock shows up in the wage premium of skilled workers. Without an increase in skill supply, skilled labor wages would rise significantly (by over 140 percent) as both of the technological shocks drive up the demand for skilled workers, thereby substantially increasing the cost of innovation.

In the absence of a rising supply of skilled labor, R&D input, measured by aggregate R&D intensity, decreases by only 0.2 percentage points (4.0 percent vs. 4.2 percent), but R&D output decreases by 24.0 percent, indicating that the productivity of R&D drops significantly. This is because in response to the increasing cost of skilled workers that is driven by the technology shocks, R&D-active firms substitute skilled workers with intermediate goods input. By doing so, R&D-active firms benefit less from the technology shocks that increase the productivity of skilled workers in the R&D sector.

The blue (lower) areas depicted in Figures 3.A and 3.B illustrate the effects of a rising skilled labor supply on the production of R&D ideas across firms with varying productivity levels. We observe that a lower innovation cost due to the lower wage of skilled labor disproportionately enhances the R&D ideas produced by the most productive firms. Consequently, without the labor supply shock, the degree of concentration of R&D ideas would decline. As shown in Figure 3.C, the expected productivity growth for highly productive firms would decrease by approximately 1 percentage point, which is significantly greater than the decline for less productive firms.

In addition, the effect of the labor supply shock is amplified by the presence of technology shocks. The difference between columns (1) and (2) in Table 8 shows the impact of labor supply shocks in the absence of technological shocks. The effect of the college expansion program on R&D choices and TFP growth is less pronounced than in the case with technological changes. The reason is that increasing the supply of skilled workers is more important when there are technology changes that drive up productivity and hence the demand for skilled workers.

5.3.2 Predicting the long-run impact

The college enrollment rate has been steady at roughly 35 percent in recent years. As young cohorts enter and old cohorts exit the labor market, the share of skilled labor in the working-age population will continue to rise in the next one or two decades. To assess the long-run impact of the college expansion program, we conduct the following exercise: we hold the skill-biased technology shock and the sector-biased technology shock at their 2018 level, and vary the share of skilled workers from 2 percent to 50 percent. While we vary the share of skilled workers, we adjust the equilibrium wages of of both types of workers to clear the labor markets. Figure 4 shows the results.

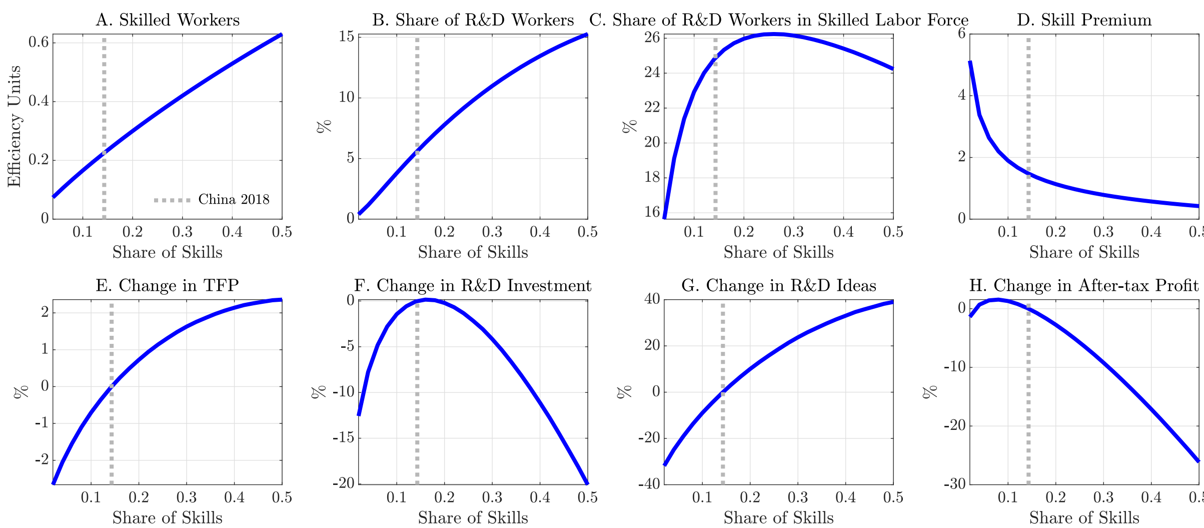

Figure 4: Predicting the Long Run Impact of Rising Share of Skilled Labor

Given our assumption that workers with above-the-threshold productivity obtain college degrees, Panel A of Figure 4 demonstrates that as the share of skilled labor grows, the rate of increase in total efficiency units provided by skilled labor diminishes. This expansion of skilled labor directly reduces the skill premium, as illustrated in Panel D. The consequent decline in the cost of skilled labor induces firms to hire more skilled workers for R&D activities (Panel B). The increase in the total number of skilled workers in the R&D sector leads to more ideas created in the R&D sector (Panel G) and ultimately enhances aggregate TFP (Panel E).

Our analysis suggests that a continued increase in the share of skilled labor could generate an additional 2 percent gain in TFP. As shown in Panel E, the marginal effects of a rising skilled labor share on TFP diminish. For instance, while increasing the skilled labor share from 10 to 20 percent raises TFP by 1.4 percent, expanding it from 40 to 50 percent yields only a 0.2 percent TFP gain. This diminishing return primarily stems from the equilibrium effect that pushes up the wage of unskilled workers. As Panel H shows, after-tax profits of firms decline sharply at an accelerating rate, reducing both their willingness and capacity to fund R&D investments.

We find a hump-shaped relationship between total R&D expenditure and the share of skilled labor, which emerges from the interplay of two countervailing forces. When the share of skilled labor is below 20 percent, the dominant effect is the reduction in R&D input costs due to the decline in the skill wage premium (Panel D), which gives firms strong incentives to invest more in R&D. However, as the skilled labor share continues to rise beyond this threshold, the dynamics shift markedly. The equilibrium effects of the rising wage of unskilled workers reduce firms' incentives to invest in R&D, despite the cost advantage. This explains why the initial positive relationship between skill share and R&D spending eventually reverses, creating an inverted-U pattern.

6 Conclusion

This paper examines the extent to which China's college expansion program, which leads to a significant increase in the supply of skilled labor, contributes to the rapid growth of firms' R&D investments and productivity between 2004 to 2018. The empirical evidence we provide suggests a crucial role for skilled labor in firms' R&D activity, and that both the labor supply shock and the technological change contribute to the surge in firms' R&D expenditure. We then construct a heterogeneous-firms model in which firms make endogenous R&D decisions under financial frictions and allocate skilled labor between the production and R&D sectors. We structurally estimate the model using firm-level data on the level and distribution of R&D, as well as macro-level data on skill prices and allocation between the production and R&D sectors.

To generate the observed growth in R&D investments, we construct three aggregate shocks: an R&D-sector-biased technological shock, a skill-biased technological shock, and a skilled labor supply shock. Our model suggests that the combination of these shocks generates a 12 percent increase in TFP, with the skilled labor supply shock accounting for approximately one-fifth of the model-predicted TFP improvement. Our model also predicts a marginal diminishing effect of an increasing supply of skilled labor in the long run because of the rising equilibrium wage of unskilled labor.

References

Aghion, Philippe, Ufuk Akcigit, Ari Hyytinen, and Otto Toivanen. 2017. "The social origins of inventors." Tech. rep., National Bureau of Economic Research. URL https://doi.org/10.3386/w24110.

Aghion, Philippe, Leah Boustan, Caroline Hoxby, and Jerome Vandenbussche. 2009. "The causal impact of education on economic growth: Evidence from US." *Brookings Papers on Economic Activity* 1 (1):1–73.

Akcigit, Ufuk, Jeremy Pearce, and Marta Prato. 2025. "Tapping into talent: Coupling education and innovation policies for economic growth." *Review of Economic Studies* 92 (2):696–736. URL https://doi.org/10.1093/restud/rdae047.

Bai, Chong-En, Chang-Tai Hsieh, and Yingyi Qian. 2006. "The return to capital in China." Tech. rep., National Bureau of Economic Research. URL https://doi.org/10.3386/w12755.

Bloom, Nicholas, John Van Reenen, and Heidi Williams. 2019. "A toolkit of policies to promote innovation." *Journal of Economic Perspectives* 33 (3):163–184. URL https://doi.org/10.1257/jep.33.3.163.

Brandt, Loren, Trevor Tombe, and Xiaodong Zhu. 2013. "Factor market distortions across time, space and sectors in China." *Review of Economic Dynamics* 16 (1):39–58. URL https://doi.org/10.1016/j.red.2012.10.002.

Buera, Francisco J and Roberto N Fattal Jaef. 2018. "The dynamics of development: Innovation and reallocation." *World Bank Policy Research Working Paper* (8505).

Cao, Linyi, Helu Jiang, Guangwei Li, and Lijun Zhu. 2024. "Haste makes waste? Quantity-based subsidies under heterogeneous innovations." *Journal of Monetary Economics* 142:103517. URL https://doi.org/10.1016/j.jmoneco.2023.09.004.

Che, Yi and Lei Zhang. 2018. "Human capital, technology adoption and firm performance: Impacts of China's higher education expansion in the late 1990s." *The Economic Journal* 128 (614):2282–2320. URL https://doi.org/10.1111/ecoj.12524.

Chen, Zhao, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu. 2021. "Notching R&D investment with corporate income tax cuts in China." *American Economic Review* 111 (7):2065–2100. URL https://doi.org/10.1257/aer.20191758.

Dai, Xiaoyong and Fang Wang. 2019. "Does the high-and new-technology enterprise program promote innovative performance? Evidence from Chinese firms." *China Economic Review* 57:101330. URL https://doi.org/10.1016/j.chieco.2019.101330.

Dang, Jianwei and Kazuyuki Motohashi. 2015. "Patent statistics: A good indicator for innovation in China? Patent subsidy program impacts on patent quality." *China Economic Review* 35:137–155. URL https://doi.org/10.1016/j.chieco.2015.03.012.

Gu, Shijun and Chengcheng Jia. 2022. "Firm dynamics and SOE transformation during China's Economic Reform." *Journal of International Money and Finance* 127:102672. URL https://doi.org/10.1016/j.jimonfin.2022.102672.

Hall, Bronwyn H, Jacques Mairesse, and Pierre Mohnen. 2010. "Measuring the Returns to R&D." In *Handbook of the Economics of Innovation*, vol. 2. Elsevier, 1033–1082. URL https://doi.org/10.1016/S0169-7218(10)02008-3.

Hsieh, Chang-Tai and Peter J. Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *The Quarterly Journal of Economics* 124 (4):1403–1448. URL https://doi.org/10.1162/qjec.2009.124.4.1403.

Hsieh, Chang-Tai and Zheng Michael Song. 2015. "Grasp the Large, Let Go of the Small:: The Transformation of the State Sector in China." *Brookings Papers on Economic Activity* (1):295–366. URL https://doi.org/10.1353/eca.2016.0005.

Hu, Albert Guangzhou and Gary H Jefferson. 2009. "A great wall of patents: What is behind China's recent patent explosion?" *Journal of Development Economics* 90 (1):57–68. URL https://doi.org/10.1016/j.jdeveco.2008.11.004.

Jia, Junxue and Guangrong Ma. 2017. "Do R&D tax incentives work? Firm-level evidence from China." *China Economic Review* 46:50–66. URL https://doi.org/10.1016/j.chieco.2017.08.012.

Karabarbounis, Marios and Patrick Macnamara. 2021. "Misallocation and financial frictions: The role of long-term financing." *Review of Economic Dynamics* 40:44–63. URL https://doi.org/10.1016/j.red.2020.09.002.

König, Michael, Kjetil Storesletten, Zheng Song, and Fabrizio Zilibotti. 2022. "From imitation to innovation: Where is all that Chinese R&D going?" *Econometrica* 90 (4):1615–1654. URL https://doi.org/10.3982/ECTA18586.

Li, Ben. 2010. "Multinational production and choice of technologies: New evidence on skill-biased technological change from China." *Economics Letters* 108 (2):181–183. URL https://doi.org/10.1016/j.econlet.2010.04.042.

Li, Xibao. 2012. "Behind the recent surge of Chinese patenting: An institutional view." *Research Policy* 41 (1):236–249. URL https://doi.org/10.1016/j.respol.2011.07.003.

Ma, Xiao. 2024. "College expansion, trade, and innovation: Evidence from China." *International Economic Review* 65 (1):315–351. URL https://doi.org/10.1111/iere.12670.

Midrigan, Virgiliu and Daniel Yi Xu. 2014. "Finance and misallocation: Evidence from plant-level data." *American Economic Review* 104 (2):422–458. URL https://doi.org/10.1257/aer.104.2.422.

Ottonello, Pablo and Thomas Winberry. 2024. "Capital, Ideas, and the Costs of Financial Frictions." Tech. rep., National Bureau of Economic Research. URL https://doi.org/10.3386/w32056.

Romer, Paul M. 2000. "Should the government subsidize supply or demand in the market for scientists and engineers?" *Innovation Policy and the Economy* 1:221–252. URL https://doi.org/10.1086/ipe.1.25056146.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti. 2011. "Growing like China." *American Economic Review* 101 (1):196–233. URL https://doi.org/10.1257/aer.101.1.196.

Tauchen, George. 1986. "Finite state Markov-chain approximations to univariate and vector autoregressions." *Economics Letters* 20 (2):177–181. URL https://doi.org/10.1016/0165-1765(86)90168-0.

Toivanen, Otto and Lotta Väänänen. 2016. "Education and invention." *Review of Economics and Statistics* 98 (2):382–396. URL https://doi.org/10.1162/REST_a_00520.

Young, Eric R. 2010. "Solving the incomplete markets model with aggregate uncertainty using the Krusell–Smith algorithm and non-stochastic simulations." *Journal of Economic Dynamics and Control* 34 (1):36–41. URL https://doi.org/10.1016/j.jedc.2008.11.010.

Appendices

A CHIP Data

This section describes our data-cleaning procedure and the definition of the normalized *Gaokao* score, and provides a summary of statistics of our sample.

Data-cleaning Procedure We keep only full-time employees. We calculate hourly wage by dividing the annual wage income (inclusive of monetary bonus and subsidies) by total working hours in the year. We calculate real wage income by adjusting nominal values to the 2007 price level using the national CPI index. We deal with outliers by excluding observations whose hourly wage income is less than 1 Chinese yuan or greater than 100 Chinese yuan in real terms.

Normalized *Gaokao* Score The raw scores of college entrance exams are not directly comparable between correspondents because the tests differ by year-province-subject. A province may also choose to write its own exam, which may have different maximum scores from the national test and from tests of other provinces. In addition, students can choose either a science-track curriculum or a humanities-track curriculum, and take different college entrance exams depending on their curriculum choices. After 2017, some provinces, such as Zhejiang Province, began to implement reform of the college entrance exam, allowing students to choose different exam subjects across different tracks.

To make *Gaokao* scores comparable between correspondents, we normalize the raw score in the following way. We first calculate the percentage score by dividing the raw score by the maximum score for each province-year-subject.[26] We then calculate the mean and standard deviation for each year-province-subject, and convert the percentage score of each correspondent to a z-score with a mean of zero and a standard deviation of one.

Summary of Statistics Table A1 provides a summary of statistics of the sample we use in the regression.

[26] There are larger cross-province variations in recent years as a growing number of provinces started to write their own tests. The maximum score is obtained from provincial education bureaus (http://www.moe.gov.cn), and some *Gaokao*-related websites such as https://www.lxbbt.com/16073.html and https://gaokao.chsi.com.cn. A small number of province-year data are missing, and we dropped observations for the missing province-year.

Table A1: Summary of Statistics

	CHIP Survey year			
Year	2002	2007	2013	2018
Age	39.29	34.44	34.53	34.65
Male (%)	54.76	59.04	56.06	56.93
Gaokao Z-score				
HS		-0.49	-0.53	-0.55
College		0.18	0.15	0.13
Hourly wage	5.73	13.02	16.32	23.22
HS	4.92	7.76	12.48	15.96
College	6.70	14.91	17.42	24.89
College and above(%)	14.24	35.45	43.27	47.45
Observations	7193	2158	3143	4839

B Educational Attainments across Age Groups

Table A2 summarizes the share of population with college and post-graduate degrees for different age groups in 2000, 2010, and 2020. The most significant change is shown in the age group of 25 to 29 years old, as regular higher education degrees are mostly obtained in one's 20s. From 2000 to 2020, the share of the population with college or higher degrees increases from around 4 percent to around 25 percent. The changes for the elder groups of the population from 2000 to 2020 are mostly due to the completion of part-time degree programs, which are arguably of less quality.

Table A2: Education Attainments across Age Groups

	College and above %			Post-graduate %		
Age	2000	2010	2020	2000	2010	2020
25-29	4.3	15.6	25.6	0.4	2.1	3.6
30-34	4.3	10.7	20.8	0.4	1.4	2.6
35-39	4.1	7.3	18.9	0.4	0.8	2.5
40-44	2.3	5.9	12.4	0.2	0.5	1.6
45-49	2.0	5.0	7.9	0.1	0.5	0.8
50-54	2.1	2.8	5.6	0.1	0.2	0.5
All	3.4	8.1	15.2	0.3	0.9	1.9

Source: China Population Census.

C Labor Productivity

This figure plots the calibrated Γ distribution of labor productivity, and the productivity cutoff for college admission in 2004 and in 2018.

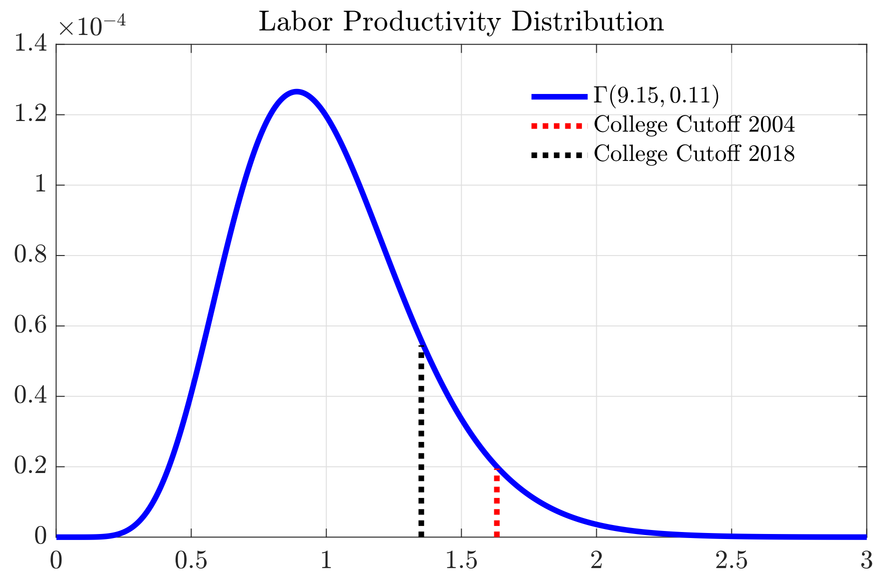

Figure 5: Labor Productivity Distribution and Cutoff for Skilled Labor

D Computational Algorithm

In this computational appendix, we first explain the method for solving the heterogeneous-firms model. Then, we describe how to compute the stationary equilibrium.

D.1 Solution Method for the Firm's Problem

In this section, we discuss how to solve the model of heterogeneous firms with endogenous R&D decisions in Sections 3.1 and 3.2.

1. Generate a discrete grid for a firm's wealth positions $G_a = \{a_1, a_2, ..., a_{Na}\}$ and productivity positions $G_z = \{z_1, z_2, ..., z_{Nz}\}$. We use 150 nodes in the grid for wealth and 48 nodes in the grid for productivity. We interpolate the function (evaluate the function outside the grid points) using piecewise linear approximation.

2. Generate a discrete grid for the household's productivity shock $G_\epsilon = \{\epsilon_1, \epsilon_2, ..., \epsilon_{N\epsilon}\}$ using Tauchen (1986) method. We use 9 nodes in the grid for productivity shocks and assume they are independently drawn from an identical discretized normal distribution in each period.

3. For each combination of wealth and productivity grid points, we solve the profit maximization problem given by Equations (9) and (10), which yields the optimal demand for capital (k), skilled labor ($n_{s,y}$) and unskilled labor (n_u).

4. Make an initial guess for the value of firms at each grid point $v_0(a,z)$.

5. Calculate the value of R&D-inactive firms $v_1^E(a,z)$ by solving Equations (6), (10), (15), and (16). We use the golden-section search method to obtain the optimal decision rules for next-period wealth (a').

6. Calculate the value of R&D-active firms $v_1^R(a,z)$ by solving the Equations (7), (17), (18), (19), and (20). We use the nested golden-section search method to jointly obtain the optimal decision rules for next-period wealth (a') and R&D ideas (x). In particular, the minimized costs of producing x units of R&D ideas can be obtained by solving Equations (11) and (12), which also yields the optimal demand for skilled labor ($n_{s,x}$) and intermediate inputs (i).

7. Compare $v_1^E(a,z)$ and $v_1^R(a,z)$ and obtain the decision rule for the R&D choices at the extensive margin. Update values for each grid point $v_1(a,z) = \max\left\{v_1^E(a,z), v_1^R(a,z)\right\}$.

8. Check convergence. If $\parallel v_1(a,z) - v_0(a,z) \parallel < 10^{-8}$, a solution is found. Otherwise, set $v_0(a,z) = v_1(a,z)$ and interate the procedure from step 5.

D.2 Computing the Stationary Equilibrium

In this section, we discuss how we simulate the stationary distribution using Young (2010) method and solve the labor market equilibrium, as described in Section 3.3.

1. Make an initial guess for the equilibrium wage rates w_s^j and w_u^j and the distribution of firms over their wealth and productivity positions μ^0. Set $j = 0$.

2. Numerically solve the heterogeneous-firms problem with endogenous R&D decisions using the algorithm described in Appendix D.1.

3. Interpolate firms' optimal policies over a finer grid. Create a big transition matrix for R&D-inactive and R&D-active firms, respectively, which requires us to know how firms' wealth and productivity evolve according to the optimal decisions and exogenous shocks.

4. Start from the initial distribution μ^0 and separate it into the groups of R&D-inactive and R&D-active firms according to their R&D decisions at the extensive margin. Obtain the new distribution μ^1 by applying the big transition matrices. If $\| \mu^1 - \mu^0 \| < 10^{-10}$, the stationary distribution is found and the iteration stops.

5. Aggregate all firms' demand for skilled and unskilled labor. Check if Equations (21) and (22) hold. If so, the stationary equilibrium is found. Otherwise, update w_s^{j+1} and w_u^{j+1} and iterate the procedure from step 2.