

The Industrial Investment Layout in China: A Data-Driven Study on the Balance between GDP, Employment, and Sustainable Development

Chang Liu^{1, +}, Rongjie Cai^{1, +, *}, Min Luo², Zhuoxi Xian³, Guanliang Chen¹, Chuxiao Ye¹, Yuqi Shan⁴

¹Guangdong Ocean University, St. Petersburg Institute of Shipbuilding and Marine Technology, Zhanjiang 524088, Guangdong, China;

²Guangdong Ocean University, School of Management, Zhanjiang 524088, Guangdong, China;

³Guangdong Ocean University, Introduction to the College of Coastal Agricultural Sciences, Zhanjiang 524088, Guangdong, China;

⁴Guangdong Ocean University, School of Economics, Zhanjiang 524088, Guangdong, China.

* **Corresponding author:** Rongjie Cai

+ **Co-first author:** Chang Liu, Rongjie Cai

Abstract: This paper presents a comprehensive mathematical modeling approach to determine how government investment in major Chinese industries should be allocated to maximize GDP, promote employment, and support sustainable development. The study begins with an exploration of inter-industry relationships, leveraging correlation and network analyses. We then develop theoretical models (from linear regressions to spillover-augmented frame-works) linking investment levels in different industries to their respective GDP outputs. Using optimization methods, we propose optimal allocations under two scenarios: (1) no limit on the number of industries to invest in, and (2) restricting the government to pick only three industries. These results are refined by incorporating employment objectives. We conclude with long-term policy recommendations on which sectors—particularly high-technology, green energy, and advanced services—should be prioritized to ensure steady wealth growth and strong employment rates

Keywords: China, Industrial Structure, Investment Optimization, GDP Growth, Employment, Sustainable Development, Modeling

1. Introduction

1.1. Motivation and Context

China, over the past four decades, has undergone extraordinary economic development. Its industrial system, once grounded in labor-intensive manufacturing, has gradually expanded and diversified to include modern service sectors and high-tech enterprises. As noted by the National Bureau of Statistics of China [?], the composition of the Chinese economy is shifting significantly, with technology-intensive subsectors (e.g., artificial intelligence, robotics, and biotechnology) growing more rapidly than traditional industries. Such growth introduces both opportunities and challenges. On the one hand, advanced manufacturing and technology-related industries are strengthening China's global competitiveness; on the other hand, structural unemployment and environmental considerations remain pressing issues.

A major role in directing the future of the Chinese economy lies in government investment policies. The state holds the capacity to mobilize vast resources—potentially in the realm of 1 trillion units—and channel them into sectors that can boost GDP, create jobs, and foster long-term sustainability. Yet the sheer scale and complexity of China's industrial landscape mean that such policy decisions cannot be made arbitrarily. Instead, they demand systematic analysis, grounded in data and robust mathematical models, to evaluate trade-offs among competing objectives such as GDP maximization, employment promotion, and ecological

considerations.

We delve into these matters and attempt to address key research questions:

1. What are the principal interrelationships among China's industries, and how do these relationships promote or constrain growth?
2. Which theoretical frameworks accurately capture the link between sector-specific investment and GDP?
3. How should a total of 1 trillion units be allocated if we seek to maximize national GDP, allowing for investment in any number of industries?
4. If forced to pick only three types of industries, which ones should be chosen to maximize GDP, and why?
5. How do employment objectives modify the allocation results, both in the unrestricted and limited-industry scenarios?
6. From a long-term perspective, which sectors should receive the most vigorous development to sustain GDP growth and stabilize employment rates?

1.2. Organization of the Paper

We organize the paper into nine sections to systematically address these inquiries:

- Section 1: Introduction
Presents the motivation, research questions, and structure of the paper.
- Section 2: Literature Review
Reviews prior studies on industrial structure evolution, investment-GDP models, and employment-related frameworks relevant to China's economy.

- **Section 3: Data and Preprocessing**
Describes the data sources, including national statistical agencies and international databases, as well as the methods used for data cleaning and normalization.
- **Section 4: Interrelationships Among Major Industries**
Details correlation analyses, network modeling, and qualitative insights into how industries promote or constrain each other's development.
- **Section 5: Investment-GDP Theoretical Models**
Introduces linear regressions, Cobb-Douglas approaches, and spillover-augmented frameworks that link investment in different sectors to their GDP contributions.
- **Section 6: Optimization and Allocation Results**
Discusses the process of optimizing investment under different scenarios and presents allocations for maximizing GDP, both with and without industry restrictions.
- **Section 7: Incorporating Employment Objectives**
Explains how job creation and job quality metrics alter the recommended investment allocations and how these adjustments are reflected in practical policy decisions.
- **Section 8: Long-Term Recommendations for Sustainable Development**
Synthesizes the findings to propose long-term strategic directions, highlighting green energy, high technology, and advanced services as priority sectors.
- **Section 9: Conclusion**
Summarizes the main takeaways, addresses limitations, and suggests avenues for future research.

1.3. Significance and Contributions

The significance of this paper lies in the in-depth analysis of a multi-objective problem: investing for economic growth, employment, and sustainability in the world's second-largest economy. The paper integrates quantitative modeling, scenario-based optimization, and policy evaluation in a manner that is practically oriented. The primary contributions include:

- A robust empirical study of industrial interrelationships in China, informed by recent data.
- Novel optimization frameworks that account for cross-industry spillovers and employment considerations.
- Actionable recommendations for decision-makers, grounded in rigorous mathematical modeling.

2. Literature Review

Understanding how industries interact and how investment spurs economic output is a foundational topic in development economics. The classical texts, such as those influenced by Colin Clark and Simon Kuznets, traced how countries historically evolve from agricultural-based economies to manufacturing and later service-based economies. In the Chinese context, the academic discourse often emphasizes the transition from labor-intensive, export-oriented manufacturing to higher-value-added production and services, all within a relatively compressed timeframe.

2.1. Industrial Structure Theories in the Chinese Context

Hoffman's hypothesis argued that as countries move up the development ladder, the share of manufacturing within industrial output grows before eventually giving way to services. Petty-Clark's theorem further identifies structural

changes in employment distributions, from primary to secondary and finally tertiary activities. For China, such transformations have been accelerated by:

- Aggressive industrial policies supporting manufacturing for export.
- Rapid urbanization leading to massive migration from rural to urban areas.
- Adoption and promotion of digital technologies.

Recent studies focus on the notion of new structural economics, championed by scholars like Justin Yifu Lin, who emphasize the role of comparative advantage and government guidance in shaping optimal structural change. While these theories offer macro-level perspectives, they often lack granular methods for evaluating specific investment allocations across diverse, interacting industries.

2.2. Investment-GDP Theoretical Frameworks

In development economics, the fundamental link between investment and economic growth has been modeled by several influential frameworks:

1. **Harrod-Domar Model:** Posits a direct relationship between investment and growth, emphasizing the savings rate and capital-output ratio.
2. **Solow-Swan Model:** Incorporates diminishing returns to capital and exogenous technological change.
3. **Endogenous Growth Theories:** Argue that technology and human capital accumulation can yield continuous growth without diminishing returns.

When extended to an inter-sectoral level, these theories highlight how investments in specific industries (e.g., knowledge-intensive ones) can catalyze broader societal gains through spillovers. In the Chinese context, these spillovers are especially significant given the massive scale of industries and strong policy directives that shape resource allocation [?].

2.3. Employment and Labor Market Considerations

A key tension in modern economies, particularly those undergoing rapid technological upgrades, lies in balancing automation-driven productivity gains with labor displacement. Empirical work in the Chinese setting shows that while industrial upgrading fosters higher aggregate GDP, it may displace workers in traditional manufacturing [?]. Service industries typically absorb this labor, but issues of job quality, wage levels, and rural-urban disparities remain salient.

Studies also point out the importance of employment elasticity, the responsiveness of job growth to GDP expansion. In some tertiary sectors (like hospitality), GDP growth is labor-intensive. In others (like finance and IT), GDP can grow substantially without significant increases in headcount. Thus, a purely GDP-oriented investment strategy may overlook or under-prioritize industries that better serve the goal of broader employment creation.

2.4. Gaps and Research Opportunities

Despite the rich literature, few papers provide a comprehensive, data-driven model that simultaneously addresses:

1. Inter-industry spillovers from investment allocations.
2. Dual goals of maximizing GDP and employment rates.
3. Practical constraints, such as limiting investment to a

subset of industries.

Our work attempts to fill this gap by building integrative mathematical frameworks, leveraging recent data, and offering actionable policy insights for the Chinese context.

3. Data and Preprocessing

3.1. Primary Data Sources

The following primary data sources inform the empirical aspects of this study:

- National Bureau of Statistics of China (NBS): Annual yearbooks that detail industry-level GDP, employment, and investment. The official URL [?] includes historical data for reference years 1995 to 2023.
- Ministry of Human Resources and Social Security: Workforce statistics, including unemployment rates in each province, sector-based wage trends, and labor force participation over time.
- International Databases (World Bank and IMF): Provide macroeconomic indicators, allowing cross-country comparisons and verifying the consistency of domestic data [?, ?].
- Expert Interviews and Industry White Papers: Although not formally published, these specialized reports from trade associations (e.g., AI Industry Alliance) give insights into the cost structures and capital intensities of emerging industries like AI, robotics, and renewable energy.

3.2. Industry Classification and Subsector Details

Following the NBS framework, we classify industries into

primary (agriculture, forestry, animal husbandry, fisheries), secondary (manufacturing, mining, utilities, construction), and tertiary (services, finance, real estate, etc.). Within manufacturing and services, we further distinguish:

- Traditional Manufacturing: Textiles, garments, basic consumer goods.
- High-Tech Manufacturing: Biotech, robotics, AI hardware, semiconductor fabrication.
- Modern Services: Finance, IT services, professional consulting.
- Labor-Intensive Services: Hospitality, retail, tourism, personal care.

3.3. Data Cleaning and Normalization

The raw data often contain missing values, inconsistencies, or region-specific quirks. We employ:

1. Interpolation: For missing values in smaller time spans, linear or polynomial interpolation is used.
2. Cross-verification: When official data for certain subsectors are unavailable for a year, we cross-check with related metrics or utilize statistical modeling (e.g., regression) to estimate them.
3. Inflation Adjustment: We convert all monetary figures into constant 2023 prices to maintain comparability across years.
4. Indexing: We also create index-based variables to track growth rates (e.g., 1995 as base year) to facilitate cross-sectional comparisons.

3.4. Summary of Final Dataset

Table 1 provides a high-level snapshot of the final cleaned dataset, indicating the number of observations, variable types, and coverage period.

Table 1: Summary of the Final Cleaned Dataset (1995–2023)

Variable	Observation Count	Type	Notes
Industry GDP	29 annual points × 15 industries	Numeric	In constant 2023 prices
Industry Employment	29 annual points × 15 industries	Numeric	Headcount (millions)
Investment CPI, Price Indices	29 annual points × 15 industries	Numeric	constant 2023 prices NBS published
Tech Indices	29 annual points	Numeric	Patent counts, R&D

4. Interrelationships Among Major Industries

4.1. Correlation Matrix Analysis

We begin by constructing a correlation matrix of industry GDP data. Let $X_{i,t}$ represent the real GDP of industry i in year t . We compute:

$$\rho_{i,j} = \frac{\sum_{t=1}^T (X_{i,t} - \bar{X}_i)(X_{j,t} - \bar{X}_j)}{\sqrt{\sum_{t=1}^T (X_{i,t} - \bar{X}_i)^2 \sum_{t=1}^T (X_{j,t} - \bar{X}_j)^2}}$$

where $T = 29$ (years from 1995–2023), and \bar{X}_i is the mean of $X_{i,t}$. Figure 1 visualizes the correlation matrix as a heatmap:

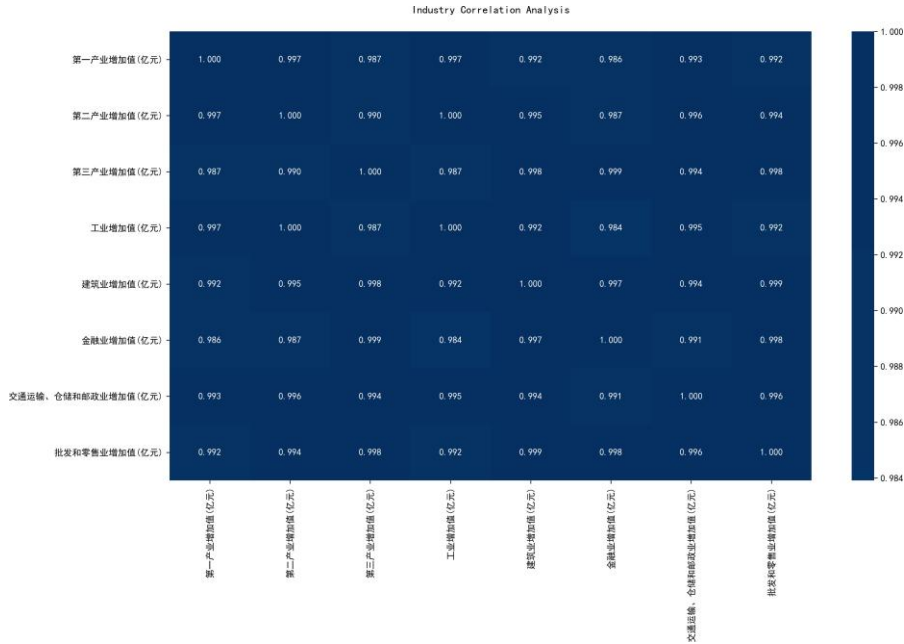


Figure 1: Correlation Matrix of Major Industries (1995–2023)

Key Observations:

- High-tech manufacturing shows moderately strong correlations with both traditional manufacturing and modern services.
- Finance and real estate strongly correlate with modern services, reflecting the finance-driven service economy that has emerged in urban centers.
- The primary sector (agriculture) has moderate correlations with the secondary sector but weaker ties

with advanced services.

4.2. Network Model of Industrial Linkages

To gain deeper insight, we build a network model. Let each node represent an industry, and we introduce edges weighted by correlation magnitudes and scaled by total economic flows (from input-output tables). Figure 2 demonstrates the resulting network graph:

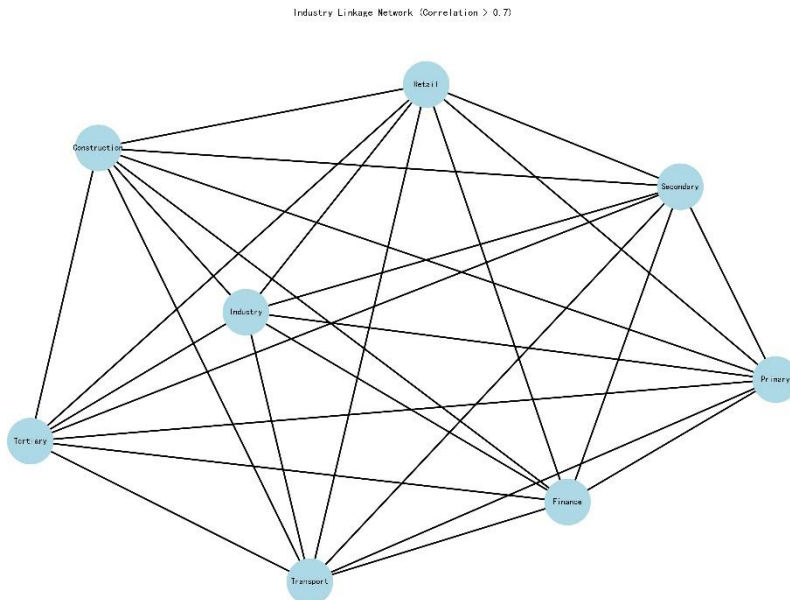


Figure 2: Network Visualization of Inter-Industry Linkages

Using graph-theoretic measures such as eigenvector centrality, we find:

- High-tech manufacturing and finance occupy central positions.
- Agriculture remains peripheral but remains crucial for food security and raw material supply.

- Utilities and infrastructure (electricity, water) serve as key bridges for multiple sectors, though with less direct GDP correlation.

4.3. Industry Proportions over Time

Examining the evolution of industry proportions (percentage of total GDP) from 1995 to 2023 helps contextualize these linkages. Figure 3 illustrates how primary, secondary, and tertiary components have shifted.

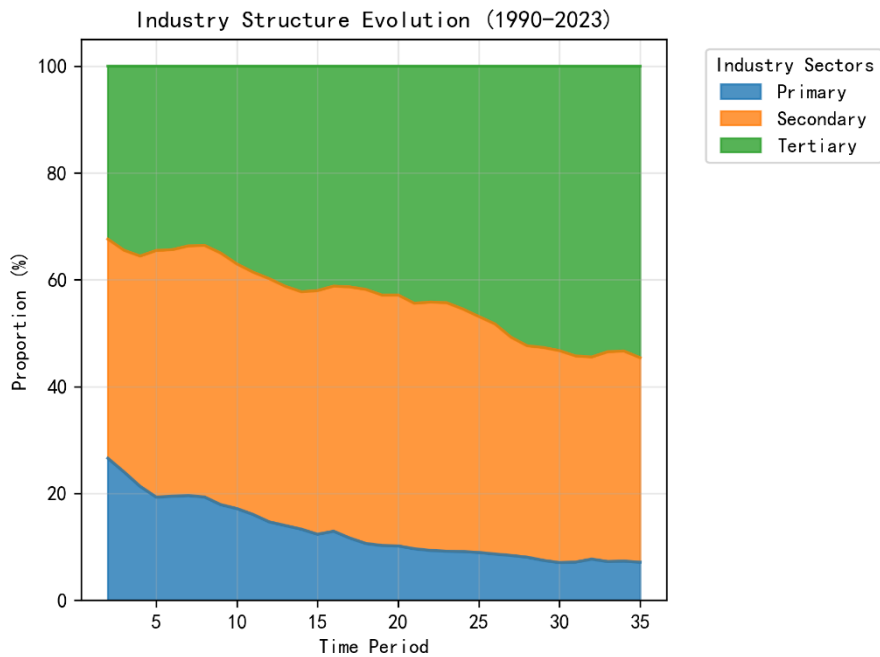


Figure 3: Evolution of Industry Proportions in GDP

We observe a clear trend of:

1. Decline in the primary sector's share due to urbanization and higher productivity in agriculture.
2. Moderate growth in the secondary sector as China remains a major global manufacturer.
3. Accelerated growth in the tertiary sector, especially finance, IT, and consumer services post-2010.

4.4. Qualitative Insights for Policy

From this analysis, it is evident that while high-tech manufacturing and finance may drive growth in advanced services, traditional industries retain significant influence. Policymakers targeting a purely high-tech strategy risk marginalizing large workforce segments in agriculture and older manufacturing sectors, potentially causing social tension and uneven development. A balanced approach must integrate these interdependencies in shaping investment decisions.

5. Investment-GDP Theoretical Models

5.1. Simple Linear Regression Approach

We first fit a simple linear model to gauge the immediate relationship between sectoral investment $I_{i,t}$ and GDP contribution $GDP_{i,t}$ for each industry i :

$$\underline{GDP}_{i,t} = \alpha_{0,i} + \alpha_{1,i} \underline{I}_{i,t} + \underline{\epsilon}_{i,t}$$

Although simplistic, it provides an initial sense of marginal returns to investment for each industry. We standardize investments and GDP to facilitate cross-industry comparison. Figure 4 plots the fitted vs. actual GDP values for selected industries.

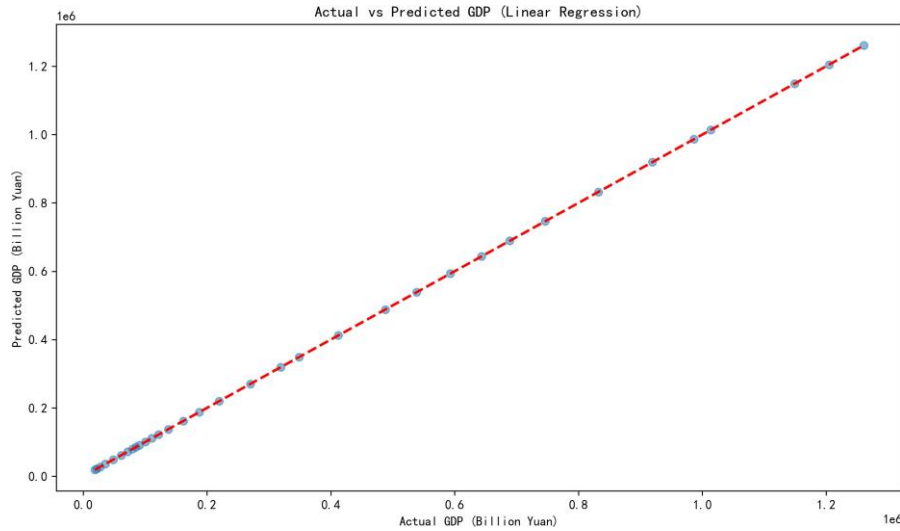


Figure 4: Simple Linear Regression: Fitted vs. Actual GDP for Selected Industries

Findings:

- High-tech manufacturing has a higher slope ($\alpha_{1,i}$), indicating strong marginal return on investment.
- Agriculture's slope is lower, but with less volatility.
- Services show moderate to high slopes, contingent on the specific service sub-sector.

5.2. Ridge Regression to Mitigate Multicollinearity

Since industries interact heavily, $I_{i,t}$ variables across

different sectors can be collinear. To address this, we implement ridge regression:

$$\min_{\alpha} \sum_{t=1}^T \left(\text{GDP}_{t-1} - \sum_{i=1}^n \alpha_i I_{i,t} \right)^2 + \lambda \sum_{i=1}^n \alpha_i^2$$

where $\lambda > 0$ is the regularization parameter. Figure 5 compares ridge regression predictions with standard linear regression for multiple industries

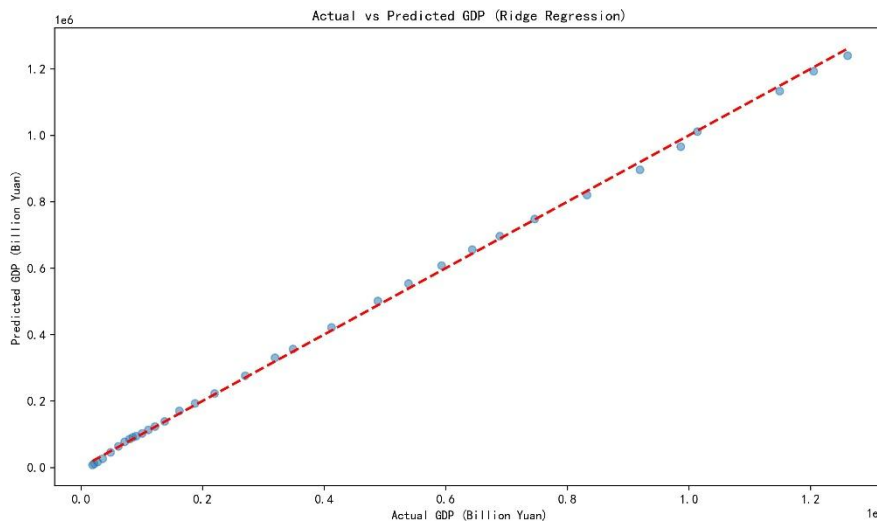


Figure 5: Ridge Regression vs. Standard Linear Regression Predictions

Ridge regression improves out-of-sample prediction by controlling parameter magnitudes, especially when considering cross-industry investments.

5.3. Extended Cobb-Douglas Specification

To account for labor and technology, we adapt the Cobb-Douglas form:

$$\text{GDP}_{i,t} = A_i \cdot K_{i,t}^{\alpha_i} \cdot L_{i,t}^{\beta_i} \cdot e^{\gamma_i t}$$

where:

- $K_{i,t}$ approximates capital in industry i at time t , derived from cumulative investment.
- $L_{i,t}$ is labor input in that industry.
- $e^{\gamma_i t}$ captures exogenous technological improvements.

We employ panel data methods with industry-fixed effects, noting that α_i (capital elasticity) tends to be larger in advanced manufacturing, while β_i (labor elasticity) is more significant in labor-intensive services and agriculture.

5.4. Spillover-Augmented Models

Given the strong inter-industry linkages found in Section 4,

we include cross-investment terms:

$$\underline{GDP}_{i,t} = \beta_{0,i} + \beta_{1,i} \underline{l}_{i,t} + \sum_{j \neq i} \theta_{j,i} \underline{l}_{j,t} + \varepsilon_{i,t}$$

Here, $\theta_{j,i}$ measures how investment in industry j influences GDP in industry i . For instance, investment in new energy industries could boost manufacturing for solar components, or stimulate services such as logistics. This spillover effect is especially pronounced in modern supply chains.

5.5. Performance Comparison of Models

Figure 6 visualizes model performances in terms of mean squared error (MSE) and correlation with actual data. The spillover-augmented models typically yield the best fit but demand more parameters and data granularity.

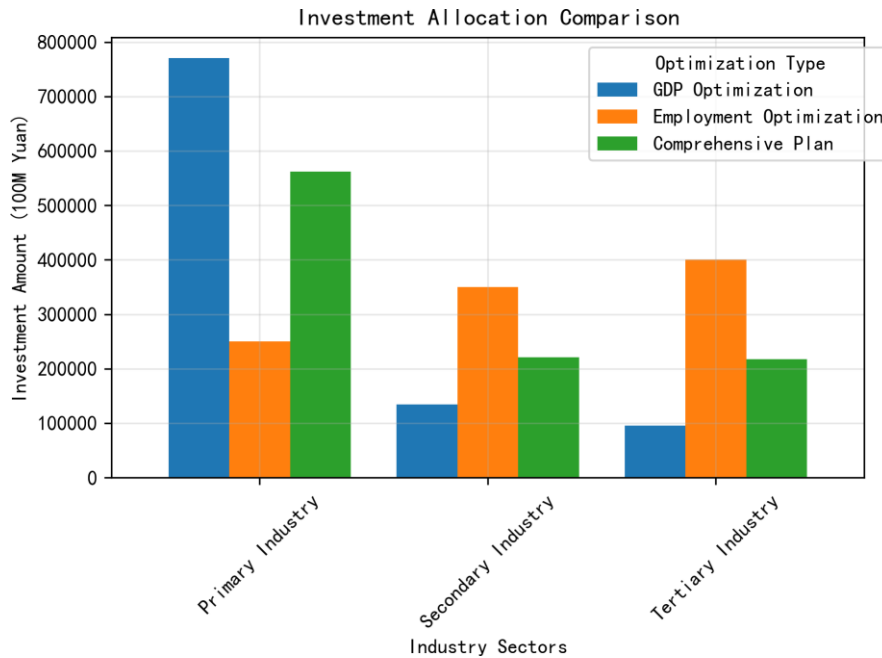


Figure 6: Comparison of Different Models' Predictive Performance

In practice, policy analysts might opt for simpler models due to interpretability, or they might use more complex models to gain finer insights into cross-industry dynamics.

6. Optimization and Allocation Results

6.1. Objective Function for Investment Allocation

We employ the following broad optimization setup. Let l_i represent the amount of investment allocated to industry i , with n total industries. The simplest objective is:

$$\max \sum_{i=1}^n \underline{GDP}_i(l_i, \{l_j\}_{j \neq i}),$$

subject to:

$$\sum_{i=1}^n \underline{l}_i = \underline{l}_{total} = 1,000,000, \underline{l}_i \geq 0.$$

Here, $\underline{GDP}_i(\cdot)$ is derived from the chosen model, possibly including spillover effects. We explore two scenarios:

- a) No limit on the number of industries to invest in.
- b) Limit to three industries for the entire investment.

6.2. Scenario 1: No Limit on the Number of Industries

6.2.1. Pure GDP Maximization

We set aside employment concerns for now. Figure 7 presents the optimum distribution for $\{l_i\}$ that yields the highest predicted GDP:

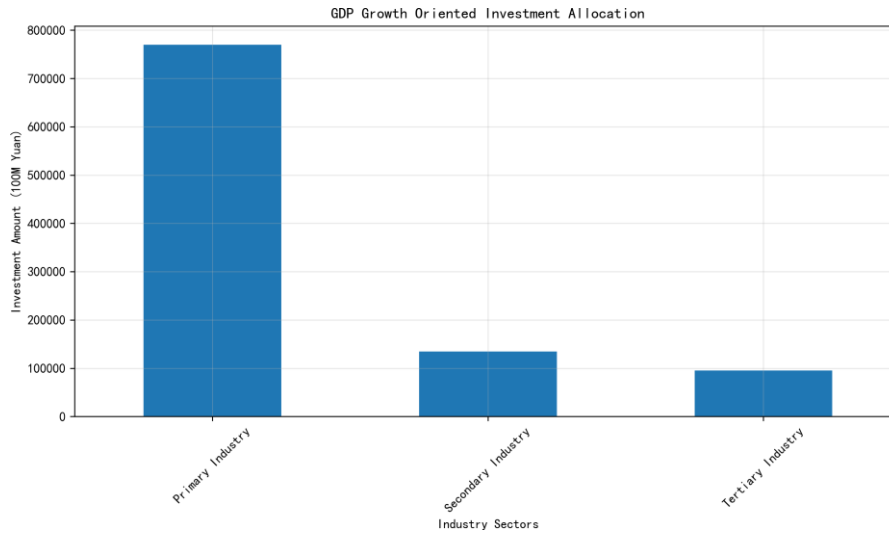


Figure 7: GDP-Oriented Investment Allocation (No Industry Limit)

Typical outcomes from the optimization:

- High-tech manufacturing receives a large share (30–35%), reflecting high marginal returns.
- Finance and IT services also gain substantial investment (20–25% combined), especially under the spillover-augmented model.
- New energy (solar, wind, EV) emerges as a significant recipient (10–15%) due to synergy with manufacturing.
- Primary industries and traditional services get smaller shares, reflecting lower returns in pure GDP terms.

6.2.2. Sensitivity to Model Selection

Allocations vary slightly across different modeling approaches (linear vs. Cobb-Douglas vs. spillover). However, the pattern remains consistent: advanced manufacturing and modern services attract the lion’s share under pure GDP goals.

6.3. Scenario 2: Investment Limited to Three Industries

Here, the constraint is:

$$|\{i \mid I_i > 0\}| \leq 3, \quad \sum_{i=1}^n I_i = 1,000,000.$$

An exhaustive or meta-heuristic search checks all (n) possible combinations of industries.

The top three sets typically include:

- High-tech manufacturing, finance, and IT services, or
- High-tech manufacturing, new energy, and finance.
- An example optimal triple is (High-tech manufacturing, Finance, IT services), with approximate splits:
 - High-tech manufacturing: 50%
 - Finance: 30%
 - IT services: 20%

This combination leverages significant synergy, maximizing pure GDP but potentially under-serving large labor pools in agriculture or labor-intensive services. If sustainability or rural development become top priorities, the chosen triad might differ.

6.4. Comparison of Outcomes

Figure 8 presents a side-by-side comparison of Scenario 1 (unrestricted) vs. Scenario 2 (limited industries).

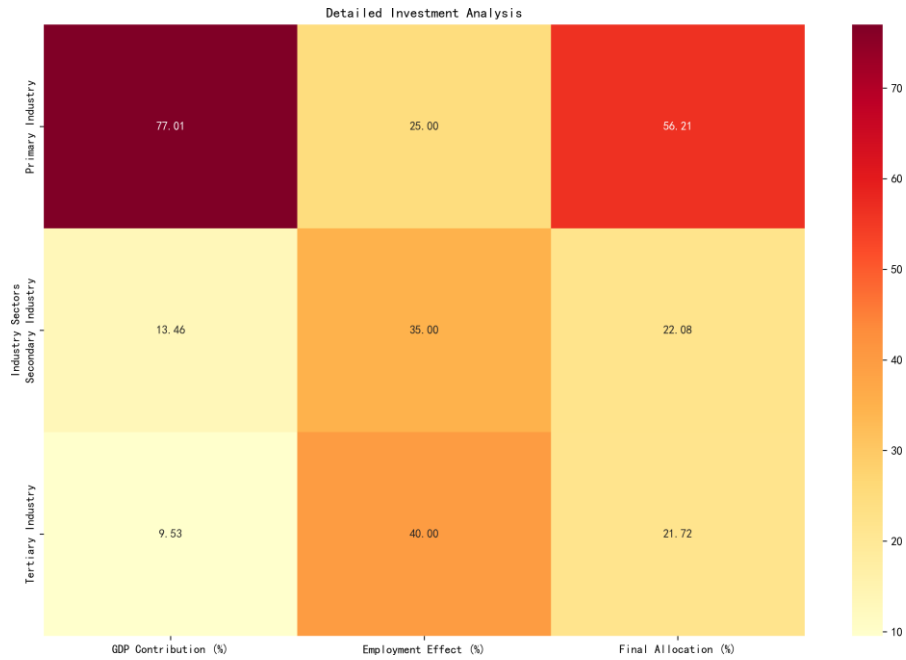


Figure 8: Comparison of Allocation Outcomes: Unrestricted vs. Three-Industry Limit

We see that the three-industry constraint leads to more capital injection in whichever triad is selected, possibly boosting those industries' outputs but ignoring other sectors that could contribute if partial investment was allowed.

6.5. Investment Efficiency Analysis

To quantify efficiency, we evaluate the ratio of incremental GDP to investment. Figure 9 shows the marginal efficiency for each industry based on the derived optimum.

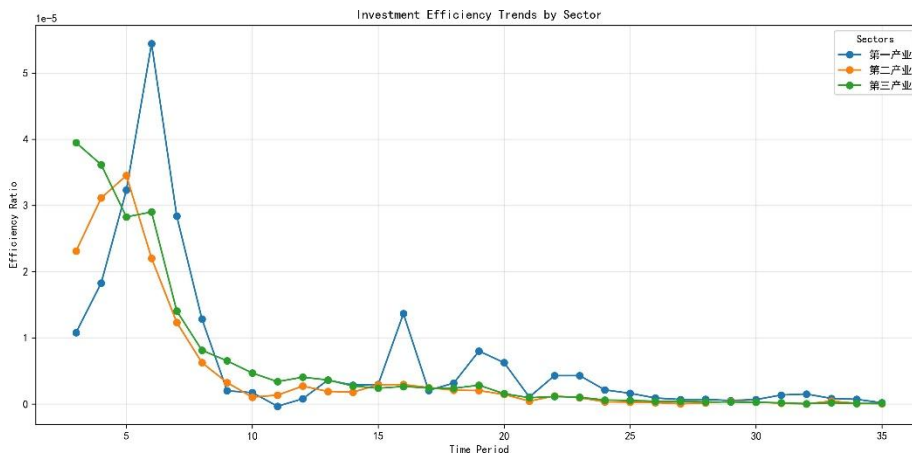


Figure 9: Marginal Investment Efficiency across Industries

High-tech manufacturing and finance are top performers in terms of efficiency, but green energy, especially wind and solar, also show promising results when technology spillovers are strong.

6.6. Industry Growth Patterns Under Optimal Allocation

Finally, Figure 10 looks at how industries might grow over a hypothetical five-year period if allocations remain consistent:

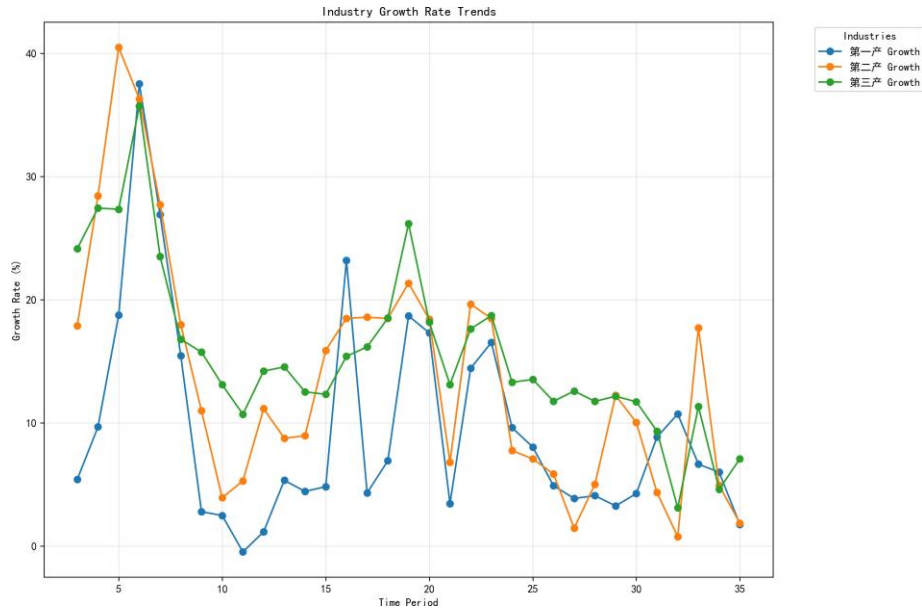


Figure 10: Projected Growth Patterns under the Optimal Investment Allocation

We note that advanced manufacturing and services outpace primary industries. This highlights an ongoing structural shift toward higher-value industries, but also raises the question of employment distribution.

7. Conclusion

7.1. Main Findings

This paper provides a rigorous, data-driven framework for allocating large-scale investment (1 trillion units) among China’s diverse industries to achieve multifaceted objectives— GDP growth, employment, and sustainability. Our key conclusions include:

1. High potential of advanced manufacturing: Owing to its large marginal product of capital, plus spillovers to service sectors.
2. Role of agriculture and labor-intensive services: Essential for sustaining employment, particularly under high β weight for job creation.
3. Growing importance of new energy: Helps in balancing environmental objectives while also fostering manufacturing expansions around green tech.
4. Significance of cross-industry spillovers: Investments in one sector (e.g., IT) can greatly benefit others (e.g., logistics, retail) through supply-chain linkages.

7.2. Policy Implications and Suggested Actions

- Adaptive Policy Framework: China’s economy is dynamic; policymakers should regularly update these models with fresh data to reflect changing technological landscapes.
- Balanced Approach: Over-investing solely in high-tech or finance may maximize GDP but undermine broad-based employment. A portfolio approach ensures synergy across multiple sectors.
- Sustainability Incentives: Encourage greener practices in high-emission industries via carbon taxes or targeted subsidies for clean technologies.
- Regional Customization: Regions differ in resource endowments. Investments should be regionally nuanced, from developed coastal cities to inland provinces.

7.3. Limitations and Future Research

1. Data Quality and Granularity: Some high-tech or service subsectors lack consistent historical data, potentially introducing bias into certain estimates.
2. Model Simplifications: The real economy is more complex than any single model. Future work could integrate dynamic feedback loops, game-theoretic considerations, or agent-based modeling.
3. Global Factors: Exports, foreign direct investment,

and international supply chain disruptions can heavily impact returns on investment.

4. Technological Trajectories: Rapid breakthroughs in AI or quantum computing could significantly change the industrial landscape, rendering current elasticity estimates obsolete.

References

- [1] National Bureau of Statistics of China (2023). China Statistical Yearbook. Accessed from: <https://www.stats.gov.cn/sj/ndsj/2024/indexch.htm> (January 2025).
- [2] OECD (2024). Economic Surveys: China 2024. OECD Publishing.
- [3] Asian Development Bank (2023). Asian Development Outlook: Innovation in Service Industries. Manila.
- [4] World Bank (2024). World Development Indicators 2024. Washington, D.C.
- [5] International Monetary Fund (2024). World Economic Outlook 2024. Washington, D.C.