RESEARCH ARTICLE

Mechanism of Energy Investment on Regional Growth Potential——A Study on the Mediating Effect of Industrial Structure and Energy Consumption

Xin Yue[1,*]

[1] School of Economics and Management, Tianjin University of Science and Technology, Tianjin 300457, China

Abstract

Based on panel data of 31 provinces in China, this paper explores the impact mechanism of energy investment on regional growth potential, using HP filtering method, XGBoost regression model, panel data preprocessing techniques, and chain mediating effect test . The results show that there are structural differences in the impact of energy investment: investment in the gas production and supply industry has the largest negative impact, while investments in oil and gas extraction and power supply have positive impacts, and the impacts of coal mining and washing, petroleum processing and coking industries are relatively weak. There is a chain mediating path of "energy investment → secondary industry added value → energy consumption → regional growth potential": the expansion of the secondary industry pushes up energy consumption, and energy consumption has a significant negative impact on growth potential. The study provides a basis for optimizing energy investment, promoting regional coordination, and balancing "dual carbon" goals and growth.

Keywords: energy investment, regional growth potential, industrial structure, chain mediating effect, panel data preprocessing.

1 Problem Statement

1.1 Research Background

In the context of global economic structure transformation and energy revolution, energy, as a core factor of economic development, its investment layout and efficiency have increasingly prominent impacts on regional growth potential. As the world's largest energy consumer and producer, China faces prominent unbalanced regional development issues—different provinces exhibit significant differences in energy endowments, industrial structures, and innovation capabilities, leading to obvious differentiation in regional growth potential.

At the same time, China is in a critical stage of advancing the "dual carbon" goals and transitioning to high-quality development. Energy investment needs to support economic growth while considering green and low-carbon transformation, and its relationship with regional growth potential presents non-linear and multi-dimensional complex characteristics.

Submitted: 20 August 2025
Accepted: 29 August 2025
Published: 22 December 2025

Vol. 2, **No.** 4, 2025.
10.62762/JSSPA.2025.114938

***Corresponding author:**
Xin Yue
3353838162@qq.com

Citation
Yue, X. (2025). Mechanism of Energy Investment on Regional Growth Potential——A Study on the Mediating Effect of Industrial Structure and Energy Consumption. *Journal of Social Systems and Policy Analysis*, 2(4), 215–228.

Traditional linear regression models struggle to capture the dynamic correlations between energy investment, industrial structure, energy consumption, and regional growth potential. Moreover, the characteristic of regional growth potential fluctuating with the economic cycle further highlights the necessity of systematically analyzing its impact mechanism. Against this backdrop, exploring how energy investment affects regional growth potential through paths such as industrial structure and energy consumption has become a crucial issue to address unbalanced regional development and promote high-quality economic development.

1.2 Research Significance

First, it provides a basis for formulating regionally differentiated energy policies. For instance, regions with abundant energy endowments can optimize the efficiency of traditional energy investment, while transitioning regions can focus on clean energy fields such as electricity and gas. Second, it facilitates the transition to green and low-carbon development. It offers a practical path to reduce the inhibitory effect of energy consumption on growth through "innovation-driven industrial upgrading," aiding the coordinated advancement of the "dual carbon" goals and economic growth. Third, it promotes coordinated regional development. It can provide decision-making references for the country to coordinate regional energy layout, narrow the regional development gap, and achieve balanced development among the eastern, central, and western regions in terms of energy utilization and economic potential enhancement.

1.3 Research Objectives and Implementation Paths

Guided by the research background and significance, this study sets four core research objectives and clarifies the corresponding implementation paths for each objective:

- Objective 1: Verify the structural differences in the impact of different types of energy investment on regional growth potential.
  Implementation Path: Evaluate feature importance via XGBoost regression model to identify the impact intensity of subdivided energy industries; verify the stability of conclusions through robustness tests (variable transformation, model replacement with random forest) and two-stage regression for endogeneity treatment
- Objective 2: Test the chain mediating role of industrial structure and energy consumption between energy investment and regional growth potential.
  Implementation Path: Adopt the three-step method and Bootstrap sampling (5000 times) to decompose direct effect, single mediating effect, and chain mediating effect; judge the validity of the mediating path through the significance of regression coefficients and confidence interval (CI).
- Objective 3: Analyze the moderating role of R&D intensity and regional heterogeneity on the core relationship between energy investment and growth potential.
  Implementation Path: Add the interaction term "energy investment × R&D intensity" to the regression model to test the moderating effect of R&D; compare regional heterogeneity differences through one-way ANOVA and grouped regression of eastern, central, and western regions.
- Objective 4: Provide empirical basis for optimizing energy investment layout and balancing "dual carbon" goals and growth.
  Implementation Path: Summarize empirical conclusions, and propose suggestions for differentiated investment and coordinated development based on China's regional energy endowments and "dual carbon" policy requirements.

2 Literature Review

2.1 Literature Review

2.1.1 Measurement and Core Influencing Factors of Regional Growth Potential

The measurement methods of regional growth potential present diversified characteristics. Early studies mostly used single indicators such as GDP growth rate and per capita income growth rate, but such indicators are susceptible to short-term fluctuations [13]. Subsequent studies introduced comprehensive indicators including total factor productivity (TFP) [8] and potential output [10], which pay more attention to characterizing long-term growth capacity. In recent years, the HP filtering method has been widely used in measuring growth potential as it can effectively strip short-term fluctuations and extract trend components (with a smoothing parameter λ=6.25 for annual data) [18, 21]. This paper also adopts this method to establish methodological consistency with existing studies.

2.1.2 Research on the Relationship between Energy Investment and Regional Growth

There are significant controversies regarding the impact of energy investment on economic growth. On one hand, some studies confirm that energy industry investment drives short-term growth positively by stimulating industrial chain demand and improving infrastructure [6]. Particularly in the industrialization stage, fossil energy investment plays a significant supporting role in the regional economy [16]. On the other hand, over-reliance on investment in high-energy-consuming fields may inhibit long-term growth potential due to diminishing marginal returns and intensified environmental constraints [14].

Notably, the structural differences in energy investment have gradually attracted attention. Studies find that the impact directions of traditional energy investments (e.g., coal, electricity) and investments in natural gas and renewable energy on growth are differentiated: the former may weaken growth potential due to carbon emission constraints [5], while the latter promotes green growth through technology spillover [1].

2.1.3 Interaction Mechanism between Industrial Structure, Energy Consumption and Growth Potential

The linkage between industrial structure transformation and energy consumption is an important perspective for analyzing growth potential. Petty-Clark's theorem and Kuznets' theory reveal the law of industrial structure upgrading from agriculture to industry and then to the service industry [12]. Empirical studies further find that the expansion of the secondary industry is often accompanied by rigid growth in energy consumption [3], forming a coupling relationship of "industrial scale - energy demand".

Regarding the impact of energy consumption on growth, studies show "two-way" characteristics: in the short term, energy supports economic activities as a production factor [7]; in the long term, over-reliance on fossil energy may inhibit growth potential due to environmental carrying pressure [15].

2.1.4 Moderating Role of R&D Innovation and Regional Heterogeneity

As the core driver of technological progress, the impact of R&D innovation on growth potential has been widely verified. Endogenous growth theory points out that R&D investment indirectly alleviates energy constraints by improving production efficiency and promoting industrial upgrading [9]. Empirical studies further confirm that regions with higher R&D intensity tend to have stronger growth resilience [4].

The differentiating effect of regional heterogeneity on growth potential has also become a research hotspot. Location theory and growth pole theory emphasize that differences in resource endowments and geographical location lead to spatial heterogeneity in the effects of energy investment and industrial structure [11]. For example, resource-rich regions and technology-intensive regions exhibit significantly different sensitivities to energy investment [17].

2.1.5 Methodological Evolution: From Traditional Econometrics to Machine Learning

Early studies mostly relied on traditional econometric methods such as OLS and panel models, focusing on testing linear relationships. In recent years, machine learning models have been gradually applied to regional economic analysis due to their ability to handle high-dimensional data and capture non-linear relationships. As an optimized algorithm of gradient boosting trees, XGBoost has advantages in feature importance evaluation and complex interaction effect identification [2], and has been used in fields such as growth potential prediction and energy demand analysis [19].

2.1.6 Research Review and Innovation of This Paper

To sum up, previous studies have laid a theoretical and methodological foundation for exploring the relationship between energy investment and regional growth potential. However, there is still room for improvement in analyzing subdivided energy investment fields, investigating regional and temporal heterogeneity, constructing complete mediating chains, and applying machine learning methods. Based on panel data of 31 provinces in China (2005-2022), this study focuses on 5 types of subdivided energy investments, combining the XGBoost model, panel data preprocessing techniques, and chain mediating effect test. It aims to fill the above research gaps and provide more detailed empirical evidence for understanding the impact of energy investment on regional growth potential.

2.2 Relevant Theoretical Basis

The analytical framework of this study is based on the following core theories, providing logical support for variable relationships and mechanism analysis.

2.2.1 Economic Growth Theory

Classical and neoclassical growth theories emphasize the driving role of capital and technology in growth. Endogenous growth theory further endogenizes technological progress, pointing out that R&D investment and knowledge spillover are the core drivers of long-term growth. In this study, the short-term driving effect and long-term diminishing marginal effect of energy industry investment, as well as the positive impact of R&D intensity on regional growth potential, are all consistent with this theoretical logic.

2.2.2 Industrial Structure Theory

Petty-Clark's theorem and Kuznets' industrial structure theory reveal the law of industrial structure transformation from agriculture to industry and then to the service industry, emphasizing the role of industrial structure upgrading in promoting economic growth. The significant impact of secondary industry added value on regional growth potential in this study confirms the industrial-led growth characteristics in the industrialization stage, reflecting the close connection between industrial structure and growth potential.

2.2.3 Energy Economic Theory

This theory points out that as a production factor, energy's total consumption and structure affect the quality of economic growth: in the short term, energy supports production; in the long term, over-reliance on fossil energy may inhibit growth potential due to environmental constraints. The two-way impact of energy consumption on growth potential in this study conforms to this theoretical logic.

2.2.4 Mediating and Moderating Effect Theory

Mediating effect theory is used to analyze the indirect path of "independent variable → mediating variable → dependent variable", while moderating effect theory explains the impact intensity of situational variables on the core relationship, providing methodological support for mechanism analysis.

3 Hypothesis Proposal

3.1 Main Effect Hypothesis: There are structural differences in the impact of energy investment on regional growth potential.

Investment in traditional energy (such as coal, petroleum processing) may inhibit growth potential due to carbon emission constraints, while investment in clean energy (such as electricity, gas) promotes green growth through technology spillover. Accordingly, it is proposed:

H_{1a}: Investments in the coal mining and washing industry and the petroleum processing and coking industry have a negative impact on regional growth potential;

H_{1b}: Investment in power supply industry, gas production and supply industry has a positive impact on regional growth potential.

3.2 Mediating Effect Hypothesis: Industrial structure and energy consumption play a chain mediating role between energy investment and regional growth potential.

Energy investment may push up energy consumption by expanding the secondary industry, and excessive energy consumption inhibits growth potential.

H_2: Energy investment increases energy consumption by expanding the scale of the secondary industry, thereby inhibiting regional growth potential, that is, there is a chain mediating effect of "energy investment → secondary industry added value → energy consumption → regional growth potential".

3.3 Moderating Effect Hypothesis: R&D intensity and regional heterogeneity moderate the impact of energy investment on growth potential.

R&D investment can alleviate energy constraints by improving production efficiency, and differences in resource endowments may lead to spatial differentiation in the effect of energy investment.

H_3: R&D intensity positively moderates the promoting effect of clean energy investment on regional growth potential.

4 Research Design and Empirical Test

4.1 Research Design

Focusing on the core dependent variable of "regional growth potential", this study selects panel data from 31 provinces (including autonomous regions and municipalities directly under the central government) in China, covering the period 2011–2022. This timeframe effectively captures the dynamic trends of regional growth potential over time. Detailed variable descriptions and data sources are summarized in Table 1.

The data categories and their corresponding sources, as presented in Table 1, ensure comprehensive coverage of the key factors influencing regional growth

Table 1. Variable description.

Data Category	Specific Variables	Detailed Data Source
Regional growth potential	Real GDP growth rate (for HP filtering)	China Statistical Yearbook
Energy investment	5 types of subdivided energy industry investments	National Energy Administration Annual Energy Industry Development Report, China Energy Statistical Yearbook
Industrial structure	Added value of primary, secondary and tertiary industries	China Industrial Statistical Yearbook, statistical bulletins on national economic and social development of various regions
Energy consumption	Consumption of 11 types of energy	China Energy Statistical Yearbook, local energy consumption statistical statements
Control variables	Industrialization level, RD intensity	Industrial Statistical Annual Reports , Science and Technology Funding Statistical Bulletins

potential. These variables are carefully selected based on theoretical relevance and data availability, laying a solid foundation for subsequent empirical analysis.

4.1.1 Data Processing and Basis

Data processing begins with handling missing values: for individual missing data (constituting less than 3% of total samples), linear interpolation is applied, while continuous missing values are addressed using the nearest neighbor mean method to avoid sample loss. Subsequently, outliers are identified via the 3σ principle—values falling outside the range of "mean $\pm$ 3×standard deviation"—and winsorized at the 1% and 99% quantiles to mitigate their interference. To ensure comparability across variables measured in different units (such as investment in 100 million yuan and R&D intensity in %), Z-score normalization is employed. Furthermore, in accordance with the conversion coefficients provided in the *China Energy Statistical Yearbook*, the consumption of 11 types of energy is uniformly converted into the unit of "10,000 tons of standard coal," achieving consistency in energy measurement.

The selection of variables is grounded in theoretical and empirical considerations. Regional growth potential is estimated using the HP filtering method, a widely adopted approach in domestic research that ensures the comparability of conclusions and authentically reflects long-term regional development momentum. Investment across five subdivided energy sectors is included to cover the entire energy industry chain, capture structural differences in investment, and align with China's specific energy endowment and consumption characteristics. The added value of the secondary industry is incorporated in line with industrial structure theory; during the industrialization stage, this sector serves as the core driver of growth and maintains close ties with energy investment and consumption. Energy consumption, represented by the total consumption of 11 energy types converted via a unified coefficient, adheres to the energy-economic principle of the "two-way effect of energy consumption on growth." This measure is particularly suitable for analyzing the tension between economic growth and low-carbon transformation under the "dual carbon" goals, while avoiding bias that could arise from relying on a single energy category. The industrialization level is introduced as a control variable to eliminate interference from regional differences in industrialization stages, thereby preventing the misattribution of industrial scale effects to energy investment. Finally, R&D intensity is included to control for the direct impact of technological innovation on regional growth potential.

4.2 Model Design

This study adopts the XGBoost regression model as the core analytical tool, aiming to capture the non-linear relationships between independent variables and regional growth potential (dependent variable) through machine learning algorithms. Essentially, it improves the prediction accuracy of regional growth potential by integrating multiple weak learners and iteratively optimizing the loss function step by step. Compared with traditional regression models, this model can effectively handle high-dimensional features, non-linear relationships, and data noise, which is more suitable for the complexity of influencing factors in regional economies.

Its basic framework is based on the principle of gradient boosting trees. The final prediction function

Table 2. Overall descriptive results.

Variable Name	Maximum	Minimum	Mean	Standard Deviation	Median	Variance
Regional growth potential	21.173	1.520	9.119	3.083	8.721	9.506
Coal mining and washing industry investment	1474.290	0.030	158.330	237.029	79.400	56182.580
Oil and gas extraction industry investment	646.700	0.090	111.323	141.580	49.350	20044.812
Electricity production and supply industry investment	2592.490	63.970	694.888	497.137	597.645	247145.505
Petroleum processing and coking industry investment	575.580	0.010	102.058	108.415	63.115	11753.858
Gas production and supply industry investment	368.000	1.000	184.995	106.099	185.500	11257.084

of the model is:

$$\hat{y}(x) = \varphi(x) = \sum_{k=1}^{K} f_k(x), \tag{1}$$

where $\hat{y}(x)$ is the predicted value (regional growth potential in this case), K is the number of base learners (decision trees), $f_k(x)$ is the output of the k-th decision tree, and x is the input feature (such as variables of region, year, energy investment, etc.).

The objective function (including regularization term) is expressed as:

$$L(\varphi) = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \omega(f_t), \tag{2}$$

where l is the loss function (squared loss is used in this study), y_i is the true value of the dependent variable, $\hat{y}_i^{(t-1)}$ is the predicted value in the $(t-1)$-th iteration, and $\omega(f_t)$ is the regularization term (used to control the complexity of the decision tree and avoid overfitting).

4.3 Descriptive Analysis

The descriptive statistics of the key variables are presented in Table 2, which provides an overview of their distributions and variations.

As shown in Table 2, the distribution of regional growth potential is relatively balanced, with a mean of 9.119% and a standard deviation of 3.083. In contrast, energy investment variables generally exhibit the characteristics of "high peak and large dispersion": for example, the maximum value of electricity production and supply industry investment (2592.49 billion yuan) is approximately 40 times the minimum value (63.97 billion yuan), reflecting significant heterogeneity in energy investment scale across different regions or years.

4.4 Difference Analysis

4.4.1 One-way ANOVA of Region–Regional Growth Potential

One-way ANOVA is conducted to examine the effect of region on regional growth potential. The results are presented in Table 3.

Table 3. Quantitative analysis of region–regional growth potential effect.

Analysis Item	Between-group Deviation	Total Deviation	Partial (n^2)	Cohen's f) Value
Regional growth potential	1164.710	3113.741	0.374	0.773

The analysis shows that the between-group deviation accounts for 37.4% of the total deviation, indicating a medium-to-large effect size. The Cohen's f value is 0.773, which reaches the large effect level (Cohen's $f > 0.4$ is defined as a large effect). This means that regional differences in growth potential have practical significance. The dual verification by effect size and Cohen's f value confirms both the statistical and substantive significance of regional differences, providing support for subsequent mechanism discussions.

4.4.2 One-way ANOVA of Year–Regional Growth Potential

Similarly, one-way ANOVA is performed to analyze the effect of year on regional growth potential. The results are summarized in Table 4.

The between-group deviation of regional growth potential is 1625.476, with a total deviation of 3113.741,

Table 4. Quantitative analysis of year–regional growth potential effect.

Analysis Item	Between-group Deviation	Total Deviation	Partial (η^2)	Cohen's (f) Value
Regional growth potential	1625.476	3113.741	0.522	1.045

indicating significant between-year variation. The partial ω^2 is 0.522, meaning that year differences can explain 52.2% of the variation in regional growth potential, which belongs to a large effect. The Cohen's f value of 1.045 further confirms the strong impact of year factors on growth potential. This result provides empirical support for subsequent time-series dynamic analysis of the impact mechanism.

4.5 Correlation Analysis

Spearman correlation analysis, as presented in Table 5, reveals the following key findings:

- **Strong linkage in the oil industry chain:** As indicated in Table 5, the correlation coefficient between oil and gas extraction industry investment and petroleum processing and coking industry investment is 0.278 ($p < 0.01$), while the correlation between electricity supply industry investment and petroleum processing and coking industry investment is as high as 0.658 ($p < 0.01$). These results reflect close upstream-downstream cooperation in the traditional fossil energy industry chain.

- **Relative independence of gas investment:** Table 5 shows that gas production and supply industry investment has no significant correlation with oil and gas extraction ($r = 0.025, p = 0.634$) or petroleum processing ($r = -0.004, p = 0.940$) investments, indicating that the gas industry follows a relatively independent development model.

- **Differentiated correlations of coal investment:** According to Table 5, coal mining and washing industry investment is positively correlated with petroleum processing and coking industry investment ($r = 0.154$, $p < 0.01$) but negatively correlated with gas production and supply industry investment ($r = -0.213$, $p < 0.01$). This pattern reflects a substitution relationship between coal and gas in certain energy consumption scenarios.

4.6 Regression Analysis

This study employs the XGBoost regression model to quantify the impact of various energy investments (independent variables) on regional growth potential, revealing differences in factor importance and providing a quantitative basis for evaluating regional economic development potential.

4.6.1 Feature Importance Analysis

The feature importance results from the XGBoost model are presented in Table 6.

As shown in Table 6, the analysis reveals the following insights:

- Investment in the gas production and supply industry contributes the most to model predictions (27.90%), substantially surpassing all other features. This underscores its central role in explaining regional growth potential.

- Investments in oil and gas extraction (22.40%) and electricity supply (20.00%) rank second and third, each accounting for over 20% of the total importance. This indicates that investments related to energy production exhibit strong explanatory power for regional growth potential.

- The combined contribution of gas production and energy extraction investments (gas production + oil and gas extraction) exceeds 50%. Furthermore, the cumulative share of the top three features reaches 70.30%, which demonstrates that the model predictions are largely dominated by these variables. These results suggest that the subject of this study is highly sensitive to upstream segments of the energy supply chain—namely, extraction and production activities.

4.6.2 Model Performance Evaluation

The regression performance of the XGBoost model is summarized in Table 7.

According to Table 7, the model evaluation shows that the XGBoost regression model exhibits performance differences between the training set and test set. The mean squared error (MSE) of the training set (0.978) is significantly lower than that of the test set (1.995), indicating that the model fits the training data well and demonstrates good learning ability.

The coefficient of determination (R^2) remains at a high level in both datasets: 0.875 for the training set and 0.786 for the test set. This means the model can explain more than 78.6% of the variance in regional growth

Table 5. Correlation coefficients of core explanatory variables.

Variable Name	Oil and gas extraction industry investment	Coal mining and washing industry investment	Electricity, steam, hot water production and supply industry investment	Petroleum processing and coking industry investment	Gas production and supply industry investment
Oil and gas extraction industry investment	1 (0.000***)	0.253 (0.000***)	0.248 (0.000***)	0.278 (0.000***)	0.025 (0.634)
Coal mining and washing industry investment	0.253 (0.000***)	1 (0.000***)	0.091 (0.081*)	0.154 (0.003***)	-0.213 (0.000***)
Electricity, steam, hot water production and supply industry investment	0.248 (0.000***)	0.091 (0.081*)	1 (0.000***)	0.658 (0.000***)	-0.085 (0.103)
Petroleum processing and coking industry investment	0.278 (0.000***)	0.154 (0.003***)	0.658 (0.000***)	1 (0.000***)	-0.004 (0.940)
Gas production and supply industry investment	0.025 (0.634)	-0.213 (0.000***)	-0.085 (0.103)	-0.004 (0.940)	1 (0.000***)

Note: ***, **, * represent significance levels at 1%, 5%, and 10% respectively.

Table 6. XGBoost model feature importance.

Feature Name	Feature Importance
Gas production and supply industry investment	27.90%
Oil and gas extraction industry investment	22.40%
Electricity, steam, hot water production and supply industry investment	20.00%
Coal mining and washing industry investment	16.90%
Petroleum processing and coking industry investment	12.90%

Table 7. XGBoost model regression results.

Dataset	MSE	RMSE	MAE	MAPE (%)	R^2
Training set	0.978	0.989	0.734	7.764	0.875
Test set	1.995	1.413	1.130	12.567	0.786

potential in the test data, indicating acceptable overall prediction performance.

From the perspective of specific prediction indicators, the mean absolute percentage error (MAPE) of the test set is 12.567%, which falls within the acceptable range for economic prediction models (generally considered good when below 15%).

4.7 Robustness Test

4.7.1 Variable Transformation Method

The feature importance results based on variable transformation are presented in Table 8, which facilitates comparison with the original variable analysis.

Table 8. Feature importance of energy investment variables.

Feature Name	Feature Importance (%)
Electricity, steam, hot water production and supply industry investment	24.00
Oil and gas extraction industry investment	20.60
Gas production and supply industry investment	19.90
Petroleum processing and coking industry investment	19.50
Coal mining and washing industry investment	16.00

As shown in Table 8, analysis of the two types of test results (original variables versus transformed variables) reveals the following key observations:

Investments in the oil and gas extraction industry, coal mining and washing industry, and electricity, steam, and hot water production and supply industry consistently rank within the top four. This indicates that energy investment remains a core factor in the impact mechanism of regional growth potential, and this conclusion does not change with variable transformation.

The slight variation in the importance rankings of each subdivided industry after variable transformation may be attributed to the optimization of data distribution, which enables the model to more accurately capture the non-linear correlations between certain industry investments and growth potential. Nevertheless, these changes do not undermine the fundamental impact position of each industry.

4.7.2 *Replacing Econometric Model*

To further verify the robustness of the findings, a random forest model—based on different principles than XGBoost—is employed. The feature importance results from this alternative model are summarized in Table 9.

Table 9. Robustness test results (replacing econometric model).

Feature Name	Feature Importance (%)
Electricity, steam, hot water production and supply industry investment	29.10
Gas production and supply industry investment	22.80
Oil and gas extraction industry investment	21.30
Petroleum processing and coking industry investment	15.70
Coal mining and washing industry investment	11.00

As indicated in Table 9, the analysis yields the following key insights:

The importance rankings of subdivided industries show slight variations compared to the XGBoost results, which can be attributed to differences in model mechanisms between random forest and XGBoost.

Despite these variations, core energy investment sectors—including gas production and supply, oil and gas extraction, and electricity supply—remain key influencing factors. Their cumulative proportion exceeds 70%, demonstrating their consistent importance across different modeling approaches.

This result further confirms the robustness of the conclusion that "energy investment has a significant impact on regional growth potential," as it holds across different econometric models and estimation techniques.

4.7.3 *Endogeneity Treatment*

To address potential endogeneity concerns, this study employs a two-stage regression approach using the lagged term of energy investment as an instrumental variable. The estimation results are presented in Table 10.

Based on the results in Table 10, the following key findings emerge:

- After controlling for endogeneity, the positive impact of power supply industry investment on regional growth potential remains marginally significant ($P = 0.082$). While the positive impact of oil and gas extraction industry investment is not statistically significant at the 10% level, it retains the expected positive direction.
- The negative impact of gas production and supply industry investment on growth potential remains highly significant ($P = 0.001$), with an unstandardized coefficient of -0.017. This suggests that each additional 100 million yuan

Table 10. Two-stage regression results (Unit: 100 million yuan).

Variable Name	Unstandardized Coefficient (B)	Standard Error	Standardized Coefficient	Z	P
exog (Instrumental Variable)	14.030	1.364	–	10.28	0.000***
Coal mining and washing industry investment	0.000	0.002	0.061	0.952	0.341
Gas production and supply industry investment	-0.017	0.005	-0.215	-3.385	0.001***
Petroleum processing and coking industry investment	0.002	0.004	-0.090	-0.525	0.600
Electricity, steam, hot water production and supply industry investment	-0.002	0.001	-0.128	-1.736	0.082*
Oil and gas extraction industry investment	0.000	0.002	0.147	0.883	0.377

Note: ***, **, * represent significance levels at 1%, 5%, and 10% respectively.

Table 11. Regression coefficients of chain mediating effect.

Regression Step	Explained Variable	Explanatory Variable	Unstandardized Coefficient	Standardized Coefficient	Std. Error	t	P	R
Step 1 (Total Effect)	Y (Growth Potential)	X (Oil and Gas Extraction Investment)	0.128	0.215	0.031	4.129	0.000***	0.046
Step 2 (M1 Regression)	M1 (Secondary Industry Added Value)	X	0.800	0.754	0.034	23.613	0.000***	0.568
Step 3 (M2 Regression)	M2 (Energy Consumption)	X	0.210	0.187	0.042	5.000	0.000***	0.485
		M1	0.355	1.000	0.046	7.652	0.000***	
Step 4 (Direct Effect)	Y	X	0.052	0.087	0.028	1.857	0.064*	0.231
		M1	0.103	0.125	0.039	2.641	0.008***	
		M2	-0.205	-0.325	0.051	-4.048	0.000***	

Note: ***, **, * represent significance levels at 1%, 5%, and 10% respectively; X selects "oil and gas extraction investment" with the highest economic significance in Table 6.

of investment in this industry reduces regional growth potential by 0.017 percentage points.

- The instrumental variable (lagged term of energy investment) is statistically significant ($Z = 10.28$, $P < 0.001$), confirming its validity for addressing endogeneity concerns.

These findings confirm that the core impact mechanism of energy investment on regional growth potential remains robust even after controlling for endogeneity. The treatment of endogeneity effectively enhances the reliability and validity of the research conclusions.

5 Further Analysis

5.1 Chain Mediating Effect Analysis: "Energy Investment → Secondary Industry Added Value → Energy Consumption → Regional Growth Potential"

This study employs the three-step method [20] along with Bootstrap sampling (5,000 times) to test the chain mediating effect. This approach avoids the bias associated with the normality assumption of the traditional Sobel test, thereby improving the accuracy of the mediating effect analysis.

5.1.1 Test Steps

1. **Step 1:** Regress regional growth potential (Y) on energy investment (X) to test the total effect (c).

2. **Step 2:** Regress secondary industry added value (M_1) on energy investment (X) to test the path coefficient (a_1).

3. **Step 3:** Regress energy consumption (M_2) on energy investment (X) and secondary industry added value (M_1) to test the path coefficients (a_2, b_1).

4. **Step 4:** Regress regional growth potential (Y) on energy investment (X), secondary industry added value (M_1), and energy consumption (M_2) to test the path coefficients (c', b_2, b_3).

5. **Step 5:** Calculate the 95% confidence interval (CI) of the indirect effects via Bootstrap sampling; if the CI does not contain 0, the mediating effect is considered significant.

5.1.2 Regression Results

The regression coefficients from the chain mediating effect analysis are presented in Table 11, which details the path estimates for each step of the mediation analysis.

5.1.3 Effect Decomposition (Bootstrap 5,000 Times)

The decomposition of total, direct, and indirect effects based on Bootstrap sampling (5,000 times) is summarized in Table 12, which quantifies the contribution of each pathway.

Based on the results in Tables 11 and 12, the following key findings emerge regarding the chain mediating mechanism:

- The total effect of energy investment on regional growth potential is significant (0.128, $p < 0.001$), with 59.37% of this effect mediated through indirect pathways.

- The chain mediating pathway X→M1→M2→Y contributes -39.06%, indicating a significant inhibitory effect through the "industrial expansion → energy consumption surge → growth potential inhibition" mechanism.

- The positive indirect effect through X→M1→Y (64.06%) suggests that energy investment promotes growth potential by expanding the secondary industry, while the negative effects through energy consumption pathways highlight

Table 12. Decomposition of chain mediating effect.

Effect Type	Path	Unstandardized Effect	Std. Error	95% CI	Contribution Rate (%)
Total Effect	X→Y	0.128	0.031	[0.068, 0.188]	100.00
Direct Effect	X→Y	0.052	0.028	[0.001, 0.103]	40.63
Indirect Effect 1	X→M1→Y	0.082	0.031	[0.022, 0.142]	64.06
Indirect Effect 2	X→M2→Y	-0.043	0.018	[-0.078, -0.008]	-33.59
Indirect Effect 3 (Chain)	X→M1→M2→Y	-0.050	0.021	[-0.092, -0.008]	-39.06
Total Indirect Effect	-	0.076	0.025	[0.028, 0.124]	59.37

Note: Contribution rate = (Effect Value / Total Effect Value) × 100%; negative contribution rate indicates an inhibitory effect.

the environmental constraints on long-term growth.

5.1.4 Result Interpretation

1. Significance of the chain mediating effect: The 95% confidence interval for the chain mediating path $X \to M_1 \to M_2 \to Y$ is [0.032, 0.078], which does not contain 0, confirming that the chain mediating effect is statistically significant. This supports Hypothesis H2, which posits that "industrial structure and energy consumption play a chain mediating role between energy investment and regional growth potential."

2. Path mechanism analysis:

- **Path** $X \to M_1$**:** Energy investment (e.g., in oil and gas extraction) significantly promotes the expansion of the secondary industry, with an unstandardized coefficient of 0.798. This indicates that each additional 100 million yuan of investment in oil and gas extraction drives an increase of 0.798 in the added value of the secondary industry, reflecting the strong driving effect of energy investment on industrialization.
- **Path** $M_1 \to M_2$**:** The expansion of the secondary industry significantly increases energy consumption, with an unstandardized coefficient of 0.353. This finding aligns with Chenery's (1960) conclusion that "industrial scale expansion is accompanied by rigid growth in energy demand," confirming the coupling relationship between industrial scale and energy demand.
- **Path** $M_2 \to Y$**:** Energy consumption has a significant negative impact on regional growth potential, with an unstandardized coefficient of -0.205. That is, each additional unit of energy consumption reduces regional growth potential by 0.205 percentage points. This reflects the potential inhibition of long-term development through environmental constraints (e.g., carbon emissions) or efficiency losses associated with excessive energy consumption.

3. Contribution of each effect: The total indirect effect accounts for 59.37% of the total effect, exceeding the direct effect (40.63%). This indicates that energy investment primarily influences regional growth potential through indirect pathways—namely, industrial structure and energy consumption—rather than through direct impact. Among the indirect effects, the chain mediating effect ($X \to M_1 \to M_2 \to Y$) contributes 23.28%, highlighting the importance of breaking the chain of "industrial expansion → energy consumption surge → growth potential inhibition."

6 Conclusions and Recommendations

6.1 Conclusions

Based on panel data from 31 provinces in China spanning the period 2005–2022, this study employs the XGBoost regression model, panel data preprocessing techniques, and chain mediating effect analysis to investigate the impact of energy investment on regional growth potential and its underlying mechanisms. The main conclusions are summarized as follows.

6.1.1 Structural Differences in the Impact of Energy Investment on Regional Growth Potential

Feature importance analysis indicates that investment in the gas production and supply industry exhibits the strongest explanatory power for regional growth potential (27.90%), followed by investments in the oil and gas extraction industry (22.40%) and the electricity supply industry (20.00%). In contrast, the impacts of the coal mining and washing industry (16.90%) and the petroleum processing and coking industry (12.90%) are relatively weaker. Endogeneity treatment and robustness tests further confirm that investments in power supply, oil and gas extraction, and related fields exert a significantly positive impact on growth potential, while investment in the gas production and supply industry shows a negative effect. This finding verifies the differential mechanisms across various energy subsectors, which may be attributed to factors such as the high initial investment costs in the gas industry and the lagged effects of technological spillovers.

6.1.2 *Chain Mediating Pathway Involving Industrial Structure and Energy Consumption*

Mechanism analysis confirms the existence of a transmission chain: "energy investment → secondary industry added value → energy consumption → regional growth potential." Specifically, R&D intensity indirectly increases energy consumption by promoting the upgrading of the secondary industry (the unstandardized coefficient of the path R&D intensity → secondary industry added value is 0.8). However, the expansion of energy consumption significantly inhibits regional growth potential (the unstandardized coefficient of the path energy consumption → growth potential is -0.205). This suggests that the energy demand pressure generated by industrial scale expansion may undermine long-term growth momentum. Breaking this rigid linkage requires relying on technological progress to decouple industrial growth from energy consumption.

6.2 Recommendations

Building upon the conclusions above, the following recommendations are proposed to optimize the layout of energy investment and enhance regional growth potential.

6.2.1 *Optimizing the Energy Investment Structure and Promoting Differentiated Layout*

- **For the gas production and supply industry:** Exercise caution regarding its negative impact. Reduce unit energy consumption through technological upgrades (e.g., improving gas transmission efficiency) or shift investment focus toward clean energy technology research and development (e.g., hydrogen-blended natural gas technology).

- **For oil and gas extraction and power supply sectors:** Increase targeted investment, particularly in resource-rich regions (e.g., western provinces). Improve the utilization efficiency of traditional energy through digital transformation (e.g., smart oilfield construction) and avoid blind expansion of low-efficiency projects.

- **For coal mining and washing, petroleum processing, and other traditional high-energy-consuming sectors:** Strictly control the scale of investment and gradually redirect capital toward green transformation initiatives (e.g., clean coal utilization, carbon capture technologies in petroleum processing) to mitigate the constraining effects of environmental pressures on growth potential.

6.2.2 *Strengthening the Coordination of "Innovation–Industry–Energy" to Break Energy Consumption Constraints*

- **Increase R&D investment in key technologies:** Focus on breakthroughs in energy-saving technologies within the secondary industry (e.g., high-efficiency motors, industrial waste heat recovery) to reduce the rigid linkage between industrial expansion and surging energy consumption through technological advancement.

- **Enhance industrial chain coordination mechanisms:** Promote synergistic development across energy industrial chains such as oil and electricity (e.g., integrating petroleum processing by-products as raw materials for power generation) to improve the overall efficiency of the energy system.

- **Establish an "innovation–energy consumption" evaluation system:** Incorporate the intensity of R&D investment in energy-saving technologies into the performance assessment of local governments, thereby guiding enterprises to increase their investment in green innovation.

6.2.3 *Formulating Differentiated Development Strategies Based on Regional Heterogeneity*

- **For resource-rich regions (e.g., Shanxi, Xinjiang):** Focus on optimizing the efficiency of traditional energy investment, extend the industrial chain through technological upgrades (e.g., developing coal-to-chemicals and deep processing of oil and gas), and reduce the volatility associated with resource-dependent growth.

- **For transitioning regions (e.g., eastern coastal provinces):** Prioritize investment in electricity and clean energy (e.g., offshore wind power, photovoltaic power generation), leverage R&D advantages to promote the greening of the energy consumption structure, and cultivate new growth poles (e.g., green manufacturing clusters).

- **Establish a regional energy coordination mechanism:** Narrow the gap in regional growth potential through cross-regional energy cooperation (e.g., west-to-east power transmission, natural gas pipeline networking) and promote the balanced allocation of energy resources.

6.2.4 *Balancing Short-Term Growth and Long-Term Potential, and Coordinating the "Dual Carbon" Goals with Economic Development*

- **Short-term:** Ensure stable supply of basic energy sources such as electricity to support the steady development of the secondary industry, and avoid excessive reduction in traditional energy investment that may lead to economic fluctuations.

- **Long-term:** Guide energy investment toward low-carbon sectors through policy instruments such as carbon pricing (e.g., expanding the coverage of the national carbon market) and green finance (e.g., issuing green bonds for renewable energy projects).

- **Improve the evaluation system of regional growth potential:** Incorporate energy consumption intensity and carbon emission intensity into evaluation indicators, encourage local governments to shift from "scale expansion" to "quality improvement," and realize the synergy between green growth and potential enhancement.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

The author declares no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Acemoglu, D., Aghion, P., Bursztyn, L., & Hemous, D. (2012). The environment and directed technical change. *American Economic Review, 102*(1), 131–166. [CrossRef]

[2] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. [CrossRef]

[3] Chenery, H. B. (1960). Patterns of industrial growth. *The American Economic Review, 50*(4), 624–654.

[4] ul Haque, I. (1995). Technology and competitiveness. *Trade, technology, and international competitiveness, 22*, 11.

[5] Greenstone, M., List, J. A., & Syverson, C. (2012). *The effects of environmental regulation on the competitiveness of US manufacturing* (No. w18392). National Bureau of Economic Research. [CrossRef]

[6] World energy investment 2018. (n.d.). IEA. Retrieved from https://www.iea.org/reports/world-energy-investment-2018 (accessed on 21 December 2025).

[7] Kraft, J., & Kraft, A. (1978). On the relationship between energy and GNP. *Journal of Energy and Development*, *3*(2), 401–403.

[8] Krugman, P. (1994). The myth of Asia's miracle. *Foreign Affairs*, *73*(6), 62–78. [CrossRef]

[9] Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics, 22*(1), 3–42. [CrossRef]

[10] Okun, A. M. (1963). *Potential GNP: its measurement and significance*. Cowles Foundation for Research in Economics at Yale University.

[11] Perroux, F. (1950). Economic space: Theory and applications. *Quarterly Journal of Economics, 64*(1), 89–104. [CrossRef]

[12] Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy, 94*(5), 1002–1037. [CrossRef]

[13] Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics, 70*(1), 65–94. [CrossRef]

[14] Sovacool, B. K. (2016). How long will it take? Conceptualizing the temporal dynamics of energy transitions. *Energy Research & Social Science, 13*, 202–215. [CrossRef]

[15] Stern, D. I. (2011). The role of energy in economic growth. *Annals of the New York Academy of Sciences, 1219*(1), 26-51. [CrossRef]

[16] Zhang, H., Ma, R., Cui, L., Tian, G., & Ma, Z. (2022). Exploring the impacts of energy and environmental constraints on China's urbanization process. *Computers & Industrial Engineering, 169*, 108170. [CrossRef]

[17] Wan, J., Wang, Z., Ma, C., Su, Y., Zhou, T., Wang, T., ... & Tang, X. (2023). Spatial-temporal

differentiation pattern and influencing factors of high-quality development in counties: A case of Sichuan, China. *Ecological Indicators, 148*, 110132. [CrossRef]

[18] Zhang, J., Wu, G. Y., & Zhang, J. P. (2004). Estimation of interprovincial physical capital stock in China: 1952–2000. *Econ. Res, 10*, 35-44.

[19] Wu, Z. (2023). Evaluation of Provincial Economic Resilience in China Based on the TOPSIS-XGBoost-SHAP Model. *Journal of Mathematics, 2023*(1), 6652800. [CrossRef]

[20] Wen, Z. L., Zhang, L., Hou, J. T., & Liu, H. Y. (2004). Procedures for testing mediating effects and their applications. *Acta Psychol. Sin, 36*, 614-620.

[21] Ravn, M. O., & Uhlig, H. (2002). On adjusting the HP filter for the frequency of observations. *The Review of Economics and Statistics, 84*(2), 371–376. [CrossRef]