



**UHASSELT**

KNOWLEDGE IN ACTION

## Faculty of Business Economics

Master of Management

### **Master's thesis**

***The impact of digital finance on green investment and the mediation effect of return on assets: A replication and robustness analysis***

**Thi Chau Loan Nguyen**

Thesis presented in fulfillment of the requirements for the degree of Master of Management, specialization Data Science

### **SUPERVISOR :**

Prof. dr. Stephan BRUNS



**UHASSELT**

KNOWLEDGE IN ACTION

**www.uhasselt.be**  
Universiteit Hasselt  
Campus Hasselt:  
Martelarenlaan 42 | 3500 Hasselt  
Campus Diepenbeek:  
Agoralaan Gebouw D | 3590 Diepenbeek

**2024**  
**2025**



# **Faculty of Business Economics**

Master of Management

## ***Master's thesis***

***The impact of digital finance on green investment and the mediation effect of return on assets: A replication and robustness analysis***

**Thi Chau Loan Nguyen**

Thesis presented in fulfillment of the requirements for the degree of Master of Management, specialization Data Science

## **SUPERVISOR :**

Prof. dr. Stephan BRUNS



# **The impact of digital finance on green investment and the mediation effect of Return on Assets:**

## **A replication and robustness analysis**

**Student name: Nguyen Thi Chau Loan**

**Program: Master of Management: Data science**

**Supervisor: Prof. Dr. Stephan BRUNS**

### **Key words:**

Digital finance

Green investment

ROA

Heavily polluting industries

### **Summary**

Nowadays, in the context of environmental pollution seriously affecting the quality of life, green investment becomes a more noticeable highlight, especially the shift in corporate strategies toward sustainable development for the economy and environment (Yuan et al., 2018). Moreover, the innovation of digital finance has brought many positive changes in promoting economic efficiency, which brings many advantages to green investment projects (Wu & Huang, 2022). Previous studies have proved a positive correlation between the level of digital finance adoption and green investment. Accordingly, green investment projects tend to be more expanded in companies with high levels of digital finance application. However, the poor return rate linked to high risk in green investment projects may adversely affect the operational performance of firms and financial institutions funding the capital, leading to certain difficulties for green investment projects in accessing the economic resources. Therefore, evaluating the financial results of these projects becomes essential.

In this paper, we replicate the result of an existing study by Ding et al., (2023) to prove the relationship between digital finance and green investment with the same approach on data of heavily polluting industries in China from 2011 to 2020, and expand our research about the financial effectiveness that digital finance has made on green investment corporates by taking Return on Assets (ROA) into account.

This study employs the R programming language to execute panel models, approaching the objectives in two stages. In the early stage, this paper examines the coefficient of digital finance and green investment with and without the inclusion of size, age, leverage, growth, stock, GDP and upgrade as control variables. To handle panel data, we keep firm, year, industry and area as fixed factors for the regressions. The results highlight a positive relationship through different models: the higher the level of digital finance application, the greater the company's green investment index. We use specification curve analysis as a robustness test for the model. By that, the main independent variable remains stable in several test groups with different variables and confirms the outcome of the original study. An additional robustness test conducted in this study involves specifications fixed effects. We sequentially remove each fixed factor from the model and eventually eliminate all fixed effects to observe the impact on the results. The coefficients of Digfin and Ginvest remain positive when area and industry fixed effects are excluded. However, they turn negative upon the removal of year, firm, and all fixed factors showing the importance of firm and year as fixed factors in panel data. This outcome partially supports the robustness of our findings and underscores the vital role of including fixed effects in the model. Consequently, the result of the original paper by Ding et al., (2023) is confirmed once again.

Previous studies demonstrate that China has emerged as the global leader in digital financial development, which creates a lot of opportunities for economic entities to access capital. However, green investment corporates still face a big barrier to acquiring financial resources due to its drawbacks. Without the driving of government strategy and the stimulation of the domestic economy, this will become a challenging problem for these entities. That is why we execute a mediation model in the second stage, in which, ROA is a financial mediator in the relationship between digital finance and green investment to evaluate the business effectiveness via the financial aspect. However, the result shows that ROA does not have statistical significance in the mediation model.

In this study, other factors that may have an impact on green investment of heavy industries in China are considered. Our analysis emphasizes that there are some key policies released by the government from 2015 to 2016 that have made significant changes in many corporations in these industries. Particularly, the Supply-Side Structural Reforms (2015) has forced entities in the steel sector to reduce production and transition toward cleaner manufacturing processes while the *Guidelines for Establishing the Green Financial System* promote green finance at a national level to promote green investment. Additionally, the tariff impact from the US–China trade war is considered as a financial constraint which negatively affects heavily polluting industries in China, leads to a potential decrease of green investment level in these industries.

Once again, the result of this study strengthens the positive coefficients between digital finance and green investment through a linear regression model with trusted robustness tests and points out the political and economic factors that could impact the findings. However, there are some existing limitations in this research. The analysis uses the existing data and focuses mainly on heavily polluting industries in China, without considering other domains and regions, leading to narrow down the

applicability of the outcomes. The result of the mediation model fails to use a financial variable (ROA) as a mediator to assess the economic efficiency of applying digital finance in green investment

For future research, we suggest expanding the samples to other economic markets and industries to enhance the applicability of the model and different financial indicators such as Profit per Share (PB), Return on Equity (ROE), Payback Period, and the Debt-to-Equity Ratio should be involved to assess the effectiveness of green investment projects. By doing that, we can contribute empirical evidence in determining the financial potential of green investment corporate, helping green investment enterprises have more opportunities to access financing sources, as well as investors make safe and correct decisions, thereby creating a “green” economic market.

## **Acknowledgements**

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Stephan Bruns, who always answered my questions, provided precise guidance, and encouraged me to pursue my ideas. At the beginning, I was confused and lacked confidence in my topic. However, with his kindness and passion, Professor Bruns patiently explained things and motivated me to go ahead. I truly appreciate his patience and the time he spent reviewing and correcting my writing, even when I repeated the same mistakes. Without his invaluable guidance and support throughout the course, I could not have completed this thesis so efficiently.

I would also like to express my sincere thanks to the other professors of the Business Economics Faculty at Hasselt University for providing valuable lessons that have contributed meaningfully to both my career path and personal development.

A heartfelt thank you goes to my classmates and friends for the thought-provoking discussions, teamwork, and wonderful memories shared during my Master's course.

Finally, I am especially grateful to my family and my friends in Vietnam for their unwavering support, patience, and belief in me. Without their encouragement and sacrifices, this achievement would not have been possible. Now, I'm confident to say loudly that my family can be proud of me.

Thank you all for being with me on this incredible journey.

Nguyen Thi Chau Loan.

## TABLE OF CONTENTS

1.	Introduction .....	4
2.	Literature review .....	5
2.1.	Main factors driving green investment .....	5
2.2.	The development of digital finance .....	6
2.3.	Using Return on Assets to evaluate corporate operation .....	7
2.4.	Hypotheses analysis .....	7
3.	Data and methodology .....	7
3.1.	Data .....	8
3.2.	Empirical approach .....	11
4.	Result and discussion .....	13
4.1.	Regression analysis .....	13
4.2.	Mediation effect analysis .....	14
4.3.	Robustness test .....	16
4.4.	Other factors impact green investment of heavy industries in China .....	17
5.	Limitations and future research .....	19
6.	Conclusion .....	19
	REFERENCES .....	21
	APPENDIX A: THE UTILIZATION OF AI TOOLS IN RESEARCH .....	24

## TABLE OF FIGURES

Figure 1.	Theoretical framework.....	12
Figure 2.	Specification curve analysis .....	16
Table 1.	Descriptive statistics .....	9
Table 2.	Correlation coefficient .....	10
Table 3.	The benchmark regression result .....	13
Table 4.	The comparison of different dimensions of digital finance.....	13
Table 5.	Mediation model result .....	15
Table 6.	Causal mediation analysis .....	15
Table 7.	Fixed effect specification analysis.....	17



## 1. Introduction

For decades, green investment has been a significant concern for society due to its positive environmental and social impacts (Hill, 2020). In the context of escalating global environmental pollution, green investment has become even more crucial, serving as an effective mechanism to address environmental challenges. However, one of the primary obstacles faced by green investment projects is the difficulty in accessing capital. Given that these projects prioritize environmental benefits over financial returns, they are more likely to attract funding from governments and non-profit organizations rather than private sector investors (Ding et al., 2023). Consequently, the cash flow generated by such projects may be insufficient to secure high-cost capital, thereby limiting their growth and development potential.

In an era where technology is shaping global trends, its positive impact on the financial system is undeniable. The rise of digital finance and FinTech has enabled businesses to enhance the quality of services, increase operating efficiency, and improve customer experience (Gomber et al., 2017). Since digital finance has gained more attention and investment, it has significantly contributed to improving business efficiency by reducing operational costs. Notably, one of the most impactful benefits of digital finance is its ability to provide a professional system that automates processes, eliminates manual operations, and safeguards stakeholders' interests by reducing managerial self-interest in decision-making (Ozili, 2021).

Furthermore, advancements in technologies such as blockchain, machine learning, and big data have enhanced information transparency, enabling banks to access more accurate financial data. This, in turn, reduces the cost of acquiring information about firms and society (Mu et al., 2023). As a result, corporations, particularly those engaged in green investment, have greater opportunities to access capital at lower costs (Wu & Huang, 2022). Moving towards a sustainable economy requires prioritizing both people and the environment as key factors in development. While people remain at the core of economic decision-making, shaping its direction and growth, the environment plays a fundamental role in all aspects of development due to its direct impact on human life, economic stability, and societal progress.

Several previous studies have shown a correlation between the level of adoption of digital finance and green investment. Specifically, in the original study, Ding et al. (2023) showed that digital finance has a positive impact on green investment of heavily polluting firms by increasing cash flow and strengthening these organizations' competitiveness in the market. In another study, Mu et al., (2023) concluded that digital finance enhances corporate ESG performance by mitigating corporate financial constraints and highlighting digital finance as a critical tool for promoting sustainable practices, especially in underdeveloped regions or smaller firms that typically face financing challenges. In addition to the benefits that digital finance brings, such as greater financial inclusion, expansion of formal financial services to non-financial sectors, and providing convenient and secure banking services for other economic institutions, there are still some challenges that the general economy must face, such as an excessive focus on transaction accounts, weak and untimely regulation, and disruption (Ozili, 2021). It is emphasized that the current challenge of green finance is the possibility of attracting private participation investors due to its several risks along with a low return rate. This is also a negative signal to financial

institutions that financially support green investment corporates. Therefore, many policies and strategies that give priority to green investment should be introduced to stimulate innovation, potentially leading to competitive advantages and improved profitability of green investment corporates (Chen & Ma, 2021).

Current studies provide a foundation for the relationship between digital finance and green investment. However, there is no information about evaluating the performance effectiveness that digital finance brings to green investment corporates. This is a significant problem because it not only impacts the business results of green investment companies but also affects financial institutions that financially support these investments. To address the existing concern, we perform a replication study and robust analysis about the research outcome of Ding et al. (2023) to strengthen the proofs about the correlation of digital finance and green investment. Additionally, this paper also contributes analytical results about the ability to use ROA to assess the business performance of green investment projects with digital finance applications. Our research questions below are therefore answered:

- Does digital finance have a positive effect on corporate green investment?
- Can ROA be the mediator between digital finance and green investment?

The structure of this study is as follows: In Section 2, Literature review and theoretical analysis are presented, followed by Data and methodology in Section 3. In Section 4, this study discusses the Results of the regression model. Challenges and limitations of the research are demonstrated in Section 5 and then, Section 6 provides a wrap-up for the full paper.

## **2. Literature review**

### **2.1. Main factors driving green investment**

Aware of the consequences of global warming and pollution, worldwide consumers have tended to express their preferences for environmentally friendly products which lead to the changes in consumption behavior. Consequently, more and more companies catch the trend and focus on investing in green projects (Chitimiea et al., 2021). Eyraud et al. (2013) define green investment as an ecological civilization idea that aims to lower air pollution and greenhouse gas emissions without reducing the production and consumption of non-energy items. Hence, by introducing environmentally friendly practices, green investments can enhance the sustainable development of a country (Kwilinski et al., 2023). However, achieving the "win-win" outcome for both ecological objectives and economic benefits is not easy (Chen & Ma, 2021). The main problems these investments have to face are a low return rate along with higher risk compared to other kinds of investments. It will limit the ability of green investment to access financial resources as well as attract new investors from private sectors (Ozili, 2021).

In their study, Chitimiea et al. (2021) highlighted key factors influencing green investment are political situations and economic development level. By issuing effective environmental policies and plans, governments can boost the applicable level of green investment, aiming at a green and sustainable economy (Kwilinski et al., 2023). It leads to the introducing of many new financial instruments such as green bond and green central banking, targeting to support projects and entities which are considered as

green investments (Ozili, 2021). Economic characteristics also have a direct effect on the financial performance of green investments projects due to emphasizing the consumption habit, productivity of corporates and carbon emission level (Chitimiea et al., 2021). Furthermore, it is illustrated that high economic growth and income level will increase environmental awareness (Eyraud et al., 2013).

Due to the significant growth of China's economy in the past 40 years, manufacturing activities cause serious environmental pollution, which mostly comes from low production cost and lack of environmental protection regulations (Chen & Ma, 2021), makes China becoming a major global polluter (Xin et al., 2022). Thus, the implementation of green investment and sustainable development become an important issue for China's government (Xin et al., 2022).

## **2.2. The development of digital finance**

In their research, Risman et al. (2021) emphasized that the stability of the financial system is dependent on financial infrastructure, leading to a phenomenon of globalization in the financial sector which is called Digital finance. Digital finance formed by the combination of finance and technology (Xin et al., 2022), by improving the delivery of traditional tasks to introducing fundamentally new business opportunities and models (Gomber et al., 2017). It includes all products and services that enable individuals and companies to obtain payments, savings, and credit facilities through the internet without having to visit bank branches or directly deal with financial service providers (Liu et al., 2022). With the breakthrough innovations it brings, digital finance allows firms to broaden their financing sources from traditional to e-finance and enhance the effectiveness of distortion of credit resource allocation in the traditional financial system (Guo et al., 2024). Furthermore, the application of big data also improves corporate information transparency, which helps to mitigate information asymmetries in financial markets (Mu et al., 2023). As a result, digital finance involves compensating for the deficiency in traditional financial services, enhancing credit availability and liquidity for firms (Liu et al., 2022). Thanks to the revolution of information technology, digital finance can take advantage of the scale effect with lower cost, which provides users with high-quality services with affordable prices. The eliminating governance mechanisms of traditional finance, digital finance also gains benefit from the behavioral footprints and accurate judging credit risks. Therefore, the risk control ability has improved significantly (Rao et al., 2022).

In China, the strict financial supervision leads to a high barrier to entry financial industry (Liu et al., 2022). Moreover, for a long time, China's bank dominated indirect financing system, credit resource is priority to large state-owned firms which limits the financing resources and makes it more difficult for small and medium entities to access capital (Guo et al., 2024). But with the development of digital finance, operating costs of financial institutions drop dramatically, opening up opportunities for other businesses to enter the finance industry (Liu et al., 2022). Although Alipay's 2004 launch marked the beginning of digital banking, the industry officially considers 2013 to be the first year of its development. In just over a decade, China has emerged as the global leader in digital financial development with mobile payment, internet banking and other digital financial ecosystems considered globally influential (Rao et al., 2022). According to the Digital Payments Report 2021 released by Statista, China is one of the biggest markets in the world in digital payment with total transactions value exceeds 2,400 billion USD in 2020. More

specifically, another report of Peking University Digital Inclusion Financial Index highlights the improvement of digital finance inclusion in 2020 is 10 times that corresponding figure in 2011. This positions China in the leading top of the world in the digital finance scale and application (Tang et al., 2020a, b).

### **2.3. Using Return on Assets to evaluate corporate operation**

Considering the inclusion of several risks and low return rate of green investment corporates (Ozili, 2021), evaluating business performance of these entities is more essential. In their research, Petersen and Schoeman (2008) demonstrated that ROA provides information about how much profits a company gains on average by each unit of its assets. High profitability (ROA) represents a company with effective business, which gains positive responses from investors (Husna & Satria, 2019).

### **2.4. Hypotheses analysis**

#### **2.4.1. The relationship between digital finance and green investment**

By reducing energy consumption, improving resource utilization efficiency and seeking new renewable energy, green investment brings many benefits to the environment, but it will increase cost burden for corporates (Chen & Ma, 2021). However, the implementation of digital finance supports mitigating financial constraints and enhancing green investment corporates performance (Mu et al., 2023). Digital finance provides effective financing channels and breaking through the geographical barriers between traditional banks and corporates which lead to capital accessibility for these entities, leading to the reduce of transaction cost and an affordable financial leverage (Liu et al., 2022). Based on this foundation, we propose the first hypothesis:

**H1:** Digital finance has a positive impact on green investment

#### **2.4.2. The impact of ROA in the relationship between digital finance and green investment**

ROA is a common metric in evaluating business performance (Petersen and Schoeman, 2008). By that way, we can evaluate the efficiency of a business in using capital. Since green investment projects contain several risks with low return rate for investors (Ozili, 2021), assessing the financial performance of these projects becomes more and more important. As digital finance can boost corporates financial performance (Mu et al., 2023), we propose the following hypothesis:

**H2a:** Digital finance positively influences ROA of green investment corporates

In this study, we choose ROA as a mediator in the relationship between digital finance and green investment. Hence, the financial performance evaluation of digital finance would be applied for green investment. We propose the following hypothesis:

**H2b:** ROA mediates the relationship between digital finance and green investment of an entity

### **3. Data and methodology**

### **3.1. Data**

The most significant source of pollution is pollution-intensive industries, especially in a big economy like China (Xin et al., 2022). Therefore, the implementation of green investment principles and the promotion of sustainable economic growth have emerged as critical concerns for both governments and enterprises. Considering the important impact of pollution-intensive industries in general and heavily polluting industries specifically on the environment, this study uses data from A-share listed companies in the heavily polluting industry in China in the period from 2011 to 2020. The sample does not include financial firms, and all continuous variables are winsorized in the 1st and 99th percentiles. Data on digital finance are sourced from the Center for Internet Finance Research at Peking University, while firm-and-city-level data are collected from the CSMAR database and the China City Statistical Yearbook. Table 1 provides us with the descriptive statistics of all variables considered. Additionally, Table 2 exhibits the correlation coefficients between variables that highlight the strong statistical relationship of some variables such as Ginvest and Size, Ginvest and Stock, Digfin and GDP, Digfin and Upgrade, ROA and Leverage with highly significant levels ( $p < 1\%$ ). The variance inflation factors of all variables are under 10, proving that there are no multicollinearity problems in the model.

**Table 1. Descriptive statistics**

<b>Variables</b>		<b>Name of variable</b>	<b>Definition</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Min</b>	<b>Max</b>
Explained variable	Ginvest	Green investment	Expresses total green investment of an entity	16.08	4.68	0.00	25.40
Explanatory variable	Digfin	Digital finance index	Digital inclusive finance index by Ant Financial	199.05	70.98	46.04	320.79
Mediating variable	Roa	Return on assets	Profitability index, measured by net profit divided by total assets	0.03	0.05	-0.23	0.27
Control variables	Size	Firm size	Measured by total assets of the firm	22.71	1.38	20.06	26.33
	Age	Firm age	Longevity of firm	2.43	0.67	0.69	3.37
	Leverage	Asset-liability ratio	Measured by gross liability divided by total assets	0.48	0.20	0.05	0.97
	Growth	Growth potential	Illustrated by the growth rate of total assets	0.11	0.22	-0.45	2.04
	Stock	Check-and-balance ownership	Expresses the ownership structure of the firm	14.70	34.41	1.00	778.16
	GDP	Economic growth level	Measured by city, GDP per capita	11.17	0.58	9.09	13.06
	Upgrade	Urban industry upgrade	Ratio of tertiary industry added value to secondary industry added value	1.07	0.56	0.14	8.75

**Table 2. Correlation coefficient**

	<b>Ginvest</b>	<b>Digfin</b>	<b>Roa</b>	<b>Size</b>	<b>Age</b>	<b>Leverage</b>	<b>Growth</b>	<b>Stock</b>	<b>GDP</b>	<b>Upgrade</b>
Ginvest	1.000									
Digfin	0.034	1.000								
Roa	-0.007	0.126	1.000							
Size	0.410***	0.139	-0.029	1.000						
Age	0.123	0.126	-0.169	0.366***	1.000					
Leverage	0.196	-0.143	-0.404***	0.471***	0.338	1.000				
Growth	-0.011	-0.026	0.246***	-0.015	-0.213	-0.069	1.000			
Stock	0.462***	-0.135	-0.090	0.078	0.089	0.137	-0.075	1.000		
GDP	0.049	0.563***	0.096	0.146	-0.018	-0.089	0.017	-0.065	1.000	
Upgrade	-0.043	-0.515***	-0.075	-0.177	-0.119	0.017	0.021	0.083	-0.310***	1.000

Notes: \*, \*\*, \*\*\* represent significance at level of 10%, 5% and 1% respectively.

### 3.2. Empirical approach

To confirm the conclusion of the original paper and test Hypothesis 1, we construct the following benchmark models:

$$Ginvest_{it} = a_0 + a_1 Digfin_{it} + v_i + \lambda_t + \gamma_j + \delta_r + \varepsilon_{ijrt}(1)$$

$$Ginvest_{it} = b_0 + b_1 Digfin_{it} + b_n Controls_{it} + v_i + \lambda_t + \gamma_j + \delta_r + \varepsilon_{ijrt}(2)$$

In which Model (1) is used to test the coefficient of Digfin and Ginvest, controlling for the impact of other variables.  $i$ ,  $t$ ,  $j$  and  $r$  denote firm, year, industry, and area, respectively, presented as fixed effect factors for empirical analysis purposes while  $\varepsilon$  illustrates the random disturbance term. Digfin is the main explanatory variable and represents the digital finance index. Ginvest is the dependent variable which shows the level of corporate green investment that we study in this research. This allows us to assess the impact of digital finance on corporate green investment while accounting for potential confounding factors.

In Model (2), Controls is added to represent a set of control variables, by that, this study can access the change of Ginvest with and without impact of other factors. By comparing the results from Models (1) and (2), we can assess the robustness of the relationship between digital finance and corporate green investment, considering the influence of additional control variables.

#### 3.2.1. Dependent variable: Green investment

According to Gu et al. (2021) and Jia et al. (2021), the explained variable is manually compiled the total investment expenditures associated with environmental protection by reviewing the annual reports of heavily polluting firms. The scope of corporate green investment includes various capital expenditures that businesses incur for environmental protection.

#### 3.2.2. Core independent variable: Digital finance

Based on the result of an existing study by Guo et al. (2019), we measure the digital finance index by using Digital Inclusive Finance Index for several reasons. Beyond its accuracy and reliability based on the development across different regions in China, this index is also constructed from Ant Financial, a leading internet finance company in China, ensuring a high level of representativeness and credibility. Additionally, the index captures a broad spectrum of Internet financial products, including third-party payments, online lending, money funds, internet insurance, and online investments, making it extensive and economically meaningful.

#### 3.2.3. Control variables

Referring to the previous model of Ding et al. (2023), this research chose these factors to represent as control variables, including firm size, firm age, financial leverage, growth capacity, equity (Stock), GDP, and upgrade level. These factors play an important role in demonstrating the internal financial potential of the enterprise as well as its ability to develop in the coming years, directly affecting the opportunities



and effectiveness of green investment. Using control variables will highlight the impact of Digfin on Ginvest through the change between the two models with and without the participation of control variables.

### 3.2.4. Mediating variable

In a previous study, Kamruzzaman (2019) mentioned that ROA is the best indicator used to measure the operating results of a company. It directly measures how efficiently a company utilizes its assets to generate profits. If a green investment corporation adopts digital financial tools, an increase in ROA could indicate that these technologies are enhancing operational efficiency and profitability. Furthermore, since green investment firms often manage high-cost, long-term sustainable projects, digital finance can optimize funding allocation and risk management, leading to better financial returns on assets.

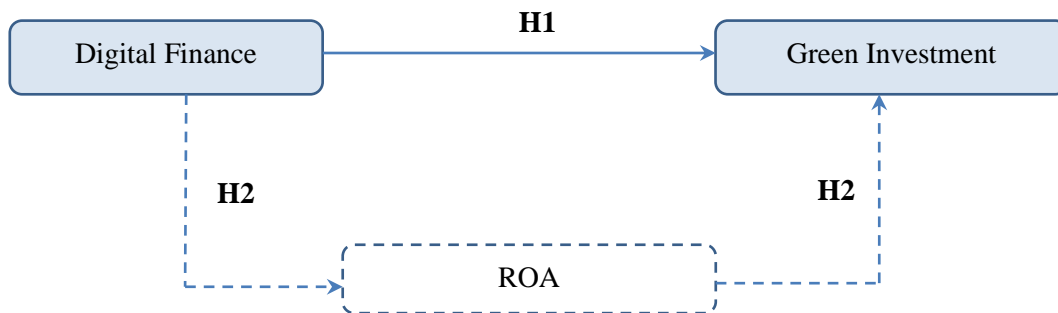
To study influential mechanisms, we use mediation analysis to test Hypothesis 2b by the formular given in Model (4). Furthermore, Model (3) will show how Digfin impacts Ginvest when ROA is defined as the mediating variable, which conducts the test for Hypothesis 2a:

$$ROA_{it} = c_0 + c_1 Digfin_{it} + c_n Controls_{it} + v_i + \lambda_t + \gamma_j + \delta_r + \varepsilon_{ijrt} (3)$$

$$Ginvest_{it} = d_0 + d_1 ROA_{it} + d_2 Digfin_{it} + d_n Controls_{it} + v_i + \lambda_t + \gamma_j + \delta_r + \varepsilon_{ijrt} (4)$$

Model (3) illustrates the correlation between Digfin and ROA, considering the impact of other control variables. According to the study by Huang and Lei (2021), green investment is a high-cost activity of a company, requiring more consideration in building investment strategies and allocating financial resources to ensure project sustainability. Therefore, ROA was taken into account as a financial indicator reflecting the profitability of enterprises to assess whether the level of digital finance application has an impact on business results. In Model (4), we perform the direct effect of Digfin on Ginvest, in which, ROA becomes the mediated variable that can support Digfin expressing its impact (Figure 1).

**Figure 1. Theoretical framework**



#### 4. Result and discussion

##### 4.1. Regression analysis

**Table 3. The benchmark regression result**

Variables	Ginvest (1)	Ginvest (2)
Digfin	0.041*** (3.647)	0.034*** (2.757)
Size		1.172*** (5.472)
Age		-0.485 (-1.087)
Leverage		-0.592 (-0.775)
Growth		0.236 (0.672)
Stock		0.004** (1.808)
GDP		0.261 (0.526)
Upgrade		-0.049 (-0.171)
Firm FE	Yes	Yes
Year FE	Yes	Yes
Industry FE	Yes	Yes
Area FE	Yes	Yes
R-squared	49.72%	50.35%
Observation	3633	3633

Notes: \*, \*\*, and \*\*\* represent significance at the levels of 10%, 5%, and 1%, respectively, and t-statistics are in parentheses.

**Table 4. The comparison of different dimensions of digital finance**

Variables	Ginvest (1a)	Ginvest (1b)	Ginvest (1c)
Digfin_cover	0.031*** (2.845)		
Digfin_usage		0.018** (2.322)	
Digfin_digitalization			0.009** (2.311)
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Area FE	Yes	Yes	Yes
R-squared	49.63%	49.58%	49.58%
Observation	3633	3633	3633

Notes: \*, \*\*, and \*\*\* represent significance at the levels of 10%, 5%, and 1%, respectively, and t-statistics are in parentheses.

According to the benchmark regression result provided in Table 3, it is highlighted that without the inclusion of control variables in Model (1), Digfin has a significant positive coefficient with Ginvest, which initially demonstrates that digital finance can enhance corporate green investment. A similar result is collected in Model (2) when control variables are included. The coefficient shows a little decrease but still illustrates the positive impact of Digfin on Ginvest. We can also see some negative parameters in Age, Leverage, and Upgrade, but with high p-values, these variables are not considered to have a statistically significant relationship with Ginvest. Therefore, these results lead to the conclusion that digital finance promotes green investment in heavily polluting firms in China. Hence, H1 is supported.

To provide more evidence to support the above hypothesis, this study has performed other regression models on other aspects of digital finance. Specifically, the results in Table 4 show the positive effects of factors such as coverage (Digfin\_cover), application level (Digfin\_usage), and level of digitalization (Digfin\_digitalization) on green investment. This once again confirms that the level of development and application of digital finance promotes corporate green investment significantly.

#### **4.2. Mediation effect analysis**

To verify H2b, a mediating model is used in this paper, in which ROA becomes the mediator to evaluate the performance effectiveness that digital finance brings to green investment. According to the regression result shown in Table 5, the coefficient of Digfin in Model (3) is negative but small with a significant level at 10%, meaning the effect of digital finance on ROA is negative but weak. Thus, H2a is not supported. In Model (4), when the effect of both digital finance, ROA, and other control variables is included, the coefficient demonstrates a positive effect but is not significant due to the high p-value (>10%) leading to the conclusion that H2b is not supported.

**Table 5. Mediation model result**

Variables	ROA	Ginvest
	(3)	(4)
Roa		0.322 (0.194)
Digfin	-2.49e-04* (-1.818)	0.034*** (2.761)
Size	5.18e-03** (2.198)	1.171*** (5.459)
Age	7.59e-03 (1.547)	-0.487 (-1.092)
Leverage	-1.85e-01*** (-21.962)	-0.533 (-0.647)
Growth	3.87e-02*** (10.024)	0.223 (0.626)
Stock	1.95e-05 (0.715)	0.004** (1.805)
GDP	5.07e-03 (0.928)	0.259 (0.522)
Upgrade	-8.31e-04 (-0.265)	-0.049 (-0.17)
Firm FE	Yes	Yes
Year FE	Yes	Yes
Industry FE	Yes	Yes
Area FE	Yes	Yes
R-squared	54.27%	50.35%
Observation	3633	3633

Notes: \*, \*\*, and \*\*\* represent significance at the levels of 10%, 5%, and 1%, respectively, and t-statistics are in parentheses.

**Table 6. Causal mediation analysis**

	Digfin	ROA	p-value
Average Causal Mediation Effect (ACME)		-3.77e-06	0.97
Average Direct Effect (ADE)	-6.52e-03		0.26

By using a two-stage regression approach as Model (3) and Model (4), this study gets the result presented in Table 6. Based on the causal mediation framework proposed by Imai et al. (2010), we

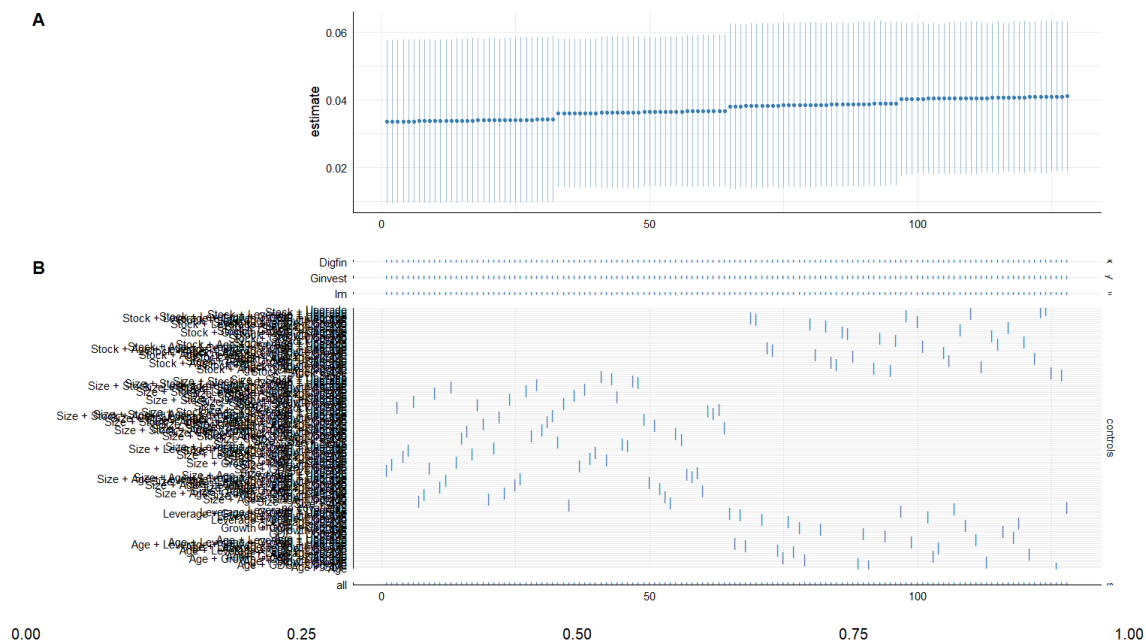
conduct a causal mediation analysis on two variables Digfin and ROA with the support of mediation package in R programming. The result emphasizes that the Average Causal Mediation Effect (ACME) is very small and statistically non-significant ( $p=0.97$ ), indicating that there is no strong evidence supporting the mediation effect of ROA in the relationship between digital finance and green investment. Similarly, the Average Direct Effect (ADE) is also non-significant ( $p=0.26$ ), suggesting that digital finance does not have a substantial direct impact on green investment through this pathway. These findings imply that ROA may not serve as a significant mediator in this context.

### 4.3. Robustness test

#### 4.3.1. Specification curve analysis

In the original study, Ding et al., (2023) have performed the Endogeneity tests for the model to take the impact of internet penetration into account. As a result, the endogeneity problem is solved and the primary finding of the paper is robust. In this paper, we choose to use specification curve analysis as a robustness test method to see if our outcome is affected by other independent variables of the model and evaluate the consistency of the result across a variety of specifications. Specifically, we test whether our results hold when different combinations of independent variables from the original model are included or excluded. To reduce potential bias and ensure comprehensive coverage, we generate and analyze 128 randomly selected model specifications. This approach enables us to observe how robust our key findings are under various model structures and confirms the stability of the estimated relationship.

**Figure 2. Specification curve analysis**



In Figure 2, it is indicated that the estimated effect of Digfin on Ginvest remains positive and relatively stable, clustering within the range of 0.03 to 0.04 across various control specifications. This consistency suggests a robust relationship between Digfin and the dependent variable.

#### 4.3.2. Fixed effects specification analysis

To evaluate the robustness and reliability of our results, we conduct a fixed effects specification analysis. The primary motivation for this test is to assess how the inclusion or exclusion of different fixed effects—such as area, industry, year, and firm—affects the estimated relationship between digital finance and green investment. Fixed effects help control unobserved heterogeneity that could bias our results. By systematically removing each fixed factor and eventually all of them, we are able to determine whether the observed impact of digital finance on green investment remains stable across different model specifications. This approach helps ensure that our findings are not driven by specific fixed effects but reflect a genuine relationship.

**Table 7. Fixed effect specification analysis**

Dropped factors	Coefficient	Std. Error
Area	0.030**	0.012
Industry	0.032***	0.012
Year	-0.003	0.003
Firm	-0.007	0.006
All factors	-0.386e-3	0.001

Notes: \*, \*\*, and \*\*\* represent significance at the levels of 10%, 5%, and 1%, respectively.

The result in Table 7 illustrates that the coefficient remains positive and statistically significant when we eliminate area and industry. But it becomes sensitive when firm and year fixed effects are excluded, which is reasonable for a panel data set. When all fixed factors are removed, the coefficient of digital finance in the model is insignificant with low confident level (high p-value), highlights the importance of incorporating fixed effects into the model. Although the positive effect of digital finance and green investment is just robust in the first two specifications, the remaining results do not invalidate the finding of this study and emphasize the crucial of fixed effects in our model.

#### 4.4. Other factors impact green investment of heavy industries in China

##### 4.4.1. Government policies relevant to heavy industries

Political and government policies always play an important role in the development of many industries. As the world's largest producer and consumer of steel, China faces increasing pressure to address the environmental challenges associated with its heavy industries, which have been linked to a range of serious problems (Zhou & Yang, 2016). In their study, Chen et al. (2018) highlight a key policy introduced in late 2015, known as the Supply-Side Structural Reforms, aimed at reducing excess capacity in the steel industry and improving air quality. This policy had a significant impact on the steel sector, forcing it to reduce production and transition toward cleaner manufacturing processes. While

some major steel companies experienced financial losses, the policy also created opportunities for businesses to invest in new technologies, clean energy, and greater operational efficiency. Such policies have become a critical driver of the shift toward sustainability within heavy industries, encouraging companies to prioritize green investments. Consequently, these transformations during this period may lead to complexities when analyzing the direct relationship between digital finance and green investment.

Additionally, in 2016, China introduced the Guidelines for Establishing the Green Financial System to promote green finance at a national level (The People's Bank of China, 2016). According to a report by the Green Finance & Development Center, prior to 2015, Chinese financial institutions had not issued any green bonds. However, by 2016, China had emerged as the world's largest issuer of green bonds. This guideline has facilitated financial support for businesses, including those in heavy industries, through various instruments such as green bonds, green banks, and carbon market mechanisms. This represents a key factor influencing corporate green investment, which has not yet been incorporated into the models proposed in this study.

#### **4.4.2. Financial constraint variables**

Financial constraints create barriers that hinder firms from accessing external financing, primarily due to high borrowing costs, low return rates, insufficient collateral, and inefficiencies in financial markets (Bond & Meghir, 1994). In their research, Xin et al. (2022) highlight that information asymmetry in inefficient financial markets poses significant challenges for heavily polluting firms seeking external funding. Similarly, Chițimiea et al. (2021) identify two major challenges associated with green investments: a relatively low rate of return and higher risks compared to other types of investments.

During the study period, several major events influenced financial constraints, notably the US-China trade war, which commenced in 2018 and significantly affected the export of large volumes of goods from heavily polluting industries in China, including iron and steel (Hasan et al., 2021). The continued imposition of tariffs by the US government on Chinese imports has further exacerbated these challenges, leading to a substantial loss of market access for many Chinese enterprises. Specifically, the United States imposed additional tariffs exceeding 25% on iron and steel, coal, and petroleum, further constraining these industries (Li et al., 2020). Consequently, the negative impact on heavily polluting industries in China suggests that the findings of this study may reflect influences that cannot be entirely isolated from these external economic shocks.

While financial constraints pose significant challenges for green investment firms in accessing capital, they also create opportunities for digital finance to play a transformative role. By lowering transaction costs, enhancing the efficiency of financing channels, and overcoming geographical barriers, digital finance improves capital accessibility for these firms, enabling greater investment in sustainable initiatives (Liu et al., 2022). By that, these factors have significant impact on the relationship between digital finance and green investment.

## **5. Limitations and future research**

While this study provides robust evidence of the correlation between digital finance and green investment, there are still certain limitations. Specifically, the analysis is based on existing data and focuses solely on heavily polluting enterprises in China, without considering other industries. This narrow scope may restrict the generalizability of the findings to different sectors or regions with distinct market characteristics. Additionally, as the study employs econometric models, concerns regarding potential endogeneity and causality remain only partially addressed. Thus, to enhance the applicability and reliability of this research, future studies may broaden the samples to include additional industries, such as construction and fast-moving consumer goods. Furthermore, data collection from firms operating in diverse economic markets, such as the European and U.S. markets, could help mitigate the impact of cluster effects.

Furthermore, the causal mediation test yielded a negative result for the Average Causal Mediation Effect of ROA in the relationship between digital finance and green investment. To gain a more comprehensive understanding of financial performance and the payback potential of these entities, we suggest future research could explore alternative indicators such as Profit per Share (PB), Return on Equity (ROE), Payback Period, and the Debt-to-Equity Ratio.

## **6. Conclusion**

In the era of technology explosion, many applications make great contributions to society. Through data collected from heavily polluting industries in China, the results of our study have emphasized the positive effect that digital finance brings to green investment. Thereby, the innovation of technology in the financial sector has created new products such as online lending, money funds, internet insurance and online investments (Ding et al., 2023), thereby improving operational efficiency and providing green investment with great opportunities to access cheap capital sources. This result is proven once again through our robust tests.

In this study, the mediation effect of ROA in the relationship between digital finance and green investment was rejected, however, this does not mean that digital finance does not bring economic efficiency to green investment corporations. We suggest that future studies can further expand this aspect by using other economic indicators.

Other factors that occurred during the period 2011 - 2020 that have a great impact on heavily industries in China are also considered in our study. In particular, the Supply-Side Structural Reforms are considered as a factor that plays a role in guiding the development of the industry, contributing to a sustainable environment by prioritizing green investment activities. In addition, the Guidelines for Establishing the Green Financial System are also a factor that creates new opportunities for digital finance to be applied more widely in the economy. From there, the relationship between digital finance



and green investment can be amplified when analyzing. Financial constraint variables are considered as internal factors coming from the characteristics of green investment companies or from the general characteristics of the Chinese economy during this period, creating certain difficulties for these companies in accessing digital finance. We also consider the impact of the US-China trade war on heavily polluting industries in China. Due to this event, many tariff policies imposed on export goods have significantly undermined the business performance of companies within the research scope, potentially influencing the analysis results presented in this study.

However, our initial objective of identifying financial factors to specifically measure the effectiveness of digital finance applications for green investment was not achieved. Therefore, we believe that future studies can continue this direction by using other variables reflecting financial results or expanding the sample to other regions and industries. Through a rigorous assessment of financial performance, future studies can contribute valuable empirical evidence in determining the financial potential of green investment corporate, helping green investment enterprises have more opportunities to access financing sources, as well as investors make safe and correct decisions, thereby creating a “green” economic market.

## REFERENCES

- Bond, S., & Meghir, C. (1994). Financial constraints and company investment. *Fiscal Studies*, 15(2), 1-18. <https://doi.org/10.1111/j.1475-5890.1994.tb00106.x>
- Chen, L., Ding, D., & Mano, R. (2018). China's capacity reduction reform and its impact on producer prices. International Monetary Fund.
- <https://www.imf.org/en/Publications/WP/Issues/2018/06/15/Chinas-Capacity-Reduction-Reform-and-Its-Impact-on-Producer-Prices-45970>.
- Chen, Y., & Ma, Y. (2021). Does green investment improve energy firm performance?. *Energy Policy*, 153, 112252. <https://doi.org/10.1016/j.enpol.2021.112252>
- Chițimiea, A., Minciu, M., Manta, A. M., Ciocoiu, C. N., & Veith, C. (2021). The drivers of green investment: A bibliometric and systematic review. *Sustainability*, 13(6), 3507. <https://doi.org/10.3390/su13063507>
- Ding, Q., Huang, J., & Chen, J. (2023). Does digital finance matter for corporate green investment? Evidence from heavily polluting industries in China. *Energy Economics*, 117, 106476. <https://doi.org/10.1016/j.eneco.2023.106476>
- Eyraud, L., Clements, B., & Wane, A. (2013). Green investment: Trends and determinants. *Energy Policy*, 60, 852–865. <https://doi.org/10.1016/j.enpol.2013.04.033>
- Gomber, P., Koch, J. A., & Siering, M. (2017). Digital finance and FinTech: Current research and future research directions. *Journal of Business Economics*, 87, 537–580. <https://doi.org/10.1007/s11573-017-0852-x>
- Green Finance & Development Center. (n.d.). Green finance trends in China (1): China's green finance policy landscape. Retrieved February 24, 2025, from <https://greenfdc.org/green-finance-trends-in-china-1-chinas-green-finance-policy-landscape/?cookie-state-change=1740426129764>
- Gu, Y., Ho, K.-C., Yan, C., & Gozgor, G. (2021). Public environmental concern, CEO turnover, and green investment: Evidence from a quasi-natural experiment in China. *Energy Economics*, 100, 105379. <https://doi.org/10.1016/j.eneco.2021.105379>
- Guo, F., Wang, J., Wang, F., Kong, T., Zhang, X., Cheng, Z. (2019). Measuring China's internet financial inclusion: index compilation and spatial characteristics. *China Econ. Quart.* 19, 1401–1418.
- Guo, F., Wang, J. Y., Wang, F., Kong, T., Zhang, X., & Cheng, Z. Y. (2021). The Peking University Digital Financial Inclusion Index of China (2011–2020). Institute of Digital Finance, Peking University. <https://idf.pku.edu.cn/docs/20210421101507614920.pdf>

- Guo, X., Wei, T., Wang, A., & Hu, H. (2024). Corporate governance effects of digital finance: Evidence from corporate tax avoidance in China. *Research in International Business and Finance*, 72, 102526. <https://doi.org/10.1016/j.ribaf.2024.102526>
- Hasan, F., Hossen, M. R., & Shahbaz, A. A. (2021). How the steel industry has been affected by the trade war between the USA and China. *International Journal of Critical Accounting*, 12(4), 348-368.
- Hill, J. (2020). Environmental, social, and governance (ESG) investing: A balanced analysis of the theory and practice of a sustainable portfolio. Academic Press. <https://doi.org/10.1016/C2018-0-02579-6>
- Huang, L., & Lei, Z. (2021). How environmental regulation affects corporate green investment: Evidence from China. *Journal of Cleaner Production*, 279, 123560. <https://doi.org/10.1016/j.jclepro.2020.123560>
- Husna, A., & Satria, I. (2019). Effects of return on asset, debt to asset ratio, current ratio, firm size, and dividend payout ratio on firm value. *International Journal of Economics and Financial Issues*, 9(5), 50–54. <https://www.econjournals.com/index.php/ijefi/article/view/8345>
- Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*, 25(1), 51–71. <https://doi.org/10.1214/10-STS321>
- Jia, F., Li, G., Lu, X., & Xie, S. (2021). CEO given names and corporate green investment. *Emerging Markets Review*, 48, 100808. <https://doi.org/10.1016/j.ememar.2020.100808>
- Kamruzzaman, M. (2019). Impact of the financial factors on return on assets (ROA). *DIU Journal of Business and Entrepreneurship*. <https://doi.org/10.36481/diujbe.v012i1.54kdown08>
- Kwilinski, A., Lyulyov, O., & Pimonenko, T. (2023). Greenfield investment as a catalyst of green economic growth. *Energies*, 16(5), 2372. <https://doi.org/10.3390/en16052372>
- Li, M., Balistreri, E. J., & Zhang, W. (2020). The US–China trade war: Tariff data and general equilibrium analysis. *Journal of Asian Economics*, 69, 101216. <https://doi.org/10.1016/j.asieco.2020.101216>
- Liu, J., Jiang, Y., Gan, S., He, L., & Zhang, Q. (2022). Can digital finance promote corporate green innovation? *Environmental Science and Pollution Research*, 29(24), 35828–35840. <https://doi.org/10.1007/s11356-022-18667-4>
- Mu, W., Liu, K., Tao, Y., & Ye, Y. (2023). Digital finance and corporate ESG. *Finance Research Letters*, 51, 103426. <https://doi.org/10.1016/j.frl.2022.103426>
- Ozili, P. K. (2018). Impact of digital finance on financial inclusion and stability. *Borsa Istanbul Review*, 18(4):329–340. <https://doi.org/10.1016/j.bir.2017.12.003>
- Ozili, P. K. (2021). Digital finance, green finance and social finance: Is there a link? *Financial Internet Quarterly*, 17(1), 1–7. <https://doi.org/10.2478/fiqf-2021-0001>

- Petersen, M. A., & Schoeman, I. (2008, July). Modeling of banking profit via return-on-assets and return-on-equity. In *Proceedings of the World Congress on Engineering* (Vol. 2, No. 1, pp. 12–37). [http://www.iaeng.org/publication/WCE2008/WCE2008\\_pp1262-1267.pdf](http://www.iaeng.org/publication/WCE2008/WCE2008_pp1262-1267.pdf)
- PsBo, C. (2016). Guidelines for establishing the green financial system. The People's Bank of China
- Rao, S., Pan, Y., He, J., & Shangguan, X. (2022). Digital finance and corporate green innovation: Quantity or quality? *Environmental Science and Pollution Research*, 29(37), 56772–56791. <https://doi.org/10.1007/s11356-022-19931-6>
- Risman, A., Mulyana, B., Silvatika, B., & Sulaeman, A. (2021). The effect of digital finance on financial stability. *Management Science Letters*, 11(7), 1979–1984. <https://doi.org/10.5267/j.msl.2021.2.007>
- Tang, K., Qiu, Y., & Zhou, D. (2020a). Does command-and-control regulation promote green innovation performance? Evidence from China's industrial enterprises. *Science of The Total Environment*, 712, 136362. <https://doi.org/10.1016/j.scitotenv.2019.136362>
- Tang, S., Wu, X. C., & Zhu, J. (2020b). Digital finance and enterprise technology innovation: Structural feature, mechanism identification and effect difference under financial. *Journal of Management World*, 36(5), 52–66+9. (In Chinese; translate the title if needed for non-Chinese audiences)
- Wu, Y., & Huang, S. (2022). The effects of digital finance and financial constraint on financial performance: Firm-level evidence from China's new energy enterprises. *Energy Economics*, 106158. <https://doi.org/10.1016/j.eneco.2022.106158>
- Xin, D., Yi, Y., & Du, J. (2022). Does digital finance promote corporate social responsibility of pollution-intensive industry? Evidence from Chinese listed companies. *Environmental Science and Pollution Research*, 29(56), 85143–85159. <https://doi.org/10.1007/s11356-022-21659-x>
- Yuan, L., Shin, K., & Managi, S. (2018). Subjective Well-being and Environmental Quality: The Impact of Air Pollution and Green Coverage in China. *Ecological Economics*. <https://doi.org/10.1016/J.ECOLECON.2018.04.033>.
- Zhou, K., & Yang, S. (2016). Emission reduction of China's steel industry: Progress and challenges. *Renewable and Sustainable Energy Reviews*, 61, 319–327. <https://doi.org/10.1016/j.rser.2016.04.003>.

## APPENDIX A: THE UTILIZATION OF AI TOOLS IN RESEARCH

### A1. List of AI Tools Used

Tool name	Purpose	Reference
ChatGPT (OpenAI)	Assisted in refining the framework code for the model and debugging errors to improve functionality and efficiency	OpenAI. (2023). ChatGPT. Retrieved from <a href="https://chat.openai.com/">https://chat.openai.com/</a>

### A2. Prompts Used with AI Tools

Below are the specific prompts entered into ChatGPT during the research process:

#### 1. Suggest me a code to run a statistical model, using plm package with fixed effect in R

##### Response:

"Here is an R script to run a fixed-effects model using the plm package. The code assumes you have panel data and want to estimate a fixed-effects model using the plm function.

```
# Load necessary libraries
```

```
library(plm)
```

```
# Load or create panel data
```

```
# Example: Simulated panel dataset
```

```
set.seed(123)
```

```
data <- data.frame(
```

```
  id = rep(1:10, each = 5), # Panel ID
```

```
  time = rep(1:5, times = 10), # Time periods
```

```
  y = rnorm(50, mean = 10, sd = 2), # Dependent variable
```

```
  x1 = rnorm(50, mean = 5, sd = 1), # Independent variable
```

```
  x2 = rnorm(50, mean = 3, sd = 1) # Another independent variable
```

```
)
```

```
# Convert to panel data format
```

```
pdata <- pdata.frame(data, index = c("id", "time"))

# Estimate fixed-effects model

model <- plm(y ~ x1 + x2, data = pdata, model = "within")

# Summarize results

summary(model)
```

Explanation:

- The dataset includes id (panel identifier), time (time period), y (dependent variable), and x1, x2 (independent variables).
- `pdata.frame()` converts the dataset into a panel format.
- `plm(y ~ x1 + x2, data = pdata, model = "within")` estimates a fixed-effects model.
- `summary(model)` displays the model results."

## **2. If my model includes 4 fixed factors, can I use `felm` package?**

### **Response:**

"Yes, you can use the `lfe` package (which provides the `felm` function) if your model includes multiple fixed effects. The `felm` function is more efficient than `plm` when handling multiple fixed effects.

Example Code:

```
# Load necessary libraries

library(lfe)

# Simulated panel dataset

set.seed(123)

data <- data.frame(

  id = rep(1:10, each = 5), # Panel ID

  time = rep(1:5, times = 10), # Time periods

  region = rep(1:2, each = 25), # Region fixed effect

  industry = sample(1:3, 50, replace = TRUE), # Industry fixed effect
```

```

y = rnorm(50, mean = 10, sd = 2), # Dependent variable

x1 = rnorm(50, mean = 5, sd = 1), # Independent variable

x2 = rnorm(50, mean = 3, sd = 1) # Another independent variable

)

# Estimate model with 4 fixed effects (id, time, region, industry)

model <- felm(y ~ x1 + x2 | id + time + region + industry, data = data)

# Summarize results

summary(model)

```

Explanation:

- `felm(y ~ x1 + x2 | id + time + region + industry, data = data):`
- `y ~ x1 + x2:` Dependent variable regressed on independent variables.
- `| id + time + region + industry:` Specifies four fixed effects.
- `summary(model):` Displays regression results.”

### **3. Generate a code in R for mediating analysis, in which ROA is the mediator.**

#### **Response:**

“Here’s an R script for conducting a mediation analysis where ROA acts as a mediator. The script uses the mediation package, which is commonly used for such analyses.

Assumptions:

- You have three different models where ROA is the mediator.
- The dataset is named data.
- The independent variable is IV, the mediator is ROA, and the dependent variable is DV.
- You are using linear regression models.

Install and Load Required Packages:

```

install.packages("mediation") # Run this if you haven't installed it yet

library(mediation)

```

Mediation Analysis Code:

```
# Load necessary packages
```

```
library(mediation)
```

```
# Model 1: IV → ROA (Mediator)
```

```
model_mediator <- lm(ROA ~ IV, data = data)
```

```
# Model 2: IV + ROA → DV
```

```
model_outcome <- lm(DV ~ IV + ROA, data = data)
```

```
# Mediation analysis
```

```
mediation_result <- mediate(model_mediator, model_outcome, treat = "IV", mediator = "ROA", boot = TRUE, sims = 500)
```

```
# Summary of the mediation results
```

```
summary(mediation_result)
```

```
# Plot the mediation effect
```

```
plot(mediation_result)
```

Explanation:

- Step 1: Fit a regression model to estimate the effect of IV on ROA.
- Step 2: Fit another regression model to estimate the effect of IV and ROA on DV.
- Step 3: Use the mediate() function to estimate the direct, indirect, and total effects.
- Step 4: Summarize and visualize the mediation results.

Interpretation:

- ACME (Average Causal Mediation Effect): The indirect effect of IV on DV through ROA.
- ADE (Average Direct Effect): The direct effect of IV on DV when controlling for ROA.
- Total Effect: Sum of ACME and ADE."

**4. Help me fix this bug "statistic(data, original, ...): mediator model is not yet implemented".**

**Response:**



"The error message "mediator model is not yet implemented" suggests that the mediate() function in the mediation package does not support certain model structures directly. Here are a few potential reasons and solutions:

Possible Causes & Solutions:

1. Mixed-effects or Panel Data Models:

- If your data includes fixed effects, you might be using lm() incorrectly.
- Instead, use linear mixed-effects models (lmer()) from the lme4 package.

2. Multilevel/Panel Data Handling:

- The mediate() function does not support plm() (used for panel data in R).
- Solution: Use lmer() or bootstrap mediation analysis instead.

**5. Help me fix this bug "Error in mediate(model\_3, model\_4, treat = "Digfin", mediator = "Roa", : NA in model coefficients; rerun models with nonsingular design matrix".**

**Response:**

You're still getting the "NA in model coefficients" error, which suggests there's a problem with collinearity, perfect separation, or missing data. Let's systematically fix the issue:

Step-by-Step Debugging

1. Check for Missing Values (NAs)

Even a single missing value can cause this issue. Run:

```
colSums(is.na(data))
```

If any variable has NAs, remove them:

```
data <- na.omit(data)
```

Now, re-run the regression models.

2. Check for Multicollinearity (VIF Test)

```
install.packages("car")
```

```
library(car)
```

```
# Check VIF for Model 3 (Roa as the dependent variable)
```

```
vif(lm(y ~ x1 + x2, data = data))
```

Fix: If VIF > 10 for any variable, remove it and re-run."

**6. I will provide my code and the errors that R returns, help me to fix the bug. Here is my code (fixed factors are year, industry and area):**

```
" > model_5 <- lm(Roa ~ Digfin + Size + Age + Leverage + Growth + Stock + GDP + Upgrade + year  
+ industry + area, data = data_selected)
```

```
> model_6 <- lm(Ginvest ~ Roa + Digfin + Size + Age + Leverage + Growth + Stock + GDP + Upgrade  
+ year + industry + area, data = data_selected)
```

```
> med_model <- mediate(model_5, model_6, treat = "Digfin", mediator = "Roa", boot = TRUE, sims =  
1000)
```

Running nonparametric bootstrap

Error in model.frame.default(Terms, newdata, na.action = na.action, xlev = object\$xlevels) : factor industry has new levels C27"

**Response:**

"The error "factor industry has new levels C27" occurs because when mediate() tries to simulate bootstrap samples, it encounters a level in industry (C27) that was not present when fitting model\_5 and model\_6. This typically happens when:

- The factor variable industry had unused levels when fitting the model.
- mediate() attempts to generate new data, but those levels were not in the original model.

Solution:

We'll convert the factors (industry, year, and area) into dummy variables using model.matrix() and include them in the model manually.

Step 1: Create Dummy Variables for industry, year, area

```
industry_dummies <- model.matrix(~ industry - 1, data = data_selected) # Remove intercept
```

```
year_dummies <- model.matrix(~ year - 1, data = data_selected)
```

```
area_dummies <- model.matrix(~ area - 1, data = data_selected)
```

These will generate matrices where each column corresponds to a dummy variable for each level of the factor.

## Step 2: Refit Models Using Dummies

Now, instead of using `factor()` directly in the model, we'll include the dummy variables we created.

```
model_5 <- lm(Roa ~ Digfin + Size + Age + Leverage + Growth + Stock + GDP + Upgrade +  
industry_dummies + year_dummies + area_dummies, data = data_selected)
```

```
model_6 <- lm(Ginvest ~ Roa + Digfin + Size + Age + Leverage + Growth + Stock + GDP + Upgrade  
+ industry_dummies + year_dummies + area_dummies, data = data_selected)
```

## Step 3: Run `mediate()` Again"

### 7. To perform a specification curve analysis, which packages need to be installed?

#### **Response:**

"To perform a Specification Curve Analysis (SCA) in R, you typically need the following packages:

- `specification_curve` – The main package for conducting specification curve analysis.
- `tidyverse` – For data manipulation (`dplyr`), visualization (`ggplot2`), and handling data frames (`tibble`).
- `broom` – To tidy model outputs.
- `purrr` – To help with mapping functions over multiple model specifications.
- `ggpubr` – For additional visualization capabilities.
- `lme4` (if using mixed-effects models) – To fit models with random effects.
- `car` (optional) – For additional statistical tools like variance inflation factor (VIF).

You can install these packages using:

```
install.packages(c("specification_curve", "tidyverse", "broom", "purrr", "ggpubr", "lme4", "car"))"
```

### 8. Please help me to check if my code is right, I want to perform a specification curve analysis for my model, my code as below:

```
"data_sorted <- data_sorted %>% arrange(stkcd, year)
```

```
specs <- setup(data = data_sorted,
```

```
  y = "Ginvest",
```

```
  x = "Digfin",
```

```

model = c("lm"),

controls = c("Size", "Stock", "Age", "Leverage"),

add_to_formula = "stkcd + year + industry + area")

summary(specs)

results <- specr(specs)

summary(results)

plot(results)"

```

### **Response:**

"Your code for performing a specification curve analysis looks mostly correct, but there are a few things to review and refine for clarity and proper execution.

Key Points to Check:

- Data Sorting:
  - You correctly sorted the data by stkcd (stock code) and year using arrange(). This is important for time-series or panel data analysis.
- Model Setup:
  - The setup() function looks good with the correct variables and model specification.
  - Ensure the stkcd, year, industry, and area variables are correctly formatted in your dataset (e.g., as factors or factors with appropriate levels).
  - If you have large numbers of unique identifiers (like stkcd, industry, area), the inclusion of these as fixed effects may be computationally expensive. Just ensure this doesn't slow down your computation significantly.
- Model Type:
  - model = c("lm") is correct for a linear model. If you want to test other types (e.g., robust standard errors), you can extend this argument, but "lm" is fine for now.
- Control Variables:
  - You've included controls like Size, Stock, Age, and Leverage, which seems appropriate. Ensure that these columns are in the dataset and are appropriately scaled.

- Specification Curve:
  - `specr(specs)` generates the specification curve analysis, which should display different model specifications based on varying combinations of the control variables or other factors.

Suggestions:

- If you find your dataset includes missing values, consider handling them (e.g., via imputation or removing incomplete cases) before running the analysis.
- Double-check if the `specificationCurve` package supports all the variables and formula syntax as you intend to use them, especially with the `add_to_formula` argument.”

### **A3. Ethical Considerations**

The author acknowledges assistance in developing the initial framework and structure of the code used in this research model of AI tool (ChatGPT). Its support in generating and debugging the initial code structure contributed to the efficiency of the model development process. However, full responsibility for modifying and adapting the code to align with the specific requirements was taken by the author.