Analytical Methodology and a Simulator for ESG-Financial Indicators Based on Causal Hypothesis Graphs

Hiroaki Ozaki[1,*], Naoya Tanahashi[1], Norio Masuda[1], Kazuo Yamada[2], Masahito Kato[2,3] and Nobuyuki Isagawa[2]

[1]Hitachi Ltd.
[2]Graduate School of Management, Kyoto University
[3]School of Management, Chukyo University

Abstract

In recent years, corporate management has shifted its focus from solely financial indicators to also include ESG indicators, which assess the environmental (E), social (S), and governance (G) aspects. However, the impact of ESG indicators on company management varies depending on several factors and is closely tied to the specific social issues each company prioritizes. Consequently, to develop a comprehensive business analysis model that incorporates both ESG and financial indicators, we developed a method for analyzing ESG-financial indicators using a causality hypothesis graph with a structural equation modeling. This method enables us to examine (1) the interrelationships between different indicators and (2) the validation of hypotheses regarding the issues that a company should prioritize. We also developed a simulator predicts future financial indicator values by comprehensively combining multiple measures. We evaluated this technology by applying it to our corporate data and present the comparative results of predicted financial indicator values with and without the implementation of a measure.

Keywords

causal hypothesis, structural equation modeling, simulation

1. Introduction

In recent years, there has been a growing emphasis on corporate social responsibility in business activities. This includes considerations such as the environmental impact of energy and resource consumption, as well as the appropriateness of employment practices. In corporate management, not only financial indicators but also ESG (Environmental, Social, and Governance) indicators, which reflect the status of the environment, society, and governance, are being given increased importance [1].

There have been numerous studies examining the relationship between ESG indicators and corporate performance [2, 3], we are working on the development of a support system to examine how companies can incorporate ESG activities into their management. When applying ESG support systems to actual management, the following challenges have been identified:

1. In cases where the interpretation of statistical analysis results does not align with the actual business situation, managers are unable to make appropriate judgments.
2. Without providing a rationale for the significance of engaging in ESG initiatives, the priority given to ESG efforts in the field may decline.
3. Merely focusing on the relevance of indicators without offering specific implementation measures does not lead to improvements in management.

To address the specific challenges mentioned above (1)(2), it is necessary to provide an analytical method that takes into account the individual characteristics of each company and demonstrates the

STAI'24: International Workshop on Sustainable Transition with AI (Collocated with the 33rd International Joint Conference on Artificial Intelligence 2024), August 05, 2024, Jeju, Republic of Korea.
*Corresponding author.
hiroaki.ozaki.yu@hitachi.com (H. Ozaki); naoya.tanahashi.oo@hitachi.com (N. Tanahashi); norio.masuda.wv@hitachi.com (N. Masuda); yamada.kazuo.7z@kyoto-u.ac.jp (K. Yamada); kato@mecl.chukyo-u.ac.jp (M. Kato); isagawa.nobuyuki.3w@kyoto-u.ac.jp (N. Isagawa)

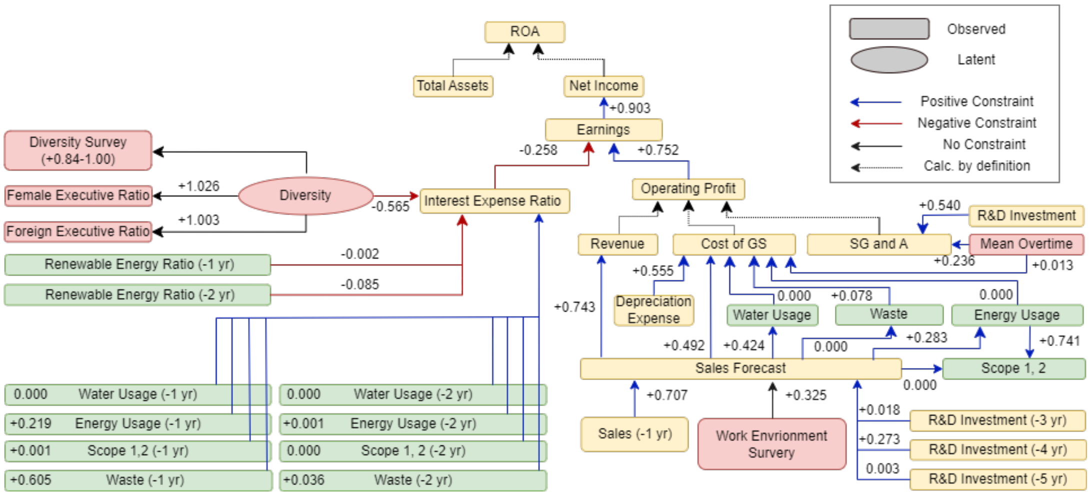


Figure 1: Causal relationship graph of ESG-financial indicator analysis results applied to actual companies. Green boxes represent E indicators, red boxes represent S or G indicators, and yellow boxes represent financial indicators. Red lines indicate positive coefficient constraints, blue lines indicate negative coefficient constraints, black lines indicate linear regression without constraints, and dashed black lines indicate values calculated by numerical computation. Diversity and labor environment are factor analyzed. Many of the relationships between financial indicators are calculated statistically, but for example, ROA (Return on Assets) is calculated from these indicators.

relationship between meaningful ESG indicators and financial indicators. Therefore, we propose a method based on a causal hypothesis constructed by experts to clarify the relationship between ESG indicators and financial indicators using structural equation modeling [4].

Furthermore, in order to address the challenges of (3), it is necessary to provide effective measures while considering the individual characteristics of each company. Therefore, in this study, based on the analysis of the relationship between ESG indicators and financial indicators using company data, we propose a simulator that predicts the future changes in financial indicators when implementing measures related to ESG.

In this paper, we discuss the effectiveness of this method by providing examples of analysis results from actual companies.

2. Causal Relation Modeling of ESG-F Indicators

As mentioned above, in order to give interpretability to the analysis of the relationship between ESG indicators and financial indicators, it is necessary to analyze the causal relationship between the indicators. Although it is possible to use statistical methods to estimate causal relationships, we considered the following points and conducted an analysis method in which experts created causal relationship hypotheses and verified them with data:

- ESG data is often available for a period of about 10 to 20 years per fiscal year, and it is difficult to obtain statistically sufficient data.
- When formulating ESG-related policies, there is a demand from a management side to verify the intended causal relationships.
- There are constraints on the causal relationships between ESG indicators and financial indicators, and it is necessary to incorporate these constraints into the analysis model.

In particular, the constraints on causal relationships mentioned above often have cases where the positive or negative coefficients are assumed in advance, even if the size of the impact relationship is unknown, such as "if the water usage increases in a company, the cost will increase accordingly". Therefore, we conducted an analysis using structural equation modeling, which incorporates manually created causal relationship hypotheses within the constraints.

2.1. Model Construction

To construct causal relationship hypotheses, a workshop was conducted with approximately 20 stakeholders, including the authors, as well as departments within the company such as finance, human resources, and sustainability management. During the workshop, over 130 ESG and financial-related indicators and more than 300 hypotheses regarding their relationships were extracted. The hypotheses of causal relationships that were created as a result were used within the scope of available empirical data. In cases where alternative data for a particular indicator existed, they were also included in the causal relationship hypothesis model. Additionally, multiple structural equation modeling iterations were performed, making adjustments to the model, such as incorporating factor analysis, to address statistical issues like multicollinearity.

2.2. ESG-Financial Indicator Model using Structural Equation Modeling

Structural Equation Modeling (SEM) can be regarded as a generalization and systematization of linear regression and factor analysis, and is useful for analyzing causal relationships. In this study, we particularly used factor analysis as a partial solution to the problem of multicollinearity in datasets. We also applied linear regression to incorporate constraints between factors. Additionally, to model the cyclic relationships between indicators, we introduced time delays for some indicators. For example, although research and development investment is an expenditure in a single fiscal year, if it contributes to sales and generates profits through business creation, it is assumed that the contribution to sales was made by past research and development investment, and this was represented by a time delay.

2.2.1. Factor Analysis

Factor analysis, which is included in structural equation modeling is a statistical analysis method for estimating the factors behind observed variables obtained through experiments or observations. For example, for multiple observed variables, if there are common factors that are common factors, an observed variable y_i can be expressed by the factor loadings a and the common factor f and the unique factor e as follows:

$$y_i = \sum_j a_j f_j + e_i \tag{1}$$

Here, the unique factor is a component specific to each item, and it only affects one observed variable, and the factor loading represents the strength of the relationship between the factor and the observed variable. In this analysis, factors were specifically set for datasets that tend to have problems with multicollinearity. For example, the results of survey questionnaires in companies have high correlations between each question, and if each question in the questionnaire is individually added to the causal relationship hypothesis model, multicollinearity problems occur in the regression analysis of structural equation modeling. In such cases, we considered the intention of the questionnaire and added factors.

2.3. Constraints

For example, in cases such as "an increase in water usage in a company leads to a corresponding increase in costs," it is believed that there is an environmental indicator of water usage and a usage fee (cost) that is expected to increase proportionally. A positive proportionality constant is assumed to exist between them. However, since accounting items that only disclose water usage fees are rare, it becomes necessary to conduct linear regression with a positive coefficient for accounting items that include water usage fees.

In this case, the analysis is performed by setting the range of possible coefficient values to be non-negative. Similarly, constraints were applied between indicators where a positive or negative relationship was assumed in advance for the purpose of conducting the analysis.

2.3.1. Calculation of the Impact of Each Indicator on Financial Indicators

The results of the analysis using structural equation modeling can be used to read the direct impact of indicators, but the impact of indicators that have causal relationships through other indicators cannot be interpreted. In this study, in particular, it is important to evaluate the impact of non-financial indicators on management indicators, but non-financial indicators rarely have direct causal relationships with management indicators. Therefore, in this technique, the impact of each indicator on financial indicators is calculated based on the causal relationship paths.

Assuming that the impact of indicator j on indicator i is to be calculated, the calculation formula is explained. In this case, there may be a direct causal relationship between indicator j and indicator i, or they may be related through other indicators. Moreover, there may be multiple causal paths between indicator j and indicator i.

In this case, the impact of indicator j on indicator i is defined as the sum of the impact through all causal paths from j to i. If there are k causal paths from indicator j to indicator i, and the direct impact of a certain indicator on another indicator in a certain path is represented as w, the impact e_k from indicator j to indicator i in path k is calculated as the product of the impact of indicators on that path, based on the characteristics of linear regression models. Therefore, if path k consists of n_k indicators, it can be expressed as follows:

$$e_k = \prod_{l=1}^{n_k} w_{k,l} \tag{2}$$

Here, $w_{k,l}$ represents the impact of the lth indicator on path k. Then, the total impact E from indicator j to indicator i is defined as the sum of all impact through the paths, so if there are K paths, it can be expressed as follows:

$$E = \sum_{k=1}^{K} e_k \tag{3}$$

Additionally, financial indicators have defined calculation formulas based on indicators. For example, operating profit is calculated by subtracting selling, general, and administrative expenses and cost of goods sold from sales. The impact on such indicators is calculated by performing calculations based on the predefined formulas for each indicator and then calculating the impact after removing the standardization of the data, such as averaging 0 and standard deviation 1.

2.4. Simulation

Future predictions in the simulator are based on the analysis results of causal hypotheses and structural equation modeling. The prediction of each indicator is executed by propagating the predicted values along the causal relationships, starting from the predicted value of the indicator that serves as the cause, to predict the value of the indicator that is influenced by the cause. This allows the non-financial indicators that change due to the applied measures to affect the financial indicators according to the intended causal relationships. The following formulas formalize the prediction methods for the five types of indicators that appear in the causal relationships.

(1) Observational variables that are not influenced by other indicators, (2) Observational variables that are influenced by other indicators, (3) Observational variables that are not influenced by other indicators but are used to calculate latent variables, (4) Latent variables, (5) Indicators calculated based on formulas from indicators within the causal relationships.

For indicators (1) and (3), which are not influenced by other indicators, the predicted value is generated by sampling from a normal distribution with the statistical properties of the data of the target indicator as follows:

$$X_{i,t} \sim \mathcal{N}(X_{i,t-1}, \sigma_{i,d}) \tag{4}$$

Here, $X_{i,t}$ represents the value of indicator i in year t, and $\sigma_{i,d}$ is the standard deviation of the data for indicator i over d years from the latest year.

For indicator (2), which is influenced by other indicators, the prediction is made using a regression equation with the other indicators as explanatory variables:

$$X_{i,t} = \sum_j a_{j,i} \cdot X_{j,t} + b_i + \epsilon_{i,d} \tag{5}$$

Here, the regression coefficients $a_{j,i}$ and the intercept b_i are the regression coefficients and intercepts for indicator i in the regression equation obtained from the structural equation modeling analysis.

The prediction of indicator (4) assumes that the predicted values of the related observational variables have already been obtained using the aforementioned prediction methods. To predict the value of indicator j based on the latent variable in indicator i, the factor score $f_{i,t}$ of the latent variable in year t is used as the explanatory variable, the factor loading $w_{i,j}$ from indicator i to indicator j is used as the regression coefficient, and the unique factor $c_{i,j}$ is used as the intercept. The regression equation for the value $\hat{X_{j,t}}$ of indicator j can be written as follows:

$$\hat{X_{j,t}} = w_{i,j} \cdot f_{i,t} + c_{i,j} \tag{6}$$

Then, using the already obtained predicted value $X_{j,t}$ of indicator j, an optimization calculation is performed to minimize the squared error between $\hat{X_{j,t}}$ and $X_{j,t}$. The obtained factor score $f_{i,t}$ is used as the predicted value of indicator (4).

$$\min \sum_j (\hat{X_{j,t}} - X_{j,t})^2 \tag{7}$$

Indicator (5) assumes financial indicators that can be calculated based on formulas, such as ROA and operating profit. For such indicators, the calculation method is defined before conducting the simulation, and the predicted values of the relevant indicators are used to calculate them.

Using the above methods, the predicted values of each indicator in the causal relationships can be obtained for a single year. By repeating this process, predictions for future years can be made.

2.4.1. Prediction Method when Applying Measures

To apply the effects of policy measures in this simulator, the settings of the measures need to be defined in advance. The settings of the measures refer to the indicators to which the measure effects are applied, the strength of the effects, and the start year of the measures, for example. The simulation is then conducted based on the settings of the measures, applying the measure effects to the predicted values of the target indicators. When a measure effect is applied to a certain indicator, the effects are propagated to other indicators that have causal relationships with that indicator. As a result, it is possible to calculate how the financial indicators have changed due to the measures.

First, let's consider the case of applying a single measure. If we denote the value of indicator i in year t before and after applying the measure as $X'_{i,t}$ and $X_{i,t}$, respectively, the equation for applying the measure effect is as follows:

$$X_{i,t} = f_{m^i}(X'_{i,t}, \Phi_{m^i,t}) \quad (\text{for } t_m \leq t < t_m + k^i_m) \tag{8}$$

Here, $\Phi_{m^i,t}$ represents the measure effect that measure m has on indicator i in year t, and f_{m^i} represents the function for applying the measure effect. t_m represents the start year of measure m. For example, if a measure m increases indicator i by 10 in year t, then $\Phi_{m^i,t} = 10$, and $f_{m^i}(X'_{i,t}, \Phi_{m^i,t}) = X'_{i,t} + \Phi_{m^i,t}$. k^i_m represents the duration of the measure effect that measure m has on indicator i. The reason for defining k^i_m is to consider that the duration of the impact from the measures differs depending on the indicator.

In this study, only indicators (1) and (3) among the five types of indicators mentioned in Section 2.1 are assumed to be affected by the measure effects. The reason for not considering the application to indicator (2) is that if f_{m^i} is a multiplication, applying the measure effect to indicators that are not defined in the measure settings would be involved in the calculation formula. For example, if the measure effect is applied to indicator (2), which is $X_{i,t}$, using multiplication, the equation becomes as follows:

$$\begin{aligned} X_{i,t} &= f_{m^i}(X'_{i,t}, \Phi_{m^i,t}) \\ &= (\sum_j a_{j,i} \cdot X_{j,t} + b_i + \epsilon_{i,d}) \cdot \Phi_{m^i,t} \\ &= (\sum_j a_{j,i} \cdot X_{j,t}) \cdot \Phi_{m^i,t} + (b_i + \epsilon_{i,d}) \cdot \Phi_{m^i,t} \end{aligned} \tag{9}$$

The first term in the above equation can be interpreted as the values of each indicator or the measure effects applied to the values or regression coefficients of each indicator. Therefore, this equation is considered inappropriate, and the application to indicator (2) is not assumed.

The reason for not considering the application to indicator (4) is that it is difficult to estimate the effects on latent variables. If you want to apply effects to indicator (4), it can be achieved by applying the measure effects to the related observational variables.

Next, let's consider the case of applying multiple measures simultaneously. When multiple measures have effects on a certain indicator, using Equation 8 would result in different prediction results depending on the order of applying the measures. Therefore, the following equation is used to calculate the change in the indicator due to each measure, and the sum of these changes is added to the value of the indicator to obtain the predicted value when the measures are applied:

$$X_{i,t} = X'_{i,t} + \sum_{m=1,m=m'}^{M} [f_{m^i}(X'_{i,t}, \Phi_{m^i,t}) - X'_{i,t}] \tag{10}$$

$$(\text{if } m' \in M \text{ and } t_{m'} \leq t < t_{m'} + k_{m'^i})$$

Here, M is the set of measures m. m' refers to the measure in M that is applied within the target year for the prediction. This allows the effects of each measure to be applied regardless of the order of application.

Indicators (1) and (3), which are not influenced by other indicators, have predicted values that are sampled from a normal distribution with the previous year's indicator value as the mean. Therefore, if measure effects are applied to these indicators, the effects will continue to be reflected in the predicted results for the subsequent years. On the other hand, there are indicators for which the effects should not continue after the impact period of the measures. Depreciation expense is one example. For example, if an investment is made in equipment with a useful life of 5 years, the depreciation expense is recorded for 5 years and then no longer recorded. To handle such measure effects, a flag $\text{flag}_{m,i}$ is introduced to indicate whether the effects should continue after the impact period. Based on this flag, the measure effects are removed from indicator i as follows:

$$X_{i,t} = X'_{i,t} + \sum_{m=1,m=m''}^{M} [\sum_{k=1}^{k_m} (g_{m^i}(X'_{i,t}, \Phi_{m^i,t_m+k-1}) - X'_{i,t})] \tag{11}$$

$$(\text{if } m'' \in M \text{ and } t = t_{m''} + k_{m''} \text{ and } \text{flag}_{m''} = \text{False})$$

m'' refers to the measure in M that has passed k_m years since the start year of the measure and has $\text{flag}_{m,i} = \text{False}$. g_{m^i} is the inverse function of f_{m^i}. Using these equations, the predicted value of the indicator when multiple measure effects are applied can be formulated as follows:

$$\begin{aligned} X_{i,t} =& X'_{i,t} + \sum_{m=1,m=m'}^{M} [f_{m^i}(X'_{i,t}, \Phi_{m^i,t}) - X'_{i,t}] \\ &+ \sum_{m=1,m=m''}^{M} [\sum_{k=1}^{k_m} (g_{m^i}(X'_{i,t}, \Phi_{m^i,t_m+k-1}) - X'_{i,t})] \end{aligned} \tag{12}$$

Category	Item	Period	Time Lag
F	Cost of Goods Sold	2004-2021	0
F	Selling, General, and Administrative Expenses	2004-2021	0
F	Revenue	2004-2021	0,1
F	Operating Profit	2004-2021	0
F	Net Income	2004-2021	0
F	Earnings Before Interest and Taxes	2004-2021	0
F	Total Assets	2004-2021	0
F	Depreciation Expense	2005-2021	0
F	Research and Development Investment	2004-2021	0,3,4,5
F	Sales Forecast	2004-2021	0
F	Interest Expense Ratio on Interest-Bearing Debt	2004-2021	0
E	Energy Consumption	2004-2021	0,1,2
E	Renewable Energy Ratio	2016-2021	0,1,2
E	Waste	2005-2021	0,1,2
E	Water Usage	2012-2021	0,1,2
E	GHG Emissions Scope 1/2	2009-2021	0,1,2
S	Average Overtime Hours	2004-2021	0
S	Survey	2018-2021	0, 1
G	Female Executive Ratio	2013-2021	0
G	Foreign Executive Ratio	2013-2021	0

Table 1
Used data and their acquisition periods. We also listed time lag settings for each item.

$$
\begin{aligned}
&(\text{if } m' \in M \text{ and } t_{m'} \leq t < t_{m'} + k_{m'^i},\\
&\text{if } m'' \in M \text{ and } t = t_{m''} + k_{m''} \text{ and flag}_{m''} = \text{False})
\end{aligned}
$$

3. Case Study

3.1. Data

Table 1 shows the actual authors' affiliation corporate data to which this technique was applied. The applied data includes ESG disclosure indicators, financial indicators, and survey data within the company. Financial indicators have been disclosed every year from 2004 to 2021, which is generally the range of data acquisition. Environmental disclosure indicators have also been obtained over a relatively long period of about 10 years. Social data such as average overtime hours can be obtained for the same period as the financial indicators, and the female and foreign officer ratios belonging to governance can be obtained from 2013. The survey data, which started in 2018, is relatively new and asks whether companies and workplaces value diversity and whether the work environment is conducive to smooth operations. The average rating given by employees was obtained. The survey used in the data analysis consisted of 14 questions, including 3 questions related to diversity and 11 questions related to the work environment.

As mentioned earlier, time delayed indicators were added to the variables to incorporate temporal causality into the structural equation modeling. In particular, time delayed indicators were set for the survey results on the work environment and research and development investment to model the time relationship between sales and sales forecasts. Furthermore, a delay of 1-2 years was set for environmental indicators, which are alternative variables for the cost of capital.

When applying structural equation modeling, the above data was normalized to have a mean of 0 and a variance of 1.

#		Indicator Name	Impact	STD
1	F	Current Net Profit	2.597	0.000
2	F	Operating Profit	1.953	0.000
3	F	Sales Forecast	1.716	0.134
4	F	Sales (-1 yr)	1.212	0.010
5	S	Working Environment	0.558	0.118
6	F	R&D Investment (-4 yr)	0.469	0.088
7	G	Diversity	0.379	0.083
8	E	Renewable Energy Ratio (-2 yr)	0.057	0.067
9	F	R&D Investment (-3 yr)	0.031	0.048
10	F	R&D Investment (-5 yr)	0.006	0.017
11	E	Renewable Energy Ratio (-1 yr)	0.000	0.050
12	E	Water Usage	0.000	0.000
13	E	Energy Usage	0.000	0.000
14	E	Water Usage (-1 yr)	0.000	0.000
15	E	GHG Scope1,2 (-2 yr)	0.000	0.000
16	E	Water Usage (-2 yr)	0.000	0.001
17	E	Energy Usage (-2 yr)	0.000	0.003
18	E	GHG Scope1,2 (-1 yr)	-0.003	0.012
19	E	Waste (-2 yr)	-0.024	0.038
20	E	Energy Usage (-1 yr)	-0.147	0.061
21	S	Average Overtime Hours	-0.337	0.096
22	E	Waste	-0.405	0.062
23	E	Waste (-1 yr)	-0.442	0.170
24	F	R&D Investment	-0.601	0.000
25	F	Interest Expense Ratio	-0.670	0.000
26	F	Depreciation Expense	-3.126	0.166

Table 2
List of averaged impacts on ROA. Standard deviations(STD) were calculated through analyzing multiple times.

3.1.1. Factor Analysis

In this study, due to the high correlation among the survey results within the company, two factors were calculated: One representing diversity and the other representing the work environment, based on the aforementioned survey data. The factor coefficients ranged from 0.77 to 1.11 for diversity and from 0.87 to 1.01 for the work environment.

3.1.2. Constraints

Figure 1 shows the causal graph and its constraints. Except for the impact of the work environment on sales forecasts, constraints were imposed on the positive or negative coefficients of the regression analysis. The red lines in the figure represent the constraint of having non-positive coefficients, and the blue lines represent the constraint of having non-negative coefficients.

3.2. SEM Analysis

3.2.1. Calculation of Impact on Financial Indicators

The calculation of the impact on financial indicators was performed using the method described in Section 2.3.1. In this case, the impact on two financial indicators, operating profit and ROA, was calculated. Operating profit was defined as "sales minus selling and administrative expenses and cost of goods sold". For ROA, it was calculated by dividing net income by total assets. However, net income was derived through statistical impact calculation using operating profit, interest expense ratio, and ordinary income, and then calculated based on the actual value of total assets.

3.2.2. Implementation and Calculation

The model was implemented using the Python implementation of structural equation modeling, semopy[1]. Structural equation modeling was applied multiple times with different seed values, and the average impact and its variance were calculated.

3.2.3. Causal Relationship Graph

Figure 1 shows the analysis results of the relationship between ESG-financial indicators based on the hypothesis of causal relationships. In the figure, only the indicators of the environmental category are shown with a one-year time lag, but data from two years ago are also used. However, the influence from two years ago was almost zero. Regarding the interest-bearing debt interest rate, which is used as a substitute variable for the cost of capital, it can be seen that the coefficient of the energy consumption and waste generation one year ago is positive, indicating that the cost of capital increases as these values increase (environmental burden increases). Therefore, it can be concluded that reducing energy consumption and waste generation is effective in obtaining funds at a lower cost from external sources. In addition, the diversity factor node has a negative coefficient with respect to the interest-bearing debt interest rate, indicating that improving diversity is beneficial for obtaining funds.

Sales forecast were analyzed using multivariate analysis with factors such as working environment and research and development investment from 3-5 years ago, aiming to represent new businesses and business development potential. As a result, it was found that the working environment and research and development investment from 4 years ago have a positive coefficient on sales. This suggests that improving the working environment and investing in research and development are effective for improving company performance. Furthermore, water usage and energy usage have positive coefficients on sales forecast, indicating that they have an increasing environmental impact on business expansion. However, it is also found that these factors do not have a significant impact on sales costs.

It was also shown that research and development investment and overtime hours in the current year lead to an increase in costs and selling, general, and administrative expenses.

3.2.4. Impact on ROA

Table 2 shows the degree of impact of each ESG indicator on ROA. This calculation utilizes the characteristics of linear regression models and takes into account the causal relationship graph. It can be seen that diversity and waste generation have a significant impact on ROA. This is due to the fact that the cost of capital has a significant influence on company operations. Furthermore, in terms of the impact through sales, research and development investment from 4 years ago and the labor environment have a positive impact on ROA.

3.3. Simulation

3.3.1. Policy Measure Setting

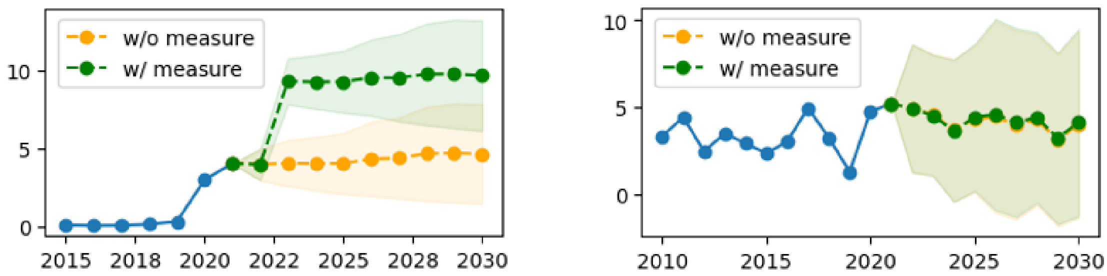
When applying structural equation modeling, the aforementioned data was normalized to have a mean of 0 and a variance of 1.

As can be seen from Figure 1, the improvement in the ratio of renewable energy in the past contributes to a decrease in the interest rate on interest-bearing debt and ultimately leads to an improvement in ROA. Therefore, in this experiment, we set "improvement in the ratio of renewable energy through investment in solar power generation facilities" as a policy and examined its impact on future ROA.

To set the policy, we calculated the cost-effectiveness of investment in solar power generation facilities. According to the materials from the Agency for Natural Resources and Energy[2], the capital cost for introducing solar power generation facilities in 2023 is estimated to be 22.3 [thousand yen/kW].

[1] https://pypi.org/project/semopy/

[2] https://www.meti.go.jp/shingikai/santeii/pdf/091_01_00.pdf



(a) Renewable Energy Ratio

(b) ROA

Figure 2: Simulation results.

Based on this, we calculated the impact on the target company using data from the fiscal year 2021, and obtained the result that "an investment of 20 billion yen will increase the ratio of renewable energy by approximately 5.26%". In addition, the useful life of a solar power generation system is specified as 9 years according to the National Tax Agency's website[3].

Based on the cost-effectiveness calculation results, we selected the ratio of renewable energy, depreciation expenses, and total assets as the indicators affected by the policy, and set the policy accordingly. The ratio of renewable energy is assumed to increase by 5.26% in the year the policy is applied, and its impact continues thereafter ($k_m = 1$, flag_m=True). Depreciation expenses and total assets are assumed to include 20 billion yen in costs for 9 years from the year the policy is applied, and the impact disappears from the 10th year ($k_m = 9$, flag_m=False). In addition, we set d in Equation 4 to 8.

In the next section, we will discuss the results of simulating the application of this policy in 2023 and conducting the simulation until 2030.

3.3.2. Experimental Results

Figure 2 shows the predicted changes in the renewable energy ratio and ROA through simulations. The simulations were run 100 times with different seed values. Figure 2 shows the average predicted values and standard deviations.

It can be observed that the renewable energy ratio increases by the start of the policy in the fiscal year 2023. Looking at the predicted values of ROA, it decreases from 4.59% in the case of no policy implementation to 4.53% when the policy is applied. This can be attributed to the fact that the increase in the renewable energy ratio has not yet affected the interest coverage ratio and at the same time, the investment costs have increased depreciation expenses and total assets.

However, by the fiscal year 2030, it is predicted that the ROA will increase to 4.14% when the policy is implemented, compared to 4.02% in the case of no policy implementation. This suggests that the positive effect of the increase in the renewable energy ratio on ROA outweighs the negative effect of investment costs.

Based on these results, it can be considered that this policy has the potential to improve ROA and is an effective measure for the target companies.

Furthermore, this method suggests that by modeling the relationship between non-financial and financial indicators and simulating how measures to improve non-financial indicators affect financial indicators, it is possible to propose ESG indicators and measures that companies should promote while considering long-term returns.

Additionally, this method suggests that it is possible to propose ESG indicators and measures that companies should promote within the balance of various financial and ESG indicators.

[3]https://www.nta.go.jp/law/shitsugi/hojin/05/12.htm

4. Conclusion

We developed a method for analyzing ESG-financial indicators using a causality hypothesis graph with a structural equation modeling. This method enables us to examine (1) the interrelationships between different indicators and (2) the validation of hypotheses regarding the issues that a company should prioritize. We also developed a simulator predicts future financial indicator values by comprehensively combining multiple policy measures. We evaluated this method by applying it to our corporate data and present the comparative results of predicted financial indicator values with and without the implementation of a measure. By these results, it is possible to consider how ESG affects management through what kind of mechanism in ESG-oriented management, using data. In the future, we plan to improve the credibility of the analysis results and consider methods for formulating ESG measures and systematization using this technology.

References

[1] D. Z. X. Huang, Environmental, social and governance (esg) activity and firm performance: a review and consolidation, Accounting & Finance 61 (2021) 335–360. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/acfi.12569. doi:https://doi.org/10.1111/acfi.12569. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/acfi.12569.

[2] A. Di Giuli, L. Kostovetsky, Are red or blue companies more likely to go green? politics and corporate social responsibility, Journal of Financial Economics 111 (2014) 158–180. URL: https://www.sciencedirect.com/science/article/pii/S0304405X13002511. doi:https://doi.org/10.1016/j.jfineco.2013.10.002.

[3] K. V. LINS, H. SERVAES, A. TAMAYO, Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis, The Journal of Finance 72 (2017) 1785–1824. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12505. doi:https://doi.org/10.1111/jofi.12505. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12505.

[4] J. F. Hair, G. T. M. Hult, C. M. Ringle, M. Sarstedt, N. P. Danks, S. Ray, An Introduction to Structural Equation Modeling, Springer International Publishing, Cham, 2021, pp. 1–29. URL: https://doi.org/10.1007/978-3-030-80519-7_1. doi:10.1007/978-3-030-80519-7_1.