ISSN　2424-0397

大阪大学経済学

第 74 巻　第 4 号
2025 年 3 月

OSAKA
ECONOMIC
PAPERS

大阪大学経済学会
大阪大学大学院経済学研究科
大阪府豊中市待兼山町

大阪大学経済学

（欧文誌名　Osaka Economic Papers）

本誌は大阪大学経済学会・大阪大学大学院経済学研究科の紀要として年４回，邦文ならびに欧文の論稿によって刊行される。

本誌の編集は，大阪大学経済学会によって選ばれた編集委員３名により行われる。編集委員は寄稿された研究成果を選定し，論文・覚書・資料および書評に類別して本誌を編集する。

大阪大学大学院経済学研究科に所属する研究者はその研究成果を本誌に寄稿することができる。なお，大阪大学大学院経済学研究科に所属しない研究者による研究成果も，大阪大学大学院経済学研究科における研究と密接な関係にあるものについては寄稿することができる。

なお，寄稿する際は「大阪大学経済学会」会員として，年会費￥4,000を納入する必要がある。

大阪大学経済学会会則

第１条　本会は大阪大学経済学会と称する。
第２条　本会は経済学，経営学の研究と発表を目的とする。
第３条　本会の事務所を大阪大学大学院経済学研究科に置く。
第４条　本会は下記の事業を行う。
　１．雑誌「大阪大学経済学」の発行（年４回）
　２．研究会及び講演会の開催（随時）
　３．その他，評議員会で適当と認めた事業
第５条　本会は下記の会員を以て組織する。
　１．普通会員（大阪大学大学院経済学研究科の教員，大阪大学の院生・学生・卒業生及び評議員会の承認を得た者）
　２．賛助会員（本会の事業を賛助する者）
第６条　会員は本会の諸事業に参加できる。
第７条　本会に下記の役員を置く。役員の任期は２年とする。
　１．会長（大阪大学大学院経済学研究科長を以ってこれに充てる）
　２．評議員（大阪大学大学院経済学研究科の教授・准教授・講師を以ってこれに充てる）
　３．雑誌編集・庶務・会計の委員若干名（評議員中より互選する）
　４．書記若干名
第８条　本会の運営はすべて評議員会の決議による。
第９条　会長は本会を代表する。
第10条　１．普通会員は会費として年額4,000円を納入するものとする。
　２．賛助会員は会費として年額10,000円以上を納入するものとする。
第11条　本会則の変更は評議員会の決議による。

大阪大学経済学会評議員

大阪大学経済学　第74巻　第4号

目　　次

論文

Impacts of Japan's Green Bond Guidelines 2020 on ESG bond issuers:
A quasi-experimental study using PSM-DID……Yuan Mingqing　1

産学間の共同研究における社会関係資本獲得を促進する方策の検討
REACH プロジェクト参加者への半構造化インタビューによる事例研究
……平丸大介・開本浩矢　32

Earnings prediction using machine learning: A survey……Yuanchao Peng　45

2024年度　学生懸賞論文　受賞作要旨……61

追悼

小泉進先生を偲んで……浦井憲　64

『大阪大学経済学』第74巻　令和6－7年　総目次……i

Impacts of Japan's Green Bond Guidelines 2020 on ESG bond issuers: A quasi-experimental study using PSM-DID*

Yuan Mingqing †

Abstract

This study investigates the impact of Japan's Green Bond Guidelines 2020 on the financial and environmental performance of firms issuing environmental, social, and governance (ESG) bonds. Using a quasi-experimental approach combining propensity score matching (PSM) and difference-in-differences (DID) methodologies, the analysis covers 795 listed Japanese firms from 2016 to 2024. Financial outcomes reveal mixed effects: firms that issued ESG bonds reduce return on assets (ROA) but increase return on equity (ROE), reflecting that revenue growth from ESG bond investment projects lags behind asset growth. Regarding environmental performance, Scope 1 and combined Scope 1 and 2 emissions exhibit significant increases, with Scope 2 and Scope 3 emissions showing no significant changes, indicating a disconnect between ESG bond issuance and actual emission reduction. ESG bonds used for refinancing exhibit negligible impacts on financial and environmental outcomes, whereas those financing new projects yield amplified financial effects. These findings identify challenges, including greenwashing risks and a lack of environmental additionality, wherein ESG bonds that refinance existing projects fail to improve environmental performance. These insights highlight the need for strengthening regulatory frameworks and policies encouraging the development of new ESG programs to enhance transparency, reporting standards, and meaningful environmental contributions.

JEL Classification: C23, Q56, Q58

Keywords: ESG bonds, Financial performance, Environmental additionality, Greenwashing, Sustainable finance

1. Introduction

Green bonds have emerged as pivotal instruments in sustainable finance, facilitating environmental, social, and governance (ESG) practices and promoting responsible business conduct (Flammer, 2020). Allen and Yago (2011) argued that green bonds based on market-based mechanisms and financial innovations can internalize

* I sincerely appreciate Professor Ono Tetsuo, Professor Onji Kazuki, and Professor Uwasu Michinori for their valuable guidance.

† Graduate Student, Graduate School of Economics, Osaka University

environmental costs into the economic decision-making processes of market participants. Previous studies have explored green bonds primarily in terms of bond premiums (Baker et al., 2018; Gianfrate and Peri, 2019; Nanayakkara and Colombage, 2019; Wang et al., 2020; Zerbib, 2019) and stock price responses (Aruga, 2024; Roslen et al., 2017; Baulkaran, 2019; Flammer, 2020; Glavas, 2020; Jakubik and Uguz, 2021). However, limited research has been directed toward exploring long-term financial and environmental performance.

ESG bonds represent a more comprehensive and emerging but rarely studied sustainable financial instrument, including green bonds, sustainability bonds, social bonds, and other specialized bonds. Issued to address specific ESG challenges, ESG bonds finance projects aimed at objectives such as climate mitigation, educational enhancement, and social welfare improvement (Japan Exchange Group, 2023). By channeling capital into sustainable development, ESG bonds serve as key mechanisms for promoting sustainable development and responsible business practices across diverse sectors.

The Ministry of the Environment of Japan publicized the "Green Bond Guidelines 2020" to promote the development of ESG bonds in 2020. The guidelines emphasize assessing and, where feasible, quantifying the environmental and social benefits generated by sustainable financial instruments. Following the publication of these guidelines, the issuance of diverse ESG bonds in Japan surged, with new green bond issuance reaching JPY 398.5 billion in 2020. Sustainability bond issuances increased to JPY 346.1 billion in 2020, compared to just JPY 25 billion in 2019.

Due to information asymmetry, external markets often lack comprehensive information, making it challenging to assess the authenticity and effectiveness of a firm's ESG practices. This information gap allows firms the opportunity to engage in strategic greenwashing (Benlemlih et al., 2022; Fatica and Panzica, 2021; Flammer, 2020; Yeow and Ng, 2021). Bond issuers may claim that the bonds are "green," without genuinely committing to environmental sustainability projects, which can lead to an overly positive reaction in stock prices. Therefore, a thorough evaluation of the long-term financial and environmental performance of firms issuing ESG bonds is essential to address greenwashing concerns and to validate the effectiveness of the Green Bond Guidelines 2020.

Despite this growth, significant gaps remain in understanding the long-term impacts of ESG bonds on financial and environmental outcomes. This study addresses key research questions: whether the guidelines incentivize ESG practices and enhance financial performance within firms. Using data on 795 listed Japanese firms, this analysis applies propensity score matching (PSM) developed by Rosenbaum and Rubin (1983) and a difference-in-differences (DID) approach to examine ESG bonds issued by these firms between January 2016 and January 2024. This study includes manually collected carbon emissions data—encompassing Scope 1, Scope 2, combined Scope 1 and 2 emissions, and Scope 3 emissions—to assess environmental performance. Returns on assets (ROA) and returns on equity (ROE) serve as indicators of long-term financial performance. This comprehensive approach provides a more detailed assessment than previous studies.

The contributions of this study are threefold. First, it addresses the research gap on the impact of ESG bonds in Japan, thereby expanding and enriching the literature on green bonds and sustainable finance. Second, it enhances understanding of the effectiveness of green bond policies at the firm level by comparing environmental and financial performance outcomes of ESG bond issuers against a comparable control group. Third, it provides the first study on the impact of the Green Bond Guidelines 2020 on ESG bond issuers, shedding light on whether the related policy guidelines incentivize investors' ESG preferences and provide insights into their effectiveness in driving sustainability.

2. Literature review and hypothesis development

The Green Bond Guidelines 2020 incorporate ESG factors across a broad spectrum of financial market participants, including issuers, borrowers, investors, financial institutions, and intermediaries. It strengthens the foundation, improves the quality, and expands the financing channels for sustainable finance in Japan. The guidelines promote rapid growth in issuers and diversified ESG bonds. However, the long-term financial impact of ESG bonds remains less explored. Zhou and Cui (2019) examined green bond issuance among Chinese listed firms, revealing positive effects on corporate profitability, operational performance, innovation, and corporate social responsibility. Globally, Flammer (2020) found the positive impact of green bonds on long-term financial performance. However, Yeow and Ng (2021) found no significant financial effect associated with green bonds.

The existing literature on the impact of ESG guidelines on corporate financial performance (CFP) intersects significantly with research on the effects of ESG and socially responsible investment (SRI)-related policies on CFP. Numerous studies suggest a positive relationship between ESG practices and financial outcomes. Quinche-Martín and Cabrera-Narváez (2020) highlighted that the reputational benefits derived from environmental innovation and CSR performance can enhance firms' market value and operational efficiency, thereby improving financial performance. Similarly, Przychodzen and Przychodzen (2015) found that firms engaged in eco-innovation and sustainability exhibit higher ROA and ROE. The green bonds advocated by the Green Bond Guidelines 2020 represent a key financial innovation, broadening access to capital and diversifying financing sources. Furthermore, the implementation of stricter standards and regulations (Albareda et al., 2007) can enhance the quality of financial products, ensuring that environmental benefits are realized alongside economic gains. For investors, this not only diversifies their portfolios and increases the range of available financial products but also enhances their ESG awareness and fosters sustainable financial practices driven by altruistic motivations (Hartzmark and Sussman, 2017; Riedl and Smeets, 2017).

Regarding the relationship between ESG and CFP, Friede et al. (2015) conducted large-scale meta-analyses and vote-count studies, revealing that approximately 63% of meta-analyses and 47% of vote-count studies report a non-negative impact. They also found that fewer than 10% indicates a negative association. Many studies support the positive effect of CSR performance on CFP (Ferrell et al., 2016; Flammer, 2015; Orlitzky, 2001; Tsai and Wu, 2022). Additionally, Bhaskaran et al. (2020), Fatemi et al. (2018), Li et al. (2018), and Yoon et al. (2018) found that ESG practices, ratings, and reporting are positively correlated with firm value.

However, some studies suggest different results. ESG bond investment serves as an emerging and vital component of SRI. Research on portfolio correlations proves a neutral SRI-CFP relationship for both institutional and private investors. This aligns with the neoclassical view of capital markets (Fama, 1970; Friedman, 1970; Fama, 1991). Schröder (2014) contended that SRI does not differ significantly from the risk-adjusted performance of conventional investments. Weston and Nnadi (2021) observed that exchange-traded funds (ETFs) adhering to ESG guidelines do not outperform traditional ETFs in market value. Revelli and Viviani (2014) also observed no clear advantage in CSR-oriented portfolios compared to traditional portfolios in financial performance. Conversely, Gavin et al. (2022) reported a negative association between ESG ratings and financial success. Chen et al. (2023) also found that ESG performance is linked to lower enhanced stock performance. Based on the literature, the following hypothesis is proposed:

Hypothesis 1: The introduction of the Green Bond Guidelines 2020 positively impacts the long-term financial performance of firms that issue ESG bonds.

The Green Bond Guidelines 2020 introduced by the Ministry of the Environment of Japan encourage the adoption of various ESG financial products, including green bonds, sustainability bonds, green loans, and sustainability-linked loans. Among these, green bonds are considered the most significant tool for financing projects aimed at achieving sustainability objectives (Ordonez-Borrallo et al., 2024). Following the guidelines, despite the proliferation of different ESG bonds, green bonds hold a dominant position within ESG bonds. ESG bonds allocate funds toward specific sustainable projects, effectively guiding the flow of capital into environmentally beneficial projects, thereby enhancing overall green performance.

However, concerns about greenwashing—a practice where firms falsely present their investments as environmentally friendly to attract eco-conscious investors while actually investing in non-environmentally beneficial projects—remain prevalent (Ministry of the Environment of Japan, 2020; Environmental Finance, 2023). Greenwashing undermines the integrity of ESG bonds and can dilute their environmental impact. The green bond market operates under a system of private governance (Fatica and Panzica, 2021). While the Green Bond Guidelines 2020 outline criteria for fund usage, management, and review, they are not legally binding. This voluntary nature of the guidelines raises concerns about the effectiveness of ESG bonds in delivering genuine environmental benefits if greenwashing is widespread. Therefore, it is crucial to examine whether the Green Bond Guidelines 2020 can substantially influence the environmental performance of firms issuing ESG bonds.

Research on green bonds' role in promoting corporate environmental performance (CEP) presents mixed findings. Ordonez-Borrallo et al. (2024) argued that green bonds can enhance corporate managers' focus on environmental performance and increase awareness of sustainability. Flammer (2020) used a market model and DID method, reporting that green bonds can lead to significant reductions in carbon emissions. Using the same approach, Yeow and Ng (2021) found that green bonds, when subject to third-party certification, positively impact environmental performance. Similarly, Benlemlih et al. (2022) reported that green bonds significantly enhance overall environmental performance, although the benefits may take a year or more to materialize.

Fatica and Panzica (2021) demonstrated that green bonds reduce total and Scope 1 emissions from non-financial firms, with greater reductions observed when refinancing bonds are excluded. This finding highlights the potential for green bonds to deliver additional environmental benefits, referred to as "additionality." In contrast, Bongaerts and Schoenmaker (2020) argued that green bonds fail to generate additionality, as they predominantly refinance existing green projects rather than fund new ones. Consequently, these bonds do not necessarily improve environmental performance. Furthermore, the decentralized issuance of green bonds reduces their liquidity, increasing financing costs and limiting the incentive for firms to pursue new environmental initiatives. Supporting this view, Wei et al. (2022) found that while green bond issuance can alleviate financial constraints, it does not lead to improved environmental performance. They emphasized that national and industry-level environmental regulations positively influence the relationship between green bond issuance and carbon performance. Given these considerations, the following hypothesis is proposed:

Hypothesis 2: The introduction of the Green Bond Guidelines 2020 positively impacts the environmental performance of firms that issue ESG bonds.

3. Research methodology

3.1 Quasi-experimental examination using PSM-DID methods

This study uses the DID model with PSM methods to explore the impacts of the Green Bond Guidelines 2020 on corporate financial and environmental performance of ESG bond issuers in Japan. The DID method designates listed firms that have issued ESG bonds as the treatment group, while other listed firms that have not issued ESG bonds serve as the control group. The study then calculates the difference in long-term financial performance between these groups before and after the implementation of the Green Bond Guidelines 2020.

The DID method does not require completely randomized assignment between treatment and control groups but instead relies on the parallel trend assumption. This assumption ensures that, in the absence of treatment, both groups would have exhibited similar trends over time, thus minimizing bias. Firms that have issued ESG bonds may have been selected based on specific factors, such as particular industry characteristics, firm size, and financial stability. This study addresses the endogeneity problem related to selection bias by employing the PSM technique originally developed by Rosenbaum and Rubin (1983) prior to applying the DID method. Following Flammer's (2020) application on financial performance and Fu et al. (2021)'s application on carbon emissions, this study estimates the policy's impact that it is uncontaminated by selection bias.

To establish a control group, this study matches firms that issued ESG bonds with firms that did not, based on criteria such as industry, age, financial characteristics, and CSR ratings. Creating comparable treatment and control groups enhances its robustness. This study uses a logit model, as specified in equation (1), to estimate parameters and predict propensity scores.

$$p(\boldsymbol{Z}_{i,t}) = \Pr(\boldsymbol{D}_i = 1 | \boldsymbol{Z}_{i,t}) = \frac{1}{1+\exp(-\boldsymbol{Z}'_{i,t}\boldsymbol{\beta})} \ . \tag{1}$$

$\boldsymbol{Z}_{i,t}$ represents matching variables that influence the likelihood of a firm issuing ESG bonds. $\boldsymbol{D}_i$ is the indicator variable that equals 1 if a firm issues ESG bonds and 0 otherwise, and $\boldsymbol{\beta}$ denotes the vector of coefficients. Both radius matching and kernel matching techniques are used to execute PSM.

The PSM estimates the probability that a firm will issue ESG bonds based on observed characteristics (propensity scores) and then matches treated firms with control firms that have similar propensity scores to create balanced sets with comparable covariate distributions (Stuart, 2010). This matching process minimizes significant differences in key variables between the treated and control groups before treatment, improving the validity of causal inferences. While PSM cannot fully eliminate biases from unobserved factors, it effectively reduces selection bias from observed factors (Rosenbaum and Rubin, 1983; Rubin and Thomas, 1992), thereby enhancing the comparability of groups and improving the accuracy of DID estimates (Becker and Ichino, 2002).

3.2 Data sources and variable definitions

The data source on ESG bonds in Japan is the website of the Ministry of the Environment of Japan[‡]. Financial data pertaining to Japanese listed firms comes from the EDINET and EOL databases, as well as Nikkei Firm Information DIGITAL. CSR ratings are from the TOYOKEIZAI annual CSR research reports. Industry-related data is sourced from the portal site of the official statistics portal of Japan, e-Stat. The manually gathered greenhouse gas (GHG) emissions data is from corporate official websites, ESG databooks, CSR reports, and annual unified reports of listed firms. Table 1 provides a detailed overview of the specific variables, their measurements, units, and sources.

Table 1. Variable specification and measurement

Variable	Indicator	Unit	Source
Total assets (TA)	Market value of total assets	Millions of yen	EOL database
Total liabilities (TL)	Market value of total liabilities	Millions of yen	EOL database
Return on assets (ROA)	Divide net income by total assets and multiply by 100	%	EDINET and EOL databases
Return on equity (ROE)	Divide net income by shareholders' equity and multiply by 100	%	EDINET and EOL databases
Corporate size ($SIZE$)	Logarithmic value of total assets		EOL database
Financial leverage (LEV)	Divide total debt by market value of total assets and multiply by 100	%	EOL database
Cashflow (CF)	Cash and cash equivalents at the end of period	Millions of yen	EOL database
Price earnings ratio (PER)	Market price per share divided by earnings per share		EOL database
Shareholders' equity ratio ($SHARE$)	Divide total shareholders' equity by total assets and multiply by 100	%	EOL database
CSR rating (CSR)	CSR rating		TOYOKEIZAI annual CSR research reports
Corporate age (AGE)	The difference between the current year and the year in which the firm was established	year	Nikkei Firm Information DIGITAL
Industry growth rates ($INDUSGR$)	Percentage change in the industry's added value from one period to the next.	%	e-Stat
Direct emissions ($Scope1$)	Direct GHG emissions from sources that are owned or controlled by the firm	kt-CO_2	Corporate official websites, ESG databooks, CSR reports, annual unified reports
Indirect emissions from energy consumption ($Scope2$)	Indirect GHG emissions through the use of electricity, heat, and steam supplied by other firms	kt-CO_2	Corporate official websites, ESG databooks, CSR reports, annual unified reports
Total of Scope 1 and Scope 2 emissions ($Scope12$)	Data for $Scope12$ were derived either from total emissions disclosed without separate $Scope1$ and $Scope2$ data or by summing individually reported $Scope1$ and $Scope2$	kt-CO_2	Corporate official websites, ESG databooks, CSR reports, annual unified reports
Indirect emissions from the value chain ($Scope3$)	All other indirect GHG emissions that occur in a firm's value chain	kt-CO_2	Corporate official websites, ESG databooks, CSR reports, annual unified reports

[‡] Data on ESG bond issuance in Japan were obtained from the Ministry of the Environment's Green Finance Portal. For further details, please see https://greenfinanceportal.env.go.jp/en/bond/issuance_data/issuance_list.html.

3.3 Estimation model

This study utilizes the Japanese government's Green Bond Guidelines 2020 policy as a quasi-natural experiment. Firms that issued ESG bonds are designated as the treatment group, while matched firms that did not issue ESG bonds serve as the control group. This study employs a DID model to assess whether the introduction of the Green Bond Guidelines 2020 can enhance both the financial performance and environmental outcomes of firms issuing ESG bonds. The empirical model is as follows:

$$Outcome\ Variable_{i,t} = \alpha + \gamma_1 ESG_{i,t} + \gamma_2 Post_{i,t} + \gamma_3 ESG \times Post + Firm\ controls_{i,t} + Industry\ controls_{\ i,t} + \lambda_t + \varepsilon_{i,t} \quad . \tag{2}$$

The subscripts i and t represent the firm and year, respectively. $ESG_{i,t}$ is a dummy variable for ESG bonds, where 1 indicates firms that issued ESG bonds and 0 indicates firms that did not. Since the guidelines were introduced on March 10, 2020, part of that fiscal year might be impacted by the policy. Therefore, the post-policy period starts in fiscal year 2020. The variable $Post_{i,t}$ is a dummy variable that takes the value of 1 for t in 2020 and beyond, and 0 for all prior periods. The interaction term $ESG \times Post$ is the DID interaction term. It takes the value of 1 for ESG bond-issuing firms after the introduction of these guidelines. A significant positive coefficient for γ_3 would indicate that the Green Bond Guidelines 2020 have a positive incentive effect on the financial performance of firms that issue ESG bonds. Control variables include firm-level control variables $Firm\ controls_{i,t}$ and industry-level control variables $Industry\ controls_{\ i,t}$. The main firm-level control variables include firm size ($SIZE_{i,t}$), total assets ($TA_{i,t}$), total liabilities ($TL_{i,t}$), financial leverage ($LEV_{i,t}$), cashflow ($CF_{i,t}$), price earnings ratio ($PER_{i,t}$), ownership capital ratio ($SHARE_{i,t}$), CSR ratings ($CSR_{i,t}$), and firm age ($AGE_{i,t}$). Since the industry fixed effects are absorbed by individual fixed effects, the time-invariant characteristics of the industry are already controlled for. Therefore, this model also controls for time-varying industry-specific factors, such as industry growth rates $INDUSGR_{i,t}$. λ_t denotes the time fixed effects.

3.4 Descriptive statistics

Table 2 summarizes 4594 observations from 795 firms, revealing significant variability in financial variables and firm characteristics related to CSR ratings and ESG bond issuance. The mean ROA is 33.96%, with the 50th percentile at 10.8%, showing a wide range of asset profitability. ROE presents a greater mean value of 59.83% and shows a higher standard deviation (SD) of 57.7. These results suggest that firms in the sample are quite profitable relative to shareholders' equity, despite significant variability. Firms are relatively mature, averaging 69.1 years, with notable differences in size, leverage, and cash flows. Industry growth rates vary widely (mean: 9.51%, SD: 34.3), with top 75th percentile values below 8.13%. Scope 3 emissions dominate environmental performance (mean: 9803, SD: 29453), highlighting the value chain's impact. CSR ratings are generally strong, with 49.72% rated AA or higher, but ESG bond issuance is limited to 16.15%.

Table 2. Descriptive statistics

Variable	Mean	SD	Min	Max	p25	p50	p75	N
TA	3.600e+06	2.300e+07	2036	3.900e+08	51429	240000	1.100e+06	4594
TL	3.100e+06	2.200e+07	676	3.700e+08	22927	100000	650000	4594
ROA	33.96	38.01	–27	146	5.400	10.80	62	4594
ROE	59.83	57.70	–374.1	204	9	45	100	4594
SIZE	12.06	1.820	7.620	17.78	10.67	11.94	13.37	3905
LEV	55.36	20.07	13.42	96.77	41.01	54.28	68.69	4594
CF	610000	5.200e+06	1	1.100e+08	4316	19954	79230	4593
PER	71.01	60.41	–9.700	266	21	54.50	109	4590
SHARE	43.31	20.74	–0.300	103	28.90	44.60	58	4594
AGE	69.10	28.70	0	145	56	72	86	4594
INDUSGR	9.51	34.30	–33.97	219.0	0.08	1.05	8.13	4594
*Scope*1	1864	7251	0	88900	6.040	51.68	243.9	1363
*Scope*2	313.5	823.2	0	7400	12.45	53.68	225.5	1374
*Scope*12	1807	6420	0.0900	92600	20.29	96.31	472	1833
*Scope*3	9803	29453	0.210	330000	514.6	2039	5013	940
CSR	Frequency		Percent					
AAA	571		12.43					
AA+	32		0.70					
AA	1681		36.59					
AA-	253		5.51					
A+	306		6.66					
A	975		21.22					
A-	391		8.51					
B	166		3.61					
BBB	54		1.18					
BBB+	105		2.29					
BBB-	16		0.35					
C	44		0.96					
Total	4594		100.00					
ESG	Frequency		Percent					
0	3852		83.85					
1	742		16.15					
Total	4594		100.00					

Notes: All variables, except for the dummy variable *ESG*, have been winsorized at the 1st and 99th percentiles. The descriptive statistics include the number of observations, mean, maximum, minimum, standard deviation, and the 25th, 50th, and 75th percentiles for each variable.

4. Results

4.1 Impacts of Green Bond Guidelines 2020 on corporate financial performance

This section reports the estimated impacts of the Green Bond Guidelines 2020 on firms' long-term financial performance, specifically measured by ROA and ROE. ROA and ROE are widely utilized financial metrics for assessing a firm's long-term financial performance (Murphy et al., 1996; Ordonez-Borrallo et al., 2024).

4.1.1 The radius matching PSM results

To mitigate selection bias, this study uses the radius matching method to apply PSM. The matching variables are selected based on their influence on the probability of a firm issuing ESG bonds. A total of ten variables are included: $SIZE$, LEV, TA, TL, CF, CSR, PER, $SHARE$, AGE, and $INDUSGR$.

Figure 1 presents the distribution of propensity scores for the treatment and control groups. The results indicate that 155 listed firms in the control group do not satisfy the common support assumption. This means that they have an exceptionally high or low probability of issuing ESG bonds. Consequently, these firms are excluded from further analysis. The final sample comprises 689 firms that meet the common support assumption, with 108 listed firms in the treatment group and 581 listed firms in the control group.

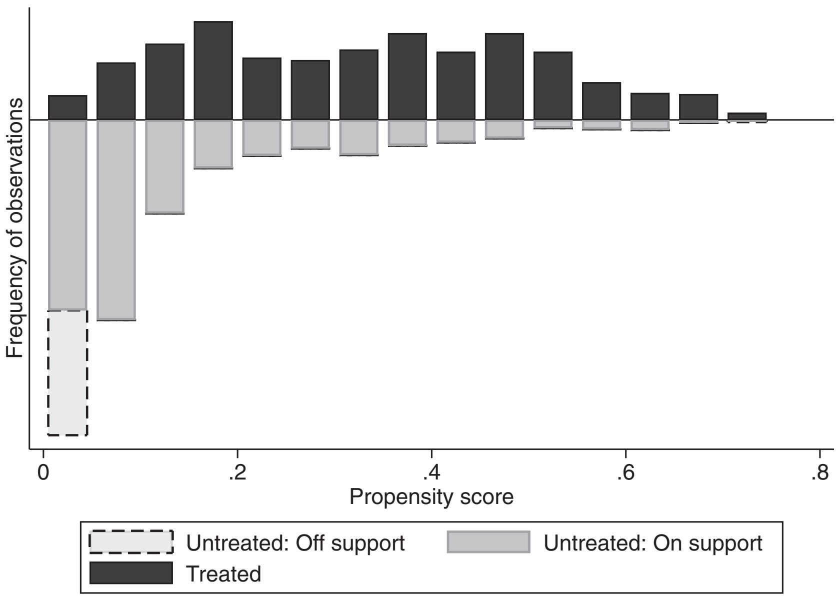
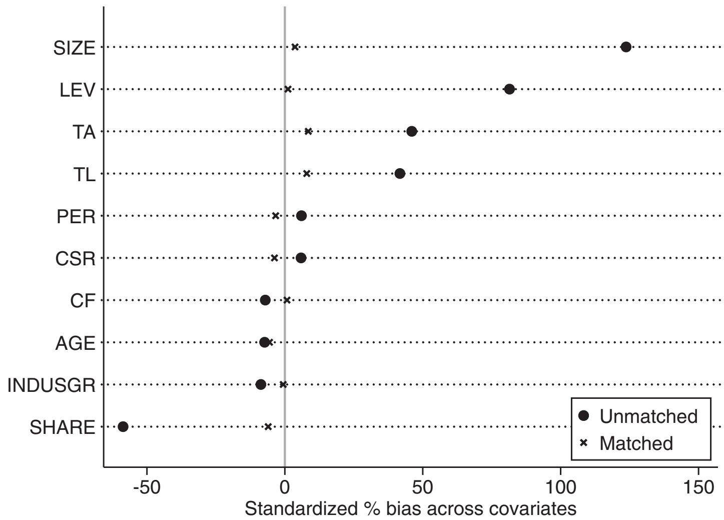


Figure 1. Propensity score distribution using radius matching

Figure 2 presents an intuitive comparison of the standardized percentage bias across various variables before and after PSM. Several variables, particularly $SIZE$, LEV, TA, and TL, exhibit substantial bias before matching. However, after applying PSM, the bias was significantly reduced. For instance, the bias in firm size was nearly eliminated, reduced to approximately 0%. We can observe similar reductions for LEV, TA, and TL. Appendix 1 and Appendix 2 show the specific balance test results of variables before and after radius matching for ROA and ROE, respectively. Overall, the bias across all variables was effectively minimized, bringing them close to zero. This visual comparison reinforces the conclusion that PSM has greatly enhanced the balance between the treatment and control groups.



Notes: The black dots represent the standardized percentage bias for each variable before matching (unmatched). The crosses (×) indicate the standardized percentage bias after matching (matched).

Figure 2. Standardized percentage bias before and after radius matching in PSM (financial impacts)

4.1.2 Parallel trend test

Figure 3 shows the parallel trend test results for ROA and ROE. This test is a crucial assumption in DID analysis, assessing whether the treatment group (firms that have issued ESG bonds) and the control group (firms that have not issued ESG bonds) exhibited similar trends in the outcome variable before the implementation

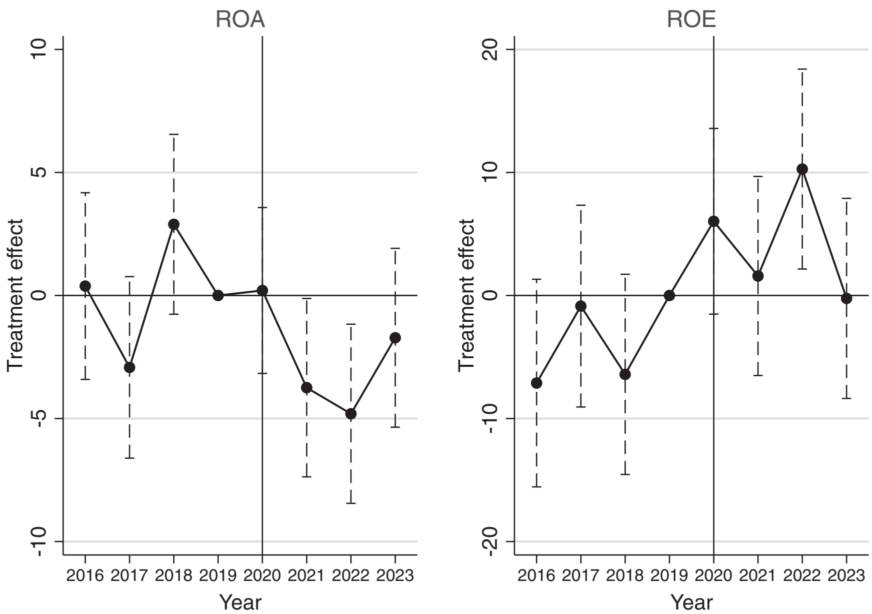


Figure 3. Parallel trend test results for financial impacts (radius PSM)

of the Green Bond Guidelines in 2020. The pre-intervention period (2016–2019) demonstrates no significant differences between the two groups, confirming that the parallel trend assumption is met. Post-2020, the treatment effect for ROA becomes significantly negative in 2021 and 2022, while the treatment effect for ROE also becomes significantly positive in 2022.

4.1.3 Baseline regression results for financial impacts

Table 3 reports the regression results for ROA and ROE with firm and time fixed effects. Column (1) suggests

Table 3. Regression analysis results for financial impacts

Variables	(1) *ROA*	(2) *ROE*
ESG	–28.964 ***	45.547 **
	(–2.84)	(1.99)
Post	–6.341	–1.843
	(–1.51)	(–0.20)
ESG × *Post*	–2.399 **	7.842 ***
	(–2.12)	(3.09)
TA	6.36e–07	–5.82e–07
	(1.39)	(–0.57)
TL	–4.27e–07	4.23e–07
	(–1.11)	(0.49)
SIZE	9.447 ***	2.791
	(4.25)	(0.56)
LEV	–0.497 ***	–0.543 ***
	(–7.32)	(–3.55)
ROE	0.133 ***	0.669 ***
	(17.30)	(17.30)
CF	0.000	–0.001
	(1.09)	(–1.53)
CSR	–0.282	0.199
	(–1.12)	(0.35)
PER	0.007 ***	0.046 ***
	(2.73)	(7.73)
SHARE	0.006 *	–0.002
	(1.88)	(–0.32)
AGE	–0.014	–0.386
	(–0.02)	(–0.30)
INDUSGR	0.017 **	0.051 ***
	(2.00)	(2.63)
Constant	–63.355	46.465
	(–1.35)	(0.44)
Observations	3766	3766
R-squared	0.172	0.149
Firm FE	YES	YES
Year FE	YES	YES
Number of firms	689	689

Notes: *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively. To maintain conciseness, the coefficients for year dummy variables are not reported in this table.

that the coefficient for the interaction term $ESG \times Post$ is negative and significant at –2.399 at the 5% level. This indicates that firms that have issued ESG bonds experience a reduction in ROA of 2.399 points, approximately a 7.1% decline from the mean ROA, after the issuance of the Green Bond Guidelines 2020, compared to firms that have not issued ESG bonds. Column (2) reveals a positive coefficient for $ESG \times Post$, with a value of 7.842, significant at the 1% level. This suggests that firms issuing ESG bonds experience a roughly 17.4% increase from the median ROE post-treatment, compared to firms that have not issued ESG bonds.

Control variables provide key insights into financial performance. As shown in Column (1), larger firms and those in faster-growing industries tend to achieve better ROA performance. In addition, higher ROE, PER, and shareholders' equity ratio are positively associated with improved ROA. Conversely, higher leverage is linked to poorer ROA performance, indicating that increased debt levels adversely affect profitability. Column (2) suggests that higher industry growth rates, ROA, and PER positively influence ROE, whereas increased leverage negatively impacts ROE performance. This reinforces the notion that increased debt levels may negatively affect a firm's profitability.

4.2 Impacts of Green Bond Guidelines 2020 on corporate environmental performance

This section investigates the impact of the Green Bond Guidelines 2020 on CEP, specifically focusing on carbon emissions as the primary indicator. To provide a comprehensive view of the effects of carbon emissions, this study assesses both direct and indirect carbon impacts based on the Greenhouse Gas Protocol, the most widely adopted GHG accounting standard globally. Scope 1 emissions ($Scope1$) represent direct GHG emissions from sources owned or controlled by the firm. Scope 2 emissions ($Scope2$) reflect the GHG emissions associated with the consumption of purchased electricity. $Scope12$ denotes the total Scope 1 and Scope 2 emissions. Scope 3 emissions ($Scope3$) encompass all other indirect GHG emissions occurring within the firm's value chain but originating from sources not owned or controlled by the firm.

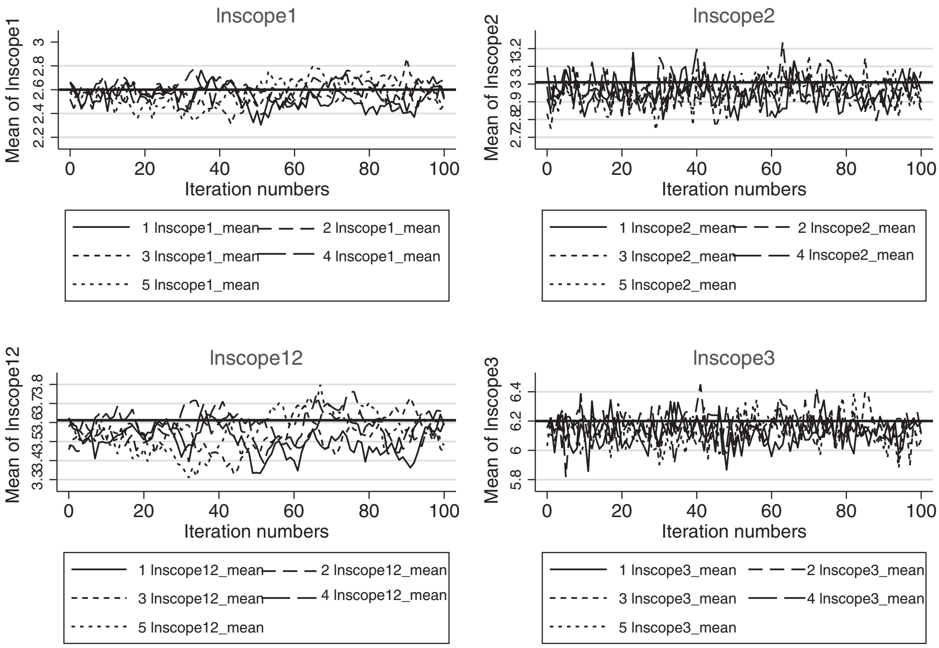
4.2.1 Addressing missing data with Multiple Imputation by Chained Equations

Given the challenges of incomplete carbon emissions data and the short timeframe for the harmonization of accounting standards, there is substantial missing data in the manually collected data for $Scope1$, $Scope2$, $Scope12$, and $Scope3$. To address these gaps, this study employs a multiple imputation method that effectively minimizes the issues associated with incomplete data and offers significant advantages over methods such as the deletion of missing values or provisional value estimation (Kofman and Sharpe, 2003). In addition, before conducting multiple imputation, the data for $Scope1$, $Scope2$, $Scope12$, and $Scope3$ were log-transformed to reduce the effects of heteroscedasticity. Their logarithmic values are denoted as $lnscope1$, $lnscope2$, $lnscope12$, and $lnscope3$, respectively.

This study employs a flexible multiple imputation method, Multiple Imputation by Chained Equations (MICE). This method assumes that missing data are missing at random, meaning the probability of missingness depends solely on observed data (Graham, 2009). The MICE process begins with a simple initial imputation, such as mean replacement, as a placeholder for each missing value (Azur et al., 2011). Each variable with missing data is then regressed on other variables in the dataset to generate predictions, which replace the initial placeholders. This iterative process continues for all variables until the model converges, achieving stability for imputation (Van Buuren, 2007). By iteratively applying regression models for each variable, MICE can flexibly handle complex data structures with suitable regression techniques (Raghunathan et al., 2000; Van Buuren, 2007). This analysis employs linear regression models for continuous variables such as $lnscope1$, $lnscope2$, $Scope12$, and $lnscope3$ to ensure accurate predictions aligned with their characteristics.

MICE offers significant advantages over single imputation methods by creating multiple imputed datasets, which better capture the uncertainty of missing data and produce more accurate standard errors (Schafer and Graham, 2002). Results from these datasets are combined using Rubin's (1987) and Schenker and Taylor's (1996) standard rules to derive final estimates. Moreover, MICE's adaptability and flexibility make it effective for complex and large-scale datasets (Azur et al., 2011), thereby enhancing the robustness of data analysis (Collins et al., 2001).

This study applies the convergence properties of MICE by performing 100 iterations with a specified random number of seed to ensure reproducibility. This approach aims to observe the aggregation trends of carbon emission estimates over these iterations. Figure 4 illustrates the mean values of five imputations for each variable ($lnscope1$, $lnscope2$, $Scope12$, and $lnscope3$) over the iterations. The horizontal black lines



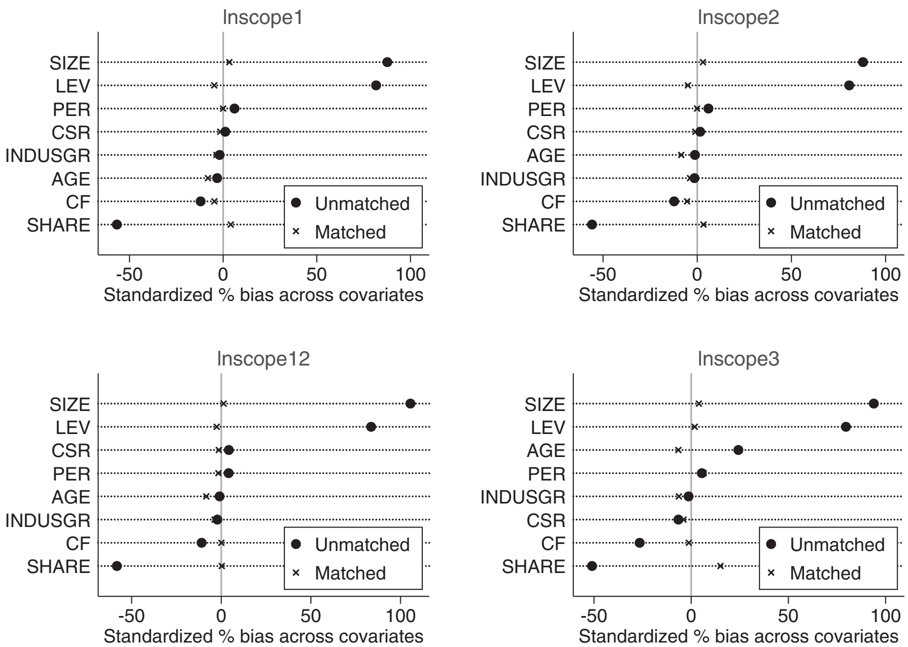
Notes:1 $lnscope1\_mean$, 2 $lnscope1\_mean$, 3 $lnscope1\_mean$, 4 $lnscope1\_mean$ and 5 $lnscope1\_mean$ denote the mean values of the five imputations for $lnscope1$. 1 $lnscope2\_mean$, 2 $lnscope2\_mean$, 3 $lnscope2\_mean$, 4 $lnscope2\_mean$ and 5 $lnscope2\_mean$ denote the mean values of the five imputations for $lnscope2$. 1 $lnscope12\_mean$, 2 $lnscope12\_mean$, 3 $lnscope12\_mean$, 4 $lnscope12\_mean$ and 5 $lnscope12\_mean$ denote the mean values of the five imputations for $lnscope12$. 1 $lnscope3\_mean$, 2 $lnscope3\_mean$, 3 $lnscope3\_mean$, 4 $lnscope3\_mean$ and 5 $lnscope3\_mean$ denote the mean values of the five imputations for $lnscope3$.

Figure 4. Convergence analysis of imputed carbon emission estimates using MICE

represent the observed mean values for each variable, providing a reference point to evaluate the stability of the imputations. As shown in Figure 4, all five chains exhibit fluctuations around the observed mean estimates of each variable. This consistent oscillation around the mean indicates convergence, suggesting that the MICE algorithm has stabilized across imputations.

4.2.2 The radius matching PSM results

Considering the correlation with carbon emission variables, this study uses eight matching variables for the PSM by excluding TA and TL. As shown in Figure 5, the standardized percentage bias for $lnscope1$, $lnscope2$, $lnscope12$, and $lnscope3$ decreases significantly after radius matching, with all variables achieving a bias reduction below 10%. This indicates a substantial improvement in the balance between the treatment and control groups, ensuring comparability.



Notes: The black dots represent the standardized percentage bias for each variable before matching (unmatched). The crosses (×) indicate the standardized percentage bias after matching (matched).

Figure 5. Standardized percentage bias before and after radius matching in PSM (environmental impacts)

4.2.3 Parallel trend test

Figure 6 presents the results of the parallel trend test conducted using the imputed values of $lnscope1$, $lnscope2$, $lnscope12$, and $lnscope3$ prior to the DID analysis. For all four variables, the treatment effects do not exhibit significant differences from zero before 2020, suggesting that the parallel trend assumption is satisfied. Post-2020 changes reflect possible treatment impacts. $lnscope1$ and $lnscope2$ show modest upward trends in treatment effects, while $lnscope12$ displays a significant increase in 2022. $lnscope3$ maintains stable and insignificant treatment effects.

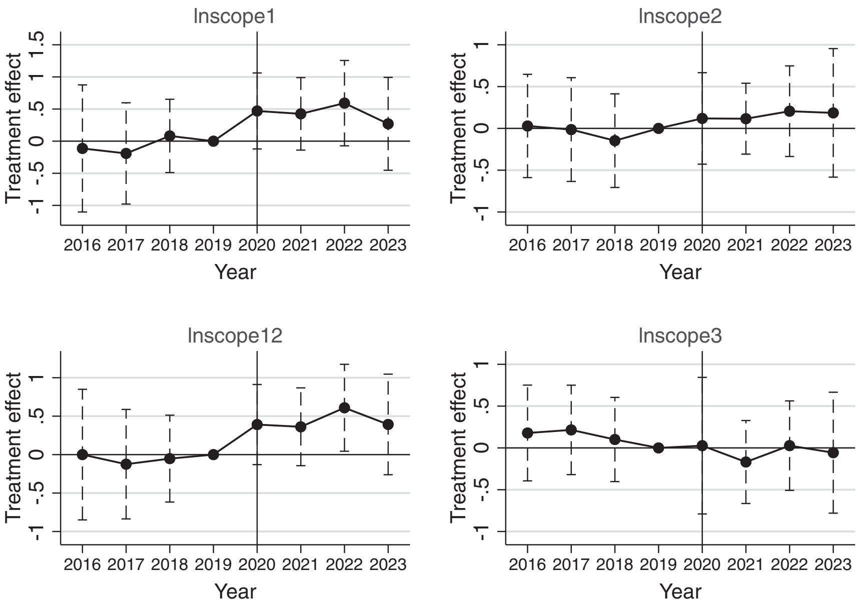


Figure 6. Parallel trend test results for environmental impacts (radius PSM)

4.2.4 Baseline regression results for environmental impacts

Table 4 presents the regression analysis results for various carbon emission scopes using fixed effects models with firm and year controls. The interaction term $ESG \times Post$ shows a statistically significant and positive effect on $lnscope1$ and $lnscope12$, with coefficients of 0.490 (Column (1)), corresponding to approximately a 6.72% increase in mean Scope 1 emissions, and 0.432 (Column (3)), indicating roughly a 5.93% increase in the mean value of combined Scope 1 and Scope 2 emissions. Both results are significant at the 1% level, suggesting that post-intervention firms experienced increased emissions in these categories. However, this treatment effect is not significant for $lnscope2$ and $lnscope3$.

Control variables reveal mixed effects: leverage (LEV) consistently shows a negative and significant impact across all emission types, indicating that more leveraged firms tend to have lower emissions. $SHARE$ positively influences emission reduction, particularly for Scope 1 emissions, while PER suggests firms with higher valuation ratios tend to reduce Scope 1 emissions.

Overall, while the 2020 Green Bond Guidelines may have promoted green bond issuance and ESG investments, they do not lead to a significant reduction in Scope 1 and combined Scope 1 and Scope 2 emissions. They have minimal impact on Scope 2 and Scope 3 emissions. This reflects the limited direct impact of ESG policies on overall environmental performance.

Table 4. Regression analysis results for environmental impacts

Variables	(1) *lnscope*1	(2) *lnscope*2	(3) *lnscope*12	(4) *lnscope*3
ESG	0.333	–0.435	0.262	0.976
	(2.337)	(1.440)	(1.925)	(2.042)
Post	0.668	–0.428	0.236	–0.010
	(0.879)	(0.703)	(0.631)	(0.626)
ESG × *Post*	0.490***	0.187	0.432***	–0.239
	(0.182)	(0.271)	(0.159)	(0.211)
SIZE	0.108	0.353	0.184	0.104
	(0.654)	(0.318)	(0.459)	(0.316)
LEV	–0.056***	–0.037***	–0.044***	–0.038**
	(–0.014)	(0.010)	(0.013)	(0.014)
ROA	0.001	0.003	–0.001	–0.005
	(0.004)	(0.003)	(0.003)	(0.004)
ROE	–0.0001	–0.006***	–0.004*	0.003*
	(0.002)	(0.001)	(0.002)	(0.001)
CF	–0.0001	0.0001**	–0.00005	–0.0001***
	(0.0001)	(0.00004)	(0.0001)	(0.00003)
CSR	0.053	0.035	0.024	0.035
	(0.055)	(0.034)	(0.056)	(0.045)
PER	–0.001**	–0.001	–0.001	–0.001
	(0.001)	(0.001)	(0.001)	(0.001)
SHARE	–0.004***	–0.001*	–0.002***	–0.002***
	(0.001)	(0.001)	(0.001)	(0.001)
AGE	–0.129	–0.020	–0.087	0.005
	(0.114)	(0.092)	(0.090)	(0.087)
INDUSGR	–0.001	0.002**	–0.002	–0.005***
	(0.002)	(0.001)	(0.002)	(0.002)
Constant	14.788	2.920	11.276	7.667
	(12.133)	(7.993)	(8.855)	(7.275)
Observations	4296	4296	4296	4296
R-squared	0.843	0.654	0.814	0.602
Number of firms	795	795	795	795
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Notes: *, ** and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively. To maintain conciseness, the coefficients for year dummy variables are not reported in this table.

5. Heterogeneity analysis

5.1 Temporal trends

This section investigates the time-varying impact of these guidelines by constructing new interaction terms, specifically $yr20 \times ESG$, $yr21 \times ESG$, $yr22 \times ESG$, and $yr23 \times ESG$, which are created by multiplying dummy variables for each year from 2020 to 2023 with the ESG variable. These interaction terms assess the temporal heterogeneity of the treatment effect.

Table 5 shows the results of temporal heterogeneity analysis for financial performance. For ROA, the coefficients of the interaction term $yr21 \times ESG$ and $yr22 \times ESG$ are significantly negative at the 5% and 1% levels, suggesting that the guidelines do not immediately impact ROA, but a dampening effect becomes evident from 2021. This highlights a lagged effect on corporate ROA performance for ESG bond-issuing firms. For ROE, both $yr20 \times ESG$ and $yr22 \times ESG$ show statistically significant positive coefficients at the 5% and 1% levels, respectively.

Table 5. Time trend analysis results for financial impacts

Variables	ROA	ROE
ESG	–28.999***	46.188**
	(–2.84)	(2.02)
$Post$	–6.754	–1.081
	(–1.61)	(–0.11)
$yr20 \times ESG$	0.145	9.241**
	(0.09)	(2.44)
$yr21 \times ESG$	–3.776**	4.951
	(–2.11)	(1.24)
$yr22 \times ESG$	–4.846***	13.651***
	(–2.69)	(3.39)
$yr23 \times ESG$	–1.747	3.221
	(–0.97)	(0.80)
Constant	–65.250	44.507
	(–1.39)	(0.42)
Observations	3766	3766
R-squared	0.174	0.150
Number of firms	689	689
Firm FE	YES	YES
Year FE	YES	YES

Notes: *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively. To maintain conciseness, the coefficients for control variables and year dummy variables are not reported in this table.

Table 6 presents the temporal heterogeneity effects of the guidelines on $lnscope1$, $lnscope2$, $lnscope12$, and $lnscope3$. The significant positive coefficients of $yr20 \times ESG$ and $yr22 \times ESG$ for lnscope1 and $lnscope12$ suggest an immediate and sustained increase in Scope 1 and combined Scope 1 and Scope 2 emissions in 2020 and 2022. In contrast, Scope 2 or Scope 3 emissions suggest no significant effects across the years.

Table 6. Time trend analysis results for environmental impacts

Variables	(1) $lnscope1$	(2) $lnscope2$	(3) $lnscope12$	(4) $lnscope3$
ESG	0.349	–0.432	0.276	0.986
	(2.332)	(1.432)	(1.924)	(2.046)
$Post$	0.695	–0.428	0.250	–0.021
	(0.898)	(0.721)	(0.638)	(0.613)
$yr20 \times ESG$	0.516**	0.150	0.436**	–0.027
	(0.263)	(0.277)	(0.211)	(0.374)
$yr21 \times ESG$	0.476	0.150	0.361	–0.386
	(0.310)	(0.302)	(0.241)	(0.273)
$yr22 \times ESG$	0.643*	0.239	0.576*	–0.260
	(0.363)	(0.365)	(0.298)	(0.267)
$yr23 \times ESG$	0.321	0.220	0.355	–0.330
	(0.347)	(0.463)	(0.282)	(0.343)
Constant	14.717	2.996	11.278	7.388
	(12.170)	(8.082)	(8.947)	(7.372)
Observations	4296	4296	4296	4296
R-squared	0.843	0.663	0.814	0.600
Number of firms	795	795	795	795
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Notes: *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively. To maintain conciseness, the coefficients for control variables and year dummy variables are not reported in this table.

5.2 Purpose-based heterogeneity analysis

This section builds on the analyses of Bongaerts and Schoenmaker (2020) and Fatica and Panzica (2021) by investigating firms that issued ESG bonds specifically for refinancing purposes, as well as those excluding ESG bonds issued for refinancing. The literature highlights a significant observation: most green bonds are predominantly used to refinance existing green projects rather than to fund new environmental initiatives. Consequently, while such bonds may allocate resources for green purposes, they generally fail to generate new environmental improvements, making it challenging to achieve environmental additionality (Bongaerts and Schoenmaker, 2020).

As illustrated in Table 7, the interaction term $ESG \times Post$ yields statistically significant coefficients for ROA (–2.933) and ROE (8.643) for firms excluding ESG bonds used for refinancing. These results are stronger than those observed in the baseline regression, highlighting the distinct financial impact of using ESG funds for new projects. ROE increases as income from these projects boosts the numerator, while ROA declines due to the rapid growth in total assets outpacing income generation, particularly in the short term. This effect becomes more pronounced for firms investing in new projects. In contrast, firms issuing ESG bonds for refinancing purposes show no significant effects on ROA and ROE, suggesting limited financial effects from such refinancing activities.

Table 7. Financial impacts of ESG Bonds: Firms issuing for refinancing vs. excluding refinancing purposes

	Firms issuing ESG bonds for refinancing purposes		Firms excluding ESG bonds issued for refinancing purposes	
Variables	*ROA*	*ROE*	*ROA*	*ROE*
ESG			–29.168***	46.492**
			(–2.82)	(2.05)
Post	–9.312	426.206**	–5.942	–2.374
	(–0.15)	(2.11)	(–1.39)	(–0.25)
ESG × *Post*	4.816	–10.321	–2.933**	8.643***
	(1.05)	(–0.69)	(–2.20)	(2.96)
TA	0.000	0.000	0.000	–0.000
	(0.62)	(0.76)	(1.54)	(–1.04)
TL	–0.000	–0.000	–0.000	0.000
	(–0.42)	(–0.72)	(–1.19)	(0.88)
SIZE	5.223	40.859*	9.394***	1.472
	(0.69)	(1.65)	(4.04)	(0.29)
LEV	–1.243***	–1.146	–0.467***	–0.497***
	(–5.07)	(–1.34)	(–6.50)	(–3.14)
ROA	0.025	0.266	0.142***	0.683***
	(1.13)	(1.13)	(17.51)	(17.51)
CF	0.000	–0.001	0.000	–0.001
	(0.12)	(–0.54)	(0.91)	(–1.19)
CSR	–0.448	4.587*	–0.240	0.004
	(–0.53)	(1.66)	(–0.91)	(0.01)
PER	0.018**	0.043*	0.006**	0.048***
	(2.42)	(1.75)	(2.15)	(7.67)
SHARE	0.003	–0.004	0.006*	0.000
	(0.49)	(–0.25)	(1.69)	(0.00)
AGE	–1.043	–62.586**	–0.020	–0.230
	(–0.12)	(–2.16)	(–0.03)	(–0.18)
INDUSGR	0.031	0.217*	0.017*	0.047**
	(0.92)	(1.95)	(1.87)	(2.37)
Constant	94.119	2,905.946*	–65.070	51.990
	(0.19)	(1.81)	(–1.34)	(0.49)
Observations	246	246	3520	3520
Number of firms	35	35	654	654
R-squared	0.257	0.131	0.176	0.159
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Notes: *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively. To maintain conciseness, the coefficients for year dummy variables are not reported in this table.

As illustrated in Table 8, the interaction term $ESG \times Post$ yields statistically significant coefficients for $lnscope1$ (2.514) and $lnscope12$ (1.911) in Panel A, which represents firms issuing ESG bonds for refinancing purposes. These results are significant at the 10% level and demonstrate coefficients notably higher than those observed in the baseline regression. However, the coefficients for Scope 2 and Scope 3 emissions are not statistically significant, suggesting limited or negligible effects on indirect and supply-chain-related emissions, aligning with baseline findings.

In contrast, firms excluding ESG bonds issued for refinancing purposes exhibit insignificant $ESG \times Post$ coefficients across all emission scopes, indicating no significant changes in environmental performance. This disparity highlights that ESG bonds used for refinancing purposes do not contribute to improved environmental performance or generate additional environmental benefits, as affirmed by Bongaerts and Schoenmaker (2020). Moreover, these findings indicate the potential risks of corporate greenwashing associated with the issuance of ESG bonds.

Table 8. Environmental impacts of ESG Bonds: Firms issuing for refinancing vs. excluding refinancing purposes

Variables	*lnscope*1	*lnscope*2	*lnscope*12	*lnscope*3
Panel A: Firms issuing ESG bonds for refinancing purposes				
Post	8.799	9.049	11.420	2.172
	(15.808)	(13.635)	(14.975)	(13.958)
ESG × *Post*	2.514*	0.165	1.911*	1.062
	(1.302)	(0.858)	(0.998)	(0.983)
SIZE	0.918	2.087	0.470	–0.184
	(2.071)	1.886)	(1.684)	(2.813)
LEV	0.0005	–0.008	0.006	–0.034
	(0.071)	(0.050)	(0.059)	(0.046)
ROA	0.003	0.012	0.002	–0.009
	(0.018)	(0.013)	(0.017)	(0.012)
ROE	–0.003	–0.007*	–0.007	0.003
	(0.006)	(0.004)	(0.005)	(0.005)
CF	–0.0001	0.0002	0.00005	–0.0001
	(0.0003)	(0.0002)	(0.0002)	(0.0001)
CSR	0.099	0.085	0.042	–0.019
	(0.204)	(0.170)	(0.154)	(0.175)
PER	0.001	–0.0003	0.002	0.001
	(0.002)	(0.001)	(0.002)	(0.001)
SHARE	–0.002	–0.001	–0.002	–0.002*
	(0.002)	(0.001)	(0.001)	(0.001)
AGE	–1.581	–1.505	–1.906	–0.522
	(2.256)	(1.959)	(2.154)	(1.949)
INDUSGR	–0.005	0.003	–0.006	–0.008
	(0.008)	(0.006)	(0.007)	(0.007)
Constant	76.609	56.038	101.553	42.151
	(12.356)	(116.338)	(119.939)	(121.470)
Observations	247	247	247	247
Number of firms	35	35	35	35
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Panel B: Firms excluding ESG bonds issued for refinancing purposes				
Post	0.685	–0.404	0.245	0.020
	(0.859)	(0.716)	(0.606)	(0.620)
ESG × *Post*	0.306	0.258	0.298	–0.308
	(0.211)	(0.253)	(0.200)	(0.234)
SIZE	0.064	0.283	0.179	0.116
	(0.640)	(0.289)	(0.445)	(0.284)
LEV	–0.062***	–0.038***	–0.048***	–0.039***
	(0.014)	(0.011)	(0.013)	(0.014)

ROA	0.001	0.003	–0.001	–0.005
	(0.004)	(0.002)	(0.003)	(0.005)
ROE	0.0003	–0.006***	–0.003	0.003
	(0.002)	(0.001)	(0.002)	(0.002)
CF	–0.0001	0.0001*	–0.00005	–0.0001***
	(0.0001)	(0.00005)	(0.0001)	(0.00004)
CSR	0.054	0.034	0.025	0.039
	(0.060)	(0.037)	(0.059)	(0.046)
PER	–0.002*	–0.001	–0.001	–0.001
	(0.001)	(0.001)	(0.001)	(0.001)
SHARE	–0.004***	–0.001	–0.003***	–0.002***
	(0.001)	(0.001)	(0.001)	(0.001)
AGE	–0.123	–0.016	–0.082	0.005
	(0.113)	(0.094)	(0.088)	(0.087)
INDUSGR	–0.001	0.003*	–0.001	–0.005**
	(0.002)	(0.001)	(0.002)	(0.002)
Constant	15.507	3.505	11.381	7.640
	(11.804)	(7.813)	(8.597)	(7.082)
Observations	4049	4049	4049	4049
Number of firms	760	760	760	760
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Notes: *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively. To maintain conciseness, the coefficients for year dummy variables are not reported in this table.

6. Robustness analysis

6.1 Placebo tests

This study conducts a placebo test by randomizing the interaction term 500 times to determine whether the estimated treatment effects significantly differed from the baseline results. Table 9 summarizes the Monte Carlo permutation test results for ROA, ROE, and carbon emissions. It assesses whether the observed treatment effect is statistically significant when compared to the randomized permutations of the treatment indicator (placebo treatment).

For ROA, the observed coefficient for $ESG \times Post$ is –2.315, indicating a negative treatment effect. The one-sided test shows that only 2 out of 500 random samples produce treatment effects more extreme than the observed effect, resulting in a p-value of 0.004. This indicates a very low probability (0.4%) that such a negative treatment effect could be obtained by chance. The two-sided p-value is 0.008, further confirming the robustness of the observed effect. For ROE, the $ESG \times Post$ coefficient is 8.092, with p-values of 0 across all tests. These results validate the robustness of the positive treatment effect.

For carbon emissions, the coefficients for $lnscope1$ and $lnscope12$ show significant treatment effects, with extremely low upper-sided and two-sided p-values for the $ESG \times Post$ coefficient. For $lnscope2$ and $lnscope3$, the small p-values suggest that the permutation effects are often more significant than the baseline results.

In conclusion, the placebo tests confirm the robustness of the treatment effects for ROA, ROE, and certain carbon emission variables, demonstrating that the observed impacts are unlikely to arise by random chance.

Table 9. Monte Carlo permutation test results

$ESG \times Post$	$ESG \times Post$ (obs)	Test	c	n	p	Standard error (p)
ROA						
Coefficient	–2.314908	lower	2	500	0.004	0.003
		upper	498	500	0.996	0.003
		two-sided			0.008	0.004
ROE						
Coefficient	8.091704	lower	500	500	1.000	0.000
		upper	0	500	0.000	0.000
		two-sided			0.000	0.000
***lnscope*1**						
Coefficient	0.370	lower	500	500	1.000	0.000
		upper	0	500	0.000	0.000
		two-sided			0.000	0.000
***lnscope*2**						
Coefficient	0.170	lower	500	500	1.000	0.000
		upper	0	500	0.000	0.000
		two-sided			0.000	0.000
***lnscope*12**						
Coefficient	0.315	lower	500	500	1.000	0.000
		upper	0	500	0.000	0.000
		two-sided			0.000	0.000
***lnscope*3**						
Coefficient	–0.323	lower	0	500	0.000	0.000
		upper	500	500	1.000	0.000
		two-sided			0.000	0.000

Notes: The $ESG \times Post$ (obs) column gives the observed values of $ESG \times Post$ from the original DID regression The “c” column represents the count of permutations where the treatment effect in the placebo samples was either smaller or larger than the observed treatment effect. The “n” column shows the total number of permutations conducted in the Monte Carlo test. The “c” column represents p-values. For a lower one-sided test, $c = \#\{\ ESG \times Post <= ESG \times Post\ (obs)\}$ and $p = p\_lower = c/n$. For an upper one-sided test, $c = \#\{\ ESG \times Post >= ESG \times Post\ (obs)\}$ and $p = p\_upper = c/n$. For two-sided test, $p = 2^*\ min\ (p\_lower, p\_upper)$.

6.2 Analysis of the kernel matching PSM results

This study tests the robustness of the results by employing the kernel matching method to apply PSM. The parallel trend tests confirm that the pre-intervention parallel trend assumption holds for all financial and environmental performance variables (Appendix 3 and Appendix 4). Table 10 displays the DID regression results, suggesting a negative treatment effect of –2.399 for ROA and a positive treatment effect of 7.842 for ROE. These results align with findings obtained through radius matching. For environmental performance, Table 10 reports significant positive effects for $lnscope1$ and $lnscope12$, while $lnscope2$ and $lnscope3$ remain insignificant. These results further verify the robustness of the initial findings.

Table 10. Regression analysis results using kernel PSM

Variables	(1) *ROA*	(2) *ROE*	(3) *lnscope*1	(4) *lnscope*2	(5) *lnscope*12	(6) *lnscope*3
ESG	–28.964***	45.547**	0.333	–0.435	0.262	0.976
	(–2.84)	(1.99)	(2.337)	(1.440)	(1.925)	(2.042)
Post	–6.341	–1.843	0.668	–0.428	0.236	–0.010
	(–1.51)	(–0.20)	(0.879)	(0.703)	(0.631)	(0.626)
$ESG \times Post$	–2.399**	7.842***	0.490***	0.187	0.432***	–0.239
	(–2.12)	(3.09)	(0.182)	(0.271)	(0.159)	(0.211)
Constant	–63.355	46.465	14.788	2.920	11.276	7.667
	(–1.35)	(0.44)	(12.133)	(7.993)	(8.855)	(7.275)
Observations	3766	3766	4296	4296	4296	4296
R-squared	0.172	0.149	0.843	0.654	0.814	0.602
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Number of firms	689	689	795	795	795	795

Notes: *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively. To maintain conciseness, the coefficients for control variables and year dummy variables are not reported in this table.

7. Discussion and conclusion

Based on data from 795 Japanese listed firms, this study examines the effects of the Green Bond Guidelines 2020 on the financial and environmental performance of these firms, employing a combination of PSM and DID methodologies.

Regarding financial performance, the findings indicate that after the implementation of the Green Bond Guidelines, firms that issued ESG bonds have significantly decreased ROA compared to those that did not. When firms issue ESG bonds, their total assets (the denominator) increase because the funds raised through the bonds are added to the balance sheet as assets. However, if the income generated by the new projects (the numerator) does not grow fast enough to match the increase in total assets, ROA will decrease. This reflects the fact that the firm's overall profitability relative to its larger asset base has declined. This outcome aligns with portfolio theory, which suggests that limitations on investment scope can negatively impact financial performance. These results are consistent with studies by Yeow and Ng (2021), Gavin et al. (2022), and Chen et al. (2023).

Conversely, the findings reveal that firms issuing ESG bonds experience significantly higher ROE after the guidelines were introduced, suggesting a positive effect on this financial metric. When firms issue ESG bonds, they often use the borrowed funds to finance projects or investments. These activities can increase the firm's net income, the numerator in the ROE formula. However, issuing bonds does not directly affect shareholders' equity (the denominator in the ROE formula) because bonds are a form of debt, not equity. As a result, an increase in net income leads to a higher ROE. This finding supports most literature on the relationship between ESG performance and CFP, including works by Ferrell et al. (2016), Flammer (2015), Orlitzky (2001), Przychodzen and Przychodzen (2015), and Tsai and Wu (2022). It underscores the role of ESG bonds in easing financing constraints.

To sum up, the Green Bond Guidelines 2020 influence firms issuing ESG bonds by affecting key financial metrics differently. It leads to an increase in net income (higher ROE), but the increase in total assets due to borrowing can outpace the income growth (lower ROA). This happens because the bonds expand the firm's asset base while keeping shareholders' equity constant, creating a divergence between these two metrics.

In addition, the financial impact depends on the purpose of the ESG bonds. For firms using ESG bonds for refinancing purposes, the effect on ROA and ROE is negligible, indicating limited financial implications. In contrast, for firms issuing ESG bonds to invest in new projects, the effects on both ROA and ROE are stronger than the baseline regression results, underscoring the significant role of new-project-driven ESG financing in shaping financial performance.

Moreover, firms operating in rapidly growing industries, as well as those with higher PER, tend to achieve better performance in terms of both ROA and ROE. In contrast, increased leverage is linked to lower ROA and ROE. The heterogeneity analysis shows that the impact of the policy guidance on CFP exhibits significant time variation. While the guidelines have an immediate positive impact on ROE in 2020, a negative impact on ROA is not observed until 2022, indicating a lagged response in ROA.

This study evaluates environmental performance through carbon emissions, focusing on Scope 1, Scope 2, $Scope12$, and Scope 3 emissions. The findings reveal that firms issuing ESG bonds do not achieve significant reductions in carbon emissions. Instead, Scope 1 and $Scope12$ emissions increase significantly, while Scope 2 and Scope 3 emissions show no significant changes when compared to firms that did not issue ESG bonds. These results suggest a disconnect between ESG bond issuance and actual emissions reduction efforts, consistent with the findings of Wei et al. (2022) and Bongaerts and Schoenmaker (2020). Control variables reveal that higher leverage reduces emissions across all types, while $SHARE$ and PER are associated with reductions in Scope 1 emissions, particularly in firms with higher ownership and valuation ratios.

Moreover, this study reveals that environmental performance varies depending on the purpose of ESG bond issuance. However, the results indicate that neither ESG bonds used for refinancing nor those allocated to new projects lead to improved environmental performance, with both failing to achieve environmental additionality.

These findings highlight several challenges associated with ESG bonds in Japan. First, there is a significant risk of greenwashing. Due to inconsistent ratings across different rating agencies for ESG bonds, it is difficult to accurately assess and verify the environmental impact of ESG bonds. This issue exacerbates the potential for firms to exaggerate or misrepresent their environmental initiatives, as noted by Benlemlih et al. (2022), Fatica and Panzica (2021), Flammer (2020), and Yeow and Ng (2021).

Second, similar to the observations by Bongaerts and Schoenmaker (2020) on green bonds, Japan's ESG bonds demonstrate a lack of additionality. These bonds often refinance existing projects, originally funded by conventional bonds, rather than financing new environmental initiatives. This approach limits their ability to deliver incremental environmental benefits, as such projects would likely continue regardless of refinancing.

Third, the decentralized issuance of green bonds reduces their liquidity, while low yields and high issuance and reporting costs further increase the effective cost of financing (Bongaerts and Schoenmaker, 2020). These challenges make ESG bonds, particularly green bonds, less appealing for funding new projects, thereby diminishing their potential to achieve additionality and meaningful environmental impact.

8. Implications

For investors, these findings reflect the importance of assessing the diverse impacts of ESG bond issuances on different financial metrics. Investors should consider not only the immediate stock market reactions but also the impacts on firms' operational performance and asset utilization. This can help make more informed investment decisions that balance financial returns with sustainability goals.

The effect of Green Bond Guidelines 2020 in reducing carbon emissions remains limited, suggesting a significant risk of greenwashing. This calls for stricter monitoring, standardization of environmental impact assessments, and measures to ensure that ESG bonds contribute meaningfully to sustainability goals. Regulators and certification bodies should enhance the scrutiny, such as reporting requirements and verification mechanisms of ESG claims for ESG bond issuers.

Moreover, the findings emphasize that ESG bonds used for refinancing fail to improve financial performance and environmental outcomes, whereas those financing new projects show greater effects on ROA and ROE. This highlights the need for targeted policies or regulatory frameworks to incentivize additionality and address inefficiencies in Japan's ESG bond market. Examples include innovative mechanisms such as green certificates (Bongaerts and Schoenmaker, 2020) and initiatives that link ESG bonds to new ESG projects.

Furthermore, many firms fail to disclose ESG data, and inconsistencies in reported emissions make it hard to accurately and comprehensively measure their environmental performance. This lack of clarity increases information gaps for investors and policymakers. To address these challenges, policymakers should standardize ESG reporting and improve transparency. Establishing uniform data requirements will improve the accuracy and comparability of disclosed information, which allows for a more accurate assessment of a firm's true environmental impact.

References

Albareda, L., Lozano, J. M., & Ysa, T. (2007) "Public Policies on Corporate Social Responsibility: The Role of Governments in Europe," *Journal of Business Ethics,* 74(4), 391–407. https://doi.org/10.1007/s10551-007-9514-1

Allen, F., & Yago, G. (2011) "Environmental Finance: Innovating to Save the Planet," *Journal of Applied Corporate Finance,* 23(3), 99–111. https://doi.org/10.1111/j.1745-6622.2011.00347.x

Aruga, K. (2024) "Are Retail Investors Willing to Buy Green Bonds? A Case for Japan," *Journal of Sustainable Finance & Investment*, 1–15. https://doi.org/10.1080/20430795.2024.2349723

Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011) "Multiple Imputation by Chained Equations: What Is It and How Does It Work?" *International Journal of Methods in Psychiatric Research*, 20(1), 40–49. https://doi.org/10.1002/mpr.329

Baker, M. P., Bergstresser, D. B., Serafeim, G., & Wurgler, J. A. (2018) "Financing the Response to Climate Change: The Pricing and Ownership of U.S. Green Bonds," *NBER Working Papers*, 25194. https://doi.org/10.2139/ssrn.3275327

Baulkaran, V. (2019) "Stock Market Reaction to Green Bond Issuance," *Journal of Asset Management*, 20(5), 331–340. https://doi.org/10.1057/s41260-018-00105-1

Becker, S. O., & Ichino, A. (2002) "Estimation of Average Treatment Effects Based on Propensity Scores," *The Stata Journal: Promoting Communications on Statistics and Stata*, 2(4), 358–377. https://doi.org/10.1177/1536867x0200200403

Benlemlih, M., Jaballah, J., & Kermiche, L. (2022) "Does Financing Strategy Accelerate Corporate Energy Transition? Evidence from Green Bonds," *Business Strategy and the Environment*, 32(1) https://doi.org/10.1002/bse.3180

Bhaskaran, R. K., Ting, I. W. K., Sukumaran, S. K., & Sumod, S. D. (2020) "Environmental, Social, and Governance Initiatives and Wealth Creation for Firms: An Empirical Examination," *Managerial and Decision Economics*, 41(5), 710–729. https://doi.org/10.1002/mde.3131

Bongaerts, D., & Schoenmaker, D. (2020) "The Next Step in Green Bond Financing," *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3389762

Chen, S., Yu, S., & Gao, P. (2023) "Environmental, Social, and Governance (ESG) Performance and Financial Outcomes: Analyzing the Impact of ESG on Financial Performance," *Journal of Environmental Management*, 345(1), 118829–118829. https://doi.org/10.1016/j.jenvman.2023.118829

Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001) "A Comparison of Inclusive and Restrictive Strategies in Modern Missing Data Procedures," *Psychological Methods*, 6(4), 330–351. https://doi.org/10.1037/1082-989x.6.4.330

Environmental Finance. (2023) *Sustainable Bonds Insight* 2023. Retrieved from https://www.environmental-finance.com/assets/files/research/sustainable-bonds-insight-2023.pdf. (Accessed October 10, 2024)

Fama, E. F. (1970) "Efficient Capital Markets: A Review of Theory and Empirical Work," *The Journal of Finance*, 25(2), 383–417. https://doi.org/10.2307/2325486

Fama, E. F. (1991) "Efficient Capital Markets: II," *The Journal of Finance*, 46(5), 1575–1617. https://doi.org/10.1111/j.1540-6261.1991.tb04636.x

Fatemi, A., Glaum, M., & Kaiser, S. (2018) "ESG Performance and Firm Value: The Moderating Role of Disclosure," *Global Finance Journal*, 38, 45–64. https://doi.org/10.1016/j.gfj.2017.03.001

Fatica, S., & Panzica, R. (2021) "Green Bonds as a Tool Against Climate Change?" *Business Strategy and the Environment*, 30(5) https://doi.org/10.1002/bse.2771

Ferrell, A., Liang, H., & Renneboog, L. (2016) "Socially Responsible Firms," *Journal of Financial Economics*, 122(3), 585–606. https://doi.org/10.1016/j.jfineco.2015.12.003

Flammer, C. (2015) "Does Corporate Social Responsibility Lead to Superior Financial Performance? A Regression Discontinuity Approach," *Management Science*, 61(11), 2549–2568. https://doi.org/10.1287/mnsc.2014.2038

Flammer, C. (2020) "Green Bonds: Effectiveness and Implications for Public Policy," *Environmental and Energy Policy and the Economy*, 1, 95–128. https://doi.org/10.1086/706794

Friede, G., Busch, T., & Bassen, A. (2015) "ESG and Financial Performance: Aggregated Evidence from More Than 2000 Empirical Studies," *Journal of Sustainable Finance & Investment*, 5(4), 210–233. https://doi.org/10.1080/20430795.2015.1118917

Friedman, M. (1970) "The Social Responsibility of Business Is to Increase Its Profits," Retrieved from https://

www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html

Fu, Y., He, C., & Luo, L. (2021) "Does the Low-Carbon City Policy Make a Difference? Empirical Evidence of the Pilot Scheme in China with DEA and PSM-DID," *Ecological Indicators*, 122, 107238. https://doi.org/10.1016/j.ecolind.2020.107238

Gavin, M., Coelho, M. T. P., McGlinch, J., & Henisz, W. J. (2022) "Pathways to Materiality: Environmental, Social & Governance (ESG) Factors and Financial Performance," *Academy of Management Proceedings*, 2022(1) https://doi.org/10.5465/ambpp.2022.16003abstract

Gianfrate, G., & Peri, M. (2019) "The Green Advantage: Exploring the Convenience of Issuing Green Bonds," *Journal of Cleaner Production*, 219, 127–135. https://doi.org/10.1016/j.jclepro.2019.02.022

Glavas, D. (2020) "How Do Stock Prices React to Green Bond Issuance Announcements?" *Finance*, 41(1), 7–51. https://doi.org/10.2139/ssrn.3279069

Graham, J. W. (2009) "Missing Data Analysis: Making It Work in the Real World," *Annual Review of Psychology*, 60(1), 549–576. https://doi.org/10.1146/annurev.psych.58.110405.085530

Hartzmark, S. M., & Sussman, A. B. (2017) "Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows," *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3016092

Jakubik, P., & Uguz, S. (2021) "Impact of Green Bond Policies on Insurers: Evidence from the European Equity Market," *Journal of Economics and Finance*, 45(2), 381–393. https://doi.org/10.1007/s12197-020-09534-4

Japan Exchange Group. (2023) ESG Bonds | *Japan Exchange Group*, Retrieved from https://www.jpx.co.jp/english/equities/products/tpbm/green-and-social-bonds/index.html (Accessed October 20, 2024)

Kofman, P., & Sharpe, I. G. (2003) "Using Multiple Imputation in the Analysis of Incomplete Observations in Finance," *Journal of Financial Econometrics*, 1(2), 216–249. https://doi.org/10.1093/jjfinec/nbg013

Li, Y., Gong, M., Zhang, X.-Y., & Koh, L. (2018) "The Impact of Environmental, Social, and Governance Disclosure on Firm Value: The Role of CEO Power," *The British Accounting Review*, 50(1), 60–75. https://doi.org/10.1016/j.bar.2017.09.007

Ministry of the Environment of Japan. (2020) *Green Bond Guidelines* 2020. Retrieved from https://www.env.go.jp/content/000042342.pdf (Accessed August 10, 2024)

Murphy, G. B., Trailer, J. W., & Hill, R. C. (1996) "Measuring Performance in Entrepreneurship Research," *Journal of Business Research*, 36(1), 15–23. https://doi.org/10.1016/0148-2963(95)00159-x

Nanayakkara, M., & Colombage, S. (2019) "Do Investors in Green Bond Market Pay a Premium? Global Evidence," *Applied Economics*, 51(40), 4425–4437. https://doi.org/10.1080/00036846.2019.1591611

Ordonez ‐ Borrallo, R., Ortiz ‐ de ‐ Mandojana, N., & Delgado ‐ Ceballos, J. (2024) "Green Bonds and Environmental Performance: The Effect of Management Attention," *Corporate Social-Responsibility and Environmental Management*. https://doi.org/10.1002/csr.2858

Orlitzky, M. (2001) "Does Firm Size Confound the Relationship Between Corporate Social Performance and Firm Financial Performance?" *Journal of Business Ethics*, 33(2), 167–180. https://doi.org/10.1023/a:1017516826427

Przychodzen, J., & Przychodzen, W. (2015) "Relationships Between Eco-Innovation and Financial Performance – Evidence from Publicly Traded Companies in Poland and Hungary," *Journal of Cleaner Production*, 90, 253–263. https://doi.org/10.1016/j.jclepro.2014.11.034

Quinche-Martín, F. L., & Cabrera-Narváez, A. (2020) "Exploring the Potential Links Between Social and Environmental Accounting and Political Ecology," *Social and Environmental Accountability Journal*, 40(1), 53–74. https://doi.org/10.1080/0969160x.2020.1730214

Raghunathan, T., Lepkowski, J., Hoewyk, J. V., & Solenberger, P. W. (2000) "A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Regression Models," *Survey Methodology*, 27(1) Retrieved from https://www.researchgate.net/publication/244959137

Revelli, C., & Viviani, J.-L. (2014) "Financial Performance of Socially Responsible Investing (SRI): What Have We Learned? A Meta-Analysis," *Business Ethics: A European Review*, 24(2), 158–185. https://doi.org/10.1111/beer.12076

Riedl, A., & Smeets, P. (2017) "Why Do Investors Hold Socially Responsible Mutual Funds?" *The Journal of Finance*, 72(6), 2505–2550. https://doi.org/10.1111/jofi.12547

Rosenbaum, P. R., & Rubin, D. B. (1983) "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Roslen, S. N. M., Yee, L. S., & Ibrahim, S. A. B. (2017) "Green Bond and Shareholders' Wealth: A Multi-Country Event Study," *International Journal of Globalization and Small Business*, 9(1), 61. https://doi.org/10.1504/ijgsb.2017.084701

Rubin, D. B. (1987) *Multiple Imputation for Nonresponse in Surveys*, John Wiley.

Rubin, D. B., & Thomas, N. (1992) "Affinely Invariant Matching Methods with Ellipsoidal Distributions," *The Annals of Statistics*, 20(2), 1079–1093. https://doi.org/10.2307/2241998

Schafer, J. L., & Graham, J. W. (2002) "Missing Data: Our View of the State of the Art," *Psychological Methods*, 7(2), 147–177. https://doi.org/10.1037/1082-989x.7.2.147

Schenker, N., & Taylor, J. M. G. (1996) "Partially Parametric Techniques for Multiple Imputation," *Computational Statistics & Data Analysis*, 22(4), 425–446. https://doi.org/10.1016/0167-9473(95)00057-7

Schröder, M. (2014) "Financial Effects of Corporate Social Responsibility: A Literature Review," *Journal of Sustainable Finance & Investment*, 4(4), 337–350. https://doi.org/10.1080/20430795.2014.971096

Stuart, E. A. (2010) "Matching Methods for Causal Inference: A Review and a Look Forward," *Statistical Science*, 25(1), 1–21. https://doi.org/10.1214/09-sts313

Tsai, H.-J., & Wu, Y. (2022) "Changes in Corporate Social Responsibility and Stock Performance," *Journal of Business Ethics*, 178, 735–755. https://doi.org/10.1007/s10551-021-04772-w

van Buuren, S. (2007) "Multiple Imputation of Discrete and Continuous Data by Fully Conditional Specification," *Statistical Methods in Medical Research*, 16(3), 219–242. https://doi.org/10.1177/0962280206074463

Wang, J., Chen, X., Li, X., Yu, J., & Zhong, R. (2020) "The Market Reaction to Green Bond Issuance: Evidence from China," *Pacific-Basin Finance Journal*, 60, 101294. https://doi.org/10.1016/j.pacfin.2020.101294

Wei, P., Li, Y., & Zhang, Y. (2022) "Corporate Green Bonds and Carbon Performance: An Economic Input–Output Life Cycle Assessment Model - Based Analysis," *Business Strategy and the Environment*, 32(6) https://doi.org/10.1002/bse.3267

Weston, P., & Nnadi, M. (2021) "Evaluation of Strategic and Financial Variables of Corporate Sustainability and ESG Policies on Corporate Finance Performance," *Journal of Sustainable Finance & Investment*, 13(2), 1–17. https://doi.org/10.1080/20430795.2021.1883984

Yeow, P. S., & Ng, T. S. (2021) "Do Green Bond Certifications Create Value? Evidence from Chinese Green Bond Market," *Journal of Sustainable Finance & Investment*, 11(4), 300–318. https://doi.org/10.1080/20430795.2020.1845775

Yoon, B., Lee, J., & Byun, R. (2018) "Does ESG Performance Enhance Firm Value? Evidence from Korea," *Sustainability*, 10(10), 3635. https://doi.org/10.3390/su10103635

Zerbib, O. D. (2019) "The Effect of Pro-Environmental Preferences on Bond Prices: Evidence from Green Bonds," *Journal of Banking & Finance*, 98, 39–60. https://doi.org/10.1016/j.jbankfin.2018.10.012

Zhou, X., & Cui, Y. (2019) "Green Bonds, Corporate Performance, and Corporate Social Responsibility," *Sustainability*, 11(23), 6881. https://doi.org/10.3390/su11236881

Appendix

Appendix 1. Balance test results of variables before and after radius matching (ROA)

Variable	Unmatched	Mean			%reduct	t-test	
	Matched	Treated	Control	%bias	|bias|	t	p>|t|
TA	U	6.1e+06	1.5e+06	46.0		15.66	0.000
	M	6.1e+06	5.3e+06	8.5	81.6	1.39	0.164
TL	U	5.1e+06	1.2e+06	41.7		14.37	0.000
	M	5.1e+06	4.3e+06	7.9	81.0	1.31	0.191
SIZE	U	14.412	12.234	123.7		28.73	0.000
	M	14.412	14.347	3.7	97.0	0.78	0.436
LEV	U	68.475	53.45	81.5		19.11	0.000
	M	68.475	68.264	1.1	98.6	0.22	0.825
CF	U	1964.9	2047.3	–7.1		–1.73	0.083
	M	1964.9	1955.9	0.8	89.1	0.15	0.883
CSR	U	4.1932	4.052	5.9		1.37	0.170
	M	4.1932	4.2836	–3.8	35.9	–0.75	0.453
PER	U	192.79	186.49	6.0		1.50	0.133
	M	192.79	196.28	–3.3	44.6	–0.62	0.535
SHARE	U	305.75	411.36	–58.7		–14.34	0.000
	M	305.75	316.64	–6.0	89.7	–1.20	0.229
AGE	U	66.999	69.176	–7.3		–1.88	0.061
	M	66.999	68.651	–5.6	24.1	–1.00	0.319
INDUSGR	U	6.1468	8.843	–8.7		–2.09	0.037
	M	6.1468	6.3327	–0.6	93.1	–0.13	0.900

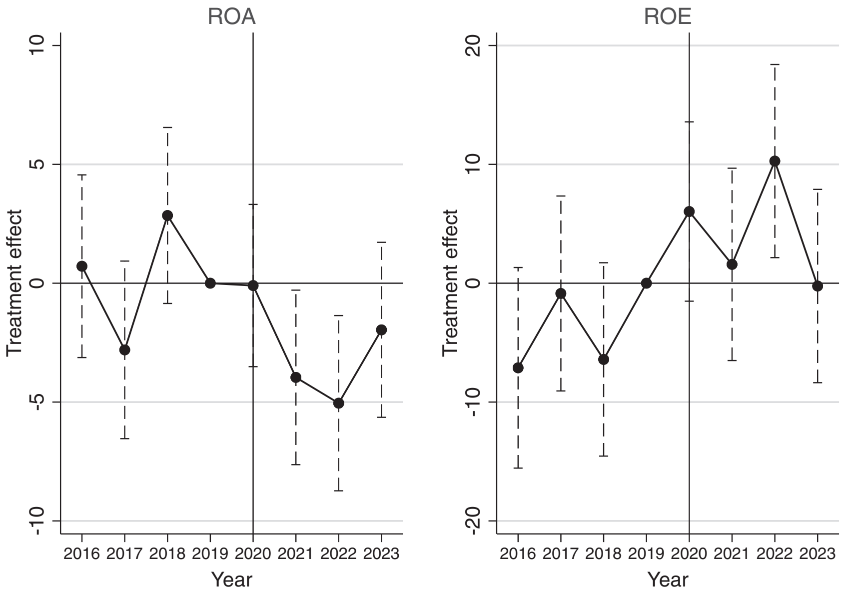
Notes: U represents unmatched, and M represents matched.

Appendix 2. Balance test results of variables before and after kernel matching (ROE)

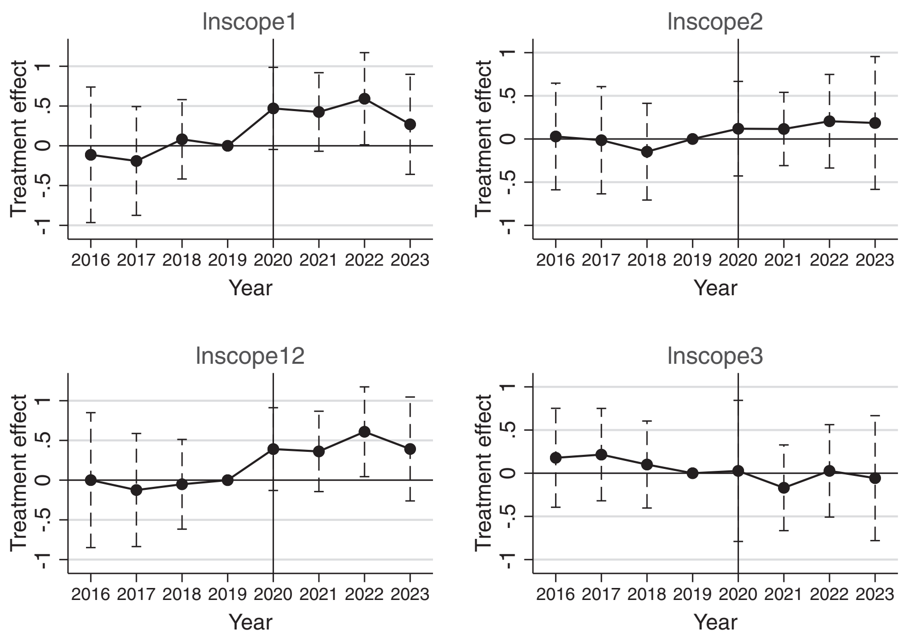
Variable	Unmatched	Mean			%reduct	t-test	
	Matched	Treated	Control	%bias	|bias|	t	p>|t|
TA	U	6.1e+06	1.5e+06	46.0		15.66	0.000
	M	6.1e+06	5.3e+06	8.4	81.8	1.37	0.170
TL	U	5.1e+06	1.2e+06	41.7		14.37	0.000
	M	5.1e+06	4.4e+06	7.8	81.2	1.29	0.197
SIZE	U	14.412	12.234	123.7		28.73	0.000
	M	14.412	14.355	3.3	97.3	0.69	0.487
LEV	U	68.475	53.45	81.5		19.11	0.000
	M	68.475	68.301	0.9	98.8	0.18	0.856
CF	U	1964.9	2047.3	–7.1		–1.73	0.083
	M	1964.9	1955.2	0.8	88.2	0.16	0.874
CSR	U	4.1932	4.052	5.9		1.37	0.170
	M	4.1932	4.2875	–3.9	33.2	–0.78	0.434
PER	U	192.79	186.49	6.0		1.50	0.133
	M	192.79	196.23	–3.3	45.4	–0.61	0.541
SHARE	U	305.75	411.36	–58.7		–14.34	0.000
	M	305.75	316.28	–5.8	90.0	–1.16	0.245
AGE	U	66.999	69.176	–7.3		–1.88	0.061
	M	66.999	68.632	–5.5	25.0	–0.99	0.325
INDUSGR	U	6.1468	8.843	–8.7		–2.09	0.037
	M	6.1468	6.3574	–0.7	92.2	–0.14	0.887

Notes: U represents unmatched, and M represents matched.

Appendix 3. Parallel trend test results for financial impacts (kernel PSM)



Appendix 4. Parallel trend test results for environmental impacts (kernel PSM)



産学間の共同研究における社会関係資本獲得を促進する方策の検討

REACH プロジェクト参加者への半構造化インタビューによる事例研究

平　丸　大　介[†]・開　本　浩　矢[‡]

要　旨

本研究は，産学間の共同研究における知識移転に重要な要素である社会関係資本の獲得を促進する方策を検討した事例研究である。対象は，大阪大学の博士後期課程に島津製作所の従業員を派遣し共同研究を行う REACH プロジェクトの参加者である。調査は，プロジェクト参加者に対する半構造化インタビューを実施した。分析の結果，研究室への貢献に「部外者」「博士学生」「専門家」の３段階があることを発見した。また，参加者の貢献の段階と出席頻度の間に同様の傾向が観察された。これらの要因として，研究室に頻繁に出席する従業員は，研究室内で対面コミュニケーションを繰り返し，研究室内での社会関係資本を獲得することで，研究室内から認められ貢献を果たせる状態にあると解釈した。本研究の貢献として，継続的な研究室への出席などでコミュニケーションの積み重ねを実現することで，社会関係資本の獲得を促進できることを示唆した。

JEL 分類：D83，O15，O32，O36

キーワード：社会関係資本，産学連携，共同研究，知識移転

1．はじめに

本研究の目的は，企業から派遣された従業員が大学の研究室内で取り組む共同研究において，より良い成果を得るために重要とされる社会関係資本が拡充されるプロセスの一端を明らかにし，その拡充を促進するための具体的な方策を検討することである。

現代社会は非常に複雑化し，変化が激しく不確実性が高い社会であり（Johansen & Euchner, 2013），このような社会環境において企業を継続的に成長させるには，企業が有する既存の知を活用するだけでなく，新たな市場や技術の探索・獲得を目指すイノベーション戦略が重要とされている（野中他，2010；窪田他，2022)。このような企業のイノベーション戦略を実行するための方法の一つとして，企業と大学の産学連携が注目され，その件数は 1990 年以降，大幅に増加している（Mirc et al., 2017；池川，2011)。

産学連携による協業には多様なプロセスが存在する（池川，2011；舟津，2022）が，産学連携の約８割は共同研究であり，企業側が産学連携に参加する動機の上位は，「事業上の重要な技術課題の解決」「大学との人的・組織的なネットワーク形成」，大学側の動機の上位は，「科学的発見，技術的知見の実用化による社会還

† 大阪大学大学院経済学研究科博士後期課程，大阪大学・島津分析イノベーション協働研究所招へい研究員，株式会社島津製作所

‡ 大阪大学大学院経済学研究科教授

元」という調査結果（長岡他，2013）から，企業と大学はいずれも共同研究を通じて知識や技術，ノウハウの伝達である知識移転を期待しており，企業側はそれに加えてアカデミアにおける人的ネットワークの形成を期待しているといえる。

企業がイノベーション創出のために期待する共同研究において，知識移転を促進するために重要な要素として共同研究に参加する関係者間での社会関係資本（Steinmo & Rasmussen, 2018; Al-Tabbaa & Ankrah, 2019）や，関係者間の信頼関係（池川，2011；長岡，2013；安田，2021；Wohlin et al., 2012）が指摘されている。これは，産学連携や共同研究は組織間の取り組みであるが，実際の研究は企業と大学のそれぞれ特定個人が参加して進められるものであり，その共同研究を通じて行われる知識移転も個人間で行われるためである。個人間における知識移転を通じて得られた様々な知見は，企業の参加者が組織内に持ち帰ることで組織のイノベーション創出に活用される。そのため，企業が共同研究に従事する個人の社会関係資本の獲得を支援することは，企業組織のイノベーション戦略上も有益であるといえる。

そこで，本研究では，知識移転に重要と考えられる共同研究に参加する個人の社会関係資本獲得を企業が促進する方策を検討する。一般に共同研究は投資的な側面があり，参加する企業は金銭的・人的な負担が発生するものの，必ずしも想定した成果を得られるとは限らない。そのような状況において，本研究により，企業が期待する知識移転を促進する方策を提示することができれば，今後の企業と大学間で実施する様々な共同研究の成果を底上げし，企業が共同研究という投資を積極的に推進することで，企業のイノベーション創出の加速に繋がると期待される。同時に，大学側は共同研究を通じて資金やニーズ・シーズ情報を取り入れることで，研究の活性化が期待される。このように，本研究の成果は企業と大学の双方へ貢献することが期待される。

本研究は事例研究であり，研究対象は大阪大学と株式会社島津製作所が推進するREACH（**RE**current & **RE**-skilling through **A**cademia and Industry **C**ollaboration for **H**igher Education） プロジェクト（以下，プロジェクト）（平丸他，2024）を通じて実施される共同研究である。このプロジェクトとは，産学連携と人材育成を目的とした取り組みであり，企業が有望視する研究領域における大学に所属するKOL（Key Opinion Leader：本研究では，特定の分野で高い専門知識と影響力を持つ研究者・専門家を指す。彼らは，専門分野における研究の方向性や政策決定に大きな影響を与えることができ，学術界や産業界において重要な役割を担う。）の元へ，従業員を博士後期課程の学生として派遣し，その研究室にて共同研究に取り組むという取り組みである。派遣された従業員は，一般的な社会人博士とは異なり，博士後期課程で取り組む共同研究が業務となるため，それぞれの研究室での研究活動に基本的にエフォート100％で取り組む。プロジェクトを通じて，企業は将来の事業展開に資する研究成果を獲得すると同時に，参加した従業員は，企業が有望視する研究領域の博士という専門家に成長することが期待される。本研究では，プロジェクトに参加する従業員へのインタビューを基に，共同研究を通じて社会関係資本が獲得されるプロセスを分析し，そのプロセスを促進する具体的な要因を検討する。

また，社会関係資本の定義は一義的ではなく多岐にわたる定義が存在する（佐藤, 2001）が，本研究では，Putnam et al.（1993）が提唱する「人々の協調行動を活発にすることによって社会の効率性を高めることのできる，『信頼』『規範』『ネットワーク』といった社会的仕組みの特徴」という定義を採用する。その理由は，知識移転に重要なものとして，共同研究に参加す

る個人間の信頼関係が指摘されており，共同研究の関係者による人的ネットワークの構築だけでなく，構築されたネットワークの中での，規範の習得や相互の信頼獲得が重要と考えたためである。

2. 先行研究

本節では，企業と大学の共同研究において，知識移転に有益な社会関係資本が獲得されるプロセスや社会関係資本の獲得を促進する方策について，先行研究での議論を示す。

Steinmo & Rasmussen（2018）は，ノルウェーの Business Innovation Program から 15 のプロジェクトに対してインタビューと公的な報告書，プレス記事などを集約して調査した。共同研究の経験が豊富な企業グループでは，研究者と目標や共通理解を共有することの重要性を意識していたが，その前提として，協力関係にある研究者とのネットワークや共通言語の獲得が完了していることを指摘し，経験が浅い企業グループでは，人間関係構築などの基本的な社会資本の獲得から必要であることが示された。Al-Tabbaa & Ankrah（2019）は，Faraday partnership initiative というイギリス政府主導の産学連携プログラムにおいて，24 件の実証研究に対してインタビューに基づき調査をした。異なる背景を持つ人々の交流には，最初から定期的なディスカッションの場や交流イベントを設定することがネットワーク構築で重要であることを示し，信頼関係は組織ではなく個人間で構築されることを指摘した。Thomas & Paul（2019）は，先行研究のレビューを通じて，大学と産業界の間の知識移転への社会関係資本の影響を理論的に検討した。社会関係資本の中で，参加者間の人間関係，信頼関係，目標の共有が重要であると述べ，参加者間の社会関係資本が知識移転に影響を与える際に，直接の影響以外に参加者間のコミュニケーションを媒介して影響を与えるというモデルを提案した。また，社会関係資本への言及はないが，知識移転を促進する具体的な方策について，酒匂・内平（2021）は，産学連携研究の事例をフレームワークによって分析する中で，知識移転に必要な負荷の解消に，企業の技術者が研究機関に籍を置くという方策が有用であることを指摘した。

これらの先行研究から，企業と大学間の共同研究において知識移転などの成果を拡大するには社会関係資本が重要であることが実証研究と理論研究の両面から示されており，ここでの社会関係資本は共同研究に参加する個人に蓄積するものであることが指摘された。そして，共同研究の成果である知識移転を促進するための具体的な方策として，研究機関に企業の人間が在籍してその場で研究に取り組む方法が示された。

しかし，既存の事例研究においては，調査対象の複数の共同研究は，その規模や参加企業にばらつきがある。そのため，獲得される社会関係資本について事例間での比較検討は困難であり，社会関係資本を獲得する要因の検討をする上で，獲得した社会関係資本の量的検討はなされていない。

本研究で対象とするプロジェクトの下で実施される共同研究は，いずれも企業から派遣された従業員 1 名が主体となり，3 年間で博士号取得に資する規模の研究を実施する取り組みであるため，各共同研究を一定程度比較することが可能と考えられ，獲得される社会関係資本の程度を検討できると考えられる。このような検討により，本研究は共同研究を通じて獲得される社会関係資本の獲得プロセスやその要因に関する新たな知見の提供を目指す。

3. 調査概要

(1) 調査対象

本研究では，プロジェクトに参加し，大学の各研究室に派遣された従業員 5 名に調査協力を

得た（以下，調査協力者）。詳細は表１に示す。

表 1. 調査協力者

	学年	研究テーマ概要
A 氏	修了	核酸医薬品分析
B 氏	D3	計測情報科学，AI
C 氏	D3	メタボロミクス分析
D 氏	D2	遺伝子治療薬分析
E 氏	D2	腸内細菌叢産生代謝物分析

筆者作成

共同研究に１年以上従事した従業員のみを選定した理由は，ある程度活動に慣れてきた状態にあると推測したためである。修了生については，2024 年時点で１名のみのため，A 氏のみとした。また，５名が派遣された各研究室は，一般的な理工系の研究室と同様に，特定の教授の下に集まった集団で構成されており，一つの場所に集まって特定分野の様々な研究を進める方式を取る。

(2) 調査方法

本研究は，５件の事例に対して，共同研究を通じて獲得された社会関係資本を検討することを目的とするため，インタビュー形式が適切であると判断した。インタビューの方式は，半構造化インタビューを採用することで，調査協力者の自由な語りを妨げないようにした。本研究の半構造化インタビューで準備したインタビューガイドについては表２に示すように７つの設問を用意した。

表 2. インタビューガイド

ガイド項目	詳　細
研究室での過ごし方	D1 と D2 以降の変化を含む
研究活動の進め方	テーマ決め，進捗会議等
研究活動に影響を与えた人物の有無	指導教官以外に，研究推進に寄与した人物
人間関係の拡大	研究室，アカデミア，社内のそれぞれに対して
研究室への貢献	研究活動以外の内容
会社への貢献	知名度向上，リクルート等
会社への違和感	社外活動による意識変化

筆者作成

各ガイド項目の狙いについては，以下に示す。派遣された研究室での生活の状況全般に関することとして，研究室での過ごし方を質問した。日常生活に加えて，活動の変遷を思い出しやすくするため，時系列に沿って D1 の研究開始直後から，D2 以降の現状について伺った。研究活動に関することでは，研究の進め方と，その研究を進める上で指導教官以外に影響を与えた人物の有無を質問した。これにより，研究を進める上で，研究室や企業の人間とどのように関わったのか，その中で特に影響を与える人物がいたのかを伺った。調査協力者からの貢献として研究室と企業の貢献をそれぞれ質問した。研究室への貢献は，互酬性を念頭におき，企業への貢献は企業へのメリットの有無を念頭に置いて伺った。また，活動場所が社外へ移ることから，企業に対する違和感についても質問した。これは，企業から長期間離れることで，企業での活動を客観的にとらえることが可能となるという想定で，企業へのメリットの一つと考えて伺った。

本研究のインタビュー調査は，第一筆者が担当し，調査協力者に対して調査同意書に基づき本研究の目的を文書と口頭で説明し，インタビューへの不参加・途中辞退が可能であること，調査協力を後日撤回できること，適切なデータ管理を約束することを説明し，倫理上の配慮を十分にした上で調査を実施した。インタビューの記録方法は音声録音とソフトウェアによる文字起こしを併用した。インタビューの所要時間はそれぞれ約１時間程度で，調査の実施時期は 2024/7/22 ～ 2024/7/30 の期間であった。

(3) 分析方法

本研究での分析は，全てのインタビューに対して，ソフトウェアで文字起こした文書を，録音データを基に修正したのちに，質的データ分析法（佐藤 2008）を参考に，コーディングにより内容の分析をした。第一次コーディングで

ラベル付けをし，第二次コーディングにて類似ラベルを集約した後に，概念生成などの検討を行った。具体的な分析作業には，MAXQDA Analytics Pro Edition (Ver.24.4.1) を用いた。

4. 分析結果

本節では，インタビューにより得られたデータをまとめ，概念抽出した結果を示す。まず，第二次コーディングにより得られたコード上位5件を表3に示す。合計で全コードの63.9％を示しており，5名のインタビューの主な要素をこれらのコードが示していると考えられる。

表 3. 調査協力者

順位	コード	出現数	割合［％］
1	人間関係	54	17.7
2	獲得したもの	47	15.7
3	研究進捗	35	11.7
4	会社貢献	34	11.4
5	研究室貢献	22	7.4
		合計	63.9

筆者作成

(1) 研究室貢献の段階と社会関係資本の程度差

研究室への貢献の度合いについて，調査した事例に共通する3段階の度合いが観察された。各段階の貢献は，調査協力者本人の意思だけではなく，研究室内の社会関係資本を一定以上獲得することが条件となることがインタビューを通じて見いだされた。

● 1段階目：部外者としての貢献

この段階は，調査協力者は，指導教官，一部の研究員（指導教官陣）と共同研究を推進する段階であり，5名全員に観察された。この段階では，研究室に対しては企業情報の提供などの貢献が見られた。共同研究の目的の共有や，議論に必要な共通言語の獲得などは達成され，指導教官陣との相互の信頼関係構築，規範の習得はされた状態であり共同研究の遂行は問題ないが，研究室内の人間の認識は，定期的に訪問する部外者だと考えられ，一般的な共同研究や社会人博士に類似した状態といえる。これらを示す発言の一部を示す。

・指導教官陣との限られたやり取り

(A氏)：一年生はコロナだったので基本リモートでしたね。なんで，その一年は本当に（研究室の）名前しか知らないし…。
(聞き手)：なるほど。でも，2021年はテーマとか色々頑張っていたじゃないですか？あれは基本的に先生とのやりとり？
(A氏)：<u>先生とのやり取りでした</u>。

(C氏)：一年目は…私あんまり研究室にはそんなには行っていなくて。…基本的には<u>指導教官とのコミュニケーションがメイン</u>なような形ですね。

(聞き手)：(研究指導は)指導教官と2人で，という感じですか？
(E氏)：あ，もう2人ですね。…きっちり<u>全部把握しているのは指導教官だけ</u>なので。

・部外者としての活動

(B氏)：<u>(学生へのアドバイス等は）してないですし</u>。世間話とか，そのまあ後は，<u>会社の紹介とかで多少興味持っていただいたりとかはあった</u>のですが，まあ結局別に。

(聞き手)：その時（=一年目）だと指導教官とはいろいろお話はしたけど，<u>他の方とはあんまりお話とか（は無かった）</u>？
(C氏)：まあそうですね，正直。

(D氏）最初はやっぱり話す人がいないのでなかなかまあ難しい。…やっぱり<u>出来上</u>

がった関係に入っていくのって難しいなと思っていた時期もありましたけど。

● **2段階目：博士学生としての貢献**

この段階は，調査協力者が学生の指導や他の学生や研究者のサポートをして研究室に貢献をする段階であり，4名に観察された。この段階では，指導教官を含めた研究室内の人間は，調査協力者の専門性を認識し適切な回答が得られることを期待する程度の信頼を得ていたと考えられ，研究室の博士後期課程の学生として認知されているといえる。同時に，調査協力者は周囲の研究内容や課題を把握することで，その研究領域が必要とする技術範囲やニーズを把握することができた。これらを示す発言の一部を示す。

・研究室からの認識

(A 氏)：クロマト※1系で困ったら連絡もらえるみたいな感じ。
(聞き手)：それは先生がクロマトで困ったら（A 氏に）聞いて，みたいな紹介をされていたのですか？
(A 氏)：あったとは思います。
※1：分析技術の一種類

(C 氏)：一年経つので何を聞けば，どう答えてくれる人かというか，どういうトラブルとか，どういうのが得意な人なのかっていうのは，（研究室に広まることが）あると思うので。

(聞き手)：CDMS※2中心にちょこちょこ質問はされるみたいな質問と分析相談？
(D 氏)：CDMS 系と物理系の話と島津の装置の話とか（の相談は），ちょくちょくある。
※2：D 氏が専門に扱う分析装置

(E 氏)：（分析相談は）めちゃめちゃされますね。めちゃめちゃ難しいことは別に聞かれないのですけど。

・研究領域の情報獲得

(A 氏)：いろんなトピックをやっているので。化学の合成化学もやっているし，バイオ系もやっているし，分析系もやっているので。いろんなノウハウというか。あ，こういうところに困るんだなみたいなのとかわかるし。

(C 氏)：自分は本当にガチガチの分析ですけれども，例えば食品関係のこととか，結構そういうところの，実際の研究のトレンドとか，研究の技術的なところっていうのは，実際に肌と肌で知るような言葉が増えたかなって，今ちょっと話していて思いました。

(D 氏)：研究室のイベントで結構いろんな方は来ますね。（分析機器メーカー複数社）のの方々とはお話はしましたけど。

(E 氏)：顧客としてっていうよりも，その中に入ることによって，こういろいろ見えてくる。例えば，中に入ることによって顧客の目線が見えてくるっていうとかもありますし。

・博士学生としての活動

(C 氏)：今年から私のテーマの延長で4回生の学生さんが卒論にしようかっていうことになったので。…研究のディレクションをしているような（指導をしている）。

(D 氏)：それで分析チームも（分析サンプルを）作るようになったほうがいいって話になって，私ともう一人のドクターの方が製造チームに教わって，それを分析チーム

どんどん教えていくっていう。

(E 氏)：今までは，その分析がメインだったので。最近になったらウェットの実験もがっつり，ドライの解析とかも入ってくるので。そういったところは，詳しそうな人にその都度聞くみたいな感じですね。…どっちかっていうと，島津の人ってよりも，完全にあのラボの中の人みたいな感じになっています。

● **3 段階目：専門家としての貢献**

この段階は，調査協力者が研究室に新たな考え方を導入したり，研究室を代表して研究室の外部に対する技術支援をしたりして研究室に貢献する段階であり，2 名に観察された。この段階では，研究室内で，調査協力者からの発信が十分に信頼され，受け入れられている状態であると考えられる。また，調査協力者を研究室の代表として外部に紹介している点も，調査協力者が十分な信頼を獲得している証左であり，調査協力者がある分野の専門家として研究室内で認知されていると考えられる。この時，調査協力者は研究室に影響力を発揮できる段階であり，企業の技術紹介等を通じて所属する研究室やその周辺に対して企業の知名度向上を実現することで，企業への貢献を実現したといえる。これらを示す発言の一部を示す。

・新たな考え方の導入

(D 氏)：先生は生物学のバックグラウンドがありますけど，私は物理学のバックグラウンドがある…。(課題に対して）生物学な思考というのか，物理学的な処理というのか…，どっちが良いかとかというのを話し合ったりとかで，それは私がどんどん決めていったりっていうふうにする（ようになっている)。

(学生からも）…物理系の話と島津の装置の話とか(相談が)，まあちょくちょくある。

(E 氏)：なんか測るってなったらバイオ系の人たちって液クロじゃなくて，ELISA が第一選択に来る。タンパクとかそういう系やったら仕方ないですけど，それ（= 液クロの方が得意な分析対象）も ELISA なんやみたいな。

(聞き手)：(E 氏が）新しい視点を導入していることになるのですかね，研究室に？

(E 氏)：そうですね。

・研究室外への技術貢献と企業の技術紹介

(E 氏)：先生のお友達の研究室で胆汁酸を測りたい，Nature に論文を出そうとしていて。で，ある胆汁酸がすごく大事みたいなったけど，結局測ってなくて困っていたところを，僕のところに話がきて測ってあげたっていう感じ。無事に（論文が）通ったみたいな。

・企業の知名度向上

(聞き手)：ちゃんと胆汁酸でデータが出てきて，(島津の）装置名も（論文に）出てくる？

(E 氏)：そうですね。

(E 氏)：先生自身も結構いろんなところで宣伝してくれているみたいで。学会で呼ばれて，こんなの（島津の分析器）使っています，とかなんかいろいろ。

・研究者としての活動

(D 氏)：最近サブサミットした論文に関しては，(指導教官が）こういう方向でいきましょうというふうに乗っかりましたね。で次（の論文を）やっていくのはかなり相互的になっているかと思います。

(E 氏)：(研究室の) MS※3 を頼りに来る人も多いので。教える時も僕が出ていくことはそんなに多くないのですけど。いつもやっている技術補助の人たちがメインで教えているのですが，なんかあったりしたら僕も入ったりしています。
(聞き手)：(E 氏が) 研究室で分析の一番上みたいな感じですか？
(E 氏)：それはそうですね。
※3：分析機器の一種類

(2) 研究室への貢献段階と出席頻度の関係

先に示した研究室への貢献段階と関係する調査協力者の具体的な行動として，研究室への出席頻度がみられる。基本的に毎日研究室へ出席する調査協力者の貢献段階は研究者の段階であり，週 1 での出席もしくはリモート参加が主体の調査協力者の貢献段階は部外者の段階であった。詳細は表 4 に示す。

貢献段階と出席頻度の時間的な前後関係として，1 年目はリモート主体で 2 年目以降に研究室へ出席するようになった 2 名の調査協力者からの話として，研究室に出席するようになった 2 年目以降に研究室との人間関係を構築し，研究室への貢献をするようになったとの説明があったため，出席頻度が時間的に先に発生し，貢献段階が変化するといえる。これを示す発言の一部を示す。

(A 氏)：(二年目に入ると) 研究室も対面が徐々に増えてきてみたいな感じですね。…ちゃんとそういうの (＝雑談など) ができてきたのが実質三年目ぐらい。

(C 氏)：二年目以降 (研究室に) 行くようになって，他の学生さんとか，あの他の先生方というところのコミュニケーションは実際まあ増えてきたっていうところあります。

表 4.　貢献度段階と研究室への出席状況

参加者	現時点の貢献段階	出席状況
A 氏	博士学生	1 年目はリモート主体で 2 年目以降は殆ど研究室に出席。一部実験を研究室外で実施。
B 氏	部外者	基本はリモートで，週 1 回程度研究室に出席。本年度はフルリモート。
C 氏	博士学生	1 年目はリモート主体で 2 年目以降は週 4 程度研究室に出席。
D 氏	研究者	ほぼ毎日研究室に出席。実験は研究室内で完結。
E 氏	研究者	ほぼ毎日研究室に出席。実験は研究室内で完結。

筆者作成

また，貢献段階が部外者段階の調査協力者について，共同研究へのモチベーション低下が原因で，出席頻度が低下し，同時に研究室への貢献が乏しいとみなすという考察も可能である。この調査協力者のモチベーションの程度をインタビューから類推したところ，学術論文は 1 本投稿済みで，2 本目を執筆中，国際学会にも報告済みであり，共同研究活動では順当な研究成果をあげていた。また，研究推進のために，企業内の人物と積極的にコミュニケーションを取る様子もうかがえたことから，共同研究へのモチベーション低下は認められないと考えられる。研究室への出席頻度が低い主な要因は，担当した研究の内容やその研究室の慣習にある。これらの検討から，研究室への出席頻度が高いほど，研究室への貢献がより高い段階を示す傾向があることが示唆された。

5. 考察

本節では，前節で示された分析結果に対する考察を行う。ここで主たる論点となるのは 2 点であり，1 点目は，研究室への貢献に段階がみられる点である。2 点目は，研究室への出席頻度と貢献段階の関係はどのような要因によるものかという点である。これらの論点について，

関連領域の知見も活用し議論を試みる。

(1) 研究室という集団への貢献段階

ここで議論する貢献とは学術的なものであり，単純な労務ではなく一定の能力が必要なため，その能力を有することを研究室の関係者が認めて，初めて貢献が可能な状態となると考えられる。これは，調査協力者が研究室という組織集団に参加し，その一員として認められていく組織社会化プロセス（高橋，1993）に該当すると考えられる。

組織社会化プロセスという視点で検討する場合，その集団で信頼を勝ち取り集団の規範を有することを周囲に認められるイベントはイニシエーション（Feldman，1977）と考えられる。理工系の研究室には，大別すると学生と研究員の二種類の立場の人間が所属しており，それぞれの立場に認められるためのイニシエーションを乗り越えた時に初めて，その立場としての貢献を研究室から期待され，それに調査協力者も答えるような振る舞いをすることになる。そのため，研究室への貢献の段階が3段階であることは，2度のイニシエーションを乗り越えると考えられる。

また，イニシエーションを乗り越えるために必要な要件は，人間関係や信頼，協調性などとされており，まさに社会関係資本の獲得である。つまり，研究室に入る人間が研究室から期待され，実施する研究室への貢献段階は，その研究室において獲得した社会関係資本の獲得の程度を評価する指標に使える可能性がある。また，今回の事例研究で対象とする研究室はいずれも大学の理工系の研究室であり，想定される学生像や研究者像はある程度共通していると考えられることから，研究室での貢献段階は，各研究室において獲得した社会関係資本の程度を比較する指標としても一定程度利用できると考えられる。

(2) 社会関係資本を介した研究室への出席頻度と研究室への貢献度段階の関係

研究室への出席頻度は，研究室にて所属する研究者や学生が調査協力者を見かけたり，両者が会話をしたりする機会に直結することから，単純接触効果（Saegert et al., 1976；富田・森川，2011）の視点に基づき考察する。稲木他（2017）は，PTA という集団において社会関係資本醸成に対する単純接触効果の影響を検討しており，単純に顔を合わせて会話するだけでは不十分であり，会う回数や時間の蓄積を経てコミュニケーションを積み重ねることにより，知り合いになること，悩みに関する相談が自然発生することが社会関係資本の獲得に繋がると指摘している。この指摘と同様のことがインタビューの中でも言及され，研究室に出席することで研究者や学生と顔を合わせ，自分の研究活動に関すること以外にも，周辺技術に関することや研究相談，時には雑談をしてコミュニケーションを積み重ねていたことが観察された。これを示す発言の一部を示す。

(A 氏)：顔と名前ぐらいは多分覚えているけど，具体的にラフな話をするかって言ったら，やっぱしなかったので。<u>ちゃんとそういうのができてきたのが実質三年目</u>ぐらい。

(C 氏)：<u>二年目以降…</u>，例えば学生さんとか，あの他で研究員さんがやっている共同研究の内容とか，研究室全体の研究のアクティビティに，自分が手を動かして実験するわけではないですけれども，<u>そういったアクティビティに結構入るようになって</u>。

(D 氏)：最初は誰が何に詳しいのか知らないので，（指導教官に）ご紹介いただいてっていうパターンが多かったですけど，<u>二年目ぐらいからは私も人間関係がかなり構</u>

> *築できてきたので，誰が何に詳しいか理解しているので，私の方から直接お声かけすることをしていますね。*

これにより，互いに信頼関係の構築や規範の習得が進み社会関係資本を獲得することに繋がると考えられる。対して，リモートでの研究活動では，自らの研究活動に関する話題のみを指導教官等の限られた人間と議論をするため，研究室での社会関係資本の獲得が進みにくいと考えられる。

上記の考察から，研究室への出席頻度が高い場合に，研究室内でのコミュニケーションを積み重ねて社会関係資本の獲得が進むことで，前述のイニシエーションを超えることが可能となり, 研究室に対する貢献に繋がると考えられる。

6. 終わりに

以上の議論により，本研究ではプロジェクトに参加し共同研究に携わる５名を対象に事例研究を行うことで，研究室に加わる人間による研究室への貢献が３段階に分類して説明できること，その貢献の段階と研究室への出席頻度は，研究室で獲得される社会関係を介して関係する示唆を得た。これにより，プロジェクトでの共同研究において社会関係資本が拡充されるプロセスとして，単純接触効果の視点に着目し，年単位で研究室に出席することで，研究室内の人間と多くのコミュニケーションを積み重ね，社会関係資本を蓄積するプロセスを提示した。このプロセスにより，一般的な共同研究や社会人博士での研究で期待されるよりも，より多くの社会関係資本を獲得できる可能性を示すことができた。

本研究で得られた知見として，共同研究を通じて社会関係資本をより多く獲得するためには，共同研究先とコミュニケーションが積み重ねること重要である可能性を示し，コミュニケーションの積み重ねを実現する具体的な方策の一つとして，共同研究先と対面での接触を拡大させるという方策が有用である可能性を示した。ただし，コミュニケーションを積み重ねる方策は多数の蓄積があり，それらの知見を活用することも期待される。また，共同研究におけるコミュニケーションは社会関係資本と知識移転の媒介変数としての機能も有する（Thomas & Paul, 2019）ことから，コミュニケーションを積み重ねることは，知識移転をより促進し共同研究の成果の拡大に貢献できると考えられる。加えて，共同研究参加者と先方研究室との間での社会関係資本が，研究室への貢献の第３段階程度まで蓄積された場合には，共同研究参加者を通じて，研究室や関連する分野の関係者に対する企業の知名度向上や，企業への信頼獲得など，その研究分野の集団と企業の間において社会関係資本のようなものを蓄積できるという，通常の共同研究では得難いメリットを享受できる可能性も示唆された。

本研究の研究対象としたプロジェクトは，研究室への出席頻度を高めやすい制度設計がされており，より多くの知識移転を実現しイノベーション創出に貢献することが期待されるプロジェクトと考えられる。

最後に, 本研究の問題点と限界について示す。まず，調査協力者のプロジェクト参加時点での専門性が派遣先の研究室に対して有用かどうかで，研究室への貢献のしやすさがばらつく可能性がある。研究室に役立つ専門性を有する人間の方が，より早期に信頼を獲得することは考えられる。

また，本研究では，企業から派遣された従業員からのインタビューのみで社会関係資本の検討をしており，獲得された社会関係資本については，調査協力者の主観によるものである。より客観的な評価を行うには，派遣を受け入れた指導教官や研究室関係者側からの調査も必要と考えられる。

加えて，本研究では，社会関係資本の程度問題を議論する上で，計測手法を採用していない。定量評価を採用できれば，研究室の貢献段階の議論をより精緻に実施できた可能性がある。今後，本研究で対象とするプロジェクトの参加者が一定数に達した段階では定量分析を行っていきたい。

最後に，第一著者自身がプロジェクトへの参加者であることから，読者が想定する客観的なインタビューがなされていない可能性がある。これについては，細心の注意を払い，客観的な分析に努めることで対応した。

参考文献

Al-Tabbaa, O., & Ankrah, S. (2019). "'Engineered' University-Industry Collaboration: A Social Capital Perspective." *European Management Review*, 16, 543–565.

Feldman, D. C. (1977). The role of initiation activities in socialization. *Human Relations*, 30 (11), 977–990.

舟津昌平（2022）.「産学連携の組織・個人・社会（性）」.『組織科学』，第 56 巻，第 4 号，50－66 頁 .

平丸大介・中西博昭・飯田順子（2024）.「産学連携を生かし戦略的な高度人材育成を目指す"REACH プロジェクト".」『生産と技術』，第 76 巻，第 3 号，62－66 頁 .

池川隆志（2011）.「オープンイノベーション時代における産学連携.」『電子情報通信学会誌』，第 94 巻，第 7 号，573－578 頁.

稲木隆一・上田菜央・扇原淳（2017）.「PTA 活動が保護者のソーシャル・キャピタル醸成に及ぼす影響：複線径路・等至性モデルによる分析.」『家庭教育研究』，第 22 巻，51－61 頁.

Johansen, B., & Euchner, J. (2013). "Navigating the VUCA World." *Research-Technology Management*, 56(1), 10–15.

窪田 祐一・劉 美玲・三矢 裕（2022）.「イノベーション戦略とマネジメント・コントロールの有効性―両利き経営のための示唆―.」『日本管理会計学会』，第 30 巻，第 1 号，3－20 頁.

Mirc, N., Rouzies, A., and Teerikangas, S. (2017). "Do Academics Actually Collaborate in the Study of Interdisciplinary Phenomena? A Look at Half a Century of Research on Mergers and Acquisitions." *European Management Review*, 14, 333–357.

長岡貞男・細野光章・赤池伸一・西村淳一（2013）.「産学連携による知識創出とイノベーションの研究―産学の共同発明者への大規模調査からの基礎的知見―.」『科学技術政策研究所』，第 221 号．https://nistep.repo.nii.ac.jp/records/4745

野中郁次郎・遠山亮子・平田透（2010）.「流れを経営する持続的イノベーション企業の動態理論.」『東洋経済新報社.』

Putnam, R. D., Leonardi, R., Nanetti, R., & Dawsonera. (1993). "Making democracy work: civic traditions in modern Italy." *Princeton University Press*.

酒匂孝之・内平直志（2021）.「概念実証の観点から見た研究成果事業化のための知識共有の分析.」『産学連携学』，第 17 巻，第 2 号，91－101 頁.

佐藤寛（2001）.「援助と社会関係資本―ソーシャルキャピタル論の可能性―.」『経済協力シリーズ』，第 194 号日本貿易振興会アジア経済研究所.

佐藤郁哉（2008）.『質的データ分析法 .』新曜社.

Saegert, S., Swap, W., & Zajonc, R. B. (1973). "Exposure, context, and interpersonal attraction." *Journal of Personality and Social Psychology*, 25(2), 234–242.

Steinmo, M., & Rasmussen, E. (2018). "The interplay of cognitive and relational social capital dimensions in university-industry collaboration: Overcoming the experience barrier." *Research Policy*, 47(10), 1964–1974.

高橋弘司（1993）.「組織社会化研究をめぐる諸問題.」『経営行動科学』，第 8 巻，第 1 号，1－22 頁.

Thomas, A., & Paul, J. (2019). "Knowledge transfer and innovation through university-industry partnership: an integrated theoretical view." *Knowledge Management Research & Practice*, 17, 436–448.

富田瑛智・森川和則（2011）.「単純接触効果研究の動向と展望 .」『大阪大学大学院人間科学研究科紀要』，第 37 巻，361－373 頁 .

Wohlin, C., Aurum, A., Angelis, L., Phillips, L., Dittrich, Y., Gorschek, T., Grahn, H., Henningsson, K., Kagstrom, S., Low, G., Rovegard, P., Tomaszewski, P., van Toorn, C., & Winter, J. (2012). "The success factors powering industry-academia collaboration." *IEEE Software*, 29(2), 67–73.

安田聡子（2021）.「産学連携の全体像の探究：公式および非公式経路から成る知識移転スペクトラム.」『研究 技術 計画』，第 36 巻，第 3 号，290－307 頁.

Study of methods to facilitate social capital acquisition in academia-industry joint research: A case study based on semi-structured interviews with REACH project participants

Daisuke Hiramaru and Hiroya Hirakimoto

This case study examines measures to promote the acquisition of social relational capital, an important element for knowledge transfer in joint research between industry and academia. The subjects were participants in the REACH project, in which Shimadzu employees were dispatched to The University of Osaka doctoral programs to carry out joint research. Semi-structured interviews were conducted with the project participants. As a result of the analysis, it was observed that there were three stages of contribution to the laboratory: ‘outsider’, ‘PhD student’ and ‘expert’. A similar trend was also observed between the stage of the participants’ contribution and the frequency of their attendance. As a result of these factors, it was considered that employees who frequently attend the laboratory are in a position to be recognized and contribute within the laboratory through face-to-face communication and by gaining social capital within the laboratory. As a practical implication of this study, it was suggested that the acquisition of social relational capital could be promoted by the accumulation of communication through continuous attendance at the laboratory.

JEL Classification: D83, O15, O32, O36

Keywords: Social capital, Academia-Industry collaboration, Joint research, Knowledge transfer

Earnings prediction using machine learning: A survey*

Yuanchao Peng†

Abstract

This survey investigates the application of machine learning (ML) techniques in predicting corporate earnings. By reviewing literature spanning 2019 to 2024, this paper aims to provide a comprehensive overview of the methodological trends, strengths, and limitations of current ML approaches in the context of earnings prediction. While most research focuses on U.S. firms, a smaller portion examines international firms, including one study on Japan. A key trend is the preference for predicting directional changes in earnings (binary classification) over actual earnings levels, as classification models leverage high-dimensional data more effectively and yield economically meaningful insights. For example, portfolios based on predicted earnings changes outperform traditional models in generating abnormal returns. This paper also points out the shortcomings of existing research, 1) there is a lack of sufficient international evidence to prove that ML is superior to traditional models, 2) most earnings forecasts are short-term, and 3) there is a lack of exploration of non-financial data or forward-looking data. These shortcomings point to promising directions for future research. Another notable trend is that large language models (LLMs) have been shown to outperform traditional methods and human analysts in predicting the direction of corporate earnings. This emerging approach demonstrates impressive predictive performance in analyzing financial ratios and trends without extensive retraining.

JEL Classification: C53, G17, M41

Keywords: Earnings prediction, Machine learning, Large language models, Classification

1. Introduction

The main purpose of this study is to investigate recent papers that apply machine learning (ML) techniques to forecast corporate earnings. This analysis covers content from leading accounting journals and working papers, most of which focus on U.S. firms. In addition, one paper (Chattopadhyay et al. 2022) analyzes both U.S. and international firms, and one paper (Yakabi et al. 2024) focuses specifically on Japanese firms. This review aims to identify general research and methodological

* I would like to express my heartfelt gratitude to my supervisor, Professor Atsushi Shiiba, at the Graduate School of Economics, Osaka University, for his valuable guidance and support throughout the writing of this paper.

† Graduate Student, Graduate School of Economics, Osaka University

trends, analyze the strengths and weaknesses of existing methods, and propose potential future research directions for the application of ML in earnings prediction.

A major finding of this study is that most papers published in leading accounting journals use ML to predict directional changes in earnings rather than predicting specific level of earnings (Chen et al. 2022; Hunt et al. 2022; Jones et al. 2023). This tendency towards classification reflects the advantages of ML algorithms in extracting information from high-dimensional data sets and achieving higher accuracy than traditional regression models. At the same time, these studies also emphasize that predicting binary outcomes is more economically meaningful than predicting actual values. Hedge portfolios based on predicted results of increasing or decreasing earnings can achieve higher abnormal returns than traditional models.

This paper also highlights the concentration of research on U.S. firms, while international evidence supporting the superiority of ML over traditional methods is relatively limited. In addition, most studies rely on historical financial statement data as input variables, while paying little attention to non-financial or forward-looking data sources, such as textual information, macroeconomic indicators, or management's perceptions on future risks and uncertainties. Therefore, integrating these alternative data types into existing models provides a promising direction for future research. I also noticed that most earnings forecasts are conducted for the short term, specifically one year ahead. Forecasting earnings in the long term (3-5 years) is still an underexplored area.

A notable emerging trend is the use of Large Language Models (LLMs), such as GPT–4, in financial analysis. A recent study by Kim et al. (2024) demonstrate that LLMs, when applied with structured and anonymized financial data, can outperform human analysts in predicting the direction of future earnings. LLMs complement both human analysts and traditional ML models, excelling in scenarios where analysts are prone to bias or disagreement. They also rival advanced ML techniques, such as artificial neural networks, in certain predictive contexts. Moreover, LLMs exhibit unique capabilities in interpreting trends and financial ratios, offering state-of-the-art performance without specialized training. This highlights their potential not only as supportive tools but as central elements in financial decision-making. The inclusion of LLMs marks a significant shift in earnings prediction research, showcasing their ability to democratize financial analysis and opening new pathways for integrating AI-driven methods into finance. These developments suggest a promising and optimistic direction for future research, which will be further discussed in Section 6.

At the end of Section 3, I provide a structured overview of recent advancements in using ML to predict future earnings in Table 3. It highlights both the diversity and evolution of methods and findings in this field. Studies from leading journals, such as *The Accounting Review* and *Journal of Accounting Research*, along with working papers, collectively demonstrate the growing preference for ML over traditional methods like logistic regression and analysts' forecasts. Decision Tree-based methods, particularly Random Forest and Gradient Boosting, emerge as high performers across various studies, often yielding superior predictive accuracy and economic benefits, such as enhanced portfolio returns. Evaluation metrics range from statistical measures (e.g., MAE, RMSE) to economic outcomes (abnormal returns), reflecting the dual focus on accuracy and practical outcome. Notably, while most studies affirm ML's advantages, exceptions like Campbell et al. (2023) underscore its limitations in specific contexts. Overall, the research captures the field's progress, showing both the promise of innovative techniques like LightGBM and LLMs and the ongoing challenges in consistently outperforming traditional methods. This survey aims to synthesize these findings, offering a comprehensive analysis of ML's role in advancing earnings prediction.

The rest of this paper is organized as follow. In Section 2, I review papers that apply ML to predict earnings directional changes. Those papers of using ML to predict level of earnings are discussed in Section 3. At the end of Section 3, I provide an overview of studies investigated in this paper. From Section 4 to Section 5, I discuss the potential challenges and opportunities of using ML to forecast future earnings. In Section 6, I present an emerging trend of using LLMs to predict earnings. Section 7 concludes and provide future research path on this topic.

2. Predicting Changes of Earnings Ratios: Classification

In this section, I will first summarize the ML methods used in this field, as well as the strengths and shortages of each method. Then I will summarize the conclusions of the main papers in this field.

2.1 Main Algorithms and Methodology

Among the classification algorithms of ML, decision tree-based algorithms have been widely used in recent years. The most popular algorithms are Random Forest and Gradient Boosting Machine. Both algorithms belong to ensemble algorithms, but they have significant differences in the construction of decision trees and the aggregation of final results. Therefore, many studies use these two algorithms at the same time to test the credibility of the results. Random forest builds multiple decision trees on different bootstrap samples of training data and randomly assigns predictor variables to each decision tree. The number of decision trees in the model and the number of predictor variables assigned to each decision tree are the two most critical parameters. The determination of these parameters usually requires a series of tests to find the best parameter combination suitable for a specific data set. This process is called parameter tuning. Random forest reduces overfitting of data by averaging the prediction results of these different decision trees and further reduces variance by reducing the influence of the main variables. The final prediction is a simple average of the predictions from each of the individual trees.

The Gradient Boosting algorithm builds decision trees sequentially, and each subsequent tree focuses on correcting the mistakes made by the previous tree. Therefore, there is a dependency between each decision tree in the Gradient Boosting algorithm, which is the biggest difference from the Random Forest. Specifically, the Gradient Boosting algorithm initially starts with a weak learner (decision tree) and iterates for each sample, calculating the residual between the predicted value of the current model and the true value. The residual represents the part that the current model failed to predict correctly. Therefore, these residuals will serve as the training target for the next decision tree. This process is iterated until the performance of the model no longer improves. In the process of training the model, the hyperparameter-learning rate is usually adjusted, which controls the degree of influence of each decision tree on the final model. A smaller learning rate makes the model more robust, but converges more slowly. The final prediction result of the Gradient Boosting model is the weighted sum of the prediction results of all decision trees. The weight is usually related to the performance of the decision tree, and the decision tree with better performance will get a larger weight. In order to prevent overfitting, the Gradient Boosting algorithm usually introduces regularization terms, such as limiting the depth of the decision tree or the number of nodes per leaf.

In recent years, many studies have used optimized Gradient Boosting algorithms to improve data processing efficiency or reduce memory usage. For example, LightGBM optimizes data processing speed and model memory consumption. XGBoost significantly improves data processing speed without reducing the reliability of the model.

I summarize the advantages and disadvantages of these decision-tree based algorithms in Table 1. From Table 1, we can see that 7 papers use the Random Forest algorithm, far more than other algorithms. A main reason is that the Random Forest effectively reduces the risk of overfitting while maintaining a high model robustness. For the same reason, 4 papers use the StochasticGBM algorithm. Table 1 also includes paper indices that indicate which studies employed these methods. For details on the papers referred to by the indices, please refer to Table 3.

Table 1: Comparison of Decision-Tree based Algorithms

Feature	LightGBM	StochasticGBM	XGBoost	GBRT	Random Forest
Data Sampling	GOSS, EFB	Random Subsampling	Full Dataset	Full Dataset	Random Subsampling
Feature Selection	Random Splits	Pre-processing	Pre-processing	Pre-processing	Random Splits
Tree Growth	Leaf-wise	Level-wise	Level-wise	Level-wise	Level-wise
Regularization	L1		L1 and L2	L1 (Optional)	
Speed	Very Fast	Moderate	Fast	Moderate	Moderate
Overfitting Risk	Higher	Lower	Moderate	Higher	Lower
Robustness	Moderate	High	High	Moderate	High
Memory Usage	Low	Moderate	Moderate to High	Moderate	Moderate to High
Interpretability	Moderate	Low	Moderate	Moderate	Moderate
Paper Index	F	C, E, I, K	D	B	A, B, C, E, H, J, K

Notes: Gradient-based One-Side Sampling (GOSS) is a technique used to speed up gradient boosting algorithms. GOSS prioritizes data points with large gradients while randomly sampling from those with smaller gradients. This selective focus ensures the algorithm learns more effectively from challenging cases while reducing the computational burden. Exclusive Feature Bundling (EFB) is a method for reducing the dimensionality of datasets with a large number of sparse features. EFB groups features that are mutually exclusive into a single "bundle." This bundling reduces the number of features without losing significant information, making the algorithm faster and more efficient while preserving predictive accuracy. For details on the papers referred to by the indices, please refer to Table 3.

2.2 Main Results of Related Papers

Many studies use ML methods to predict the direction of earnings changes. One of the reasons is that a large number of studies construct hedge portfolios based on the predicted direction of earnings changes.

Chen et al. (2022) use Random Forest and Stochastic Gradient Boosting to forecast the direction of one-year-ahead earnings change. They obtain detailed financial ratios from XBRL filings and apply a large set of variables including 4,000 distinct financial items with their current, lagged and percentage changes value. They solve the class imbalance problem (earnings increase sample outnumber decrease sample) by adjusting earnings changes for the average change in EPS (Earnings Per Share) over the past four years. They obtain 3,610 earnings increase samples and 4,539 earnings decrease samples during year 2012 to 2018. Instead of using standard cross-validation, the study uses a rolling sample splitting approach that training and validation samples are gradually shift forward in time. This approach ensures that predictions rely only on the most recent data without backward-looking biases. Their models achieved an area under the curve[1] (AUC) between 67.52% and 68.66%, significantly outperforming random walk model (50%). The annual size-adjusted returns to hedge portfolios formed based on predictions range from 5.02% to 9.74%. The superior performance compared to traditional logistic regression models and analyst forecasts is attributed to both the nonlinear interactions captured by ML and the use of more detailed financial data. These findings underscore the value of ensemble learning and detailed financial data for binary earnings change predictions.

Jones et al. (2023) uses Gradient Boosting Machine to predict next period change in profitability based on a model proposed by Penman and Zhang (2004). Changes in profitability is defined as the difference between return on net operating assets (RNOA) at year t+1 with RNOA at year t. To avoid look-ahead bias, the dataset is divided into seven distinct training and test periods, ensuring that no future data from the test samples influences the training process. They find that Gradient Boosting Machine and Random Forests, consistently outperformed traditional models across various metrics (R^2, MAE[2], RMSE[3]). They identified both asset turnover and profit margin (components of the DuPont decomposition) as strong predictors, contradicting to the results of prior research. The study also found that the PZ model's key variables (e.g., growth in net operating assets and RNOA) remained robust predictors even in high-dimensional settings. The research suggests that while ML models enhance interpretability and accuracy through nonlinear interactions and high-order effects, they may not always translate to superior economic returns in portfolio applications compared to traditional regression models. Future research is encouraged to explore when ML's predictive gains lead to economic benefits.

---

[1] The Area Under the Curve (AUC) is a metric used to evaluate the performance of binary classification models. It is derived from the Receiver Operating Characteristic (ROC) curve, which plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at various classification thresholds. AUC ranges from 0 to 1. The closer the AUC is to 1, the better the model's ability to separate positive and negative classes.

[2] MAE represents the average absolute difference between predicted and actual values.

[3] RMSE represents the square root of the average squared differences between predicted and actual values.

Hunt et al. (2022) evaluate ML's effectiveness in predicting the direction of earnings changes and its utility in returns prediction. They found that while Elastic Net Regression did not outperform traditional Stepwise Logit models, Random Forest significantly improved out-of-sample prediction accuracy across all subsamples and time periods. Additionally, trading strategies based on Random Forest predictions yield higher abnormal returns than those based on other models, suggesting a practical advantage in financial contexts. The study advocates for exploring other nonparametric methods (e.g., Neural Networks, Support Vector Machines[4]) as they may offer even better performance. Random Forest's flexibility and ability to handle raw data without preprocessing (like standardization) highlight its utility for practical applications in predicting binary earnings outcomes.

I also surveyed two working papers on this topic. Cui et al. (2020) evaluated the application of LightGBM combining the dimension reduction technique-Principal Component Analysis (PCA) for forecasting directional changes of earnings. Their study compared the model's performance against analysts' consensus estimates and traditional logistic regression models. While the proposed model outperformed logistic regression in prediction accuracy and computational speed, it fell short of matching the performance of analysts, who benefit from broader information, including qualitative and potentially insider insights that are difficult to quantify. The study highlights the limitations of relying solely on structured data from public databases like Compustat and Thomson Reuters but emphasizes the potential for improvement. The authors suggest that incorporating non-quantitative data through advanced techniques such as Natural Language Processing (NLP) could enable the model to extract valuable insights from market news and textual disclosures.

Anand et al. (2019) investigate the effectiveness of classification trees in generating out-of-sample profitability forecasts. Using data from U.S. firms (1963-2017), the study evaluates directional changes in five profitability measures: return on equity (ROE), return on assets (ROA), return on net operating assets (RNOA), cash flow from operations (CFO), and free cash flow (FCF). The ML method achieves classification accuracies between 57% and 64%-significantly better than the random walk's 50%. Notably, its performance remains stable over a five-year forecast horizon. The study finds higher classification accuracy for cash flow measures (CFO, FCF), especially when accruals are included, compared to earnings-based measures (ROE, ROA, RNOA). However, in extreme portfolios of conditioning variables, earnings-based measures often outperform cash flow measures, indicating that no single profitability metric is superior under all conditions.

Although, most of the studies are using samples from the U.S., there is one study provides international evidence from Japan. Yakabi et al. (2024) examines the predictability of the direction of future earnings changes using ML techniques applied to Japanese companies' financial data based on the methodology of Chen et al. (2022). They find that Random Forest and Gradient Boosting outperformed Logistic Regression in terms of prediction accuracy and portfolio return. The abnormal

[4] Support Vector Machines (SVM) is a supervised machine learning algorithm used for classification and regression tasks. The primary goal of SVM is to find the best decision boundary (or hyperplane) that separates data points of different classes in a feature space. For specific explanation, please refer to: https://link.springer.com/chapter/10.1007/978-1-4899-7641-3_9.

returns generated by portfolios based on ML model predictions are statistically significant, indicating that the market does not fully incorporate information available in the financial statements. This finding challenges the efficient market hypothesis. They use 62 financial indicators as features, derived from Japanese companies' financial statements. Predictive performance is evaluated using the area under the ROC curve (AUC) and abnormal returns from hedge portfolios constructed based on the predictions. They also conduct a preliminary analysis using a Large Language Model (LLM), specifically GPT–4, to assess its potential in predicting earnings changes. The LLM (GPT–4) showed mixed results. While achieving a lower AUC compared to other models, it generated the highest abnormal return (AR). This suggests the LLM might provide valuable insights by incorporating qualitative factors alongside quantitative data, though further research is needed to confirm its reliability.

3. Predicting Level of Earnings Ratios: Regression

3.1 Main Algorithms and Methodology

The OLS, LASSO, RIDGE are the most popular algorithms in this task. The OLS model aims to estimate parameters by minimizing the sum of squared differences between observed and predicted values:

$$\beta^{OLS} = \arg\min_{\beta} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 .$$

When the number of parameters increases, OLS is prone to overfitting the model in-sample, leading to poor predictive performance out-of-sample (Chattopadhyay et al. 2022). To address this problem, penalized models, also referred to as "Shrinkage" methods, are designed to give the highest weights to a subset of predictors that demonstrate the strongest predictive power. RIDGE minimizes the sum of squared deviations while adding a penalty proportional to the square of the coefficients' magnitudes.

$$\beta^{RIDGE} = \arg\min_{\beta} \left\{ \sum_{i=1}^{N} \left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \right\}.$$

LASSO also minimizes the sum of squared deviations but adds a penalty proportional to the absolute values of the coefficients.

$$\beta^{LASSO} = \arg\min_{\beta} \left\{ \sum_{i=1}^{N} \left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\}.$$

Elastic Net combines the penalty of LASSO (L1) and RIDGE (L2) to enable it to handle high-dimensional data. At the same time, as the complexity of the model increases, the interpretability of the results decreases. From Table 2, we can see that there is no preference of specific model among papers investigated here. In general, there is no best model that is suitable for all situations. We need to select the most proper algorithm according to different data and purposes.

K-Nearest Neighbors (KNN) is a simple yet powerful supervised learning algorithm used for classification and regression tasks. For a given input data point, KNN calculates its distance to all the data points in the training dataset. Then, it identifies the *k* closest data points (neighbors) based on a distance metric (e.g., Euclidean, Manhattan, or Minkowski distance). Lastly, KNN predicts the output

Table 2: Comparison of OLS-based Algorithms

Feature	OLS	LASSO	RIDGE	Elastic Net
Penalty	None	L1	L2	Combination of L1 and L2
Feature Selection	No	Yes	No	Yes
Multicollinearity	Sensitive	Robust	Robust	Robust
Interpretability	High	High	High	Moderate
Flexibility	Low	Moderate	Moderate	High
Paper Index	B, C, I	B, C, D	B, C, D	B, H

Notes: For details on the papers referred to by the indices, please refer to Table 3.

by averaging the values of the k neighbors. One of the key steps in KNN is choosing the proper value of k, which determines the number of neighbors considered for making predictions. A small k can make the algorithm sensitive to noise, while a large k can dilute the influence of nearby neighbors, making predictions less specific. In recent study, Easton et al. (2024) introduced this method into earnings prediction task.

Artificial Neural Networks (ANNs) is another powerful ML methods that can handle high-dimensional dataset and complex relationships between features. However, ANNs are often criticized for being difficult to interpret compared to simpler models like linear regression or decision trees. Despite the fact that there is only one study used this method in earnings prediction, I will still briefly introduce how ANNs works. ANNs typically consist of three types of layers: 1) the input layer, 2) hidden layers, and 3) the output layer. Information flows from the input layer through the hidden layers to the output layer. Each neuron in the layers computes a weighted sum of its inputs:

$$z = \sum_{i=1}^{n} w_i x_i + b,$$

where w_i is the weight of i-th input, x_i is the value of i-th input, and b is the bias term. The weighted sum z is passed through an activation function to determine the neuron's output:

$$a = \mathrm{f}(z).$$

The activation functions include Sigmoid: $\sigma(z) = \frac{1}{1+e^{-z}}$, ReLU: $\mathrm{ReLU}(z) = \max(0, z)$, and Tanh: $\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$. Activation functions impact how gradients are calculated and propagated. Each activation function has its advantages and disadvantages. For example, Sigmoid and Tanh may cause gradients to shrink (vanishing gradient problem), and ReLU ensures gradients flow effectively for positive z, improving training in deep networks. The output a becomes the input for the next layer or the final prediction. a is compared to the true value using a loss function, which quantifies prediction errors. The algorithm calculates the gradient of the loss with respect to the weights and biases using the chain rule. Gradients are propagated backward through the network to update weights and biases. Weights and biases are adjusted to minimize the loss:

$$w_i = w_i - \eta \frac{\partial \text{Loss}}{\partial w_i},$$

where η is the learning rate, controlling the step size of updates. Lastly, the above steps are repeated across multiple iterations until the loss converges to a satisfactory level.

3.2 Main Results of Related Papers

I find that relatively few studies use ML to predict specific level of earnings. Even though it is more challenging to provide accurate prediction of exact level of earnings than to predict the direction of earnings changes. However, this does not mean that predicting level of earnings is meaningless or impossible. Several studies has provide insights on this topic.

Easton et al. (2024) utilize KNN approach to predicting one-year-ahead earnings. They developed a simple and effective method to predict the future earnings of a target firm by identifying the firms with similar history earnings. For instance, the Euclidean distance between firm *i*'s M-year earnings history ending in year *t* and firm *j*'s M-year earnings history ending in year *s* is calculated as:

$$\text{Distance}^{M}_{(i,t,j,s)} = \sqrt{\sum_{m=1}^{M} \left(EARN_{i,t-m+1} - EARN_{j,s-m+1}\right)^2}.$$

This method is based on the assumption that historical earnings serve as a reliable indicator of future performance; firms with similar past performance are likely to exhibit similar future performance. The earnings prediction for the subject firm-year is derived from the median of the lead earnings observed among its identified nearest neighbors. Easton et al. (2024) advocate for the simplicity of the KNN algorithm in forecasting corporate earnings, drawing on comparisons with firms that have similar current and past earnings. They assert that a simpler forecasting method is easier to interpret, modify, and less prone to overfitting. Their findings indicate that KNN significantly outperforms more complex models, such as advanced KNN variations, random walk models, and existing regression models in terms of accuracy. Moreover, KNN forecasts of longer-term earnings per share (EPS) and aggregate EPS were found to be more precise than those generated by professional analysts. A distinct advantage of the KNN algorithm is its ability to self-assess accuracy through the Mean Absolute Deviation[5] (MAD) metric, which effectively predicts forecast accuracy and provides investors with an indication of reliability. Their research underscores the notion that increasing the number of variables or extending the historical data scope does not enhance forecast accuracy. Instead, it affirms that recent earnings history is a robust predictor of future earnings when contextualized appropriately. This study highlights the practical value of a simple, comparable-firm-based method for earnings prediction, advocating for both simplicity and the careful selection of relevant historical data.

[5] In Easton et al. (2024), MAD is the median of the absolute values of the differences between each nearest neighbor's lead earnings and the median of the nearest neighbors' lead earnings.

Cao and You (2024) find non-linear ML models significantly outperform traditional earnings prediction models. This improved performance is largely due to the models' capacity to identify economically important predictors and capture nuanced, nonlinear relationships within financial data. The information collected from ML forecasts has substantial economic value for investors, as it also demonstrates predictive power for future stock returns. The study further compares ML-based forecasts with analysts' consensus forecasts, noting that ML models perform comparably to analyst forecasts over a one-year horizon and surpass them over longer periods. Moreover, ML models provide incremental information beyond analyst forecasts, even when analysts have access to comprehensive financial statement data. The ML models also help in detecting optimistic biases in analysts' predictions, supporting investors' need for objective data analysis. Overall, their results underscore that ML is an effective tool for deriving relevant insights for investors from financial statements and highlight the continued value of fundamental analysis. This reinforces the potential for ML in financial analysis by enabling sophisticated pattern recognition and data utilization for improved earnings prediction.

Campbell et al. (2023) conclude that while ML methods have the potential to improve earnings forecasts, their effectiveness is highly dependent on model specification choices. Specifically, they found that 90% of ML models evaluated did not outperform analysts' forecasts. However, the best-performing ML forecasts consistently correct for predictable analyst biases related to past errors and stock prices, leading to statistically significant accuracy improvements, particularly for small-cap firms and over longer forecasting horizons. Additionally, the study reveals that investors' earnings expectations, as reflected in stock prices, partially account for these biases but do not fully correct them, with price realizations often lagging by up to nine months. Overall, the findings indicate that the most accurate ML forecasts can mitigate predictable biases in analyst forecasts and align more closely with investors' expectations, particularly for large-cap firms with significant institutional ownership.

Van Binsbergen et al. (2023) conclude that the pricing of assets heavily relies on earnings forecasts, which are often upward-biased. They introduce a novel ML forecasting algorithm that is statistically optimal and resistant to variable selection bias, demonstrating its effectiveness in out-of-sample contexts compared to traditional linear forecasts. This new benchmark serves not only as a valuable input for asset pricing but also as a real-time tool for evaluating analyst earnings forecast biases over time and across different stocks. Their analysis reveals significant variation in these biases, and they find that stocks with the most upward-biased earnings forecasts tend to experience lower future returns, while those with downward biases generate higher returns. This suggests that analysts' forecast errors can significantly influence asset prices.

Chattopadhyay et al. (2022) explore the effectiveness of ML techniques in forecasting future earnings and estimating implied cost of capital (ICC). The study evaluates three ML models-LASSO regression, RIDGE regression, and Extreme Gradient Boosting (XGBoost). They compare the results with the random walk model, the Hou et al. (2012)'s model, and the earnings persistence and residual income models (Li and Mohanram, 2014). Additionally, the ML models are benchmarked against a simple linear model with an augmented set of predictors. They are the first study that investigate both U.S. and international firms. In the U.S. sample, they find that XGBoost generates

Table 3: Overview of Machine Learning Applications in Earnings Prediction

Author (Year)	Journal	Conclusion	ML method	Evaluation Metric	Paper Index
Anand et al. (2019)	Working Paper	ML achieves classification accuracies ranging from 57-64%, better than the random walk.	Random Forest	Out-of-sample Accuracy	A
Campbell et al. (2023)	Working Paper	90% of ML models evaluated did not outperform analysts' forecasts	OLS, LASSO, RIDGE, Elastic Net, Random Forest, GBRT	MAE, MSE	B
Cao and You (2024)	Financial Analysts Journal	ML models generate more accurate forecasts than state-of-the-art earnings prediction models	OLS, LASSO, RIDGE, Random Forest, StochasticGBM, ANNs	MAFE	C
Chattopadhyay et al. (2022)	Working Paper	Nonlinear tree-based models outperform extant models over different horizons	LASSO, RIDGE, XGBoost	MAFE	D
Chen et al. (2022)	Journal of Accounting Research	ML methods outperform logistic regressions and analysts' forecasts	Random Forest, StochasticGBM	ROC, AUC, Hedge portfolio return	E
Cui et al. (2020)	Atlantic Economic Journal	ML models outperform logistic regression but unable to surpass analysts	LightGBM	Out-of-sample Accuracy	F
Easton et al. (2024)	The Accounting Review	KNN forecasts are more accurate than extant approaches for one-, two-, and three-year-ahead earnings	KNN	MAFE, MSE, MAD	G
Hunt et al. (2022)	Accounting Horizons	Random forest provides better out-of-sample accuracy and higher abnormal returns	Stepwise Logit Regression, Random Forest, Elastic Net	Out-of-sample accuracy, Abnormal return	H
Jone et al. (2023)	Contemporary Accounting Research	ML methods predict out of sample better than traditional regression methods	StochasticGBM, OLS	MSE, RMSE, MAPE, R^2	I
Van Binsbergen et al. (2023)	The Review of Financial Studies	Introduce a novel ML algorithm resistant to variable selection bias, demonstrating its effectiveness in out-of-sample contexts compared to traditional linear forecasts	Random Forest	Forecasting Bias	J
Yakabi et al. (2024)	Working Paper	Random Forest and Gradient Boosting outperformed Logistic Regression in terms of prediction accuracy and portfolio return	Random Forest, StochasticGBM, LLMs	ROC, AUC, Hedge portfolio return	K

the most accurate forecasts, particularly for small firms and firms with volatile earnings. However, the improvements in forecast accuracy are modest. For the international sample, XGBoost demonstrates significantly superior performance, highlighting its robustness in settings with sparse coverage and volatile earnings. ICC tests corroborate these findings, with XGBoost consistently outperforming other models, especially for international firms where traditional cross-sectional models underperform. The paper highlights methodological contributions by demonstrating XGBoost's ability to deliver accurate forecasts with relatively low computational demands compared to other ML models like Random Forest or Gradient Boosting. However, they also acknowledge limitations, noting that their analysis relies on a static set of explanatory variables. Future research could enhance forecast accuracy by incorporating non-financial and market-based signals. The findings emphasize the potential of ML models in advancing earnings forecasting and ICC estimation.

4. Challenges in using ML to Predict Future Earnings

Although ML models have been proven to be more efficient and accurate than traditional models in many fields, their application in earnings prediction still faces many challenges.

i. Data Quality and Complexity

ML models require extensive and detailed data to accurately predict earnings. Studies (e.g., Chen et al., 2022) leverage large sets of detailed financial information, yet this dependency can introduce challenges related to data collection, preparation, and ensuring consistency across time periods and different firms. Complex and high-dimensional data also increase the likelihood of overfitting, especially with certain algorithms. Although decision tree-based ML algorithms are good at processing high-dimensional data, it is not easy for researchers to collect, clean, and integrate such high-dimensional data. In order to obtain more abundant predictive variables, more observations need to be sacrificed most of the time. Therefore, when using such algorithms for profit forecasting, how to ensure that the entire data processing process is controllable is still a challenge for researchers. Many ML models are trained on historical data and assume that past relationships will hold in the future. In rapidly changing markets, this reliance on historical patterns may not always yield accurate predictions.

ii. Model Selection and Overfitting Risks

Although ML models can capture complex, nonlinear relationships (as noted by Cao and You, 2024), this complexity can also make models prone to overfitting. When there are too many predictors or a high-dimensional feature space, as in the study by Jones et al. (2023), models might capture noise instead of true patterns, reducing predictive accuracy in out-of-sample tests. Different studies highlight the effectiveness of various models (e.g., Random Forests, Gradient Boosting Machine, KNN), but there is no one-size-fits-all solution. Choosing the most appropriate model is challenging, as performance can vary based on the dataset, feature selection, and model hyperparameters, which requiring extensive testing and optimization.

iii. Interpretability and Transparency

Machine learning models, especially nonparametric ones like ANNs and ensemble methods, often lack transparency and can be difficult to interpret for stakeholders. As explained in Jones et al., (2023), ML models uncover complex interactions among predictors that may not be straightforward to explain.

iv. Practical Application and Economic Value

While ML models may yield more accurate predictions than traditional models, translating these predictions into economically significant gains is not guaranteed (Jones et al., 2023). For example, in portfolio return analysis, ML predictions did not always result in superior abnormal returns compared to traditional regression-based models. This limitation indicates that improved forecast accuracy does not always correlate with better investment outcomes. Financial markets are dynamic, and earnings predictors may change in relevance over time. While ML models can adapt to changes, the robustness of these models across different economic conditions and market cycles remains a concern, as highlighted by studies like Hunt et al. (2022), which suggest the need for ongoing refinement and testing over time.

5. Opportunities of using ML to Predict Future Earnings

Challenges also mean opportunities. Next, I will discuss where future research can be carried out.

i. Exploration of New Data Types

When exist studies mainly focused on financial data, future studies could investigate the inclusion of alternative data sources such as text sentiment, customer reviews, and macroeconomic indicators. Such data has shown promise in other areas of finance and could enrich earnings forecasts by capturing more dimensions of market and firm sentiment. Future research can explore techniques to analyze these data types, perhaps using natural language processing (NLP) alongside traditional financial metrics.

ii. Forecasting Earnings in Longer Horizon

Many current models focus on short-term (one-year-ahead) earnings predictions. Research can explore ML methods in a longer forecast horizons, which allowing analysts to consider broader economic cycles.

iii. Incorporate Forward-looking Information

Most studies reviewed in this paper rely on historical financial ratio to predict future earnings and earnings changes. With more and more firms disclose their perceptions of future risk and opportunities in their annual report, researchers should utilize those forward-looking information to strengthen the prediction of earnings. The most promising approach might lie in hybrid models that combine both historical financial ratios and forward-looking information. By integrating structured historical data with unstructured forward-looking disclosures, such models can leverage the consistency of past performance data and the adaptability of current expectations.

6. Emerging Trend in Earnings Prediction: Large Language Models

With the development of Large Language Models (Hereafter, LLMs), many studies incorporate this method into financial analysis. [6]

Kim et al. (2024) investigate the capabilities of LLMs, such as GPT–4, in financial statement analysis. The study provides LLMs with structured, anonymized financial statements and uses a Chain-of-Thought prompting technique to simulate the analytical process of human financial experts, excluding any narrative inputs. Specifically, they evaluate the performance of LLMs in predicting the direction of future earnings using a two-step approach. First, corporate financial statements are anonymized and standardized to eliminate potential bias from the model's memory of specific companies or time periods. Company names are removed, years are replaced with labels (e.g., t and t–1), and financial statements are reformatted to align with Compustat's balancing model, ensuring consistency across firm-years. In the second step, they employs carefully designed prompts to guide the LLMs in performing financial analysis. Alongside a simple prompt, a Chain-of-Thought[7] (CoT) prompt is introduced to emulate the analytical process of human financial experts. The CoT prompt directs the model to identify trends in financial statement line items, compute key financial ratios (e.g., operating efficiency, liquidity, leverage), synthesize this information, and predict whether next year's earnings will increase or decrease. This structured prompting effectively mirrors the reasoning process used by professional analysts, enabling the LLMs to simulate complex financial analysis tasks.

Kim et al. (2024) demonstrate that LLMs can outperform analysts in predicting the direction of future earnings, particularly in scenarios where analysts are prone to bias or disagreement. LLMs complement both human analysts and ML models. They perform better than humans when additional narrative context is unnecessary and outperform quantitative ML models in areas like analyzing loss-making firms, showing "human-like" qualities. Conversely, they exhibit "machine-like" tendencies by excelling with larger firms. Surprisingly, GPT–4's performance rivals advanced ML models like ANNs and exceeds them in certain contexts. Additionally, the narrative analysis generated by the model adds substantial informational value.

Kim et al. (2024) highlights GPT–4's ability to derive insights from trends and financial ratios, emphasizing its broad reasoning capabilities over memory-based performance. A trading strategy based on GPT–4's predictions outperformed strategies using traditional ML models, yielding higher Sharpe ratios and alphas. The findings suggest that general-purpose LLMs can democratize financial analysis by offering state-of-the-art performance without specialized training. While LLMs have the potential to act as central elements in financial decision-making, the study calls for further exploration into the broader implications of AI-driven financial analysis. At the end of their study, they also commented that while LLMs can mimic human reasoning through chain-of-thought prompts, the underlying mechanics of their decision-making are not always clear, particularly when predicting

[6] Recent studies using LLMs to imply a wide range of tasks, including summarization of complex disclosures, sentiment analysis, information extraction, report generation, compliance verification, etc. Please refer to Kim et al. (2024) for comprehensive review of studies on this topic.

[7] Chain-of-thought in LLMs refers to a technique that involves prompting the model to break down complex reasoning tasks into a series of intermediate steps, mimicking human-like logical reasoning.

complex financial outcomes. It remains unclear which specific elements within the prompt are essential for achieving great performance (Kim et al. 2024).

7. Conclusion

This study reviewed recent accounting literature to explore the application of ML in earnings prediction. The findings indicate that most studies concentrate on using ML to predict changes in earnings, with portfolio returns based on these predictions significantly outperforming those derived from traditional regression methods. In contrast, fewer studies focus on forecasting exact earnings levels, likely due to the complexity and lower predictive accuracy of such tasks in prior research. This divergence highlights a critical debate on whether predicting directional changes provides more economic value than forecasting specific level of earnings.

The use of ML for earnings prediction is not without challenges. Key obstacles include data quality and consistency issues, the risk of overfitting in high-dimensional datasets, and the limited interpretability of complex models. Moreover, translating improved prediction accuracy into economic gains remains an open question, warranting further investigation.

Despite these challenges, this field presents exciting opportunities for future research. One can leverage new data types, such as textual information from corporate disclosures or macroeconomic indicators, to enhance model performance. Additionally, extending forecasting horizons to incorporate long-term trends and integrating forward-looking information, such as management forecasts and risk disclosures, could further boost predictive accuracy and relevance.

An emerging and promising trend in earnings prediction research is the application of Large Language Models (LLMs), such as GPT–4. Kim et al. (2024) shown that LLMs can outperform human analysts in predicting earnings direction, particularly in scenarios prone to analyst biases or disagreements. LLMs also complement traditional ML models, excelling in specific contexts like analyzing loss-making firms or larger companies. Their ability to process structured financial data and derive insights without specialized training underscores their potential as transformative tools in financial analysis. Overall, LLMs represent a significant innovation in earnings prediction, offering a path to democratize financial analysis and bridge gaps in traditional methods. By addressing existing challenges and exploring these new opportunities, the integration of advanced ML techniques and LLMs could fundamentally reshape the landscape of earnings prediction and financial decision-making.

References

[1] Anand, V., Brunner, R., Ikegwu, K., & Sougiannis, T. (2019). Predicting Profitability using Machine Learning. *SSRN Journal.* 3466478.

[2] Campbell, J. L., Ham, H., Lu, Z., & Wood, K. (2023). Expectations Matter: When (not) to Use Machine Learning Earnings Forecasts. *SSRN Journal.* 4495297.

[3] Cao, K., & You, H. (2024). Fundamental Analysis via Machine Learning. *Financial Analysts Journal*, 80(2), 74–98.

[4] Chattopadhyay, A., Fang, B., & Mohanram, P. (2022). Machine Learning, Earnings Forecasting, and Implied Cost of Capital – US and International Evidence. Working Paper.

[5] Chen, X., Cho, Y. H. (Tony), Dou, Y., & Lev, B. (2022). Predicting Future Earnings Changes Using Machine Learning and Detailed Financial Data. *Journal of Accounting Research*, 60(2), 467–515.

[6] Cui, X., Xu, Z., & Zhou, Y. (2020). Using Machine Learning to Forecast Future Earnings. *Atlantic Economic Journal*, 48, 543–545.

[7] Easton, P. D., Kapons, M. M., & Monahan, S. J. (2024). Forecasting Earnings Using k-Nearest Neighbors. *The Accounting Review*, 99(3), 115–140.

[8] Hou, K., Van Dijk, M. A., & Zhang, Y. (2012). The Implied Cost of Capital: A New Approach. *Journal of Accounting and Economics*, 53(3), 504–526.

[9] Hunt, J. O. S., Myers, J. N., & Myers, L. A. (2022). Improving Earnings Predictions with Machine Learning. *Accounting Horizons*, 36(1), 131–149.

[10] Jones, S., Moser, W. J., & Wieland, M. M. (2023). Machine Learning and the Prediction of Changes in Profitability. *Contemporary Accounting Research*, 40(4), 2643–2672.

[11] Kim, A., Muhn, M., & Nikolaev, V. (2024). Financial Statement Analysis with Large Language Models. *arXiv preprint arXiv: 2407.17866*.

[12] Penman, S. H., & Zhang, X. (2004). Modeling Sustainable Earnings and P/E Ratios using Financial Statement Information. Working Paper, Columbia University and University of California, Berkeley.

[13] Van Binsbergen, J. H., Han, X., & Lopez-Lira, A. (2023). Man versus Machine Learning: The Term Structure of Earnings Expectations and Conditional Biases. *The Review of Financial Studies*, 36(6), 2361–2396.

[14] Yakabi, K., Kuroki, Y., & Nakagawa, K. (2024). Predicting Earnings Change Using Machine Learning with Data from Japanese Companies. *Jsaisigtwo.* Fin–033, 68–75. (In Japanese)

【2024年度　学生懸賞論文受賞作　最優秀賞要旨】

芥川賞・直木賞を受賞したことによる話題量への効果と，その効果の決定要因について

大西咲樂

本稿の目的は，芥川賞・直木賞の受賞が，消費者によって行われる口コミやレビューの話題量に与える影響と，それを決定する要因について検討することである。話題量とは，具体的には，芥川賞・直木賞の受賞日から１カ月後までのレビュー数を，受賞日より前のレビュー数で割ることにより求めたものである。これは，自分が買って読んだ本について，自分の感情や思いを表現するため，または他者に情報を与えるためなどの動機から，個人が外部に向けてその本に関する何らかの情報を発信する意欲が，芥川賞・直木賞の受賞というイベントによってどれほど大きくなるかを測定したものである。本稿ではこれを「受賞後発信拡大率」と呼ぶ。

１章では，芥川賞・直木賞を受賞した書籍に関して受賞後発信拡大率を求め，芥川賞を受賞した書籍の方が, 直木賞を受賞した書籍と比べ, 受賞後発信拡大率が大きいことが示された。相対的に知名度が低い作家に与えられる芥川賞を受賞した書籍の方が，相対的に知名度が高い作家に与えられる直木賞を受賞した書籍に比べ，発信拡大率は大きくなったのである。ここから2章では，同じ芥川賞内において「作家の事前の知名度が低いほど，芥川賞受賞後の発信拡大率は大きくなる」という仮説を検証した。しかし，仮説とは逆に，「作家の事前の知名度が高いほど，芥川賞受賞後の発信拡大率は大きくなる」という結果が得られた。なぜこのような結果が得られたのかについて，またここから得られる示唆については最後に説明する。

【2024年度　学生懸賞論文受賞作　優秀賞要旨】

ミールプランへの加入が大学生の欠食に及ぼす影響の分析

村田拡介

大学生の欠食に対応するべく大学生協はミールプラン制度を運用している。しかし，ミールプランと欠食の関係は明らかになっていない。そこで，本研究ではミールプランへの加入が大学生の欠食に及ぼす影響を分析した。第 58 回学生生活実態調査および大学情報サイト「マナビジョン」のデータを用い，ミールプラン加入が学生の食事摂取に及ぼす効果を推定した。ただし，ミールプランが任意加入であるために生じるセレクションバイアスに対処するべく，操作変数法による推定を行った。推定の結果，ミールプラン加入は１日３食以上の食事をとることに有意な正の効果をもたらしており，欠食を減少させていることが明らかとなった。また，ミールプラン加入によって食事をとる学生が有意に増加している時間帯と有意に減少している時間帯の双方が存在した。これは，生協食堂の営業時間を背景として，ミールプランが適切な食習慣の形成にも寄与している可能性を示唆する。

【2024年度　学生懸賞論文受賞作　優秀賞要旨】

相対的な収入が生活満足度に与える影響及び，男性の中年期における生活満足の減少要因の分析

三谷康太　大野創　後藤啓介　井上京大

本稿の目的は，ミッドライフクライシスと呼ばれる主に男性が中年期において幸福度（生活満足度）が最も低くなるという現象について，その要因を分析することである。本稿では健康状況，家庭状況，経済状況の３つを生活満足度に影響を与える要因であると仮定し，若年期，中年期，高齢期に前述の要因がそれぞれどのように生活満足度に作用するかを分析した。その結果，健康状況は生活満足度に影響を与えないものの，家族状況と経済状況では中年期のみに観測される特徴があった。具体的に，経済格差は中年期において低所得者の生活満足度を減少させることが分かった。また，家庭状況における配偶者の存在が生活満足度に与える影響は中年期に最も低くなることもミッドライフクライシスの原因だと推測される。

【2024年度　学生懸賞論文受賞作　優秀賞要旨】

自転車道リニューアル工事が自転車道の使用と通行の速さに与える影響

豊泉有理

大阪大学豊中キャンパス正門入口は，多くの学生が通学のため自転車で走行しており，専用の自転車道が設置されている。しかしながら，路面表示の劣化や，オートバイの不正入構を阻止するためのH型自転車ゲート等の障害によって，利便性が損なわれており，自転車利用者に対する自転車道の使用割合は低いものであった。この問題に対処するため，大阪大学では自転車道のリニューアル工事を行った。

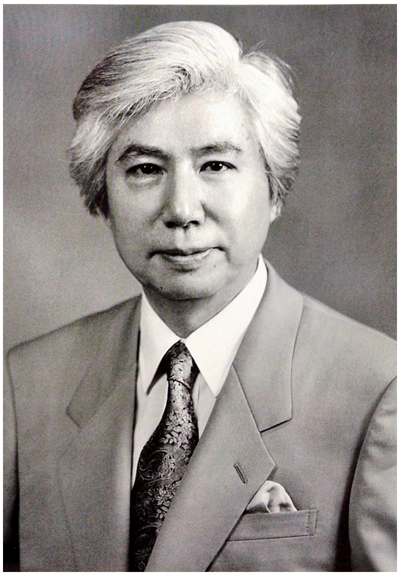
本稿では，自転車道整備が，自転車利用者の行動に与える影響を検証した。また，その際施工された加速体感型矢羽根型路面表示について，仕掛学の観点からその特徴を分析した。工事の段階に伴って自転車道使用割合と走行スピードを観察し，重回帰分析の結果，H型自転車ゲート撤去，自転車道の色分け，文字表示，加速体感型矢羽根型路面表示には自転車道使用割合を上昇させる効果があり，またH型ゲート撤去には走行スピードを上昇させる効果があることが示された。また，整備の効果は一定期間の後も持続することが示された。

【2024年度　学生懸賞論文受賞作　優秀賞要旨】

バスケットボールにおける審判の判定と検証要求権保持との関係

神橋龍ノ介　河合求真

本研究では，不作為バイアスに影響を与える要因を，NBAのプレーに対する審判の判定とその正誤のデータ等を用いて検証したものである。研究結果として，僅差かつ試合終了間際で判定の正誤が後日検証・公表されるような状況での，審判がファール判定をすべきであるのにファール判定をしないという不作為バイアスの大きさは全体で20.5%であることを示した。また，不作為バイアスを発生させる2つの要因を実証的に明らかにした。第一は，観客がいるその場で判定の正誤が検証される可能性があるという要因（検証回避要因）であり，不作為バイアスを2.1%ポイント大きくしていた。第二は，観客にその場で見られているという要因（目線要因）であり，不作為バイアスを5.9%ポイント大きくしていた。



小泉進先生を偲んで

1981年4月，筆者が大阪大学経済学部経済学科に入学した当時，小泉進先生は本学の経済学部長を務めておられ，そしてまた経済学部生にとって最も重要な必修科目である『経済原論』を担当しておられました。その意味で，筆者にとって小泉先生は初めての「大学の先生」の記憶であり，また初めて精緻な「経済学理論」に触れた記憶そのものでもあります。加えて，自身その後大変お世話になった永谷裕昭先生，久我清先生といった純粋理論経済学の先生方の，その出身ゼミの先生でもあります。そのようなことから，今日の経済学研究科スタッフを代表し大阪大学経済学での追悼記事を書かせていただくということに，甚だ僭越と畏れながらも，以下小泉先生のご業績を振り返りつつ，筆者がその役割を務めさせていただきますこと，お許しいただきたく思います。

小泉進先生は1929年8月に神戸市でお生まれになり，1949年3月に灘高等学校を卒業後，大阪大学法経学部に進学され，1953年3月に卒業されました。卒業後すぐに大阪大学法経学部の助手，1954年にアメリカ・ミシガン大学での研究助手ならびに大学院への入学，1958年6月にはミシガン大学からPh.D.学位を授与され，同年8月に大阪大学経済学部の講師として復帰されました。その後1960年に助教授，1964年にはペンシルベニア大学ワートン・スクールの客員准教授として渡米，1年間の任期を務められ，帰国後1969年4月に大阪大学経済学部教授に就任されました。1979年に経済学部長，1981年までその職務を果たされ，1993年3月に大阪大学を定年退職された後は帝塚山大学に移られ，大阪大学および帝塚山大学の名誉教授の称号が授与されています。

小泉先生は，計量経済学モデルの発展に貢献されました。先生のご研究は，政府支出などの政策手段がどの程度経済成長に影響を与えるかを考察するもので，政府が経済をある程度コントロールできるというケインズ経済学的な発想がその基盤となります。小泉先生が本格的に研

究生活に入られた1950-60年代は，まさに市場経済を計量化し，予測に基づいて不況の回避や経済成長を促進する方法を模索しようとする世界的な大きな流れの下にありました。そうした中，小泉先生もまた日本経済の動向を分析し，政策効果を検証する研究を行われました。特に師であるLawrence Klein先生が米国で行っていたような計量モデルを活用した研究に注力され，経済予測や政策の効果分析に取り組んでおられたことが，小泉先生の研究業績の重要な側面となっています。その成果の一部は，1964年International Economic Review誌に掲載された論文“A Priori Information and Time Series Analysis: A Review Article”等に結実しております。

小泉進先生は1979年7月から1981年7月まで大阪大学経済学部長として，優秀な人材を集める強力な経済学部を目指して活動されました。冒頭にも述べた筆者の大学入学時期にも当たっていますが，その当時，本学の理論政策系では林敏彦氏，中谷巌氏，本間正明氏，植田和男氏，井堀利宏氏や林文夫氏，猪木武徳氏，蝋山昌一氏など，極めて強力な人材が集まっておりました。優れた若手人材の早期の教授昇進を可能とするため，小泉先生は大講座制の導入が不可欠と判断され，総長の山村雄一氏にも協力を依頼。結果，大学側の全面的なバックアップのもと文部省との協議が成功し，全国的にも最早期における大講座制の採用が可能となりました。

小泉先生は，教育活動にもまた非常に熱心に取り組まれました。大学紛争の激しかった1960年代末から70年代初め，教員は単なる研究者ではなく教育者としての責任を強く自覚すべきと考えるようになられたと伺います。小泉先生の『経済原論』は，岩波の『価格理論I』と『所得分析』を教科書とし，科学としての経済学の方法論を導入としながら，Edgeworth Boxから入った厳密なミクロ理論と，古典派およびケインズ理論の対比に基づく実践的マクロ理論の優れた展開および調和という，近代経済学の入門講義として当時に望まれる最高レベルのものであったと思います。小泉先生の原論を出発点に，林敏彦先生のミクロ，中谷巌先生のマクロ，本間正明先生の財政政策，植田和男先生の金融政策と続いた往時の理論政策系カリキュラムは，今日にして思えば圧巻そのものであり，その時代に大阪大学で経済学理論を学ぶことができたことの幸運を思わずにおれません。同時にその後の本学部に於いて，当時に匹敵する魅力あるカリキュラムが実現できたかどうか。この30年は大学を取り巻く事情も大きく変わりましたが，そのことを差し引いても，もし力及ばずであるとするならば，その責任の一端を担う者として，改めて自己を不甲斐なく思わずに居れぬ次第です。

研究・行政・教育の面から大阪大学経済学部発展へ多大な貢献をなされた小泉先生に感謝を申し上げつつ，先生のご冥福を心よりお祈り申し上げます。

（浦井　憲　大阪大学大学院経済学研究科教授）

『大阪大学経済学』　第 74 巻　令和 6 － 7 年

総　　目　　次

論　題	著　者	巻　号	年　月	頁
論　文				
戦間期在ベルリン日本人鉄道職員―「ドイツ経験」は何をもたらしたか	鳩　澤　歩	74 － 1･2･3	R. 6. 12	1 － 22
中国における LED 照明産業の拡大過程と日系企業の知財戦略に関する一考察：日亜化学工業の事例とその特色	魏　晶　京・許　衛　東	74 － 1･2･3	R. 6. 12	23 － 45
Impacts of Japan's Green Bond Guidelines 2020 on ESG bond issuers: A quasi-experimental study using PSM-DID	Yuan Mingqing	74 － 4	R. 7. 3	1 － 31
産学間の共同研究における社会関係資本獲得を促進する方策の検討 REACH プロジェクト参加者への半構造化インタビューによる事例研究	平　丸　大　介・開　本　浩　矢	74 － 4	R. 7. 3	32 － 44
Earnings prediction using machine learning: A survey	Yuanchao Peng	74 － 4	R. 7. 3	45 － 60
彙　報				
学会消息		74 － 1･2･3	R. 6. 12	46 － 64
2024 年度　学生懸賞論文　受賞作要旨		74 － 4	R. 7. 3	61 － 63
追悼：小泉進先生を偲んで	浦　井　憲	74 － 4	R. 7. 3	64 － 65
『大阪大学経済学』第 74 巻　令和 6 － 7 年　総目次		74 － 4	R. 7. 3	i

Editorial Policy

The Osaka Daigaku Keizaigaku (English title, Osaka Economic Papers) is published quarterly by the Economic Society of Osaka University and the Graduate School of Economics, Osaka University. The articles may be either in Japanese or in Western languages.

The Journal shall be under the editorial direction of an editorial board of three persons chosen from members of the Graduate School of Economics of Osaka University. The editorial board shall select papers for publication from submissions and classify them into the following categories: articles, notes, data, and book reviews.

Researchers who belong to the Graduate School of Economics of Osaka University may submit their studies for publication to this journal. Those who do not belong to the Graduate School may also publish their papers in this journal, if their contribution is closely related to research being undertaken in the Graduate School of Economics of Osaka University.

In the case of contributed manuscripts, the author should be a member of the Economic Society of Osaka University, who has paid the yearly membership fee of 4,000 yen.

大阪大学経済学　第74巻　第４号（通巻240号）
令和７年３月発行

編集兼発行人	〒560-0043	豊中市待兼山町１番７号	佐々木　　勝
印　刷　所	〒530-0043	大阪市北区天満１丁目９番19号	株式会社NPCコーポレーション
発　行　所	〒560-0043	豊中市待兼山町１番７号	大阪大学経済学会・大阪大学大学院経済学研究科

tel　06-6850-5270　fax　06-6850-5270
振替　00940-2-19842

OSAKA ECONOMIC PAPERS

Vol. 74 No. 4 March 2025

Table of Contents

Articles

Impacts of Japan's Green Bond Guidelines 2020 on ESG bond issuers:
A quasi-experimental study using PSM-DID ········ Yuan Mingqing 1

Study of methods to facilitate social capital acquisition in academia-industry joint research:
A case study based on semi-structured interviews with REACH project participants
········ Daisuke Hiramaru and Hiroya Hirakimoto 32

Earnings prediction using machine learning: A survey ········ Yuanchao Peng 45

Abstracts of Prize-Winning Papers in the Students Essay Contest, 2024 ········ 61

In memory of Professor Susumu Koizumi ········ Ken Urai 64

Index to Volume 74 (2024–2025) ········ i

THE ECONOMIC SOCIETY OF OSAKA UNIVERSITY
GRADUATE SCHOOL OF ECONOMICS, OSAKA UNIVERSITY
TOYONAKA, OSAKA, JAPAN