ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-64-84
UDC 336.63,519.237.5(045)
JEL C23, C26, G30

Endogeneity Problem in Corporate Finance: Theory and Practice

Z.V. Selezneva, M.S. Evdokimova
National Research University Higher School of Economics, Moscow, Russia

ABSTRACT

Endogeneity can cause a significant bias in the coefficient estimation, up to the change in sign. It leads to controversial research results, which also makes it difficult to adequately test individual hypotheses and theories in corporate finance (CF). For practitioners, such as company valuation consultants, these model problems interrupt obtaining the most reliable estimates in the interests of the customer. The **aim** of this study is to review an endogeneity problem in CF and ways to solve a problem of endogeneity. We will illustrate the methods found in the systematic review with an empirical example. The paper provides the reasons for this problem from an econometric point of view and with examples from the CF and econometric methods of dealing with it. As a **result** of a systematic literature review, we have shown that dynamics panel models, in particular the Blundell-Bond **method**, are mostly used to consider endogeneity in CF studies. We have verified empirically the conclusion made in the framework of the literature review. To detect the endogeneity, we used the Hausman test, the endogeneity test, and the analysis of the correlation matrix, including the saved regression residuals. Eliminating step-by-step endogeneity, we **concluded** that the Blundell-Bond method is not always the optimal one for dealing with endogeneity in CF, as well as regression with a fixed effect. It was revealed that the two-stage least squares method (IV 2SLS) is the most appropriate method for the cost of capital model estimation eliminating endogeneity. In addition, the estimates of the cost of capital model, which analyzes the impact of non-financial reporting, have been improved.

Keywords: corporate finance; endogeneity; instrumental variables; panel data regression analysis; cost of capital; non-financial reporting

For citation: **Selezneva Z.V., Evdokimova M.S. Endogeneity problem in corporate finance: Theory and practice.** ***Finance: Theory and Practice.*** **2022;26(3):64-84. DOI: 10.26794/2587-5671-2022-26-3-64-84**

INTRODUCTION

Corporate finance (hereinafter referred to as “CF”) as an independent empirical science has been developing for seventy years [1]. It is difficult to find theoretical papers in this area over the past two decades without elements of empirical research. At the same time, the intensive use of various econometric models in the CF.

In this area, hypotheses have traditionally been tested using linear regression. However, the complexity of the CF study object of a company suggests the potential mutual influence between several factors [2]. And the limited data collected suggests that linear regression may be misspecified, missing important factors. In other words, studies in CF are accompanied by endogeneity.

There has been an ongoing discussion about the presence of endogeneity in the CF models for about 9 years. Wintoki and Linck [2] note that researchers mainly rely on two possible sources of endogeneity and use models for panel data, in particular, the fixed effects model, to overcome it. Other scientists [3] have shown that most CF research is conducted without regard to potential endogeneity. Several studies [4, 5] use simulated data to show which methods are superior in accuracy to estimates of fixed effects (FE) models in the presence of endogeneity. They show that FE models may not be sufficient to deal with endogeneity. Although, as Banik notes [6], this approach is common in modern scientific research.

The presence of endogeneity in models is a violation of the premise of the Gauss-Markov theorem on the independence of regression residuals from explanatory variables, which causes bias and inconsistency in estimates. From a practical point of view, this has a negative effect:

- applicability of estimation models: results obtained with ‘naive’ estimating, such as the least squares method, do not provide unbiased estimates in the presence of endogeneity. As a result, the actual influence of the independent variables may be underestimated or overestimated. For example, in the paper of Molina [7], taking into account endogeneity leads to an increase in the negative impact of financial leverage on the credit rating by 3 times, and in the article by Chen et al. [8] after evaluating the model taking into account endogeneity, the influence of the CEO Duality became insignificant for the firm's performance;

• interpretability of models. One of the manifestations of endogeneity is that it is impossible to establish causality, which can be caused by several reasons:

– reverse causality. Initially, the explanatory variable (X) is assumed to influence or precede the explanatory variable (Y). But the analysis reveals otherwise. A similar question is raised in many studies in various areas of CF [9–11];

– simultaneity. For example, Harada and Ngyen [12] note that the ownership structure and financial policy of the company are determined simultaneously;

– dependence of the explained and explanatory variables on the same factor not taken into account in the model (CMV). For example, Molina [7] proved that the financial leverage and credit rating of a company depend on the unobservable fundamental risk of the company;

• comparability of research results. Flannery and Hankins [5] note that regression estimates for the same models and similar data vary significantly from study to study. This is also observed in modern research. So, for example, the ambiguity of the impact of non-financial reporting on the performance of the company is noted by both Russian researchers: Zhukova and Melikova [13]; Martynova [14]; Polyakov et al. [15] and foreign ones: Zahid et al. [16];

• accuracy of predictive models, which is important, for example, for analysts when predicting the value of assets both in the financial market [17] and profitability in the real sector [18], and for representatives of the state [19];

• stability of models. Models become sensitive to the introduction of additional factors or changes in the sample, which indicates their instability. For example, Coles et al. [20] demonstrate that a 10% change in parameters leads to a change in the sign of the endogenous variable in the model without considering endogeneity;

• reliability of conclusions and adequacy of measures taken on the basis of the results of the regression analysis: decisions based on the conclusions of the model may not be optimal. Li [21] showed that the model without endogeneity evaluates the impact of CEO compensation on the firm's performance (Tobin's Q) as positive, but with it — negative.

Summarizing the above, the researcher runs the risk of encountering an incorrectly estimated model due to endogeneity. At the same time, the estimation bias in such models can vary from insignificant to huge. The estimates may even change sign or be insignificant.

In this article, we systematize information about potential sources of endogeneity and methods to overcome it in CF research. Like Flannery and Hankins [5], we are primarily interested in the structuring of information about endogeneity in CF studies, how researchers define it and how they deal with it.

Unlike previously cited articles, this paper provides descriptions of potential sources of endogeneity. Examples from different areas of CF research are also given. In addition, we will illustrate one way of overcoming endogeneity by examining the impact of companies' additional non-financial reporting about their activities on the firm's performance. This is a relatively new topic in the field of CF. The Journal of Corporate Finance published 56 articles on this topic, 33 of which were published in the last 5 years.[1] Within the framework of this topic, methods for determining endogeneity and ways to deal with it are demonstrated on a sample of 663 companies in the BRICS countries in 2007–2016. During model evaluation, it was shown that one of the most reliable methods

[1] The search was carried out using the keywords "non-financial report", "CSR", "ESG", "IR", "sustainability" in all publications of the journal in Scopus.

for estimating endogeneity regressions, the Blundell-Bond method, is not the default method for working with endogeneity, like the individual fixed effect model. From a statistical point of view, it was possible to overcome endogeneity in the model only by using the two-stage least squares method. This shows that researchers should choose the best estimation method using appropriate econometric instruments.

Further, this article is structured as follows: in the section "Endogeneity" this problem is considered as such, and the applicability of methods to deal with it in CF is analyzed; The section "Causes of endogeneity and ways to deal with it" is devoted to the systematization of potential sources of endogeneity in CF and methods to overcome it, depending on its source. The third section, "Endogeneity in panel data models", contains a table describing the advantages and disadvantages of the most popular methods for dealing with endogeneity using panel data. The fourth section "Testing and dealing with endogeneity using the example of a model for assessing the impact of non-financial information disclosure on the cost of capital in the BRICS countries" is devoted to an empirical test of the applicability of the above methods. The last section contains conclusions from the theoretical and empirical parts of the paper.

ENDOGENEITY

Before moving on to the real reasons for the inconsistent and/or biased results of model building with endogeneity, it is worth noting that it may not have been present in many research papers. Wintoki et al. [2] cited articles in which endogeneity was present, but there were no ways to overcome it. However, they might have gotten this impression because of the style of presenting the methodology in the articles. As in all branches of economics, econometric models serve as a tool in CF, a detailed description of which is quite often neglected. This is because the main research result usually consists in the refutation or confirmation of the hypotheses posed by the researcher in the field of CF, and not in the application of the econometric model.[2] In such cases, the authors omit "extra" details, because they are forced to present the main material in a text of a limited size. The description of important econometric details depends on technical factors (requirements for the text length in journals) and accepted norms among scientists from different fields of economics. The issue of detailing the description of methodologies deserves to be noted in this article and a separate discussion. Nevertheless, the insufficiently detailed description of the methodology prompted the authors to focus on one of the important problems in CF — endogeneity.

By definition, endogeneity is the presence of a non-zero conditional expectation of the regression residual for a particular explanatory factor. The second more familiar definition is the presence of non-zero covariance between the explanatory factor and the regression error.

How does this occur in CF? The company solves complex problems. Decision-making on each of them is influenced by both external and internal corporate factors. As a result, there is a simultaneous influence of internal factors on each other, which is one of the reasons for endogeneity in CF models — simultaneity. We consider this and other reasons for a specific example. Often the issue of endogeneity is raised in capital structure studies, which is one of the main topics in CF. The CEO power can influence the capital structure. This is a relatively new explanatory factor that endogeneity may entail.

The reasons for this may be the following:

1) it is a hard-to-measure quantitative value, which may remain an omitted variable;

2) a proxy can be invented for it; however, it can be inaccurate and will reflect the CEO power with a big error. Or the proxy will take into account not only the influence of the CEO in the company (measurement error);

3) if researchers decide to use peer review to measure CEO power and other indicators, then a specific measurement error may arise — common-method variance (CMV);

[2] An exception is the papers in which the econometric model is applied for the first time within the framework of the task. In such cases, the authors usually describe each step in detail.

4) during the study, simultaneity may take place. In CF, it can arise, for example, when studying the factors affecting the absolute liquidity ratio. Trade-off theory states that financial leverage and liquidity ratio are determined simultaneously.

The presence of one of the four sources of endogeneity makes the estimates of the coefficients biased and inconsistent, which also makes it difficult to interpret the model as a whole. Moreover, the inconsistency gets worse as the number of observations increases. And large samples (more than a thousand observations) are characteristic of studies in the field of CF published during the last decade.

In turn, Grisser and Hadlock [22] emphasize that a large number of studies in the field of CF do not test the model for endogeneity (in a strong or weak form), as well as the application of methods to deal with this problem. Gippel et al. [3, p. 144] note that only three out of 30 studies use tests to detect endogeneity and methods for leveling it. Most of the studies simply mention its possible presence. Although the volume of articles on CP with the mention of the word "endogeneity" has increased, a similar phenomenon and incorrect ways to deal with endogeneity continue to be observed not only in Australian articles but also in articles around the world [23–25].

However, in the last 5–7 years, CF researchers have come across the problem of endogeneity much more often. This can be caused by several reasons:

- researchers [5, 16] recognize that endogeneity may be one of the possible causes of the problem of inconsistency and incomparability of results;
- the number of articles devoted to the problem of endogeneity has increased;
- the growing interest in regressions to evaluate panel data, which have become widely used to work with CF endogeneity.

If we look at methodological econometric papers published around the same time as Gippel et al. [3], they are rarely cited in empirical studies on CF[3] (*Fig.*).

A small number of citations may be due to the fact that they were published in journals not devoted to the topic of CF: over the past three years, methodological articles [22, 26–29] from CF journals have been cited more often than articles from econometric journals [30–34].

This article contributes to methodological research on endogeneity in CF and adds to the existing articles as follows:

1. This article lists possible sources of endogeneity in CF studies.
2. This article highlights tests for endogeneity and ways to deal with it.
3. Like Flannery, Hankins [5] and Campbell, Nagel [35], we illustrate with real data the possible bias of the study results that does not take into account endogeneity. In contrast to these works, this illustration was made as part of solving the problem of assessing the impact of issuing non-financial reporting on the cost of equity of companies based on data of 663 companies from the BRICS countries for 2007–2016.

CAUSES OF ENDOGENEITY AND WAYS TO DEAL WITH IT

As mentioned above, there are 4 causes for endogeneity: measurement error variables, common method variance, omitted variables, and simultaneity. Endogeneity due to any of these causes can be detected using the Hausman test [36], also called the Durbin-Wu-Hausman (DWH) test, and the Wooldridge test [37].

The Hausman test [36] consists in comparing the reliability of estimates. In the case of endogeneity, the least squares method (LSM) estimates and instrumental variables (IV) are compared with the endogenous variable. If there is no difference between them in terms of consistency, then the null hypothesis is not rejected, the variable is exogenous. The DWH test is presented differently by the authors,

[3] We refer to citations for papers specifically on the method, and not to the result of applying the method. For example, Dang et al. (2015) [26] are cited in 67 papers. Most of the citations are related to the degree of influence of SOA (speed of adjustment) on the level of leverage. Only in some papers, researchers refer to the effectiveness of applying one or another method, depending on the presence of endogeneity and other problems in the data.

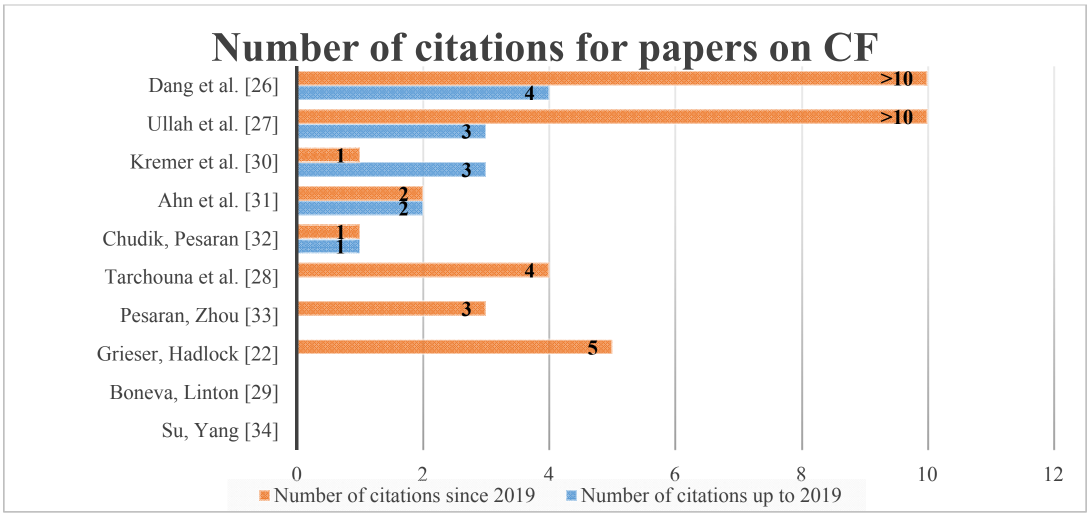

Fig. 1. **Citation rate of methodological publications on the use of panel models with endogeneity**

Source: authors' analysis of publications in Scopus.

however, if the errors are homoscedastic, the options of the three investigators are equivalent. In turn, the Wooldridge test [37] is the right instrument to test for endogeneity in the presence of heteroscedasticity. The null hypothesis is similar to the null hypothesis of the Hausman test.

Measurement errors

Measurement errors can be defined as the use of error-measured regressors instead of real regressors. In CF, measurement errors can be associated with:

- source data errors (for example, incorrectly loaded data from financial statements);
- errors in data aggregation (for example, irrelevant peer companies were selected to calculate the discount factor; an error in the formula when calculating the average capitalization for the industry);
- errors related to measurement methods (for example, CEO skills are poorly measured by their total work experience).

Measurement errors can be avoided by some common preventive measures to check the obtained data. This is controlled at the stage of uploading data and performing calculations to obtain exogenous factors. The control of the factor at the stage of unloading is carried out by checking the methodology for calculating the variable in the data source. Checking one's own calculations is also a method of dealing with measurement errors. However, such preventive measures do not guarantee the absence of measurement error. Several methods for dealing with data containing measurement errors have been proposed by Zhang et al. [38] and Qin et al. [39]. They are grouped as follows:

- methods that directly correct the bias of naive estimates: the naive estimator correction method, simulation extrapolation (SIMEX);
- correction methods based on the likelihood function, generalized linear models;
- methods based on unbiased evaluation functions;
- method of instrumental variables.

It is worth remembering that these methods have limitations in application. Instrumental variables may not be available. When implementing SIMEX, it may be difficult to select an extrapolation function. Correction methods based on the likelihood function require assumptions about the true distribution of factors and measurement errors. The construction of unbiased estimating functions under given assumptions about the distribution of factors is also a non-trivial problem. In addition to these methods, there are other robust methods for estimating regression equations [20, 21], which are improved methods for estimating the parameters of generalized linear models.

In financial economics, there is extensive practice in evaluating models with errors.

Aggregate information on this practice can be found in Chen et al. [40]. In the financial economy, measurement errors are reduced with the help of:

- calculation of estimates of the generalized method of moments (GMM);
- Bayesian estimates;
- three stage least squares method (3SLS);
- linear structural relations equations (LISREL);
- a particular case of LISREL is the Multiple Indicator, Multiple Cause (MIMIC) model.

Common method variance

The common method variance, CMV, is detailed by Wall et al. [41]. In fact, CMV is a pseudo-connection between variables, which is due to the presence of a factor — the source of the values of these variables. If a researcher were to conduct a survey among experts regarding the company performance on various parameters, the pessimistic experts would underestimate the estimates for all parameters. This problem is mainly related to personal data. CF researchers potentially rarely encounter such a source of endogeneity. If it is necessary to eliminate such bias in models, one should turn to articles in psychology or sociology [41–43]. Approaches to the compilation of survey materials and their processing are described in various sources, the most cited among which is the article by Podsakoff et al. [42]. Methods for checking CMV and some ways to deal with it are described in articles [41, 43].

Omitted variables

Omitted variables are one variation of the model specification error. The main ways to deal with the omitted variable error are to change the specification of the model and apply the IV method. A specification change can be expressed both in the transition to a completely alternative specification and in the transition to structural equation models (SEM).

Radically changing the model specification is an expensive way to solve the problem of omitted variables. This method will require a lot of time and resources. However, the main obstacle in its implementation is usually the lack of theory of using alternative model specifications. As for instrumental variables, these are variables that should have two properties: strongly correlated with the endogenous explanatory factor and weakly correlated with the model error. Although finding the right instrument is not a trivial task, this method is quite popular in articles. Over the past 5 years, the Journal of Corporate Finance has published about 270 articles on CF using the IV method.[4] This is partly due to the fact that there are some universal tools: endogenous variable lags or high-order moments. Instrumental variables can also be searched in new databases: specific industry and regional indicators can be found in the Observatory of Economic Complexity database,[5] CEO information can be found in the BoardEx of Management Diagnostics Limited,[6] Standard & Poor's ExecuComp database, etc.

Estimation methods with instrumental variables are usually instrumental variable method (IV), which includes Jackknife corrections (Jackknife IV estimates, described in the paper by Carrasco et al. [44]), two-stage least squares (2SLS, described in the paper by Schaffer [45]) and the generalized method of moments (GMM, described in the paper by Schaffer [45]), which is recommended for potentially overdetermined models. There is also a three-stage least squares method (3SLS, described in the paper by Qian, Schmidt [46]), the application of which will change the specification of the model to a system of externally uncoupled equations (SUR). Fixed effect models have become popular as a method of solving the omitted variable problem associated with each company. They clearly include the individual effect for each firm. However, this method is not a panacea, as described in more detail by Campbell, Nagel [35].

[4] Search results for articles with the keyword "instrumental variable" in the Journal of Corporate Finance. URL: https://www.sciencedirect.com/search?qs=%22instrumental%20variable%22&pub=Journal%20of%20Corporate%20Finance&cid=271687&years=2021%2C2020%2C2019%2C2018%2C2017&lastSelectedFacet=years (accessed on 06.07.2021).

[5] URL: https://oec.world/en/ (accessed on 06.07.2021).

[6] URL: https://corp.boardex.com/ (accessed on 06.07.2021).

Simultaneity

This problem is related to the fact that not one endogenous variable appears simultaneously in the model, but two or more. In other words, in one unit of time, both the explained variable and some explanatory factors are determined. This is also manifested in an unobvious causality between endogenous factors. In CF, simultaneity arises when considering the following questions:

- non-financial reports add value to the firm or high-value firms issue non-financial reports;
- the owners of the company set the level of credit risk the company accepts. Or can a certain level of the company's credit risk attract new owner investors?

As with the previous problem, lag variable models and systems of simultaneous equations such as vector error correction models (VECM) or vector autoregression (VAR) can be used to work with simultaneity).

The inclusion of an endogenous explanatory variable in the lag model is a fairly popular and effective option for dealing with endogeneity. However, some of its applications do not eliminate the bias caused by simultaneity but exacerbate it. Reed [47] showed that in the process of overcoming endogeneity, researchers evaluated models simultaneously, although there was no evidence for this. He also demonstrated that endogenous factor lags are better used as instrumental variables. They showed that lags are an effective instrument if they are sufficiently correlated with the variable being explained, and lag values are not included in the initially estimated equation. In other words, if the equation was originally estimated:

$$Y_t = \alpha + bX_t + cX_{t-1} + \varepsilon_t, \qquad (1)$$

where Y_t, X_t — time series, $\varepsilon_t \sim N(0;\sigma_y)$, and the coefficients *b* and *c* can be equal to zero. Many researchers have replaced[7] specification with:

$$Y_t = \alpha + \beta X_{t-1} + \text{error term}. \qquad (2)$$

But there is no reason to replace with, because X may not precede Y [29, p. 898–902]. And one should evaluate (3):

$$Y_t = \alpha + \beta X_t + \text{error term}, \qquad (3)$$

using the IV method. In this case, the lag of the endogenous explanatory factor (X_{t-1}). will be a strong instrument. These results are also confirmed for panel models.

ENDOGENEITY IN PANEL DATA MODELS

In the course of developing models for panel data, methods for their evaluation were developed that make it possible to successfully deal with endogeneity, in particular, a model with a fixed effect (FE). Also, the method of instrumental variables was adapted for panel data. At the end of the 20th century, Arellano Bond and Blundell Bond [48] developed methods of the same name for estimating dynamic panel models (Arellano-Bond — AB and Blundell-Bond — BB). In addition to these methods, methodologies proposed in other studies can work successfully with endogenous variables [49, 50]. These 4 methods (FE, IV, AB, BB) were comprehensively compared by Flannery, Hankins [5]. They tested which of the dynamic panel regression estimates gave the smallest root mean square error (RMSE). The study was conducted on real and simulated data on CF, while the simulated data was created with the following imperfections: unbalanced panel, heterogeneity, and various types of endogeneity. The results of the study showed that the BB methodology in most cases gave the lowest RMSE values (*Table 1*).

In addition, BB proved to be effective in heterogeneity as well. Flannery, Hankins [5] and Dang et al. [26] showed that better estimates are obtained only when using the bootstrap technique for a fixed effect model. Nevertheless, the BB method is currently in demand among researchers [6, 52–54].

However, it is always worth sticking to common sense and not relying on the most popular methods. First, researchers must be sure that they are using the best model available today. Just as in assessing

[7] Some of them are listed in Reed, 2015, p. 897–898.

Table 1

Comparison of methodologies for panel models' estimation used in CF

Methods	Description
OLS	Not applicable
FE	The coefficient accuracy is low when working with wide (short) panels. Exogenous regressors are estimated quite accurately in the presence of regression lags, and the coefficients for lags are estimated biased. However, in the presence of second-level lags, it shows more reliable results than the Arellano-Bond and Blundell-Bond methods. Ineffective in the presence of weakly changing explanatory factors
LSDVC [51]	Despite the semantic identity of the coefficients in the case of FE models and the consistency of estimates with an unbalanced panel, this option is not suitable for estimating panel models with endogenous variables
IV (2SLS)	Two-stage least squares method is the traditional way to evaluate regressions to deal with endogeneity. Required prerequisites: homoscedasticity, including conditional, and does not imply work in systems of equations with dependent errors (simultaneous systems). Potentially can be used to deal with omitted variables, simultaneity, and measurement errors
Arellano-Bond	Refers to GMM estimates, which, together with the specification, allow model evaluation with endogenous variables. Perfect for evaluating models with instruments — lags. If the lag values of the explanatory factors are weak instruments, the estimates are biased
Blundell-Bond	Refers to GMM estimates, which, together with the specification, allow model evaluation with endogenous variables. Similar to the AB approach, it gives inconsistent estimates in the presence of second-order autocorrelation. Shows more effective results in the presence of weakly variable explanatory factors. Other things being equal, it is invariant to the degree of endogeneity, in contrast to AB, and is the preferred method of assessment, because it gives the lowest RMSE in the presence of different sources of endogeneity
Long difference [49]	Unlike the previously mentioned approaches, lags of the second and higher order can be used as instruments. The authors suggest choosing the largest of the available lags. The method showed average results in the presence of endogeneity and other issues
Four period differencing [50]	Unlike Hahn et al. [49], it is recommended to use not the largest available lag. The method showed average results in the presence of endogeneity and other issues

Source: authors' analysis [5].

it, is it reasonable to use a very sophisticated instrument to deal with endogeneity when a relatively simple problem is being solved? Should one immediately take an advanced method that has already been tested by someone, or should one first check the alternatives? The question is open whether at least it is worth showing in the study if the problem of endogeneity was discovered, if it really exists, and what methods were used to overcome it. And what if one managed to overcome it, and the tests say that everything is in order? Maybe not entirely. It is possible that there was more than one source of endogeneity. Based on this principle, in the next section, we will show that the Blundell-Bond method is not always the preferred method for estimating panel regressions in CF. Although often used without comparison with alternative methods.

TESTING AND DEALING WITH ENDOGENEITY USING THE EXAMPLE OF A MODEL FOR ASSESSING THE IMPACT OF NON-FINANCIAL INFORMATION DISCLOSURE ON THE COST OF CAPITAL IN THE BRICS COUNTRIES

As an example of testing and dealing with endogeneity, we use a model for assessing the impact of non-financial information disclosure on the cost of equity. To do this, we use a sample of 663 companies from the BRICS countries from 2007–2016 (balanced panel). The topic of assessing the impact of non-financial information disclosure, including corporate social responsibility (CSR) and environmental, social, and corporate governance (ESG), began to gain popularity in the previous decade and is still relevant today. This is evidenced by numerous articles on this subject [55–64]. Non-financial information is disclosed in non-financial reports (hereinafter referred to as NFR). The sample includes the release of NFR according to the most popular standards: the Global Reporting Initiative (GRI)[8] and the International Integrated Reporting Council,[9] which recently included the International Integrated Reporting Committee (IIRC)[10] and the Sustainability Reporting Standards Board (SASB).[11]

This topic was chosen as an illustrative example, as many researchers in this field [64, 65] drew attention to the presence of an endogeneity problem in models. A meaningful interpretation of the endogeneity problem is: could the companies that issued NFR raise capital at a lower rate, or did companies that had access to cheaper equity issue NFR?

Model

Based on a number of studies [66–70], a model for assessing the impact of information disclosure in non-financial reporting on the cost of capital was formed, the description of the variables of which is presented in *Table 2*.

$$COE_{i,t} = \beta_0 + \beta_1 NFR_{i,t} + \beta_2 Size_{i,t} + \beta_3 M/B_{i,t} + \\ + \beta_4 Lev_{i,t} + \beta_5 LTG_{i,t} + \beta_6 ROA_{i,t} + \sum_{j-1} Industry_j + \\ + \sum_{t-1} Year_t + \sum_{k-1} Country_k + \varepsilon_{i,t}.$$

Sample

The sample contains 6630 observations of 663 companies over 10 years: 2007–2016. The data was collected from the Bloomberg database and a GRI database.[12] Descriptive statistics of variables are presented in *Table 3*.

The number of observations for the variable long-term growth rate (LTG) is 1326 fewer (663 observations each for 2007 and 2008), since data for 3 years were used in the calculation. The correlation matrix is presented in the *Appendix* (*Table 1*). *Tables 2* and *3* of the *Appendix* show the distribution of companies by country and industry, respectively. Financial sector companies were excluded from the sample due to their excellent balance sheet structure and additional government regulation.

Model evaluation and endogeneity testing

The models were evaluated using the STATA14 data analysis software package. The first step in the econometric analysis is the choice of a model estimation method. Let us compare the end-to-end model (pool) estimated by the least squares method without taking into account the individual effect with the estimates obtained by the generalized least squares method (GLM) with random (RE) and individual fixed (FE) effects (*Table 4*).

In the FE model, the inclusion of variables that do not change over time is impossible, since they are taken into account in the individual fixed error.

Based on the results of Breusch-Pagan [73], Hausman [36], and the F-test of the FE model [74], the FE model was chosen. The test results are presented in *Table 5*.

The Breusch-Pagan test demonstrates that the null hypothesis is rejected and the inclusion of individual random error makes the estimates more accurate, in other words, the RE model is preferred to the pooled model. According to the results of the Hausman test, the main hypothesis is rejected (p-value < 0.01): the FE model describes the data better than the RE model. The chosen model is significant because the null hypothesis F of the test is rejected. Note that fixed effects

[8] Global Reporting Initiative. URL: https://www.globalreporting.org/ (accessed on 22.07.2021).

[9] Value Reporting Foundation. URL: https://www.valuereportingfoundation.org/ (accessed on 22.07.2021).

[10] International Integrated Reporting Council. URL: https://integratedreporting.org/ (accessed on 22.07.2021).

[11] Accounting Standards Board. URL: https://www.sasb.org/ (accessed on 22.07.2021).

[12] URL: https://www.globalreporting.org/reporting-support/reporting-tools/sustainability-disclosure-database/ (accessed on 22.07.2021).

Table 2

Description of the variables used in the model

Variable	Description	Impact
$COE_{i,t}$	A company's cost of equity is calculated using the capital asset pricing model (CAPM [71]): $COE = r_f + \beta * (r_m - r_f)$, where r_f – the risk-free rate measured as a yield on 10-year government bonds, r_m — market return; $\beta = \frac{cov(COE, r_m)}{var(r_m)}$ – a measure of systematic risk	
$NFR_{i,t}$	The binary variable NFR is 1 if the company issued an NFR in the current year, and 0 otherwise	–
$Size_{i,t}$	Natural logarithm of a company's total assets at the end of the year	–
$M / B_{i,t}$	The market-to-book ratio of the company at the end of the year	–
$Lev_{i,t}$	Financial leverage, measured as the debt-to-equity ratio at the end of the year	+
$LTG_{i,t}$	Long-term growth rate calculated as the average of 3-year sales growth rates	–
$ROA_{i,t}$	Return on assets, calculated as the company's net profit-to-assets ratio at the end of the year	–
$Industry_j$	Binary variable for industries	
$Country_k$	Binary variable for countries	
$Year_t$	Binary variable for years	

Source: Evdokimova M.S., Kuzubov S.A. [72].

models are a widely used way of estimating panel data in CF.

The obtained results indirectly show the presence of endogeneity: the FE model turned out to be the best, i.e. the usual error ($\varepsilon_{i,t}$) is not enough to correctly estimate the model. Further analysis using a modified Wald test revealed heteroscedasticity, which we will correct in the further analysis. Multicollinearity and spatial autocorrelation were not found. However, according to the Wooldridge test [75], there is serial autocorrelation. The feasible generalized least squares (FGLS) method can deal with this problem.

To test the hypothesis that the cost of capital decreases *after* the release of NFR the second model (*Table 6*) includes the NFR lag, which turned out to be significant. Therefore, the issuance of NFR leads to a reduction in the cost of equity, and not vice versa. Lags of higher order are insignificant. Note that financial leverage is significant only at the 15% significance level (p-value = 0.12). In the course of clarifying the reasons for the insignificance of this factor, a correlation was found between the residuals of the model and NFR equal to 0.34, which indicates the endogeneity of NFR. We also performed an endogeneity test, the null hypothesis of which is that the variable is exogenous. According to the results of this test, the endogeneity of NFR was again confirmed (*Table 7*).

Table 3

Descriptive statistics

Variables	Obser.	Average	Stand. error	Min.	Max.
COE	6.630	11.88	2.95	1.73	21.87
NFR	6.630	0.15	0.35	0	1
Size	6.630	6.75	1.38	2.86	11.48
ROA	6.630	5.03	6.47	−89.16	65.38
Leverage	6.630	0.35	0.55	0	11.95
LTG	5.304	7.83	11.24	−38.01	138.76
M/B	6.630	2.61	1.75	0.05	9.96

Source: authors' calculations.

Table 4

Estimates of pool model and GLS models with random (RE) and individual fixed effects (FE)

VARIABLES	pool	RE	FE
	Cost of equity		
NFR	−0.144	−0.452***	−0.730***
	(0.0917)	(0.103)	(0.117)
Size	0.0216	−0.0676	−0.551***
	(0.0263)	(0.0420)	(0.0950)
ROA	−0.0458***	−0.0245***	−0.0146***
	(0.00525)	(0.00526)	(0.00559)
Leverage	0.0658***	0.0845***	0.118***
	(0.0232)	(0.0303)	(0.0392)
LTG	−0.0155***	−0.00921***	−0.00555**
	(0.00280)	(0.00260)	(0.00268)
M/B	0.0871***	−0.139***	−0.266***
	(0.0200)	(0.0218)	(0.0240)
Dummies for years	Yes	Yes	Yes
Dummies for countries	Yes	Yes	No
Dummies for industries	Yes	Yes	No
	(0.117)	(0.0945)	(0.101)
Constant	11.47***	12.27***	14.51***
	(0.199)	(0.303)	(0.593)
Observations	5.304	5.304	5.304
R-squared	0.366		0.450
Number of Id		663	663

Note: Standard errors in parentheses; *** p < 0.01, ** p < 0.05, * p < 0.1.

Source: authors' calculations.

Table 5

Results of Breusch-Pagan, Hausman tests and F-test of FE model

Test	(1) Breusch-Pagan	(2) Hausman	(3) F-test of FE model
H0	No individual random effect	Complex error model with $cov(u_i, X_{it}) = 0$ – RE correctly specified	No individual fixed effect
Statistics	$\bar{\chi}^2(1)$ = 2222.33	$\chi^2(6)$ = 175.95	F (13,4628) = 290.77
P-value	0.0000	0.0000	0.0000

Source: authors' calculations.

Table 6

FGLS FE estimation taking into account serial autocorrelation and heteroscedasticity

VARIABLES	FGLS,	
	(1) NFR	(2) lag NFR
	Cost of equity	
NFR	−0.202**	
	(0.0952)	
L.NFR		−0.253**
		(0.103)
Size	0.109***	0.111***
	(0.0264)	(0.0260)
ROA	−0.0698***	−0.0699***
	(0.00531)	(0.00530)
Leverage	0.0368	0.0364
	(0.0237)	(0.0237)
LTG	−0.0194***	−0.0194***
	(0.00291)	(0.00291)
M/B	0.101***	0.101***
	(0.0199)	(0.0199)
Dummies for years	Yes	Yes
Constant	10.23***	10.21***
	(0.191)	(0.190)
Observations	5.304	5.304
Number of Id	663	663

Note: Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: authors' calculations.

Table 7

NFR endogeneity test

NFR endogeneity test statistics	10.634
P-value	0.0011

Source: authors' calculations.

DEALING WITH ENDOGENEITY

To deal with the NFR endogeneity, we turn to the method of instrumental variables (IV), the essence of which is to select instruments that are correlated with the endogenous variable, but not correlated with the residuals of the model. The search for instruments was carried out using the analysis of the extended correlation matrix, including the saved residuals of the FGLS model. The results of model estimation by the method of instrumental variables with NFR lag and financial leverage lag used as instruments are presented in *Table 8*.

It is worth noting that the IV method on panel data, when using a model with individual fixed effects, does not allow adding a constant to the model. Because of this, the value of the coefficients for binary variables for years has increased. To assess the quality of instruments, we have to perform several tests:

1. Underidentification test, its null hypothesis: instruments have insufficient explanatory power to predict endogenous variables [76].

2. Weak identification test checks whether instrumental variables explain the endogenous variable. Instruments are considered strong if the Cragg-Donald Wald F statistic exceeds all critical Stock-Yogo values [77].

3. Hansen-Sargan's J overidentification test [76]. Null hypothesis: instruments are valid (not correlated with errors).

In our case, all three tests were successfully passed, therefore, the instruments are strong (*Table 9*).

However, we established a second-order autocorrelation (AR (2)), to overcome this we included the first and second lags of the cost of equity in the instrumental model, which solved the problem of endogeneity. The results of the evaluation of the final model are presented in *Table 10*.

The instruments for the NFR are the NFR lag, the financial leverage lag, and the cost of equity lag used in the first stage of the 2SLS valuation. All three tests for this specification were successfully passed (*Table 11*), the instruments are strong, and the problem of endogeneity is solved (*Table 12*).

The correlation between the lags of the cost of capital and the residuals of the final model does not exceed 0.01 (*Table 13*).

In the presence of serial autocorrelation and endogeneity, dynamic models of AB and BB or LSDVC estimates and the method of long differences (LD) can be used. However, the necessity of their application is justified by the

Table 8

IV FE estimation taking into account heteroscedasticity and endogeneity

VARIABLES	IV, Cost of equity
NFR	−1.303***
	(0.223)
Size	−0.455***
	(0.0886)
ROA	−0.0160**
	(0.00671)
Leverage	0.110**
	(0.0440)
LTG	−0.00656***
	(0.00254)
M/B	−0.259***
	(0.0244)
Dummies for years	Yes
Observations	5.304
Number of Id	663
R-squared	0.446

Note: Robust standard errors in parentheses; *** p < 0.01, ** p < 0.05, * p < 0.1.

Source: authors' calculations.

Table 9

Test results for evaluating the quality of instruments

Test	Underidentification test	Weak identification test	Overidentification test
Statistics	Kleibergen-Paap rk LM statistic = 290.956	Cragg-Donald Wald F statistic = 1090.522	Hansen J statistic = 0.363
P-value	0.000		0.5467
Stock-Yogo critical values:			
10% maximal IV size		19.53	
15% maximal IV size		11.59	
20% maximal IV size		8.75	
25% maximal IV size		7.25	

Source: authors' calculations.

presence of a non-zero correlation between the lags of the explained variable and the residuals of the model, which is not observed in our model. In other words, in the case of our application of AB, BB or LD at once, without gradually solving all the problems that violate the Gauss-Markov assumptions, we would never get unbiased coefficient estimates. Going to any of these models without testing the assumptions and characteristics of the simpler models would show that they are irrelevant. The BB estimates bias can be seen in *Table 14*.

Considering that the average value of the cost of capital in the analyzed sample is 11.88 percentage points, its reduction from NFR issuance ranges from 0.73 and 0.78 percentage points (6% of the average) in FE and final models, respectively, up to 1.3 percentage points (11% of the average) in the IV model without taking into account autocorrelation. The estimates differ almost by a factor of 2, which is essential for interpreting the results obtained and using them in practice.

Since no other endogenous variables were identified, the analysis of the illustrative example is completed on the final model, the results of which are given in *Table 10*.

Table 10

Final model: IV FE estimation taking into account heteroscedasticity, endogeneity and AR (2)

VARIABLES	IV with COE lags, Cost of equity
NFR	−0.784***
	(0.264)
L.COE	0.182***
	(0.0197)
L2.COE	−0.0970***
	(0.0145)
Size	−0.479***
	(0.131)
ROA	−0.0157**
	(0.00734)
Leverage	0.107**
	(0.0494)
LTG	−0.00337
	(0.00281)
M/B	−0.250***
	(0.0272)
Dummies for years	Yes
Observations	4.641
Number of Id	663
R-squared	0.446

Source: authors' calculations.

Note: Robust standard errors in parentheses; *** p < 0.01, ** p < 0.05, * p < 0.1.

CONCLUSIONS

Due to the complexity of company organization, CF models can potentially always contain

Table 11

Test results for evaluating the quality of instruments

Test	Underidentification test	Weak identification test	Overidentification test
Statistics	Kleibergen-Paap rk LM statistic = 209.207	Cragg-Donald Wald F statistic = 710.499	Hansen J statistic = 0.672
P-value	0.000		0.4123
Stock-Yogo critical values:			
10% maximal IV size		19.53	
15% maximal IV size		11.59	
20% maximal IV size		8.75	
25% maximal IV size		7.25	

Source: authors' calculations.

endogeneity. This paper continues the discussion about the presence of endogeneity in these models and how to deal with it. In contrast to previous works on this topic, our article provides examples of potential sources in different areas of research in the field of CF and considers various ways to deal with endogeneity.

In the course of a systematic review of the literature, we found that many researchers recommend using regressions to evaluate dynamic panel models, as they are the most successful in leveling the effects of various kinds of endogeneity. In particular, according to the results of the review, the Blundell-Bond method is the most appropriate.

However, when conducting an empirical study using methods to detect and overcome endogeneity, we found that this method may not be sufficient to eliminate endogeneity. As part of an empirical study, we tested the impact of NFR on the cost of capital of 633 companies from the BRICS countries from 2007–2016. Our result is consistent with previously published research on this topic: NFR has a negative impact on the cost of capital. However, this result was obtained using the method of instrumental variables on panel data (IV 2SLS). And the generalized least squares method with an individual fixed effect (FGLS FE), as well as BB, was not enough to overcome endogeneity.

According to the estimates obtained in this study, we found a non-critical change in the coefficient for the endogenous variable between FE and the final IV models, however, if we settled on the 2SLS IV estimate without taking into account autocorrelation, the coefficient would be 2 times larger. According

Table 12

NFR endogeneity test

NFR endogeneity test statistics	2.148
P-value	0.1428

Source: authors' calculations.

Table 13

Correlation of potentially endogenous variables with the final model's residuals

Variables	COE	L.COE	L2.COE	NFR	Residuals
COE	1.00				
L.COE	0.49***	1.00			
L2.COE	0.21***	0.47***	1.00		
NFR	−0.09**	−0.01***	0.09***	1.00	
Residuals	0.58***	0.00	0.00	0.02	1.00

Source: authors' calculations.

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 14

Beta coefficients of NFR variable in the estimated models

Model	FE	FGLS	IV	BB	IV final
NFR beta	−0.730*** (0.117)	−0.202** (0.0952)	−1.303*** (0.223)	−1.018*** (0.217)	−0.784*** (0.223)

Source: authors' calculations.

Note: Standard errors and robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

to the review of the literature, this is not the worst manifestation of endogeneity, since the absence of a correction for it can lead to an even greater coefficients bias, insignificance, and even a change in sign.

The results of our study will be primarily useful both for experienced and novice researchers of the cost of capital, and for CF researchers in general. Also, our study may be of interest to practitioners in the field of business valuation, and investors to evaluate and apply models that have a high predictive ability due to objective coefficient estimates. Also, this article can be useful when conducting seminars on CF or econometrics.

Research on the problem of endogeneity has great potential for development both from the point of view of researchers in the field of econometrics involved in the development of new models that take into account endogeneity and from the side of CF practitioners, many of whom are faced with the problem of choosing the best way to evaluate the model with endogeneity.

ACKNOWLEDGEMENTS

This paper is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE University). National Research University Higher School of Economics, Moscow, Russia.

REFERENCES

1. Tucker G.M. On problems of corporate finance. *The American Journal of Economics and Sociology*. 1948;7(2):235–236. DOI: 10.1111/J.1536–7150.1948.TB 00679.x
2. Wintoki M.B., Linck J.S., Netter J.M. Endogeneity and the dynamics of internal corporate governance. *Journal of Financial Economics*. 2012;105(3):581–606. DOI: 10.1016/j.jfineco.2012.03.005
3. Gippel J., Smith T., Zhu Y. Endogeneity in accounting and finance research: Natural experiments as a state-of-the-art solution. *Abacus*. 2015;51(2):143–168. DOI: 10.1111/abac.12048
4. Barros L.A.B.C., Bergmann D.R., Henrique Castro F., da Silveira A.D.M. Endogeneity in panel data regressions: Methodological guidance for corporate finance researchers. *Revista Brasileira de Gestao de Negocios*. 2020;22:437–461. DOI: 10.7819/rbgn.v22i0.4059
5. Flannery M.J., Hankins K.W. Estimating dynamic panel models in corporate finance. *Journal of Corporate Finance*. 2013;19:1–19. DOI: 10.1016/j.jcorpfin.2012.09.004
6. Banik A., Chatterjee C. Ownership pattern and governance-performance relation: Evidence from an emerging economy. *Global Business Review*. 2021;22(2):422–441. DOI: 10.1177/0972150920966699
7. Molina C.A. Are firms underleveraged? An examination of the effect of leverage on default probabilities. *The Journal of Finance*. 2005;60(3):1427–1459. DOI: 10.1111/j.1540–6261.2005.00766.x
8. Chen C.-W., Lin J.B., Yi B. CEO duality and firm performance: An endogenous issue. *Corporate Ownership and Control*. 2008;6(1):58–65. DOI: 10.22495/cocv6i1p6
9. Poletti Hughes J. R&D and dividend payments as determinants of corporate value in the UK: Empirical evidence after controlling for endogeneity. *International Journal of Managerial Finance*. 2008;4(1):76–91. DOI: 10.1108/17439130810837393
10. Zhou T., Li W.-A. Board governance and managerial risk taking: Dynamic analysis. *The Chinese Economy*. 2016;49(2):60–80. DOI: 10.1080/10971475.2016.1142823
11. Malik Q.A., Hussain S., Ullah N., Waheed A., Naeem M., Mansoor M. Simultaneous equations and endogeneity in corporate finance: The linkage between institutional ownership and corporate financial performance. *The Journal of Asian Finance, Economics and Business*. 2021;8(3):69–77. DOI: 10.13106/jafeb.2021.vol8.no3.0069
12. Harada K., Nguyen P. Ownership concentration and dividend policy in Japan. *Managerial Finance*. 2011;37(4):362–379. DOI: 10.1108/03074351111115313
13. Zhukova N. Yu., Melikova A.E. Corporate social responsibility: Strengthening brand value and affecting company's financial performance. *Finance: Theory and Practice*. 2021;25(1):84–102. DOI: 10.26794/2587–5671–2021–25–1–84–102

14. Martynova M. Climate disclosure as a factor of investment attractiveness of Russian companies. *Moskovskii economicheskii zhurnal = Moscow Economic Journal*. 2021(5):365–379. (In Russ.). DOI: 10.24411/2413–046X-2021–10277
15. Polyakov K.L., Polyakova M.V., Samoylenko S.V. Modeling the impact of the debt load on efficiency of business entities. *Voprosy statistiki*. 2016;(9):17–29. (In Russ.). DOI: 10.34023/2313–6383–2016–0–9–17–29
16. Zahid M., Rahman H.U., Khan M., Ali W., Shad F. Addressing endogeneity by proposing novel instrumental variables in the nexus of sustainability reporting and firm financial performance: A step-by-step procedure for non-experts. *Business Strategy and the Environment*. 2020;29(8):3086–3103. DOI: 10.1002/bse.2559
17. Cheng Q., Lo K. Insider trading and voluntary disclosures. *Journal of Accounting Research*. 2006;44(5):815–848. DOI: 10.1111/j.1475–679X.2006.00222.x
18. Gul F.A., Hutchinson M., Lai K.M.Y. Gender-diverse boards and properties of analyst earnings forecasts. *Accounting Horizons*. 2013;27(3):511–538. DOI: 10.2308/acch-50486
19. Kim C.-J., Nelson C.R. Estimation of a forward-looking monetary policy rule: A time-varying parameter model using ex post data. *Journal of Monetary Economics*. 2006;53(8):1949–1966. DOI: 10.1016/j.jmoneco.2005.10.017
20. Coles J.L., Lemmon M.L., Meschke J.F. Structural models and endogeneity in corporate finance: The link between managerial ownership and corporate performance. *Journal of Financial Economics*. 2012;103(1):149–168. DOI: 10.1016/j.jfineco.2011.04.002
21. Li F. Endogeneity in CEO power: A survey and experiment. *Investment Analysts Journal*. 2016;45(3):149–162. DOI: 10.1080/10293523.2016.1151985
22. Grieser W.D., Hadlock C.J. Panel-data estimation in finance: Testable assumptions and parameter (in) consistency. *Journal of Financial and Quantitative Analysis*. 2019;54(1):1–29. DOI: 10.1017/S 0022109018000996
23. Kirshin I.A. An empirical analysis of determinant factors of the company's capital structure. *Finansy: teoriya i praktika = Finance: Theory and Practice*. 2017;21(2):106–112. (In Russ.). DOI: 10.26794/2587–5671–2017–21–2–106–112
24. Mohammad W.M.W., Wasiuzzaman S. Environmental, social and governance (ESG) disclosure, competitive advantage and performance of firms in Malaysia. *Cleaner Environmental Systems*. 2021;2:100015. DOI: 10.1016/J.CESYS.2021.100015
25. Alareeni B.A., Hamdan A. ESG impact on performance of US S&P 500-listed firms. *Corporate Governance*. 2020;20(7):1409–1428. DOI: 10.1108/CG-06–2020–0258
26. Dang V.A., Kim M., Shin Y. In search of robust methods for dynamic panel data models in empirical corporate finance. *Journal of Banking & Finance*. 2015;53:84–98. DOI: 10.1016/j.jbankfin.2014.12.009
27. Ullah S., Akhtar P., Zaefarian G. Dealing with endogeneity bias: The generalized method of moments (GMM) for panel data. *Industrial Marketing Management*. 2018;71:69–78. DOI: 10.1016/j.indmarman.2017.11.010
28. Tarchouna A., Jarraya B., Bouri A. How to explain non-performing loans by many corporate governance variables simultaneously? A corporate governance index is built to US commercial banks. *Research in International Business and Finance*. 2017;42:645–657. DOI: 10.1016/j.ribaf.2017.07.008
29. Boneva L., Linton O. A discrete-choice model for large heterogeneous panels with interactive fixed effects with an application to the determinants of corporate bond issuance. *Journal of Applied Econometrics*. 2017;32(7):1226–1243. DOI: 10.1002/jae.2568
30. Kremer S., Bick A., Nautz D. Inflation and growth: New evidence from a dynamic panel threshold analysis. *Empirical Economics*. 2013;44(2):861–878. DOI: 10.1007/s00181–012–0553–9
31. Ahn S.C., Lee Y.H., Schmidt P. Panel data models with multiple time-varying individual effects. *Journal of Econometrics*. 2013;174(1):1–14. DOI: 10.1016/j.jeconom.2012.12.002
32. Chudik A., Pesaran M.H. Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics*. 2015;188(2):393–420. DOI: 10.1016/j.jeconom.2015.03.007
33. Pesaran M.H., Zhou Q. Estimation of time-invariant effects in static panel data models. *Econometric Reviews*. 2018;37(10):1137–1171. DOI: 10.1080/07474938.2016.1222225
34. Su L., Yang Z. QML estimation of dynamic panel data models with spatial errors. *Journal of Econometrics*. 2015;185(1):230–258. DOI: 10.1016/j.jeconom.2014.11.002

35. Campbell R.C., Nagel G.L. Private information and limitations of Heckman's estimator in banking and corporate finance research. *Journal of Empirical Finance*. 2016;37:186–195. DOI: 10.1016/j.jempfin.2016.03.007
36. Hausman J.A. Specification tests in econometrics. *Econometrica*. 1978;46(6):1251–1271. DOI: 10.2307/1913827
37. Wooldridge J.M. Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics*. 1995;68(1):115–132. DOI: 10.1016/0304–4076(94)01645-G
38. Zhang Y., Qin G., Zhu Z., Zhang J. Robust estimation in linear regression models for longitudinal data with covariate measurement errors and outliers. *Journal of Multivariate Analysis*. 2018;168:261–275. DOI: 10.1016/j.jmva.2018.07.015
39. Qin G., Zhang J., Zhu Z. Simultaneous mean and covariance estimation of partially linear models for longitudinal data with missing responses and covariate measurement error. *Computational Statistics & Data Analysis*. 2016;96:24–39. DOI: 10.1016/j.csda.2015.11.001
40. Chen H.-Y., Lee A.C., Lee C.-F. Alternative errors-in-variables models and their applications in finance research. *Quarterly Review of Economics and Finance*. 2015;58:213–227. DOI: 10.1016/j.qref.2014.12.002
41. Wall T.D., Michie J., Patterson M., Wood S.J., Sheehan M., Clegg C.W., West M. On the validity of subjective measures of company performance. *Personnel Psychology*. 2004;57(1):95–118. DOI: 10.1111/j.1744–6570.2004.tb02485.x
42. Podsakoff P.M., MacKenzie S.B., Lee J.-Y., Podsakoff N.P. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*. 2003;88(5):879–903. DOI: 10.1037/0021–9010.88.5.879
43. Tehseen S., Ramayah T., Sajilan S. Testing and controlling for common method variance: A review of available methods. *Journal of Management Sciences*. 2017;4(2):146–175. DOI: 10.20547/jms.2014.1704202
44. Carrasco M., Doukali M. Efficient estimation using regularized jackknife IV estimator. *Annals of Economics and Statistics*. 2017;(128):109–149. DOI: 10.15609/annaeconstat2009.128.0109
45. Schaffer M.E. XTIVREG2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models. 2020.
46. Qian H., Schmidt P. The asymptotic equivalence between the iterated improved 2sls estimator and the 3sls estimator. *Econometric Reviews*. 1997;16(4):441–457. DOI: 10.1080/07474939708800398
47. Reed W.R. On the practice of lagging variables to avoid simultaneity. *Oxford Bulletin of Economics and Statistics*. 2015;77(6):897–905. DOI: 10.1111/obes.12088
48. Roodman D.M. How to do Xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*. 2009;9(1):86–136. DOI: 10.1177/1536867X0900900106
49. Hahn J., Hausman J., Kuersteiner G. Long difference instrumental variables estimation for dynamic panel models with fixed effects. *Journal of Econometrics*. 2007;140(2):574–617. DOI: 10.1016/j.jeconom.2006.07.005
50. Huang R., Ritter J.R. Testing theories of capital structure and estimating the speed of adjustment. *Journal of Financial and Quantitative Analysis*. 2009;44(2):237–271. DOI: 10.1017/S 0022109009090152
51. Bruno G.S.F. Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models. *Economics Letters*. 2005;87(3):361–366. DOI: 10.1016/j.econlet.2005.01.005
52. Ezeani E., Salem R., Kwabi F., Boutaine K., Bilal, Komal B. Board monitoring and capital structure dynamics: evidence from bank-based economies. *Review of Quantitative Finance and Accounting.* 2022;58(2):473–498. DOI: 10.1007/S 11156–021–01000–4
53. Lartey T., Danso A., Boateng A. Co-opted boards and capital structure dynamics. *International Review of Financial Analysis*. 2021;77:101824. DOI: 10.1016/J.IRFA.2021.101824
54. Badayi S.A., Matemilola B.T., Bany-Ariffin A.N., Theng L.W. Does corporate social responsibility influence firm probability of default? *International Journal of Finance & Economics*. 2021;26(3):3377–3395. DOI: 10.1002/IJFE.1966
55. Schadewitz H., Niskala M. Communication via responsibility reporting and its effect on firm value in Finland. *Corporate Social Responsibility and Environmental Management*. 2010;17(2):96–106. DOI: 10.1002/csr.234
56. de Klerk M., de Villiers C. The value relevance of corporate responsibility reporting: South African evidence. *Meditari Accountancy Research*. 2012;20(1):21–38. DOI: 10.1108/10222521211234200

57. Iatridis G.E. Environmental disclosure quality: Evidence on environmental performance, corporate governance and value relevance. *Emerging Markets Review*. 2013;14:55–75. DOI: 10.1016/j.ememar.2012.11.003
58. Fatemi A., Fooladi I., Tehranian H. Valuation effects of corporate social responsibility. *Journal of Banking & Finance*. 2015;59:182–192. DOI: 10.1016/j.jbankfin.2015.04.028
59. Oshika T., Saka C. Sustainability KPIs for integrated reporting. *Social Responsibility Journal*. 2017;13(3):625–642. DOI: 10.1108/SRJ-07–2016–0122
60. de Villiers C., Venter E.R., Hsiao P.-C.K. Integrated reporting: background, measurement issues, approaches and an agenda for future research. *Accounting & Finance*. 2017;57(4):937–959. DOI: 10.1111/acfi.12246
61. García-Sánchez I.-M., Noguera-Gámez L. Integrated information and the cost of capital. *International Business Review*. 2017;26(5):959–975. DOI: 10.1016/j.ibusrev.2017.03.004
62. Albuquerque R., Koskinen Y., Zhang C. Corporate social responsibility and firm risk: Theory and empirical evidence. *Management Science*. 2019;65(10):4451–4469. DOI: 10.1287/mnsc.2018.3043
63. Bhuiyan M.B.U., Nguyen T.H.N. Impact of CSR on cost of debt and cost of capital: Australian evidence. *Social Responsibility Journal*. 2020;16(3):419–430. DOI: 10.1108/SRJ-08–2018–0208
64. Yeh C.-C., Lin F., Wang T.-S., Wu C.-M. Does corporate social responsibility affect cost of capital in China? *Asia Pacific Management Review*. 2020;25(1):1–12. DOI: 10.1016/j.apmrv.2019.04.001
65. Manchiraju H., Rajgopal S. Does corporate social responsibility (CSR) create shareholder value? Evidence from the Indian Companies Act 2013. *Journal of Accounting Research*. 2017;55(5):1257–1300. DOI: 10.1111/1475–679X.12174
66. Dhaliwal D.S., Li O.Z., Tsang A., Yang Y.G. Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review*. 2011;86(1):59–100. DOI: 10.2308/accr.00000005
67. El Ghoul S., Guedhami O., Kwok C.C.Y., Mishra D.R. Does corporate social responsibility affect the cost of capital? *Journal of Banking & Finance*. 2011;35(9):2388–2406. DOI: 10.1016/j.jbankfin.2011.02.007
68. Boujelbene M.A., Affes H. The impact of intellectual capital disclosure on cost of equity capital: A case of French firms. *Journal of Economics, Finance and Administrative Science*. 2013;18(34):45–53. DOI: 10.1016/S 2077–1886(13)70022–2
69. Zhou S., Simnett R., Green W. Does integrated reporting matter to the capital market? *Abacus*. 2017;53(1):94–132. DOI: 10.1111/abac.12104
70. Fama E.F., French K.R. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*. 1993;33(1):3–56. DOI: 10.1016/0304–405X(93)90023–5
71. Fama E.F., French K.R. The cross-section of expected stock returns. *The Journal of Finance*. 1992;47(2):427–465. DOI: 10.1111/j.1540–6261.1992.tb04398.x
72. Evdokimova M.S., Kuzubov S.A. Non-financial reporting and the cost of capital in BRICS countries. Higher School of Economics. Basic Research Program Working Paper. 2021;(83/FE). URL: https://wp.hse.ru/data/2021/06/25/1430000555/83FE 2021.pdf
73. Breush T.S., Pagan A.R. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*. 1979;47(5):1287–1294. DOI: 10.2307/1911963
74. Park H.M. Practical guides to panel data modeling: A step by step analysis using stata. Niigata: International University of Japan; 2011. 53 p. URL: https://www.iuj.ac.jp/faculty/kucc625/method/panel/panel_iuj.pdf
75. Drukker D.M. Testing for serial correlation in linear panel-data models. *The Stata Journal*. 2003;3(2):168–177. DOI: 10.1177/1536867X0300300206
76. Baum C.F., Schaffer M.E., Stillman S. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *The Stata Journal*. 2007;7(4):465–506. DOI: 10.1177/1536867X0800700402
77. Stock J.H., Yogo M. Testing for weak instruments in linear IV regression. In: Andrews D.W.K., Stock J.H., eds. Identification and inference for econometric models: Essays in honor of Thomas Rothenberg. Cambridge, New York: Cambridge University Press; 2005:80–108. DOI: 10.1017/CBO9780511614491.006

Appendix

Table 1

Correlation matrix

Variables	COE	NFR	Size	ROA	Leverage	LTG	M/B
COE	1.00						
NFR	−0.03***	1.00					
Size	0.04***	0.43***	1.00				
ROA	−0.15***	0.09***	−0.03**	1.00			
Leverage	0.06***	0.05***	0.29***	−0.28***	1.00		
LTG	−0.09***	0.03**	0.08***	0.32***	−0.01**	1.00	
M/B	−0.03***	−0.05***	−0.25***	0.24***	−0.043	0.12***	1.00

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: authors' calculations.

Table 2

Companies' distribution across countries

Country	Number of companies
Brazil	15
China	437
India	163
Russia	11
South Africa	37
Total	663

Source: authors' calculations.

Table 3

Companies' distribution across industries

GICS code	Industry	Number of companies
10	Energy	17
15	Materials	133
20	Industry	177
25	Consumer Discretionary	115
30	Consumer Staples	58
35	Healthcare	54

Table 3 (continued)

GICS code	Industry	Number of companies
45	Information Technology	53
50	Communication Services	4
55	Utilities	39
	Diversified Corporations	13
	Total	663

Source: authors' calculations.

ABOUT THE AUTHORS

Zinaida V. Selezneva — Research Assistant, Financial Engineering and Risk Management Laboratory, PhD student, School of Finance, National Research University Higher School of Economics, Moscow, Russia
http://orcid.org/0000-0002-0487-4181
zseleznyova@hse.ru

Mariia S. Evdokimova — Lecturer and PhD student, School of Finance, National Research University Higher School of Economics, Moscow, Russia
http://orcid.org/0000-0002-7915-4736
evdokimovamary@mail.ru

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 30.07.2021; revised on 15.08.2021 and accepted for publication on 17.12.2021.
The authors read and approved the final version of the manuscript.