



Vetting of Bloomberg's ESG Governance ISS: QualityScore [GQS™]: Discriminant Testing

¹ Edward J. Lusk, ² Osamuyimen Omorogbe-Akpata, ³ Mia Wells

¹ Emeritus, The Wharton School, University of Pennsylvania, Pennsylvania, USA

Emeritus, School of Economics & Business, SUNY: Plattsburgh, Plattsburgh, USA

Emeritus, International School of Management: Otto-von-Guericke, Magdeburg, Germany

² Departemnt of Accounting & Business Administration, School of Economics & Business, SUNY:
Plattsburgh, Plattsburgh, USA

³ Assistant Director, Lippincott Library of the Wharton School, University of Pennsylvania, Pennsylvania,
USA

Abstract: Recently, vetting investigations were conducted of the Institutional Shareholder Services [ISS] taxonomy whereby firms are assigned to Corporate Governance-Risk[CGR] decile-groups:[GQS[1] the lowest CGR-group and GQS[10] the highest CGR-group] based upon ISS:GovernanceQualityScores: (GQS™). These results are available in the Environment, Social, and Governance [ESG™]-platform offered by Bloomberg™ Professional Services. These vetting investigations indicated: (i) there was no evidence that the ISS-taxonomy used, as a driver, the reported magnitude of the accounts of the Balance Sheet or Income Statement, (ii) for the GQS[1]-group the ISS-taxonomy assignment seemed to align with Managing of the Revenue at the margin while that of the GQS(10) aligned with Net Asset management. In the second study, a group of experts were given two groups of GQS firms, 10 from each of the two-polar groups. They used the Bloomberg ANR®-platform and any other information that they desired to assign these 20-firms into two groups. There was no evidence that their assignment aligned with the ISS-CGR-assignment. These studies affirmed that the ISS-CGR assignment is not likely driven by the firm's periodic reported performance profile. Current Study We offer an extension of the vetting of the ISS-assignment to determine if there is discriminant evidence that a set of 21-financial variables from the Balance Sheet, Income Statement, Cash Flow Statement, and the Market Reported Annual Market Price are aligned with the reported ISS-triage of the firm into one of the ISS:CGR-polar groups. This will address a question begged by the study where there was alignment with the Revenue or the Net Asset Management basis for the Firms in the ISS;polar-groups. Results We elected to use a single BBT-Account Panel matched with the ISS-Polar Assignments to form the Discriminant 2×2 Classification-profile. This allows the computation of: (i) the percentage of the Misclassifications, (ii) inferential measures, and (iii) the R² Entropy measure. We used these inferential measures to profile the discrimination of the BBT-Panel vis-à-vis that of the ISS-Assignment. Interestingly, among the 21 Discriminant Classification Matrices there were NO cases where the Nulls of no effect of these inferential measures could be rejected in favor of alignment of the Panels with the ISS-taxonomic assignment. Summary There was no Individual Panel over the 21 Accounts that indicated that the Accounts of these BBT-Panels could have been the likely driver of the ISS-assignment.

Keywords: GQS[1] and GQS[10] Polar Assignments, Financial Taxonomy Homeomorphs

1. Introduction

1.1 Overview

There is a myriad of dimensions of risk that are in play in purchasing stock of an organization as an investment. This is why stock market trading exchanges were “invented” and shortly thereafter these trading stock exchanges were: Monitored, Evaluated, and Controlled through the legislative authority of the government as part of “protection and

assurance of the Public's General Welfare." This process is on-going and, in most countries, adjustments are frequently made due to the current Monitoring- and Evaluation-intel collected regarding trading exchange activities. This is needed to give assurance to those who wish to trade their resources—currency—for stock certificates—pieces of paper—issued by an organization. The history of the US-stock exchange before and after the 14 Oct 1929 Stock Market Crash is fascinating reading and is evidence of the need for control in market trading environments. At the risk of belaboring the point, consider the recent history of the USA. Circa 1986 through 1995 there was the Savings and Loan (S&L) Crisis, then the 1990s due to the internet and the surreal returns of the dot.coms where many of the “brick” firms such as Enron, Inc., Qwest Communications International, Inc. and HealthSouth, Inc. to mention only a few, engaged in criminal defalcations over the span of the 1990s. Even after the creation of the PCAOB, circa 2007 Lehman Bros. LLP created the subprime debacle. All of this activity was ongoing while the SEC, the safeguard/watchdog of the integrity of the US-trading markets, was comfortably asleep that the switch. The record of reliable assurance provided by governmental agencies charge with protecting investor interests so as to shore up confidence in the US-trading markets is best characterized as being executed somewhere in the interval: {Sloppiness to Incompetence Born of Ineptitude}. To fill the assurance-vacuum there has been a proliferation of firms that have offered firm-evaluation services. In the Financial Services Group: Morningstar, <<https://www.morningstar.com/>>, Reuters, <<https://www.reuters.com/>>, Charles Schwab, <<https://www.schwab.com/>> and Bloomberg, <<https://www.bloomberg.com/>> are leaders in providing multidimensional characterizations of the longitudinal profile of firms. Also, many of the Investment Banking Firms: Goldman Sachs <<https://www.goldmansachs.com/>>, Morgan Stanley <<https://www.morganstanley.com/>>, and JPMorgan Chase <<https://www.jpmorganchase.com/>> are the IB-sector leaders and have advisory fee-for-services largely focused on IPOs and providing information on the incomprehensible worlds of Indexes and other trading derivatives. Finally, in the self-help context the SEC has a Cornucopia of data that is free of charge and as of circa 2019 offers a simple and well-designed GUI-link to datasets of listed organizations accessible through: Electronic Data Gathering, Analysis, and Retrieval system [EDGAR]. In this vein, there are the following very richly endowed data-sources: WRDS <<https://wrds-www.wharton.upenn.edu/>>, S&P[COMPUSTAT] <<https://www.marketplace.spglobal.com/>>, and CRSP <<https://crsp.org/>>. They are world-class but very pricey sources of data. Most of the above sources offer data and/or financial market risk advisory platforms; those few that do offer Corporate Governance Risk [CGR] calibration offer platforms that are more or less in developmental phases.

However, recently there has been more than a passing interest in CGR as a latent but important driver of Risk in all of its incarnations the tentacles of which eventually impact the market risk. See (Zopounidis, Garefalakis, Lemonakis and Passas (2020)). The interest in CGR seems to have been awakened from the Public Company Accounting Oversight Board's [PCAOB] interest in Management's System of Internal Control over Financial Reporting [ICoFR]. In fact, the SEC also has joined the ICoFR-bandwagon that was a criterial corporate management control variable introduced in the 1990 by the Committee of Sponsoring Organizations of the Treadway-commission. The PCAOB, in reviewing the corporate Enron-esque defalcations of the 1990s, latched on to ICoFR as a necessary aspect of the audit needed to regain the confidence of the investing sector in the veracity of the information being generated as part of the audit and communicated through theegis of the SEC. Thus, the first set of the Audit rules promulgated by the PCAOB, AS2, required a second audit opinion that addressed the adequacy of management's system of ICoFR; this is called the COSO-opinion. The SEC has gone “proactive” and now through the PCAOB's Audit Standards requires that the LLPs audit the firm's ICoFR and based upon these audit-results, the SEC issues Deficiency, Significant Deficiency and Weakness public information on the firm's attention to providing “self-control” through ICoFR and so logically but only tacitly impacting CGR. See[PCAOBⁱⁱ]. Given this historical context, consider now the Risky-Business of not giving Corporate Governance its due Props.

1.2 Corporate Governance Risk Profile circa 2020

As Risk is the bane of investor, and Risk derives from the lack of attention to the ICoFR that is enabled by inattention to effective control of Corporate Governance, then financial advisory services that do not usually offer useful evaluation of the likely nature of CGR will not be competitive with those firms that offer useful advisory intel re: CGR—the Smithian Invisible Hand of competitive survival. In this context, we recommend the excellent research report of Huber and Comstock (2017). They have offered a well-documented and useful analysis of firms that are offering CGR-analytics. They note p. 30-31

“Most international and domestic public (and many private) companies are being evaluated and rated on their environmental, social and governance (ESG) performance by various third-party providers of reports and ratings. Institutional investors, asset managers, financial institutions and other stakeholders are increasingly relying on these reports and ratings to assess and measure company ESG performance over time and as compared to peers.

There are currently numerous ESG data providers, a summary of each of which is beyond the scope of this article, but some well-known third-party ESG report and ratings providers include: (1) Bloomberg ESG Data Service, (2) Corporate Knights Global 100, (3) DowJones Sustainability Index (DJSI), (4) Institutional Shareholder Services (ISS), (5) MSCI ESG Research, (6) RepRisk, (7) Sustainalytics Company ESG Reports, and (8) Thomson Reuters ESG Research Data. The purpose of this article is to provide a snapshot of some ESG report and ratings providers. It is not meant to be a comprehensive overview of all such providers. “

The Huber and Comstock report was the progenitor of the studies of Lusk and Wells (2021a) and Lusk and Wells (2021b) that have offered vetting of the integration of the (4) Institutional Shareholder Services (ISS) as part of the Bloomberg ESG Data Service. It is significant that this merger or acquisition took place. Bloomberg is ubiquitous and is the gold standard of Market related intel. That Bloomberg decided, circa 2018 to offer ISS coding speaks volumes of the critical impact of CGR given the plethora of variables that are part of Bloomberg’s ESG platform—to wit Bloomberg’s invitation to ISS was likely a synergistic-boon for both organizations. With this as the historical tapestry, consider the research agenda.

2. Research Agenda

Having given contextual support for the importance of assessing the nature and quality of corporate governance, we will focus on the vetting of the Bloomberg and ISS merger focusing on the veracity of the ISS-risk calibration. Specifically, we will:

- i. Detail the nature of the ISS-CGR scoring platform, discuss the nature of vetting-intel, and detail the focus of vetting the ISS-platform offered by Bloomberg,
- ii. Review the CGR-literature and examine in detail the Lusk and Wells (2021a) and Lusk and Wells (2021b) studies that have provided initial vetting-intel of the ISS-platform,
- iii. Introduce the next vetting examination of the ISS-platform where we will use as the inferential context Discriminant Analysis,
- iv. Detail the inferential context of Discriminant Analysis as well as the computation of the sample-size needed for the vetting,
- v. Discuss the Datasets, Variables and Testing-protocol that are used in the vetting-discrimination,
- vi. Present the vetting results and summarize their analysis, and
- vii. Offer the extensions of this vetting-investigation.

3. Bloomberg’s ESG Governance ISS

The ISS-platform has been undergoing developmental modification for circa a decade. We will note a few of the salient aspects that are taken from the ISS-Protocol that is on-line at: [ESG Fund Rating - ISS \(issgovernance.com\)](https://www.issgovernance.com/ESG-Fund-Rating-ISS) and may be downloaded as a PDF: ISS ESG Governance QualityScore[29Oct2020]. The nature of Bloomberg’s ISS:CGR-protocol is interesting; no ISS-scoring details are given. These computational details are “under the ISS-hood”. Each firm is evaluated and a Governance Quality Score [GQS-score] is created by ISS; this GQS-score is used to dynamically assign a particular firm to one of ten ordered CGR-groups. For example, it is noted:

“ISS ESG Governance QualityScore (GQS) is a data-driven scoring and screening solution designed to help institutional investors monitor portfolio company governance. At both an overall company level and along topical classifications covering Board Structure, Compensation, Shareholder Rights, and Audit and Risk Oversight, scores indicate relative governance quality supported by factor-level data. That data, in turn, is critical to the scoring assessment, while historical scores and underlying reasons prompting scoring changes provide greater context and trending analysis to understand a company’s approach to governance over time. [p.4]”

Further, the manual offers:

Employs robust governance data and attributes. Governance attributes are categorized under four topical categories: Board Structure, Shareholder Rights and Takeover Defenses, Compensation/Remuneration, and Audit and Risk Oversight. Governance QualityScore calls upon a library of more than 230 governance factors across the coverage universe, of which up to 127 are used

for any one company (defined by region). Governance QualityScore highlights both potentially shareholder-adverse practices at a company, as well as mitigating factors that help tell a more complete story. The underlying dataset is updated on an ongoing basis as company disclosures are filed, providing the most-timely data available in the marketplace. [p.4]

Finally, as a summary it is noted that:

Presents at-a-glance governance rankings relative to index and region. Governance QualityScore features company-level decile scores, presented as integers from 1 through 10, plus underlying category scores using the same scale that together provide a clear understanding of the drivers of a company's governance risk. A score in the 1st decile indicates higher quality and relatively lower governance risk, and, conversely, a score in the 10th decile indicates relatively lower quality and higher governance risk. These scores provide an at-a-glance view of each company's governance risk relative to their index and region. The individual factor breakdown takes a regional approach in evaluating and scoring companies, to allow for company-level comparisons within markets where corporate governance practices are similar. [p.5]

Thus, the salient-aspects-summary of the ISS-scoring platform: (i) The ISS is focused on Corporate Governance Risk (ii) The ISS-assignment is dynamic, (iii) Changes in GQS-group assignment are made to insure currency, (iv) ISS-protocols use a copious amount of data most of which has veracity-screens, and (v) The collected data is filtered through numerous mathematical and statistical inferential protocols to arrive at a scoring assignment into decile CGR-groups that can be accessed through the Bloomberg[ESG[ISS]]-platform. In this context, the following vetting test is begged:

It is a fool's errand to try to analytically reverse-engineer the ten GQS-Profiles to approximate the weighting-schema(s) used by ISS in forming their firm-assignments to the ten ordered decile-groups. In its stead, however, vetting analyses would be useful to determine if the firms in the polar-decile groups: {EGS[ISS[GQS[1]] and EGS[ISS[GQS[10]]} seem to be therein assigned based ONLY upon their single GAAP-reported variables. If such GAAP-alignment were to be the case, this would undermine the credibility of the veracity of the ISS-assignment protocols. It would be fool-hearty to believe that there is a simple ordered mapping of a single derived GAAP-variable of firm-data that is homographic to the ISS-groups. Thus, such alignment would be a fatal-chink in the ISS-armor.

4. Research Reports on ISS-vetting Analyses

4.1 Overview

Vetting studies are critical to enhance inference validation. Vetting simply poses a logical population expectation that IF not validated would raise an inferential-alert re: any inference may not pertain to the expected population from which the sample should have been taken. In this context, we will present the results of two vetting analyses of {EGS[ISS[GQS[1]] and EGS[ISS[GQS[10]]}.

4.1 Vetting the ISS-assignment Protocol

The vetting issue underlying the ISS-assignment is: If the ISS-assignment of firms to GQS[1] and GQS[10] tacitly was aligned with the monetary values of the GAAP-variables that would raise questions as to the logic of categorization of firms based upon Corporate Government Risk. Thus, there were two studies that were formed to provide inferential information by using a random sample of firms in: GQS[1] and GQS[10] and then to use GAAP-financial information to determine if the ISS-assignment was inferentially replicated using only this GAAP-account information.

4.2 Vetting using Monetary Values

Lusk and Wells (2021a) report[Paraphrasing]:

The GQS-platform offers a data-driven approach to scoring and screening designed to help investors monitor Corporate-Governance-Risk [CGR] activities so as to better inform their decision-making regarding CGR. Our vetting addresses the inferential question: Is there a logical reason to reject the belief that the set of GQS-assignment protocols are not well formed thus creating Governance-risk-groupings that have no intra-group coherence and so exhibit no inter-group differentiability with respect to CGR? This question intends to offer a vetting test using monetary values taken from PCAOB-audited accounts of firms randomly selected from the GQS[1] and GQS[10] polar ISS-groups. IF the logic of an ISS-coding assignment follows, in the main, monetary values taken from the GAAP-reported values taken from the Balance Sheet and the Income Statements then we would fail to reject the belief that the ISS-Risk protocol is not well-formed supporting the indication that it would be very unlikely that there would be CGR intra-group coherence and so then logically it is doubtful that there would be inter-group differentiability with respect to CGR. If we can reject that the

ISS-CGR-assignment is simply aligned along the monetary-values then this would lead to rejecting that the ISS-protocol is not well-formed and so not put into question the credibility of the logic of the ISS-assignment.

Point of Clarification True, this vetting test logic seems to some extent “contorted”. This is due to the necessity to context the question in the Null-testing FPE-formulation. In a simplification, vetting poses a question for the expected population that if found to be the case would cast doubt that that population characterization could be true. For example, the vetting question says if the ISS-assignment protocol just uses the magnitude of the GAAP-monetary values of the firms that would cast doubt on the quality of the ISS-assignment as just single GAAP-monetary values are not likely to be a surrogate for Corporate Governance Risk.

For the Lusk and Wells (2021a) study they found no evidence that the ISS-protocol was driven by the magnitude of the monetary values. However, after converting the magnitude of the account-values to ratios—e.g., using the Current Ratio rather than the magnitude of the Current Assets— they found that ISS Group-GQS[1] was differentiable from that of ISS-GQS[10] on the management of the Revenue at the Margin dimension of the firm’s decision imperatives. Additionally, they found that ISS Group-GQS[10] was differentiable from that of ISS-GQS[1] on the management of the firm’s Asset[Net] Management. In summary, these results are sufficient to suggest that the ISS:GQS-assignment protocols seem to be well-formed and capable of offering useful differentiation.

4.3 Expert Classification Vetting Study

To examine the possibility that the Lusk and Wells (2021b) vetting missed the possibility that their design created latent-signals that were “embedded” in the financials overall that queued the ISS-factor assignment protocol-set, they created a test of this latency-effect by using Experts to assign firms to differentiated binary-groups. Thus, their simple vetting test used the financials—effectively the GAAP-market-profile—of a random sample of firms from the [GQS [1] and GQS [10]] to determine if the market or financial profile of these firms would suggest a taxonomic assignment triage that mirrored that of the ISS. Thus, to effect such a vetting test Lusk and Wells (2021b) randomly selected 20 firms—ten each from the polar decile-groups [GQS [1] and GQS [10]]. These firms were profiled by the Bloomberg Analyst Recommendations [ANR[®]]. The ANR-PDF-captures were not identified as to the GQS-group to which they were assigned by ISS. These 20 ANR-profiles were given to nine-volunteers with advanced expertise in market-related discipline areas, and they were asked to: (i) sort the 20-firms into two groups of equal-size, and (ii) note their assignment logic. The vetting results testing logic then was: (i) if the majority of these individuals replicated in the main the ISS-groups then the vetting would fail to reject that there was not sufficient evidence that the CGR was the assignment logic, or (i) if there was no alignment between the Expert-testing triage and the ISS-polar groups then this would rationalize rejection that the CGR was not the driver of the ISS-assignments. The inferential results were very clear. There was no inferential evidence overall for the assignments made by the volunteers that there were sufficient numbers of triage-matches to the ISS [1]-group to reject the Null of Chance of 50% as the actual assignment of the experimental subjects. These results suggest that the latent-Revenue at the Margin and -Asset[Net] possibility proposed for Lusk and Wells (2021a) was not likely the latent “driver” that queued the ISS-assignment—to wit this does not raise a question re: the ISS-assignment protocol and so is a valuable vetting indication that ISS-Corporate: Governance: Risk-assignments are not surrogate-holomorphs to the relative ANR-profit-profiles.

5. Current Vetting: Discriminant Screening

5.1 Overview

As a further test of the latent-financial-signaling that may have queued the ISS-assignment, we offer the following vetting test. Discriminant Analysis is a screening technique for testing relationships that may be “drivers” of the assignment where there is a protocol or natural taxonomic mapping. For example, ISS assigns firms using its scoring-system to one the decile-GQS-groups. We have a copious amount of GAAP-information on these firms. Discriminant Analysis can be used to answer the query:

For the ESG[ISS[GQS [1] and GQS [10]]]-binary set of firms created to provide maximum differentiation between the firms assigned to GQS[1]—the set of firms with the lowest-CGRisk—and those firms assigned to GQS[10]—the firms scored as representing the highest-CGRisk—are there GAAP-account relationships that align with the ISS-assignment scoring.

5.2 The Nature of Discriminant Analysis:[DA]

To better understand how DA is a useful analytic-screening-tool for the question posed above, a examination of a few of the DA-component platforms would be most instructive. Assume for the purposes of illustration: There are two datasets GQS[1] and GQS[10]; each has 15-Firms in their GQS-category-set. Each firm contributes four GAAP-derived account information each from The Balance Sheet, The Income Statement and The Cash Flow Statement, and each of these GAAP-account panels have 20-audited and recorded values. Thus, the total information for the ISS-polar-groups has 7,200 values: $2 \times \{15 \times 4 \times 3 \times 20\}$. The simplest and most defensible Discriminant Protocol tests the disaggregates of this collection. An example of a particular DA is:

Assume that we have selected the Current Ratio [CR] from the Balance Sheet. In this case, the DA will be:

DA[GQS[1:[CR, 15×20]] and GQS[10:[CR, 15×20]]]

Description: In this case, there are 300 Current Ratios in total contributed by firms in GQS[1] and the same number for GQS[10]. The DA will compute the centroid measures for these two collections. These create probability zones around the Means/Centroid of the two groups: GQS[1] and GQS[10]. Then the DA takes each of the Actual 300 CR-Data values for each of the ISS-polar-groups and computes the likelihood that the Actual CR-point under examination is closest in likelihood/probability to one of the two computed Centroids. If the selected CR-Point is closest to the centroid of: GQS[1] then that CR-point is assigned by the DA to GQS[1] otherwise that CR-point is assigned by the DA to GQS[10]. This is done for each of the 600-CR values. This creates a Classification Matrix. Assume after all the assignments are made that the Classification Matrix is:

Table 1: Illustration of the DA Classification Matrix

	ISS→GQS(1)	ISS→GQS(10)	Totals	Inference
DA→GQS(1)	239	<i>61</i>	300	Entropy R^2
DA→GQS(10)	<i>48</i>	252	300	Misclassified%
Totals	287	313	600	Power

Discussion The main-diagonal is the correct-classification-zone where the a priori ISS-assignment is aligned with that of the DA-empirical assignment protocol [Bolded]. The other two-cells, shaded, are where there was a “misclassification: where the DA-assignment was NOT aligned with the ISS-a priori assignment. Recognize that the benchmark is the a priori ISS-assignment. This does not mean that the ISS-assignment was “correct” with respect to the Firm’s CGR. It just means that the DA, based upon the totality of the ISS-assignments, found an empirically justifiable probability to make an assignment based upon the location of the Centroid. Also, given the DA-classification matrix, there are three “inferential” measures that are the usual fare in guiding decision-maker judgments:

1. The Entropy R^2 [ER²] is rather complicated and there is some contentious inferential debate as to the exact meaning of the FPE of the Nullⁱⁱⁱ. However, in a practice-context, ER² is very often used in the same sense as the OLS-regression R². The reason that it is labeled as an entropy-measure is that it reflects the nature of the order or lack thereof that results from the dataset classified by in the DA relative to the a priori ISS-assignments. If there was a true state of entropy i.e., chaos then there is no order and the ER² would be 0 and all the cells in Table 1 would have the same proportions or number of entries—i.e., no information or the state of stasis Re: entropy. If at the other extreme there was a perfect DA-classification so that only the main-diagonals were to be filled, in that context, the ER² would be 1 indicating perfect order in classification. With all the usual OLS-Regression caveats. Very often if $ER^2 > 0.5^5$ where: 0.5^5 is the usual Harman factor-cutoff for uniqueness in the Varimax-factor rotation-model, this is a strong indication of DA-alignment with the ISS-assignment, $ER^2 < [1 - 0.5^5]$ is an indication of a lack of interesting alignment and in-between these values are the interesting-zone. In this case, the ER² is 53.9% and this places these DA-results in the interesting zone.
2. The misclassification percentage is very often used to create Binary-percentage confidence intervals or to test the association using a test-of-proportions. In this case, for Table 1 the miscalculations are: 109 [48 + 61]. The misclassification percentage is: 18.2%: [109/600].
3. The Power will be discussed anon.

Point of Information. Often Wilk’s Lambda [Wilk’s(λ)] is used in the inference profile of the DA-results. However, it usually follows in trajectory the ER² and is redundant to the Power, for this reason we will not profile the DA-results using Wilk’s(λ).

Having provided the above context, this would lead to the assessment of the sample-size that we would propose for the DA-vetting test of the ISS-assignments as well as the inferential tests. There are a number of inferential contexts in play

in this study. We have opted for the Binary tests in the Percentage assignment for the DA-assignment classification matrix. This will be discussed following.

5.3 Sample Size Computations

We will be computing a vetting check for the DA. The vetting test will address the accrual veracity of the DA-dataset for the ISS-polar-groups. For this vetting, we will take two random-samples of each of the DA-datasets of the sample-size recommended by using the Wang and Chow (2007) [W&C] non-directional test for two-population percentages. If these random samples test to have different DA-misclassification percentages for which the FPE-Null had a p-value <5%, we reject the Null and select the random-set that had the lowest misclassification percentage between the two random-sets as the sampling realization. This is conservative as it is biased to not rejecting the Null of *Ho*. Otherwise, we will use the first random-selection for the test of the DA-assignment. Consider now these computational protocols.

5.4 Wang and Chow Sample-size

For the inference between two sampled populations of percentages, we have the following sample-size formulation due to Wang and Chow (2007):

$$n = (Z_{\alpha/2} + Z_{\beta})^2 * [(P_1 \times (1 - P_1) + P_2 \times (1 - P_2)) / (P_1 - P_2)^2]$$

Where: $Z_{\alpha/2}$ is the z-calibration for the non-directional α -False Positive Error risk to give the level of confidence of $1 - \alpha$; Z_{β} is the z-calibration for the directional β -False Negative Error risk for the Power of the test, P_1 is the expectation for DA[ISS[GQS[1]]] and P_2 is the expectation for DA[ISS[GQS[10]]]. These expectations are only the values that are used to form the sample-size for a misclassification non-directional test-effect of $[Abs(P_1 - P_2)]$.

Using the following calibrations:

Degree of Confidence 95%, FPE $\equiv \alpha = 5\%$ [1-95%] thus, the non-directional $Z_{\alpha/2} = 1.96$; FNE [$\beta = 20\%$] and Power $\equiv 80\%$ [1-20%] and $Z_{\beta} = 0.842$; the assumption for the test-effect is: ISS[GQS[1]] = 50% and ISS[GQS[10]] = 38% gives the test-effect of 12%. *Rationale*: In a related test of similar datasets, the median of the misclassifications was 43.6%. Thus, we decide to use 50% and 38% which gives the average 44% and an effect size of 12%. Thus, the sample-size is:

$$265 = (1.96 + 0.842)^2 * [(50\% \times (50\%) + 38\% \times (62\%)) / (50\% - 38\%)^2]$$

Discussion of the Vetting Test The sample-size of 265 indicates the random-selection from the DA-Account datasets. The DA-vetting protocol is: For each of these DA-sets, (i) we will take a random sample of 265-DA data-points for a particular GAAP-account, (ii) run the DA {ISS[GQS[1]] v. ISS[GQS[10]]} for that GAAP-account, (iii) record the percentage of misclassification reported in DA-results profile, (iv) take another random-sample of 265-DA data-points, repeat steps (ii) and (iii). This will give two-sample percentages from the same GAAP-account of the misclassifications:

Random Sample I : Percent of DA-misclassifications: P%[I], and

Random Sample II : Percent of DA-misclassifications: P%[II]

Summary Assuming that the above FPE- or p-value-risk as computed is ≥ 0.05 , we will fail to reject the Null that P%[I] and P%[II] are not different and so the P%[I]-dataset will qualify as a representative sample from a vetted population; otherwise, we will select the dataset with the smallest misclassification percentage for the DA-test.

5.5 Inference Test of the Qualified-DA

In this *a priori* stage, we will state the toggle-Null-form of the test of the study hypothesis. **Alert**: This Null proposed is a composite hypothesis-Null that will have a simple Power-index using the Operating Characteristic Function [OCF]. See Tamhane and Dunlop[2000[Sec.6.3.4[p.214]]]:

Ho: The Percentage of the DA-misclassifications in the vetted-population of the GAAP-accounts selected for the DA: {ISS[GQS[1]] v. ISS[GQS[10]]} analysis is not greater than 10%. We have selected 10% as the Maximum of the “fail to reject *Ho*-point” as this seems a not unreasonable overall indication of ISS and DA alignment. The decision to reject *Ho* tacitly indicates that one accepts *Ha*: That there is likely evidence that the misclassifications in the population are >10%. The associated test implication of *Ha* is that the ISS-assignment protocol was not using the GAAP-measures to assign firms to the two polar groups: ISS[GQS[1]] and ISS[GQS[10]].

Conditioned Protocol for the Rejection of *Ho* in favor of *Ha*:

- 1.) Judgmental a-priori FPE: Condition: The FPE-risk for rejection of *Ho* is < 15%
- 2.) Judgmental a-priori FNE: Condition: Given π_{Ho} under the judgmental a-priori FPE: Condition, the FNE-risk—failing to reject *Ho* when it is not likely that the population-misclassifications are $\leq 10\%$ --given the actual percentage of misclassifications experimentally measured, is 15%.
- 3.) This then gives the Dual-Conditioning of the FPE and FNE risks of 15%-each.

Discussion: The judgmental a-priori FPE: Condition under the Test for Ho Assume that the sample size is 265; computations:

FPE: Condition Using as the Ho-Standard Error: $s_e = 0.0184: \left[\frac{10\% \times (1-10\%)}{265} \right]^{0.5}$, the percentage of misclassifications for which the directional FPE-risk = 15% is 11.91% or assumed to be: $[\pi_{Ho} \equiv 12\%]$. For example, following are these computations:

The z_{cal} that produces a RightHandSide FPE-risk of 15% is:

$$FPE_{z_{cal}} \equiv T.INV.2T(15\% \times 2, 10000) = 1.036487 \equiv -T.INV(15\%, 10000)$$

Thus, the test-against-value $[\pi_{Ho}]$ vis-à-vis Ho of 10% is:

$$[s_e \times FPE_{z_{cal}} + Ho] \text{ or } [0.0184 \times 1.036487 + 10\%] = 11.91 \text{ or } = 12\%$$

Computational Check: $T.DIST.RT(1.036487, 10000) = 15\%$

Discussion: This suggests that if the True state of nature is that there are 10% DA-misclassifications in the population, the chance of finding a misclassification of 11.91% or 12% would happen 15% of the time under random sampling—i.e., the FPE- or α - or Type 1- or p-value-risk of being wrong and accepting Ha is 15%. Thus, if one rejects Ho and so is operating under the inferential belief that Ha is the true state of nature they are “Ignoring the Odds” that 15% of the time one could observe misclassifications of 12% or greater in the population where the misclassifications in the sampled population are =10%. Usually, a FPE of 15% is not sufficiently high to reject Ho with assurance—however, it is sort of suggestive and so merits consideration.

FNE: Condition In this case, assuming: (i) the misclassification under the FPE-condition of Ho of 12%, (ii) the standard error under Ho is 0.0184, and (iii) an actual percentage of misclassifications noted as: π_{FNE} , then using a toggle-point of the FNE of 15% or Power $[1-15\%]$ gives, in calculation:

FNE : LeftHandSide $N(0,1)$ for $\left[\frac{12\% - \pi_{FNE}}{0.0184} \right] \rightarrow \pi_{FNE} = 13.91013\%$ or 14% misclassifications as the boundary-value. In this case, following are these computations:

The z_{cal} that produces a FNE of 15% $[FNE_{z_{cal}}]$ is: $T.INV(15\%, 10000) = -1.036487$

Thus, the test-against-value $[\pi_{FNE}]$ vis-à-vis π_{Ho} of 12% is:

$$- [(s_e \times FNE_{z_{cal}}) - \pi_{Ho}] \text{ or } - (0.0184 \times -1.036487) - 12\% = 13.91013\% \text{ or } = 14\%$$

Finally, the Power of this FNE-testing frame is $(1 - FNE)$ or 85% $(1 - 15\%)$.

Computational Check: $T.DIST(-1.036487, 10000, TRUE) = 15\%$

Discussion: In this case, the conditioned Rejection-region of Ho using the OCF is simply any realization where the sample size is $\Rightarrow 265$ and the percentage of miscalculations is $> 14\%$. In this case, the FPE-Condition is redundant as 14% is $> 12\%$ given the FNE- condition.

5.6 Illustration

At this point an illustration of these two stages is in order. We will use one of the datasets in the study. Assume that we have the *Current Ratios* [CR] of 30 firms assigned by ISS to GQS[1] and 31 firms assigned by ISS to GQS[10]. In total there were 339 CRs recorded for the GQS[1] firms and 481 CRs recorded for the GQS[10] firms. The random-sampling selected $n=122$ datapoints in ISS[GQS[1]] and $n=164$ data points in ISS[GQS[10]]—in total 286 for the first random sample. For the second random sample in total there were $n=114$ datapoints in ISS[GQS[1]] and $n=135$ datapoints in ISS[GQS[10]]—in total 249 for the second random sample. For the vetting phase, the first random sample had 39.9% misclassifications and the second random sample had 43.8% misclassifications. The p-value for this difference 3.9% $|\text{Abs}[39.9\% - 43.8\%]|$ was 0.36 as this is > 0.05 , we fail to reject the Null of no difference and thus the first random-sample moves to the DA-inference testing phase.

DA-Inference test of H1 The classification matrix for the first-random-sample was:

Table 2: Current Ratio DA-Random Sample

	ISS→GQS(1)	ISS→GQS(10)	Totals	Inference
DA→GQS(1)	8	114	122	R² Entropy [0.007]
DA→GQS(10)	0	164	164	Misclassified% [39.9%]
Totals	8	278	286	Power > 85%

There were 114 misclassifications [shaded-cells] and this gives a percentage of 39.9% [114/286]. The FPE[p-value] of this result is:

Standard Error: 0.01774: $\left[\frac{[10\% \times (1-10\%)]}{286} \right]^{0.5}$

$z_{cal} = [16.9] = \left[\frac{39.9\% - 10\%}{0.01774} \right]$, and

FPE[p-value = <0.0001] which is < than the Conditioned Rejection value of 15%.

AND

The Power [1-FNE] > 85% as 39.9% > than 14%. Specifically,

Standard Error: 0.01774 as above and

$z_{cal} = [-15.78] = \left[\frac{12\% - 39.9\%}{0.01774} \right]$ and T.DIST(-15.78,10000,TRUE) = <0.0001 is a FNE< than 15%; and so Power >85%.

This is an illustration of the use of the boundary values presented above where the only DA-test needed is to determine if the misclassifications percentage is $\geq 14\%$ thus rejecting H_0 in favor of H_a that there is not likely ISS-alignment with the BBT-test panels as the FPE and the FNE of the actual results are both < 15%. This greatly simplifies the testing conditions to screening only for the misclassification > 14%.

In this case, for this example there is clear evidence that the population percentage of misclassifications is greater than 14% and the inferential power of this result is > 85%. Further, the lack of alignment of ISS with the values of the Current Ratios is also indicated by the fact that the Entropy R2 was 0.007 and this is < [1- 0.5⁵].

Summary Indication There is no inferential evidence that the Current Ratios are used as a major driver of the ISS classification into the polar-groups: ESG[ISS[GQS [1] and GQS [10]]].

6. Testing the Veracity of the ISS-Assignment Protocol

6.1 The Vetting-Variables Testing-Screens

In the testing of the ISS-assignment, the inference of which is driven by the OCF-testing re: H_0 , we will examine GAAP-accounts from the four financial statements/platforms that are offered by Bloomberg: Balance Sheet: Income Statement: Cash Flow Statement and Firm Price/Valuation Platform. We sampled the following GAAP-information as follows;

Balance Sheet

X₁ Current Ratio [SK&FO^{iv}][CUR_RATIO] Defined as: [Current Assets[BS_CUR_ASSET_REPORT] / Current Liabilities[BS_CUR_LIAB]]

X₂ Tangible Common Equity Ratio [SK&FO] TCE_RATIO] Defined as: Measure of financial strength which shows the tangible value of equity as a percentage of tangible assets. Both the total assets and the common equity are adjusted for the amount of intangible assets such as goodwill, licenses, trademarks, copyrights, etc. Total assets are not adjusted based on risk. Calculated as: (Tangible Common Equity / Tangible Assets) * 100

X₃ Current Payables / Current Liabilities[SK&FO] [(ACCT_PAYABLE_&_ACCRUALS_DETAILED)/CL]

X₄ Current Assets / Operating Cash Collection [Days][FO] [BS_CUR_ASSET_REPORT/ CASH_CONVERSION_CYCLE[CCC]] A measure of the proportion CA relative to the total time needed to convert resources to Cash Metric which expresses the length of time, in days, that it takes for a company to convert resource inputs into cash flows. Calculated as: [Inventory Turnover Days + Account Receivable Turnover Days - Accounts Payable Turnover Days]. For example, if CA =100 and the number of days in the CCC is 50 days this suggests that it would take TWO CCC to replace the depletion of the Cash.

X₅ TA /TA [SK&FO] [BS_TOT_ASSET / BS_TOT_LIAB2]

X₆ Annual Percent To Realize Cash Collection[CCC] as: CASH_CONVERSION_CYCLE/365.25

X₇ Current Assets [SK&FO][Vetting]

Income Statement

X₈ Gross Margin[SK&FO] [GROSS_MARGIN] defined as: Gross margin represents the percent of total sales revenue that the company retains after incurring the direct costs associated with producing the goods and services sold by a company. Calculated as: ((Net Sales - Cost of Goods Sold) * 100 / Net Sales).

X₉ Operating Margin [FO][OPER_MARGIN] defined as: Ratio used to measure a company's pricing strategy and operating efficiency, in percentage. Calculated as: (Operating Income (Losses) / TotalRevenue) * 100 where **Operating Income**[Losses] [IS_OPER_INC] Calculated as: ([Operating Income (Losses)] / Net Revenue) * 100

X₁₀ Profit Margin [SK&FO] [PROF_MARGIN] Measuring the company's profitability, this ratio is the comparison of how much of the revenue incurred during the period was retained in income. Calculated as: (Net Income / Revenue) * 100

X₁₁ Gross Profit [SK&FO][GROSS_PROFIT] : Company's revenues less its cost of goods sold. Calculated as: Revenues - Cost of Goods Sold [Vetting]

X₁₂ EBITA [SK&FO] [EBITA] This measure calculates earnings before interest, taxes and amortization. Calculated as: EBITDA - Depreciation Expense where EBITDA is an indicator of a company's financial performance which is essentially net income with interest, taxes, depreciation, and amortization added back to it, and can be used to analyze and compare profitability between companies and industries because it eliminates the effects of financing and accounting decisions. EBITDA is calculated as: Operating Income + Depreciation & Amortization + Operating Lease Rental Expense Adjustment. [Vetting]

X₁₃ EBITA / GrossProfit[FO] [EBITA / GROSS_Profit].

Cash Flow

For the Cash Flow there are three measures: Cash Flow from: **Operations** or **Investing Activities** or **Financing Activities**. To create a ratio measure, we took the Maximum of these three as their benchmark. For example, for *ADTN* the Cash Flow from: **Operations**[105.2] or **Investing Activities**[-39.9] or **Financing Activities**[-21.5]. Thus, the benchmarked values used for the DA were:

Operations[100%: [105.2/105.2] or
Investing Activities[-038%: [-39.9/105.2] or
Financing Activities[-20%: [21.5/105.2]

X₁₄ In this case, we will use ALL three of these values {**Operations & Investing & Financing**} as single DA-variable-screen.

The respective definitions are:

Cash From Operations [CF_CASH_FROM_OPER] defined as: Total amount of cash a company generates from its operation. The effect of Changes in Non-cash Working Capital on Cash from Operations can be either positive or negative. Decrease in current assets or increase in current liabilities, increases Cash from Operations; while an increase in current assets or decrease in current liabilities, decreases Cash from Operations. Generally calculated as: Net Income + Depreciation & Amortization + Other Noncash Adjustments + Changes in Non-cash Working Capital

Cash from Investing Activities [CF_CASH_FROM_INV_ACT] defined as Sum of Disposal of Fixed Assets, Capital Expenditures, Decrease in Investments, Increase in Investments, and Other Investing Activities.

Cashflow from Financing Activities: CFF_ACTIVITIES_DETAILED defined as Cash flow from financing is the cash from all financing activities, such as dividends paid, repayments of borrowings, repurchase of equity and other financing activities. Calculated as: Dividends Paid + Proceeds from Repayments of Borrowings Detailed + Proceeds from Repurchase of Equity Detailed + Net Cash from Discontinued Operations Financing + Other Financing Activities (Excl Foreign Exch)

X₁₅ Trailing 12M EBITDA Margin[SK & FO]: [EBITDA_MARGIN] Percentage margin of trailing 12-month Earnings Before Interest Taxes Depreciation and Amortization (EBITDA) divided by the trailing 12 month Sales. Computed as: (Trailing 12-month EBITDA / Trailing 12-month Sales) * 100.

X₁₆ Free Cash Flow/Basic Shr: [FO] [FREE_CASH_FLOW_PER_SH] Measure of a company's financial flexibility that is determined by dividing free cash flow by the weighted number of shares outstanding. This measure serves as a proxy for measuring changes in earnings per share. Calculated as: Free Cash Flow / Weighted Number of Shares Outstanding,

X₁₇ Max-Value is the largest of the three Cash Flow values [Vetting]

Value in the Trading Markets

X₁₈ Correlation In this case we report the Pearson Product Moment correlation of the Panel of Yearly Stock Prices with the Time-Index. We used this as a screen as if the ABS[PPM] is less than 25% that firm was eliminated from the Value DA.

X₁₉ Stock Valuation Change [CHG_PCT_PERIOD] [FO]: For example, for UFCS the 12-month ending was for 2002 and 2003 was 16.725 and 20.18 respectively. The Change Percentage was calculated as [(20.18 -16.725)/ 16.725]×100 = 20.7%

X₂₀ HISTORICAL_MARKET_CAP [FO] [Vetting Test]

X₂₁ Max v. Min Change is computed as: [(MaxPrice -MinPrice)/ MaxPrice]×100

These are the variables that will be used to test the vetting conjecture that the ISS-GQS-assignment protocol uses GAAP-account information to form the polar-decile CGR-profiles. Of course, if this were to be the inferential reality, this would call into question to credibly or operational veracity of the ISS-protocols as fine tuned by a multiplicity of Governance Risk-related factors of the corporation. Also, of note, there are five magnitude GAAP-account variables that are vestiges of the vetting tests conducted by Lusk and Wells (2021a). This is then an additional veracity-test of that research report.

6.2 Selection of the Firms

We have selected 31 firms that were scored and assigned to ISS[GQS[1] and 31 firms so reported as members of ISS[GQS[10] as reported by Bloomberg as of 31Oct 2020. One of the GQS[1]-firms was eliminated due a profound lack of data. For logical control this was the same accrual-list as was used in Lusk and Wells (2020a,b); however, the firm information was collected from November 2021 through December 2021. These firms are reported in Appendix A. For each firm there were 20 years downloaded for each of the 21 GAAP variables as reported above. If more than 10 Panel-years were missing that Variable-dataset was eliminated from the accrual. The download of Panel-sets was {FY2001 through FY 2020}. This produced for each GAAP-Panel 30- or 31-Firms and 20-Years and two ISS-groups GQS[1] and GQS[10] or *on the order* of 25,000 total Points for the two-ISS groups: $2 \times [30 \times 21 \times 20]$ possible. There was a short fall of 8.3% due the sampling reality that not all the firms had the particular GAAP-accounts as screened from the BBTs and/or that the Panel-size was slightly less than 20 years.

The Discriminant Analysis protocol used the *SAS*TM:platform^v. We used the standard linear canonical form of the DA and will use the toggle-point of **14%** to screen the results as presented in the inference section. We will also report the Entropy R^2 and the indication of the boundary value for the FNE. As there is clearly multi-collinearity that compromises the inference for multi-variable screenings also see [Soekarno and Kinanthi (2020)] we will use, with one exception, single-panels formed for ISS[GQS[1] and ISSGQS[10]. This will generate the results-matrix as presented in Table 3. Additionally, we will note the GAAP-accounts that are used magnitude-vetting checks as used in Lusk and Wells (2021a), and finally, should there be a random-sampling failure for any of the Panels we will report that event.

6.3 Results and Inference-Profile

All of the DA-test results are presented in Table 3:

Table 3: Study Results DA-misclassifications with respect to the ISS-classification Protocol

GAAP:Variables	%Misclassifications	95%CI	Aligned: Power	ER ²
Balance Sheet				
X_1	39.9%	[34.2% : 45.5%]	Not Aligned: >85%	0.007
X_2	56.6%	[50.9% : 62.4%]	Not Aligned: >85%	<0
X_3	51.0%	[45.3% : 56.8%]	Not Aligned: >85%	<0
X_4	55.2%	[49.5% : 60.1%]	Not Aligned: >85%	<0
X_5	50.1%	[45.3% : 56.8%]	Not Aligned: >85%	<0.
X_6	41.6%	[35.9% : 47.3%]	Not Aligned: >85%	<0.002
X_7 Vetting	36.0%	[30.5% : 41.6%]	Not Aligned: >85%	0.08
Income Statement				
X_8	39.5%	[33.2% : 45.8%]	Not Aligned: >85%	<0
X_9	45.2%	[38.9% : 51.5%]	Not Aligned: >85%	<0
X_{10}	40.7%	[34.4% : 47.0%]	Not Aligned: >85%	<0
X_{11} Vetting	35.1%	[29.1% : 41.2%]	Not Aligned: >85%	0.09
X_{12} Vetting	33.2%	[27.3% : 39.1%]	Not Aligned: >85%	0.11
X_{13}	41.2%	[34.7% : 47.7%]	Not Aligned: >85%	<0
Cash Flow Statement				
X_{14}	43.7%	[37.9% : 49.4%]	Not Aligned: >85%	0.006
X_{15}	38.3%	[32.3% : 44.3%]	Not Aligned: >85%	0.02
X_{16}	53.3%	[47.0% : 59.6%]	Not Aligned: >85%	<0
X_{17} Vetting	43.4%	[37.6% : 49.1%]	Not Aligned: >85%	0.02
Value in the Trading Markets				
X_{18}	42.4%	[29.8% : 55.0%]	Not Aligned: >85%	0.04
X_{19}	46.8%	[40.7% : 52.8%]	Not Aligned: >85%	<0
X_{20} Vetting	38.9%	[32.7% : 44.1%]	Not Aligned: >85%	0.4
X_{21}	42.4%	[36.6% : 48.3%]	Not Aligned: >85%	0.2
Average, n=21	43.5%	[40.6% : 46.5%]	N/A	N/A

The inferential results are very clear and consistent with the two Lusk andWells studies. As simple vetting indications: (i) all the random test-screens, excepting three, did not produce a non-directional FPE vetting p-values that were < 0.05. For these three, we selected the smallest of the misclassification-percentages for the study. For the five-vetting checks in Table 3, there was no evidence that the magnitude of these GAAP-accounts were aligned with the ISS-assignment to the

polar-ISS[GQS]-groups. As for the results of the 16-Calibrated GAAP-accounts where the scale was selected to eliminate the magnitude-effect, there was NO instance where these calibrated accounts exhibited an inferential alignment with the ISS-Assignment. This is clear from:

- (i) the magnitude of the misclassifications all of which were $> 14\%$,
- (ii) (ii) the ER^2 all of which were in the $[1-.5^{.5}]$ -zone, and
- (iii) (iii) the Power $[1-FNE]$ for all the trials was $> 85\%$.

In this case, the rejection of *Ho* is clearly indicated and this then offers as the likely State of Nature *Ha*: There is no evidence that the major driver of the ISS-assignment protocol that created the polar-decile groups: ISS[GQS[1]] and ISS[GQS[10]] was the GAAP-reported data as reported by the BBTs.

7. Summary and Outlook

7.1 Summary

The results from Lusk and Wells (2020a,b) and the above DA-profiles offer the same message: The Vetting-conclusion of these three analyses is that the ISS-GQS-assignment to the decile polar groups is not aligned with the GAAP-reported data. In this sense, the latent result, not actually tested, is:

that Corporate Governance Risk is relative ethereal—will-o'-the-wisp-esque in nature. We risk this characterization, even in the light of the indication offered by Lusk and Wells (2021a) that ISS[GQS[1]] is a collection of firms inferentially characterized by management of Revenue while ISS[GQS[10]] are firms concerned with Asset management, because it is curious that in the final analysis firms with excellent management of CGR are not somehow differentiated in the trading markets from the set of firms that are at the polar distance from the BEST group—i.e., the ISS[GQS[10]]—the highest CGR-firms. So, this seems to beg the question—*What are the benefits of paying attention to managing the CGR?*

In this regard, we read with interest the research report of Huber and DiGabriele (2021) who note:

Corporate governance has been the subject of dozens, if not hundreds, of books and articles in legal, accounting, finance, and economic literature since at least 1932. Disclosure has also been the subject of dozens, if not hundreds, of books and articles in legal, accounting, finance, and economic literature, but interest in the subject is a more recent phenomenon. It is important therefore to understand the purpose, scope, limitations, and meaning of corporate governance. It is equally important to understand the purpose, scope, and limitations of the effective transparency of information, i.e., disclosure, for publicly listed companies, including what information is disclosed, how it is disclosed, and why it is disclosed.

Yes, certainly it is the case that the more intel that is available the more an analyst can partition the accrual dataset and perhaps tease out differences in the firms so partitioned. In fact, this is the exact tact taken by Dorfeitner, Kreuzer and Sparrer (2020) who offer:

This paper is the first one investigating positive screened portfolios dependent on the controversies score, which measures the amount of ESG-based controversies a company has faced. The calculations based on the Fama and French (2015) five-factor model show that there is still potential for an investor to achieve a significant outperformance. Even though a value-weighted investing strategy does not show any significant over- or underperformance and therefore confirms many of the previous literature findings (see: Halbritter and Dorfeitner (2015), we can find some noteworthy results.

7.2 Outlook

In our paper, we did not see the benefit of intra-group-portioning the ISS-firms in our accrual-set as there is not sufficient range in the misclassifications to hope for an inferential result of differences over the partition(s). Additionally, it is important to realize that the ISS-coding is integrated in the ESG of Bloomberg. We did not ISS-block the analysis over the various six ESG-dimensions of the scoring of the ESG one measure of which would have been the ISS-grouping. This we suggest is a valuable next step.

Appendix

Table A1: ISS Bloomberg-GQS[1] Assigned Firms

AMGN	BECN	CYH	DDD	DE	EIX	ELS	FFG
FISV	FOSL	GNRC	GTS	HRB	HSII	IVAC	KLAC
KRA	LEG	LNG	LRN	MLP	MRLN	MTG	MTX
MWA	PEAK	ROL	TPX	TXT	UFCS		

Table A2: ISS Bloomberg-GQS[10] Assigned Firms

ADTN	ADUS	ASTE	ASUR	ATEC	AZPN	BRKR	BWA
CODA	CONN	CVGI	DZSI	FCEL	HTH	INOD	KONP
KWR	LWAY	MDP	MNTX	NGS	NWL	ORA	RICK
SHYF	STCN	TAP	TEN	TNAV	TRXC	TSLA	

Acknowledgments

Appreciation is due to: Prof. Dr. H. Wright, Boston University: Department of Mathematics and Statistics, Mr. Frank Heilig: Strategic Risk-Management, Volkswagen Leasing GmbH, Braunschweig, Germany, The Faculty of the School of Business and Economics: SUNY: Plattsburgh for the SBE Workshop Series: in particular: Prof. Dr. Kameliia Petrova, for their careful reading, helpful comments, and suggestions.

References

- Dorfeitner, G., Kreuzer, C. & Sparrer, C. (2020). ESG controversies and controversial ESG: About silent saints and small sinners. *Journal of Asset Management*, 21,393–412 <<https://doi.org/10.1057/s41260-020-00178-x>>
- Fama, E. & K. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics*, 116, 1–22.
- Fraser, L. & Ormiston, A. (2013). *Understanding Financial Statement Analysis*, 10th Edition, Pearson: ISBN-13: 978-0-13-265506-4.
- Halbritter, G., & G. Dorfeitner. (2015). The wages of social responsibility—Where are they? A critical review of ESG investing. *Review of Financial Economics*, 26, 25–35.
- Huber-M, B. & Comstock, M. (2017). ESG Reports and rating: What they are, why they matter? *Corporate Governance Advisory*, 25, 1-10.
- Huber, W.D., DiGabriele, J. A. (2021). Corporate governance and disclosure: purpose, scope, and limitations *International Journal of Disclosure and Governance*, 18, 153-160. <<https://doi.org/10.1057/s41310-021-00103-7>>
- Lusk, E. & Wells, M. (2021a). Vetting of Bloomberg’s ESG Governance ISS: QualityScore[GQS™]. *ABR*, 9, 15-42. <<https://doi.org/10.14738/abr.94.9952>>.
- Lusk, E. & Wells, M. (2021b). Bloomberg’s ESG Governance ISS: QualityScore[GQS™]: Vetting results of a taxonomic sorting trial. *I. J. Scientific & Management Research*, 4, 97-105. <https://ijsmr.in/volume-4-issue-5-september-2021/>
- Soekarno, S. & Kinanthi, E. (2020). Discriminant function analysis to distinguish the performance of Information and Communication Technology (ICT) companies: A study of U.S. companies listed in U.S. Stock Market, *The Asian Journal of Technology Management*, 13, 113-128.
- Tamhane, A. & Dunlop, D. (2000). *Statistics and Data Analysis*. Prentice Hall: ISBN: 0-13-744426-5.
- Wang, H. & Chow, S.-C. (2007). Sample size calculation for comparing proportions. Test for equality: *Wiley encyclopedia of clinical trials*. <<https://doi.org/10.1002/9780471462422.eoc005>>

-
- Zopounidis, C., Garefalakis, A., Lemonakis, C. & Passas, I. (2020). Environmental, social and corporate governance framework for corporate disclosure: A multicriteria dimension analysis approach. *Management Decision*, 58, 2473-2496. <<https://doi.org/10.1108/MD-10-2019-1341>>

ⁱ Here is the active URL-link to EDGAR: <<https://www.sec.gov/edgar/searchedgar/companysearch.html>>

ⁱⁱ See the PCAOB-link:https://pcaobus.org/oversight/standards/archived-standards/pre-reorganized-auditing-standards-interpretations/details/Auditing_Standard_5 [Auditing Standard No. 5 [An Audit of Internal Control Over Financial Reporting that is Integrated with an Audit of Financial Statements [APPENDIX A – Definitions]]A11. A **significant deficiency** is a deficiency, or a combination of deficiencies, in internal control over financial reporting that is less severe than a material weakness, yet important enough to merit attention by those responsible for oversight of the company's financial reporting.}]

ⁱⁱⁱ The ER^2 can have a negative value when the DA-model selected does not follow, in major sections of the Panel, the linear-trend of the data. In this case, whenever the ER^2 is <0 this is by definition a clear indication of a lack of DA:ISS alignment.

^{iv} We have selected the GAAP-accounts that are, in spirit, the standard-accounts often used in ferreting out the essential nature of the firm in a competitive market. A standard text, the we have used in our Audit and Assurance courses over the years; Fraser & Ormiston (2013) and the recent research report of Soekarno & Kinanthi (2020) were invaluable in selecting the GAAP-variable set for testing the vetting the ISS-assignment protocol. If our 21-GAAP-variables were, in nature, in the genre-set of the Soekarno & Kinanthi or in that of Fraser & Ormiston we noted that indication as (SK or FO).

^v $SASTM[JMP^{TM}[Analysis\ Multivariate[v.13][Discriminant]]]$