Journal of Law, Finance, and Accounting, 2020, 5: 307–336

Sense and Nonsense in ESG Ratings

Ingo Walter*

Department of Finance, Stern School of Business, New York University, New York, NY 10012, USA; iwalter@stern.nyu.edu

ABSTRACT

Concerns about the future of the natural environment, prevailing social conditions, and governance of private and public institutions inspire today's ESG movement. This paper proposes a heuristic that can be useful in examining the ESG-scoring issue. We begin with a social control diagnostic covering business activities – one that addresses the interests and actions of various stakeholders in the system. We examine its dynamics in the context of economic, social, and political pressures, including various initiatives to set standards against which business conduct may be calibrated. We evaluate efforts to create metrics that reflect normative improvements in ESG outcomes and performance scoring against them. We assess the industrial organization of the ESG ratings industry and review key empirical studies of ESG-driven investing. We conclude with policy recommendations intended to alleviate existing shortcomings in ESG ratings and improve their role in capital allocation and corporate governance.

Keywords: Environmental degradation, social impact, sustainability, corporate governance, corporate ratings, target funds, investment fund performance

JEL Codes: B55, D18, D30, D62, D68, D74, F53, H89, I30, K29, K38, K42, K49, Z128

1 Introduction

Like gravity, market outcomes are a powerful guide to what makes things better or worse – for calibrating the costs and benefits of private decisions

*I am grateful to John Chambers, Patrizio Cox, Fernando Falbo, Lebogang Mahlare, Roy Smith, Paul Tice, Tensie Whalen and Robert Whitelaw for comments on previous drafts and for research support to the New York University Global Institute for Advanced Study.

ISSN 2380-5005; DOI 10.1561/108.00000049
©2020 I. Walter

Electronic copy available at: https://ssrn.com/abstract=3696718

and public policies that touch on economic performance, resource allocation, and growth. They are granular, and ultimately aligned with human nature in defining economic welfare. In trying to assess the consequences of actions that disturb market outcomes, it is generally good to be humble, and the law of unintended consequences is never very far away.

Today's environment-social-governance (ESG) movement reflects a waterfall of impact initiatives aimed at remediating damage to the natural environment, improving social justice, and reconfiguring the process of governance in both the public and private sectors. The ESG model has value, but it also has conceptual and practical flaws, notably in the hands of those looking for user-friendly calibrations and metrics without bothering to think too hard about them.

Concerns about each of today's ESG impact targets have simmered for decades and even centuries. Issues flare up from time to time as high-profile excesses occupy the public consciousness and need to be handled. Most of the time they are brought to a tolerable resolution and problem ESG behavior gets changed, sometimes after contentious public debate. Not everybody is happy with the relocated goalposts. But things tend to move along, and the market-driven system adapts without being mortally wounded.

So, what is new about many of today's ESG initiatives? One objective seems to be to ramp up precision on what is meant by each of the three letters, what is measurable and actionable, and what deviations from market-driven conditions are tolerable and sustainable. Some comments on each of the three ESG clusters:

Creating sustainability metrics and estimating remediation costs involving the *natural environment* remains a daunting task despite the ability to draw on natural science. Environmental resources involve costs and benefits that are usually hard to measure, unevenly distributed, and often cross political borders. Nevertheless, *E-scoring* is in vogue, notably focusing on climate change.

Creating the kinds of *social impact metrics* needed for *S-scoring* is even tougher. Credibility requires that highly granular (even interpersonal) welfare comparisons have to be drawn, and they have to be defended as legitimate. People and communities are different, so welfare comparisons among them are conceptually challenged. Economists have faced this legitimacy problem forever, but are resigned to living with it in the world of policy analysis – either requiring an array of assumptions, caveats, and footnotes, or accepting second-best outcomes (Lipsey and Lancaster, 1956). S-scoring in the ESG context takes this problem to a new level.

Governance debates have been vigorously productive for decades. We know a lot more than we did a century ago about incentives, interest alignments, fiduciary responsibilities, transparency, and accountability – among the key terms in the governance vocabulary.[1] So *G-scoring* may have a somewhat

[1] For a survey, see Smith and Walter (2006).

Electronic copy available at: https://ssrn.com/abstract=3696718

firmer foundation than its partner-ESG components. But it remains a work in progress, judging from the frequency of governance eruptions and the diversity of governance systems around the world.[2]

Composite ESG-scoring and its use as a guide to capital allocation, business strategy, and public policy seems a stretch. Maybe it is. In the meantime, however, a cottage industry of scorers and activists has formed among shareholders, investment managers, executives, social activists, and a host of commentators to create definitive-sounding metrics. Alphanumeric displays and rankings tend to command disproportionate attention in most people's minds.

The following section proposes a heuristic for sorting out the elements of a social control platform that captures the key normative targets of the ESG movement from the perspective of modern business firms and investment funds that mostly operate under the classic conditions of market discipline. The interactive dynamics of the diagnostic are examined in the following sections of the paper, leading to a discussion of the ESG rating process itself – attempts to identify, calibrate, and condense corporate and investor conduct across a wide spectrum of behaviors against embedded normative benchmarks. The industrial economics of the ESG ratings business is discussed next, as is evidence from empirical studies of ESG-driven portfolio returns. The paper concludes with workable reforms to improve the process and enhance the value of ESG analytics.

2 The Social Control Platform

Market forces governing corporate control and shareholder value are supposed to ensure that agency conflicts are kept in check and that waste, inefficiency, and bad management are best reflected through share price formation in well-functioning, continuous-time capital markets. The broad outlines are deceptively simple. Share-price erosion eventually tends to attract "active" financial investors or the management teams of other firms who think they can do better. They are willing to put their capital down and seek control. Incumbent management can respond with array of financial and operational moves to fend off a change in control, which may or may not succeed. Both offense and defense may ultimately put forward some of the same strategies and tactics to boost performance.

Who wins contests for control in well-functioning markets is less important than improvements in the performance of firms subject to market discipline. Whether the presumptive performance gains are sustainable is a question for another day. In the moment, the market decides. This story has to be adapted

[2]Promoting the Chinese blend of a semi-directed market economy under a thick blanket of political totalitarianism and aggressive surveillance is a tough sell in much of the rest of the world.

Electronic copy available at: https://ssrn.com/abstract=3696718

Figure 1: The External Control Platform.

to institutional settings with inefficient capital markets, unequal control rights, block-holders such as families, government ownership, and other institutional factors. The devil is always in the details.

Consider a generic business firm in the center of Figure 1 – call it the "social control platform" (SCP). The core mission of enterprise managers and boards of directors is to deliver an economically valuable and sustainable business model and execute at high levels of performance in a competitive marketplace.

This means consistent pursuit of revenues and attention to costs and efficiencies, balanced by a well-resourced and sharp eye on the array of risks to which the firm is exposed – risks that are driven in good part by the firm's own strategy and its execution. Some of these costs and risks are economic and financial – and are therefore relatively tractable in terms of calibration, measurement and pricing. Others are much more subtle, harder to measure and more difficult to build into management and investor decisions.

Figure 1 depicts this SCP heuristic as functional segments of a web that surrounds the operation of a market-driven firm within a given political-economic setting. Each segment of the web has its own history and differs across geographies.

We know that virtually all forms of commercial activity are limited, influenced, or explicitly controlled by external constraints in one form or another, and that the control mechanism is dynamic over time. The idea is that these constraint "domains" governing market-driven conduct will balance the distribution of power, countervail excesses, and help resolve conflicting interests

Electronic copy available at: https://ssrn.com/abstract=3696718

among firms and their customers, employees, shareholders, and other stakeholders. It forms an uncomfortable equilibrium, but one that is understood, tolerated, and continually tested over time.

3 Systemic Stresses

Cracks in SCP may show up from time to time and lead to behavior that tests the limits. In the early 2000s, for example, corporate governance failures in the United States helped create major bankruptcies affecting shareholders and all kinds of other stakeholders. Although the number of cases was small (only about 20 out of the more than 15,000 companies filing annual reports with the Securities and Exchange Commission at the time), the impact was substantial. Four of the 10 largest bankruptcies from 1980 to 2004 occurred in less than a year.[3] The failures and the underlying corporate scandals helped deflate the "tech bubble" of the late 1990s and may have influenced the dynamics of future stock market cycles.

The corporate misconduct of the time showed great variety, but it highlighted a common thread: weaknesses in the corporate governance domain of the social constraint platform in the USA. On the surface, it appeared that proper governance mechanisms were in place. Financials were audited by top accounting firms. Boards of directors seemed to follow sound governance principles. Nonexecutive board members were deemed sufficiently independent in carrying out their duties of care and loyalty to shareholders. But in retrospect, governance mechanisms in place at the time either did not function as intended or were suppressed by embedded conflicts of interest.[4]

So the pendulum swung. The internal and external governance failures that surfaced – and the economic losses they helped bring about – triggered calls for new ways to control bad governance, something that market discipline had evidently failed to achieve. One result was the Sarbanes-Oxley Act of 2002 (SOX), which set new or expanded requirements for boards of public companies, senior management, and public accounting firms. It added criminal penalties for some kinds of misconduct and required the SEC to create regulations defining more precisely how public corporations had to comply with the law.

Sarbanes-Oxley added a new element to the SCP depicted in Figure 1. There were plenty of complaints about SOX. But in the end, the US social control platform was modified and made more robust. Even so, the issues that triggered SOX pale in comparison to the hair-raising financial market practices of the 1920s, which led to the creation of the SEC in 1933 and

[3] Enron in December 2001, Global Crossing in January 2002, Adelphia in June 2002, and WorldCom in July 2002.

[4] For an overview of the governance benchmarks and failures at the time, see Smith and Walter (2006).

Electronic copy available at: https://ssrn.com/abstract=3696718

passage of its enforcement mandate in 1934. A tough series of Congressional hearings produced far-reaching and durable financial reforms – hearings that reportedly brought tears to the eyes of the legendary J.P. Morgan himself (Pecora Commission, 1932).

In practice, the compliance tools that policymakers have at their disposal range from "fitness and properness" criteria under which enterprises may be established, allowed to continue to operate, or be shut down; line-of-business constraints on what types of activities they may undertake; and what kinds of management and governance practices must be applied. The toolkit can leave its own footprints on firms and markets. And it can be problematic when there are jurisdictional conflicts, or when markets evolve rapidly and the rules get one or two steps behind.

These basic parameters of a sensible "inner core" of the SCP depicted in Figure 1 seem uncontroversial. Nevertheless, they are a product of a political dynamic within which enterprises themselves play a key role. The overall challenge is to create and sustain an SCP that blends efficiency, growth, fairness, robustness, and competitiveness. Each of these terms is subject to interpretation, and so is the degree of alignment among diverse stakeholders. How should efforts to pursue "the public interest" be targeted? What does "sustainability" mean? What distinguishes "good" control initiatives from "bad" ones? What is the balance between self-regulatory codes of practice and formal statutory rules and enforcement? Should the key elements of the SCP be more detailed and compliance-based or more general and principles-based?

Regulators constantly face the risk that permeable controls will result in costly failures, but also that excessive regulation will result in significant opportunity costs – or business flight to other, more friendly control regimes. Since any foregone gains can only be measured in terms of damage that *did not occur* and regulatory costs that were successfully *avoided*, the argumentation is inevitably centered on "what if" hypotheticals and counterfactuals. And there is the issue of public attitude toward systemic risk – itself conditioned by past policy failures. So, there are no definitive answers with respect to the optimum design of control platforms. There are only "better" and "worse" solutions, as perceived by the public to whom the policymakers are ultimately responsible.

4 Weaker Signals on the Periphery – Expectations and Values

The previous section focused on the inner "regulation and compliance" core of the SCP depicted in Figure 1. Rules of the game are formed at the social and political level and find their way into the legislative and regulatory framework within which the market functions. But laws and regulations governing the market conduct of firms are themselves rooted in people's expectations as to what is appropriate and inappropriate, and these in turn are driven by

Electronic copy available at: https://ssrn.com/abstract=3696718

embedded values. Most are quite basic. They deal with lying, cheating, and stealing, with trust and honor, or with what is considered right and wrong, responsible and irresponsible. They can be found, for example, in religion and broadly held ethical values. They form defining benchmarks against which conduct is ultimately measured – often well beyond the mandates of explicit regulation.

So, it is possible for an individual enterprise (or managers of a pool of financial assets) to be in full compliance with regulation and law yet run seriously afoul of what is broadly considered responsible and honorable. Conduct that is in violation of implied and sometimes nuanced social contracts can trigger severe and tangible losses in enterprise value. These reputational and conduct dimensions populate the outer rings of the SCP heuristic depicted in Figure 1. They are arguably much more difficult for managers, boards, and investors to understand and address. *Predatory*, *irresponsible*, and *unethical* are bad words when describing behavior in markets, and they are not easy to overcome. There may be plenty of dissonance in how these words are defined and interpreted, but their collective impact can be enough to get investor attention. This is where much of the ESG action is concentrated.

To elaborate, consider Figure 2, a cross-section of the SCP in Figure 1. The firm operates at the center of the social control platform, comprising multiple layers. The two innermost depict the external constraint and compliance process just discussed, and the legal environment in which it is rooted. Not so in the next layer of the SCP, where enterprise conduct is neither in violation of regulatory constraints or legislation yet is deemed to be in conflict with social norms and expectations. Consequences may include customer defections that damage revenues or create supply chain stresses. Investors may flee to safer harbors. Or the events of the day may turn out to be a precursor to tighter regulatory scrutiny, civil litigation, or legislative remedies down the road. So a defensible business case can be made for management "preemption" or "voluntarism" in addressing such issues. Revenues can be protected, future "retrofit" costs avoided, and reputational risks reduced.

There can be considerable slippage between the periphery of Figure 2 and the inner core of public policy that transforms those expectations into rules and constraints on firm conduct. Perhaps a firm goes too far. There is outrage. It turns out that the offending firm is not a rogue outlier. Instead, its conduct simply reflects "industry practice." Confluence of opinion on the issue then informs public policy. The system reacts, and another set of constraints on firm conduct emerges in the inner core of the SCP in the form of legislation, regulation, and enforcement.

In short, it is incumbent upon management, boards, and fiduciaries who carry control rights on behalf of shareholders under duties of care and loyalty to execute a defensible balance between market performance and the dynamics of the social control platform, and to manage the associated risks. The con-

Electronic copy available at: https://ssrn.com/abstract=3696718

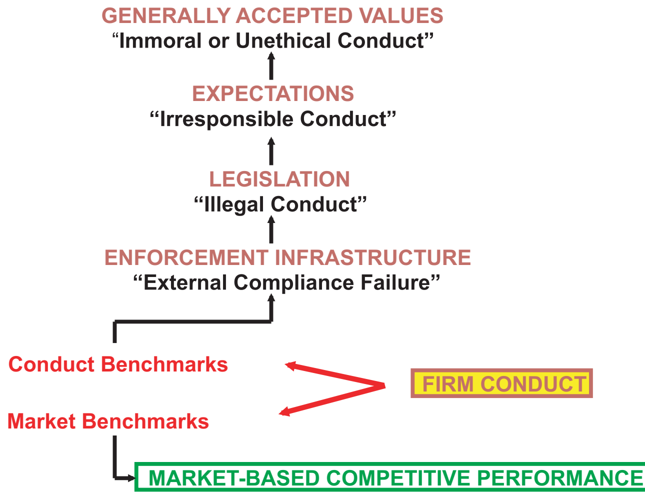

Figure 2: Performance Gaps, Competition & Conflict

temporary ESG movement is intended to help do just that. It repackages what many well-managed and sustainably successful companies have consistently tried to do, under a new-sounding rubric claimed to be better suited to today's social ecosystem.[5] But identifying, calibrating, and then scoring specific ESG conduct benchmarks is a heroic exercise, one with a great deal to prove.

5 Political Economy and the Social Control Platform

Business firms and investors are active stakeholders in the environment that emerges from the interaction between market drivers and the social control platform discussed here. The constraint-setting process is by nature political. It reflects perceptions of self-interest and the "common good" among diverse constituencies that are ultimately reconciled through the political process.

Lobbying is usually the dominant tool for securing more favorable SCP rules of engagement – and for tilting the social control dynamic in favor of well-organized, politically powerful groups. This can involve payments in cash or in kind, "educating" public officials and their staffs, distributing perks, or supporting reelection through campaign contributions. Anecdotes abound of paid advocates drafting rules that govern industries like financial services –

[5]Message to shareholders and asset managers: "I think investing without considering material ESG issues is breaking the fiduciary duties investors have – anyone who is doing that should be out of a job or in jail" (Karp, 2019).

Electronic copy available at: https://ssrn.com/abstract=3696718

proposed rules that are sometimes adopted essentially intact (even verbatim) by legislators and regulators.

Lobbying advocates argue that their efforts help public officials understand the complexities of SCP issues affecting a firm or industry and reflect the collective views of the electorate. But in the public mind, lobbying is often vilified as a parasite attached to the political underbelly – one that can divert the contours of the SCP away from the public interest to serve special interests with big resources and high and concentrated stakes in the game. Prominent examples include spates of deregulation and reregulation in industries like banking, automotive, pharma, tort law, and private-equity taxation. Collateral damage? Lost efficiency or sacrificed fairness. Sometimes both.

Beyond lobbying lies bribery of government officials to corrupt the inner contours of the SCP. Bribery can usually be traced to market imperfections created by regulation itself. The problem is that market imperfections are themselves necessary in order to make the social control platform function as intended and ultimately serve the public interest.

So, lobbying, rent-seeking and bribery reside in the shadows, ready to undermine and contort the SCP and disrupt an already fragile equilibrium (Baum *et al.*, 2019) What are the options for resolving tensions between the SCP and market-based competitive drivers?

6 Reconciling ESG Norms and Market Practice

Misconduct and governance failures continue to arise despite the long history of the SCP control and compliance architecture – notwithstanding periodic allegations of corporate irresponsibility and ethical failures and their adverse effects on the value of business franchises. One option is "codes of conduct" or "codes of best practice," which can be applied across multiple tiers of the SCP.

Standards of appropriate behavior may cover various conduct domains, possibly preempting formal regulatory constraints. Adherence to industry-level or more general standards of conduct applied by self-regulatory organizations (SROs), established at the industry or functional level, can be cheaper and more efficient than government regulation. They can also be easier to adapt to changing circumstances. This works to the general benefit, as long as codes are drafted and monitored effectively by the relevant stakeholders. SROs are part of the constraint system in many countries, perhaps led by the United Kingdom, which has a long tradition of principles-based codes of governance and behavior.

At the other extreme of the conduct-code spectrum is China's Social Credit System, launched by the Chinese Communist Party in a regional beta test in 2001 and 2002 to bring economic and social behavior in line with the Party's decisions as to what China's SCP ought to look like. Taking full

Electronic copy available at: https://ssrn.com/abstract=3696718

effect in 2020 and backed by China's growing mass-surveillance infrastructure, based on big-data analytics and artificial intelligence, it represents a unified national reputation system intended to standardize the assessment of citizens' and businesses' economic and social conduct. It is built around a single system-wide social credit score for each citizen and business. The conduct inputs (behavioral debits and credits) and their weights are determined by the Communist Party.

China's tech-driven Social Credit System provides high-velocity feedback and adjustments in the social credit score for an individual or enterprise. This in turn triggers rewards or punishments such as loss of employment, travel restrictions, access to credit, and a range of other carrots and sticks. Sometimes justified by the alleged tendency of Chinese individuals, businesses, and regional governments to ignore, violate, or work around rules that don't suit them, the Social Credit System is designed to create SCP order and enforcement around a single, monolithic vision dictated by the Party.

How will China's Social Credit System work out? Perhaps not too well if market behavior and human nature ultimately emerge to undermine the system, as they have done historically in most alternative visions of social order. Or maybe it will be considered better than nothing in an otherwise chaotic Chinese context. But the model is unlikely to travel well as China extends its influence globally. With this as an extreme example, in today's ESG world it seems codes of conduct are in vogue.

The United Nations Principles for Responsible Investment (PRI) – launched under the auspices of the UN Environment Program (UNEP) in 2005 – stands out among global ESG codes of conduct. It had accumulated about 2,400 signatories by the end of 2019.[6] Signatories are "... required to provide a signed declaration on company letterhead to commit to ESG issues in their investment analysis and decisions, to promote the PRIs within the investment industry, and to publicly report on their progress toward implementing the principles."[7]

The UN PRI set up a (not particularly granular) litmus test against which SCP conduct can be calibrated. It strives to compress some of the "weak-signal" outer reaches of Figure 1 toward the more familiar regulatory/compliance center of the diagram. Its credibility rests on the presumptive consensus embodied by the signatories and on the UN Charter itself.[8]

[6]Specifically: "(1) We will incorporate environmental, social and corporate governance (ESG) issues into investment analysis and decision-making processes; (2) We will be active owners and incorporate ESG issues into our ownership policies and practices; (3) We will seek appropriate disclosure on ESG issues by the entities in which we invest; (4) We will promote acceptance and implementation of the principles within the investment industry; (5) We will work together to enhance our effectiveness in implementing the principles. (6) We will each report on our activities and progress towards implementing the principles."

[7]https://www.investopedia.com/terms/u/un-principles-responsible-investment-pri.asp

[8]Among investment funds, limited evidence so far suggests that only select signatories

Electronic copy available at: https://ssrn.com/abstract=3696718

There is no shortage of other codes of conduct, including the 2003 IFC-sponsored Equator Principles, intended as an operational risk management framework for calibrating and addressing environmental and social problems in project finance – notably in developing countries.[9] Another example is the Poseidon Principles, put forward in 2019 through the International Maritime Organization (IMO), part of the UN Economic and Social Council dealing with the global shipping industry.[10] If well implemented and carefully targeted, such "functional" conduct initiatives can be helpful in enabling firms and investors to focus on key ESG issues in a targeted range of activities and make decisions with reference to what appear to be consensus views on responsibility and sustainability.

In 2018, UNEP put forward its Principles for Responsible Banking, which took effect at the end of September 2019. The key pillars encouraged signatory banks to align their operations at "strategic, portfolio, and transactional levels" to prevailing social and environmental commitments, particularly those expressed in the Paris Climate Agreement and the UN's Sustainable Development Goals (SDGs).[11]

At a national level, the US Business Roundtable – a corporate grouping with about 200 members that together generated over $7 trillion in revenues in 2019 – issued a "Statement of Purpose of a Corporation" in 2000. It downgraded its traditional focus on shareholder value and fiduciary responsibility to the lowest priority.[12] Evidently the members of the Business Roundtable decided

make visible changes to ESG while most are using the PRI to attract capital" (Kim and Yoon, 2020).

[9]The Equator Principles cover project financing components such as corporate loans and bridge loans, guarantees and standby facilities, vendor financing, etc., as well as project advisory services. By early 2019, 96 major financial firms involved in the majority of emerging-market project financings had signed the Equator Principles and pledged to adhere to its due diligence, monitoring, and decision-making standards. https://equator-principles.com/

[10]The IMO purports to make a serious effort to cut back emissions by at least 50% by 2050. Slowing ship speeds has been an early area of discussion because higher speeds of diesel-powered vessels generate disproportionately high levels of various pollutants. A key objective is to encourage financial institutions to align their ship finance portfolios to the objective of shipping decarbonization. The Principles are applicable to lenders, lessors, and financial guarantors globally. Eleven banks heavily engaged in ship finance were signatories of the IMO declaration by the end of 2019, all of them European firms except Citigroup.

[11]The PRBs try to define what "responsible conduct" means and set global conduct parameters. The goal is a framework through which banks can align their strategy, portfolios, and business practices with the social and environmental objectives embedded in the PRBs. Banks are obliged to periodically disclose their positive and negative social, economic, and environmental impacts, and publish verifiable targets in accordance with the ambitions of the Paris Climate Agreement and the SDGs. Additionally, they must assure stakeholder engagement, accountability, and transparency. Banks that fail to meet their announced targets or fail to adhere to the PRBs will lose their signatory status.

[12]In order of priority they are: "(1) Investing in our employees. This starts with compensating them fairly and providing important benefits. It also includes supporting them through training and education that help develop new skills for a rapidly changing

Electronic copy available at: https://ssrn.com/abstract=3696718

that ESG issues had become serious enough to require a blanket response while leaving the details for later. The context was growing populism, fed by important deregulation in the inner core of the American SCP. It coincided with increasingly vocal populist concerns about the ability of market-based approaches to attack specific SCP domains such as the natural environment, medical care, racial discrimination, and wealth distribution.[13] There were plenty of questions about what "inclusive capitalism" actually means and what the unintended consequences might be.

Observers debated whether the Business Roundtable exercise in the face of growing ESG pressure was a defensive, well-timed public relations exercise or a substantive change in the consensus view of the mission of the modern corporation. Key elements were unilateral, voluntary implementation by corporate signatories and the absence of enforcement mechanisms. Less than a year later, the Business Roundtable principles underwent an unexpected stress test with the onset of the Covid-19 pandemic and the systemic racism debate of 2020 – and related government policy responses. Anecdotal evidence suggests that the principles made little difference in layoffs, employee benefits, dividends, share repurchases, and management compensation (Goodman, 2020). Given the high profile of the Business Roundtable principles and the severity of damage to corporate stakeholders that may long be remembered, it is arguable whether the whole exercise was in the interests of the signatory firms.

Besides the fog of well-intentioned codes and pronouncements, there have been some moderately credible ESG standard-setting efforts in the regulatory infrastructure. For example, the Sustainability Accounting Standards Board (SASB) was launched in 2011, with the support of several foundations, as an initiative intended to evolve ESG principles in the direction of the US Federal Accounting Standards Board (FASB). It created benchmarks aimed at helping businesses report "financially material" ESG issues. In 2018 it announced ESG guidelines covering 77 industries. As of March 2020, 130 companies had filed SASB reports in full or in part. Another ESG reporting effort, the Global Reporting Initiative (GRI), was launched in 1997 as an alternative focus on ESG impact-materiality rather than financial-materiality. Companies have been advised to report against both sets of benchmarks (Polk, 2020).

The SASB and GRI standards themselves, how they were determined, and how they must be reported are predictably controversial. SASB-based stock benchmarks created by index vendors like Bloomberg and compliance pressures

world. We foster diversity and inclusion, dignity, and respect; (2) Dealing fairly and ethically with our suppliers. We are dedicated to serving as good partners to the other companies, large and small, that help us meet our missions; (3) Supporting the communities in which we work. We respect the people in our communities and protect the environment by embracing sustainable practices across our businesses; (4) Generating long-term value for shareholders, who provide the capital that allows companies to invest, grow and innovate."

[13] https://opportunity.businessroundtable.org/ourcommitment/

Electronic copy available at: https://ssrn.com/abstract=3696718

applied by major institutional investors like BlackRock and Vanguard add to the complexity. The SASB objective is presumably to eventually integrate its standards into the Form 10-K that must be filed by US public companies with the SEC.[14] This would represent a quantum leap in SASB legitimacy, but it would also impose on the financial economy reporting standards that are far less definitive or generally accepted than accounting standards.[15] The GRI standards have the additional problem of calibrating complex and often obscure ESG outcomes.

7 Developing ESG Metrics and Tracking Investment Effects

The ESG movement – its appeal and its shortcomings – can be easily rooted in the SCP heuristic discussed in the first part of this paper and summarized in Figure 3. It purports to identify and gauge the importance of "desirable" states of the world with respect to the natural environment, social conditions, and governance practices, and create a set of benchmarks against which the conduct of business firms – and those who hold control rights as investors or fiduciaries – can be gauged. Assessment of enterprise or investment-fund conduct is then calibrated against these benchmarks, usually producing some sort of alphanumeric scoring to identify degrees of underperformance. Patterns of underperformance are purported to reflect management and governance shortcomings, which in turn affect the market's view of business firms' performance in a world of shifting conduct goalposts, changes in investor preferences, and attitude toward risk.

Most of these ESG descriptors, discussed below, can be considered legitimate goals within the framework of the social control platform discussed here.[16] Each can be helpful by extending the social constraint process beyond the regulatory- and compliance-focused inner core of the SCP toward the preemptive adaptation and alignment with stakeholder expectations and

[14] https://www.sasb.org/

[15] Meantime, a cottage industry has emerged. Alternative-assets data provider Prequin works with SASB to create ESG indicators for its clients. SASB sells its industrial classification system and analytics to asset managers which then create ESG products. A cohort of ESG other standard-setters compete for attention. To be safe, many companies report under several of the leading standards. Ultimate sanctions for not filing or misfiling SASB or other ESG reports will presumably be consequential. How and to what degree remains unclear. The big issue is what is supposed to be reported, to whom, in what form, and what it actually means, given that many of the ESG objectives may themselves not be credible or generally accepted. Equally predictably, the major accounting firms have rushed to offer their "sustainability accounting services," hoping they will learn more about sustainability accounting and reporting as they go along. Coherence may be a long time in coming, and the key underlying issues will continue to be vigorously debated when they get there.

[16] State Street Global Advisors launched a diversity-driven ETF See https://us.spdrs.com/en/etf/spdr-ssga-gender-diversity-index-etf-SHE

Electronic copy available at: https://ssrn.com/abstract=3696718

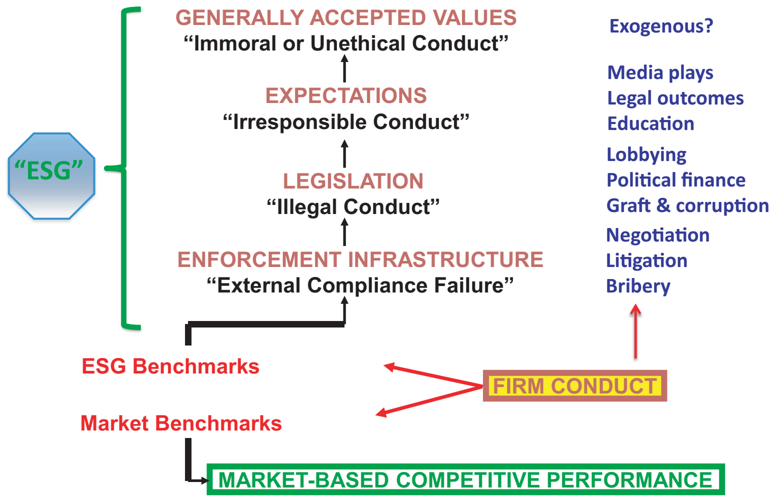

Figure 3: Firm Response Options and ESG

values at the periphery of the SCP platform. Weak signals from the periphery may be amplified by well-structured ESG assessments before they reach management and board levels. Indeed, they may already be reflected in share prices. Together, adverse ESG signals and stock price movements may trigger preemptive corrections in a firm's business strategy and practice, avoiding more serious consequences down the road.

Perhaps the most visible private-sector example of external ESG pressure through corporate monitoring and control is BlackRock – the world's largest fund manager with over $7.4 trillion of assets under management in mid-2020. The firm has attempted to create a distinctive brand for itself as an arbiter of corporate conduct. Because clients are considered long-term investors, ESG risks and opportunities can significantly affect long-horizon returns. The embedded assumption is that the best companies in which BlackRock invests consistently demonstrate operational excellence and transparency, so that shareholders and other stakeholders have enough information to understand the drivers of competitive performance and risks.[17]

[17]BlackRock claims to build mutual understanding with portfolio companies and asks pertinent ESG questions, rather than prescribing solutions. It claims to engage with about 1,500 companies a year at the executive and board levels on a range of ESG-related issues. It also seeks the views of large shareholders and consultants. A key aim is to help inform its own proxy voting and asset allocation decisions. Along with Vanguard (a cooperative), BlackRock (a listed company) is one of the world's largest holders of proxy votes, mainly due to the size of its footprint in the market for passive funds. Divesting individual companies on ESG grounds confounds its need to maintain index weights, so proxy voting is considered the broadest available form of engagement. BlackRock annually votes at more than 15,000

Electronic copy available at: https://ssrn.com/abstract=3696718

The firm exerts pressure on corporate boards by applying its own normative SCP views through its control of proxy votes – over two thirds of US corporate control rights were exercised by institutional investors in 2000. BlackRock, State Street, and Vanguard alone held 20% of listed US shares in 2020. They and other major fiduciaries have based their ESG benchmarks and performance indicators either on their own homework or by following advisory firms like Institutional Shareholder Services (ISS) or Glass-Lewis. Voting rights attached to shares apply both to actively managed and passive (index) funds. So institutional investors have become important control agents as ESG issues have risen to board-level significance. They are in a position (without a public mandate) to impose their own views on boards of directors, whether or not they reflect the views of the ultimate shareowners – a fact that itself has serious fiduciary implications.

For example, in March 2020, BlackRock threatened to punish the directors of corporations that did not meet to meet its expectations on corporate conduct issues, placing ESG at the center of its investment process. It asked for evidence of meaningful progress from companies on "sustainability" issues, which had been raised in 2019 by CEO Larry Fink as central to BlackRock's stewardship and investment businesses.[18] This was an effort to position BlackRock as a leader in "sustainable" investing after criticism that the firm had failed to exercise its governance power on issues such as climate change. In announcing the specifics of its intensified stewardship focus going forward, BlackRock raised the possibility of voting each year against the longest-serving non-executive board member of offending companies.[19]

Such investment-fund-based efforts to encourage ESG "improvements" is based on active due diligence in each of the three ESG dimensions, applying calibrations and weights that align with preferences embedded in their fiduciary mandates, interacting with firms at the right level on issues of concern, applying control rights where necessary, and ultimately making decisions on portfolio allocations. Performance forensics are presumably candid enough to limit agency problems and satisfy clients who can otherwise apply political pressure or move their assets.

shareholder meetings on over 130,000 proposals. How exactly this massive workload is actually carried out is not obvious. BlackRock's starting position is to support management and allow time to address or resolve ESG issues that may arise. Proxy votes against management proposals are confined to cases of corporate unresponsiveness. This seems a sensible, preemptive sequence, although hidden behind the curtain is the nature of the targeted issues, the standards applied, the underlying shareholder mandate, the metrics used, and the decision process.

[18] https://www.blackrock.com/corporate/investor-relations/larry-fink-ceo-letter

[19] "BlackRock To Target Companies on Governance Despite Coronavirus," Financial Times, 18 March 2020 at https://www.ft.com/content/03d73458-6876-11ea-a3c9-1fe6fedcca75

Electronic copy available at: https://ssrn.com/abstract=3696718

Things get more complicated when we shift from in-house application of ESG criteria to the actual business of for-profit ESG rating services. At issue is the quantitative enormity of the task, the integrity and transparency of the rating process, the application of criteria weights, and the display metrics that are applied.

A key question is whether the cloud of normative and positive ESG issues can be concentrated into defensible descriptors, which would be based on information and interpretation that successfully sweeps up the key variables or proxies and rates them, weights them, adjusts them for sector and country specifics, and produces some sort of alphanumeric output that is useful in asset pricing and selection. Such "ratings manufacturing" issues are compounded by the diversity of business models among ESG raters themselves, conflicts of interest that inevitably arise, as well as overreliance, misinterpretation, and herding behavior on the part of ESG ratings users. Importantly, manufacturing ESG ratings should be transparent as to data sources and identification, quantification, calibration, weighting, aggregation, and especially the ability to replicate. These are attributes that dominate defensible social science research. Any shortcomings will continue to cast a cloud over the whole exercise, one that will not easily dissipate.

8 The ESG Rating Industry

Despite all of the conceptual and practical problems of creating credible ESG benchmarks and indexes, an array of incumbents and new entrants to the ratings business have taken up the opportunity. The market for ESG ratings was estimated at $617 million in revenues in 2019, with growth foreseen to an estimated $1 billion by 2021 – including revenues from selling ESG indexes. Europe was estimated to account for 60% of ESG spending, compared to 33% in the USA and 7% in Asia. The majority of ESG rating purchases (59%) came from asset managers, 19% from sell-side firms, 12% from direct asset owners, 6% from consultants and advisers, and 4% from corporate entities (Bradford, 2020).

The push toward ESG scoring has added a new element to the global investment ratings industry, joining credit rating agencies (CRAs) dominated today by Moody's, S&P Global, and Fitch Ratings. CRAs trace their histories back to the 1890s and have endured multiple economic and credit cycles. They deal with a much narrower and more tractable set of issues – advising lenders and investors on the likelihood that borrowers will be willing and able to service their debts "in full and on time." In other sectors, firms like Morningstar and Thomson Reuters Lipper cover the performance of managed funds. RiskMetrics covers market risk exposure. Institutional Shareholder Services (ISS) and Glass Lewis advise on institutional investor proxy votes,

Electronic copy available at: https://ssrn.com/abstract=3696718

Environmental	Social	Governance
Airquality output	Access and affordability of product or service	Accounting and audit process
Biodiversityimpacts	Consumer rights	Board composition
Carbon Footprint	Corporate philanthropy	Business ethics
Climate change resiliency	Customer relations	Compliance
Energy consumption	Data security and customer privacy	Executive remuneration
Environmental policy	Diversityissues	Lobbying and political contributions
Fresh water use	Employee engagement	Ownership structure
Ground water depletion	Fair disclosure and labelling	Reporting and disclosure
Impacts on the cryosphere	Health and safety of communities	Shareholder rights
Impacts on the food supply	Human capital management	Succession planning
Land use	Human rights	Transparency
Natural resource management	Product quality and safety	Voting procedures
Ocean productivity and acidification	Responsible R&D	Advisory committee powers and composition
Regulatory and legal risks	Stakeholder and community relations	Client alignment and fee structure
Supply chain management	Supply chain management	Fund governance
Vulnerabilityto extreme weather		
Waste and hazardous materials management		

Source: Elyse Douglas, Tracy Van Holt and Tensie Whelan, "Responsible Investing: Guide to ESG Data Providers and Relevant Trends," Journal of Environmental Investing, 8, No. 1 (2017).

Figure 4: Some Common ESG Indicators

although both have come under fire from time to time. Governance Metrics International and The Corporate Library have also been visible through activist work on corporate governance.

Ratings form a large and profitable industry. Among ESG ratings and analytics services available to investors, suppliers range from small boutiques and NGOs to ESG units of the dominant credit-rating agencies and data vendors; and from sustainability generalists to industry-focused raters and issue specialists, including the ESG performance of sovereign states.

What is to be rated? The columns in Figure 4 list a small sample of conduct calibrations from which ESG ratings tend to be drawn, triaged, concentrated, and published as aggregate descriptors.[20]

There are many more, some of which significantly extend the types of conduct to be evaluated, and some that partially or wholly duplicate others. Expanding these columns into granular, normative profiles would vastly expand the number of entries and the complexity of the rating task.

[20]Investopedia at https://www.investopedia.com/terms/e/environmental-social-and-governance-esg-criteria.asp

Electronic copy available at: https://ssrn.com/abstract=3696718

	Indeces				Ratings				Ranking
Providers	ESG	E	S	G	ESG	E	S	G	ESG
Market Data[1]									
Bloomberg	x	x	x		x	x	x	x	
FTSE Russell	x	x			x	x	x	x	
MSCI	x	x	x		x	x	x	x	
Thomson Reuters	x	x	x	x	x	x	x	x	
ESG Exclusive[2]									
Arabesque					x	x	x	x	
Covalence	x				x	x	x	x	x
CSR Hub					x	x	x	x	
Ethos				x	x	x	x	x	
Inrate		x			x	x	x	x	
Oekom Research	x				x	x	x	x	x
Robeco SAM	x	x			x				x
Sustainalytics	x	x	x	x	x	x	x	x	x
VigeoEIRIS	x				x			x	x
Specialized[3]									
ISS/IW Financial					x			x	x
CDP		x				x			
RepRisk					x				x

Notes: 1. All market data providers offer ESG research, ratings and indices as a subset of their product and service offerings. 2. Focus solely on ESG Research, Ratings and Analysis. 3. Focus solely on one or more aspects of ESG but *not* all three. Source: Responsible Investing: Guide to ESG Data Providers and Relevant Trends – NYU Stern.

Figure 5: Coverage of Seventeen ESG Ratings Providers

Many of the listed considerations – such as “ownership structure” or “land use” or “responsible R&D” – are meaningless in the absence of careful impact identification and specification. Resolving these kinds of issues is a prerequisite to serious aggregation, assessment, and justification for business strategy and policy. Rating output that is definitive as a guide to management practice and capital allocation – in the balance between financial and social performance – seems like a heroic objective.

The basic issue of “what’s the question?” is compounded by the sources and quality of ESG inputs. They range from self-reporting and surveys to publicly available data, key-word scraping of media traffic, and inferences from stock market performance. There are plenty of identification, circularity, and causality problems. Profitable, high-growth-rated firms tend to score highly on multiple criteria, while struggling competitors tend to be pummeled on just about everything. There is also the competitive structure of the industry itself. Figure 5 depicts some of the major ESG rating providers, suggesting fertile ground for M&A activity as the industry consolidates (Douglas *et al.*, 2017).

All of the big market-data vendors like Bloomberg and Thomson Reuters offer ESG research, ratings, and indexes as a unit of their core global information and transaction services. So do the major credit rating agencies. The research-only, or “ESG-exclusive,” raters limit themselves to scoring and analytics sold to the investor community and have no other businesses that might trigger conflicts of interest. Not all focus on the three ESG components or release both indexes and ratings.

Electronic copy available at: https://ssrn.com/abstract=3696718

Each of the ratings players is in some direct or indirect way paid by investors, so they should be able to avoid the issuer-pays problem that continues to prevail among the credit rating agencies. One hopes that market discipline and securities regulation will ferret out corrupt "ratings-for-sale" operators. Some ratings services also provide consulting to clients on how to improve their ESG scores, so they run into the same audit/consulting problems that have plagued the big accounting firms from time to time. Dealing with these conflicts can be difficult and expensive.

Most of the major ESG raters cover thousands of entities using hundreds of indicators often grouped into "key issues." There is limited transparency among the indicators, proxies, algorithms, or qualitative proprietary assessment techniques applied. Presumably, this is the competitive "secret sauce" that provides value to ESG-sensitive investors. Approaches range from computer-driven models to analyst impressions and questionnaire-based evaluations. Hybrid approaches further erode transparency in the search for relevance. Application of artificial intelligence is on the way. With respect to the outputs, the alphanumeric scalars used by the various ratings suppliers are diverse, have relatively short track records, and are subject to periodic re-basing and revision. This confounds both longitudinal and cross-sectional rating comparisons and back-testing.[21]

"Specialized" raters and many others concentrate more narrowly on one or another of the individual ESG components or on specific sectors, which eases the enormity of the task and can help bolster credibility of the results – although limiting their certification power and market share.

Certainly, compared to credit ratings, ESG ratings seem conceptually and practically challenged. Manufacturing granular and credible performance metrics and weights from each of the SCP domains in Figure 1 – extending from the relatively definitive core out to the fuzzier periphery of a multidimensional construct – stresses the imagination.

Glossing over interpersonal preferences (or simply assuming uniformity in preferences) likewise puts credibility to the test. It helps erode confidence in the whole exercise, and it will never go away. It can be partially finessed by deferring to conduct guidelines such as the UN PRI and its companions, which aim to incorporate environmental, social, and governance factors into investment decisions and better manage ESG risk.[22] But it is still a long way to the equivalent of NRSRO certification with teeth that applies to US

[21]This is of natural concern to capital market regulators. In a 2020 statement, Jay Clayton, chair of the SEC, commented on the imprecision of aggregate ESG ratings, "I have not seen circumstances where combining an analysis of E, S and G together, across a broad range of companies, for example with a "rating" or 'score' – particularly since a rating or score would facilitate meaningful investment analysis that was not significantly over-inclusive or imprecise." https://www.ft.com/content/2c662135-4fd3-4c1b-9597-2c6f8f17faed

[22]See https://www.unpri.org/

Electronic copy available at: https://ssrn.com/abstract=3696718

credit rating agencies. S&P Global's acquisition of RobecoSAM's ESG ratings business and analyst teams in 2019 is noteworthy – one rationale being that ESG "performance," however wobbly, ultimately may have a bearing on an obligor's willingness and ability to meet debt service obligations. The same is true of the Moody's acquisition of French ESG rating agency Vigeo Eiris. Capitalizing on this synergy is doubtless a work in progress.[23]

Given conditions on the ground, it follows that corporate and investor initiatives reliant on ESG ratings can be inordinately problematic. As noted, it all starts with the design, assignment, and weighting of ESG "performance" benchmarks. The terms under which a firm relates to each of the SCP's domains are complex, diverse, and interactive – and often subjective, arguable, and imbued with value judgments. Causality can be unclear. ESG goalposts can change over time, sometimes dramatically, and they can be geographically varied. Overcoming these issues and landing on a credible approach to ESG scoring is a real challenge, especially given the "weak signals" from the ethical and moral periphery of the SCP.

Rating vendors usually aspire to be the "gold standard." But it's hard to make gold out of a turnip. C*aveat emptor* rules. But even the most jaded skeptics and curmudgeons should pay some attention. In a land of voodoo, it's usually good to know something about voodoo.

9 ESG Ratings, Economic Performance and Capital Allocation

If the "stylized facts" discussed in this paper describe what actually happens in markets, then a good case can be made for ESG-driven institutional asset-management strategies and the incorporation of ESG considerations in individual investor behavior. Roughly $20.6 trillion was invested in 307 US ESG-sensitive funds identified by Morningstar in 2019, with total assets under management of $119.3 billion at the end of March 2020 in 305 U.S. "sustainable" funds.[24]

Investment managers specifically concerned about particular ESG impact issues may be willing to suffer the opportunity costs of possible total return shortfalls and still meet their fiduciary duties to ESG-sensitive clients. The high price that ESG investors might be willing to pay (the low risk-adjusted return that they may be willing to take) creates a tangible capital allocation

[23] For example, a BNP Paribas loan to JetBlue Airways that was linked to the ESG score in early 2020 amended an existing $550 million senior secured revolving credit facility and included a sustainability-linked provision for the airline to align its strategic initiatives with its ESG goals and objectives. At https://www.google.com/search?q=BNP+Paribas+loan+to+JetBlue+Airways&oq=BNP+Paribas+loan+to+JetBlue+Airways{&}aqs=chrome..69i57.3046j0j7&sourceid=chrome&ie=UTF-8

[24] Financial Times, May 28, 2020, at https://www.ft.com/content/2c662135-4fd3- 4c1b-9597-2c6f8f17faed

Electronic copy available at: https://ssrn.com/abstract=3696718

effect in line with their values. The reduced expected return is reflected in a lower cost of capital for the corresponding firms.[25]

From the perspective of mainstream corporate finance and asset pricing, it is reasonable to believe that firms incorporating ESG factors may see a positive effect on their long-term cash flows and lower operational risk from damaging confrontations with elements of the social control platform. If the market understands and prices these cash flow and risk effects, there should be zero impact on risk-adjusted returns. If the market underestimates the importance of these effects (or cannot measure them), then "good" ESG firms will outperform other firms on a risk-adjusted basis. And if there is abnormally high demand for ESG-linked shares, this will cut their risk-adjusted returns relative to non-ESG-linked shares, although socially conscientious investors may be happy to pay that price. Among the various perspectives on ESG criteria and the ways in which companies and investment funds are rated, there are two justifications for the SCP approach put forward in this paper.

The first issue is how the adoption of ESG principles affects economic and financial performance. The literature on this issue is already extensive. But since ESG definitions are not unambiguous and there are many ways to measure performance, there is no conclusive evidence so far on the impact on investors. One may argue that achieving high ESG scores reflects better performance and lower risks, stronger reputation, and consumer preferences, among other benefits. On the other hand, one may argue that higher ESG ratings are associated with higher capital and operating costs and increased exposure to ESG controversies. The debate remains open, with empirical results reported on both sides using existing ESG scoring metrics (Cornell, 2020).

Perhaps the most exhaustive review in this matter is Friede *et al.* (2015). The paper uses both meta-analyses and vote-count methodologies considering 2200 individual academic primary studies that focus on at least one of the three ESG pillars. It concludes that approximately 90% of the studies find a non-negative relation between ESG and financial performance, of which 47.9% of vote-count studies and 62.6% of meta-analyses find a positive relationship.

More recent studies have focused on the role of risk and "materiality." For instance, Dunn *et al.* (2018) consider the role of ESG criteria in risk exposure. To measure companies' exposure to environmental, social, and governance issues, the study uses the MSCI ESG dataset covering 5,000 companies based in a number of countries.

[25] Fund expenses are part of this issue, since they represent a drag on investment returns, which is essentially certain. One would expect actively managed funds wrapped in an ESG label to have a higher expense ratio than non-ESG funds because assembling ESG funds may involve additional research and other costs, and because investors are willing accept lower returns net of expenses in order to satisfy their ESG objectives. This should apply to passive funds as well. For example, Vanguard's ESG exchange traded fund (ETF) charged an expense ratio of 0.12% in 2020, four times what it charged on its S&P 500 fund. https://www.wsj.com/articles/bloomberg-sells-sustainability-but-buyer-beware-11583193439

Electronic copy available at: https://ssrn.com/abstract=3696718

The dataset accounts for 97% of the MSCI market cap and spans the period January 2007 to December 2015. The paper covers stocks in the Russell 3000 index for the US, MSCI World ex US for international developed countries, and MSCI Emerging for emerging markets. It reports forward-looking (*ex ante*) risk estimates from the Barra's GEM2L risk model. The study uses simple linear regressions controlling for company size and other stock characteristics and computes fixed effects to capture specific variation in risk over time and across sectors and countries. The key finding is that ESG ratings are negatively correlated with a given stock's risks for most of the specifications applied. Considering the more general specification, the study reports significant coefficients of −0.0117, −0.070 and −0.004 for total risk, stock-specific risk, and beta, respectively. The study also finds that Social and Governance indexes are more strongly correlated than Environmental indexes.

In a sample of 4,000 listed companies during the period from 2000 to 2014, Garvey *et al.* (2016) examine 18 "common ESG criteria." Across the dataset of 60,000 company/year observations, ESG-positive firms that encountered no problems show a total incremental return of +0.6% in the event year. ESG-positive firms which did encounter significant problems showed total return of −6.4% in the event year. Non-ESG firms with no problems showed a cumulative total return of +3.2% in the event year, and non-ESG companies with problems showed a cumulative total return of −4.0% in the event year.

Khan (2019) proposes a novel ESG score adopting the framework of "materiality" and studies the impact on stock returns. The new ESG measure introduces two considerations. First, it examines cross-industry variation in material ESG issues. In this framework an ESG issue is "material" if it affects financial performance and shareholder value. Note that this ESG perspective, instead of focusing on environmental and social externalities, is focused on shareholder interests. Second, it incorporates the cross-country variation in governance, taking into account country differences in ownership dispersion, institutional strength and legal tradition. Using a constructed ESG score combining companies' material E, S, and G scores over the 2013 to 2017 period and controlling for style, time, and sector differences, the study finds a top-quartile to bottom-quartile return[26] spread of 36 bps monthly.

The second issue involves the scoring metrics themselves. How do ESG rating agencies build their reported scores? Here there is greater consensus regarding performance of the ESG rating agencies facing important challenges in achieving a consistent methodology. As noted, they remain far less transparent than the output of credit rating agencies. Methodologies and indicators used to compute scores are often opaque, and results are mostly proprietary and are not in the public domain due to their 'investor-pays' business model.

[26]Performance is gross of transaction costs and does not reflect the impact of potential manager-imposed constraints.

Electronic copy available at: https://ssrn.com/abstract=3696718

The lack of common rating criteria is reflected in the very low correlations between different rating agencies.[27]

As ESG ratings have become more relevant in supplying information inputs for investors, confusion and skepticism about their reliability have emerged as key concerns. The main issues are that (a) most of the rating agencies basically use self-reported data to construct their metrics, and (b) there are no generally accepted guidelines as to how to compile, normalize, weight, and process ESG data. This lack of standardization is responsible for the significant variation between ESG scores across the competing rating agencies (Huber and Comstock, 2017). This well-known issue has been the focus of several recent studies.

In a 2018 paper, Doyle (2018) highlights the case of Bank of America. While the bank holding company's RepRisk score (CCC)[28] falls well below average, its Sustainalytics score (70)[29] is well above the average. According to Doyle, there is evidence for three major biases in ESG ratings.

- Higher scores are attributed to larger firms. More highly capitalized companies[30] show systematically higher ESG ratings, with an average value of 64 for mega-cap firms and 46 for micro-cap firms.

- Companies in countries with more disclosure regulation show higher scores, with an average value of 66 for Europe (most regulated) and 50 for North America (less regulated).

- The oversimplification of industry weighting and company alignment lead to biased ratings that do not properly account for company-specific risks. For instance, the utilities industry shows a 61.3 average ESG score, the highest value among all the sectors, while the lowest is 48 for the healthcare sector.

According to Doyle, this is not surprising given the unequal information availability result in that "... large, mature, dividend-focused companies like utilities score better than less mature companies that focus on reinvesting through R&D."[31]

Similar points are raised by Kotsantonis and Serafeim (2019). The study focuses on four limitations in the available data that potentially fuel the

[27]See, for example, "Heavy Flows into ESG Funds Raises Questions Over ratings." Financial Times, 3 March 2020 at https://www.ft.com/content/0bd9d2ea-5c15-11ea- 8033-fa40a0d65a98

[28]Rating scale: AAA-D.

[29]Rating Scale: 0–100.

[30]"Market capitalization is the market value of a company's outstanding shares calculated by multiplying the stock price by the total number of outstanding shares" Doyle (2018).

[31]Doyle (2018) at https://accfcorpgov.org/wp-content/uploads/2018/07/ACCF_RatingsESGReport.pdf, p. 5.

Electronic copy available at: https://ssrn.com/abstract=3696718

divergence across rating agencies: (1) Different firms have non-comparable metrics attributed to the same issue. The paper analyzes 50 large companies across various sectors and reports 20 different ways to report employee health and safety issues as well as different units of measure; (2) Since scores are computed in comparisons between different firms in a given peer group, a firm's performance will depend on the performance of the peer group (assigned by the rating agency) in the aggregate; and (3) There are wide differences in methodologies to incorporate missing data.

Among the 50 companies examined, only half had a health and safety policy, and about 15% reported incident rates and workplace fatalities. Employee health and safety is a material ESG issue for 9 out of the 11 sectors covered by the ESG analysts, who must fill the gaps using alternative approaches that may lead to quite different results. For example, analysts used several methods[32] to estimate 2017 employee turnover at Lufthansa, resulting in scores ranging from −8.9% to +7.1%. Moreover, the more data are disclosed, the greater tends to be to the variation across ESG ratings.

The latter issue is the subject of a study by Serafeim and Sikochi (2019), which uses fixed-effects OLS models in a sample of 30,700 firm-year observations from 2004 to 2016 and finds evidence of a positive relationship between disagreements among ESG scores and the degree of disclosure. When ESG disclosure increases by one standard deviation, ESG disagreement[33] increases by 1.41. The average ESG disagreement in the sample is 12.32, so this appears to be significant.[34] The intuition is that, as more information becomes more available, rating companies have wider discretion to use their own criteria and consequently assign different scores.

Such findings are consistent with a study by Berg *et al.* (2019). Using several techniques to disaggregate the divergence between five ESG rating agencies[35] on a balanced sample of 823 firms in 2014, the study finds that correlations between scores are on average 0.61 – far below correlations among credit rating agencies at 0.99. Among the three ESG pillars, the lowest average correlations appear in the G domain (0.38) while the highest appear in the E domain (0.65). The rating agency KLD showed the lowest correlations with the other raters with a correlation of −0.01 (for example, with the Asset4 rater) for the G domain. Measurement divergence (the use of different indicators to measure the same attribute), explains more than 50% of the overall divergence, while scope (the set of attributes in the scores) and weight (the relative importance of attributes) combined are slightly less important.

[32]Rules-based, Regression method, Predictive Mean Matching.

[33]*"The standard deviation of ESG ratings that a firm received for year t's ESG performance"* from MSCI IVA, TR Asset4 and Sustainalytics.

[34]"*The firm's ESG disclosure score* (range from 0.1 to 100) *for the ESG report pertaining to year t's performance"* from Bloomberg.

[35]KLD, Sustainalytics, Vigeo-Eiris, Asset4, and RobecoSAM.

Electronic copy available at: https://ssrn.com/abstract=3696718

In short, while rating agencies are trying to measure the same attributes, they employ different indicators and assign significantly different weights.

Based on the state of the literature, empirical work that meets accepted research standards will have to await much better calibrated and standardized ESG performance indicators, backed by strong intuition and consensus that they measure what they are intended to measure.[36]

Investors ought to be able to use ESG ratings as an effective way to screen and reduce reputational risk or advocate ESG principles. But ESG rating agencies have intractable problems delivering on these objectives. As Doyle (2018) notes, Volkswagen AG and Wells Fargo NA were awarded ESG scores well above their respective peers during years when serious internal failures were well underway, yet the ratings were not revised downward until after the scandals went public in 2015 and 2017, respectively.

ESG ratings also seem susceptible to gaming, including "greenwashing" – preemptive but meaningless alterations in strategies, tactics or operations designed to extract higher ESG scores, broaden access to ESG-sensitive investors, and extract higher share valuations. This creates a new form of financial fraud that can leverage an insider's understanding of the rating process and sophistication in covering tracks. The more ESG is taken seriously in asset pricing, the more vulnerable it becomes to fraudulent practices and to misuse of information. ESG ratings contain new information and can move markets. So ESG raters must be held to the same standards that apply to other rating services. As the credit rating agencies have shown in recent history, this is not easy.

10 Summary, Reforms, and Policy Recommendations

How well does ESG scoring align with a firm's strategic positioning and execution – especially when the radial components of the SCP in Figure 1 are themselves diverse and subject to change over time? How can armies of analysts

[36] A more conceptual take on ESG scoring was put forward by Porter *et al.* (2019). The authors raise the concept of "shared value" as a superior approach to ESG scoring. They claim that corporate managers should avoid ESG standards that are not material to corporate performance or consensus social progress criteria but should instead recognize their importance to long-term competitive advantage. Running an innovative and profitable business model is the best way to make a positive social impact – the central idea is shared value, identifying the connections between social issues and competitive strategy. The authors claim that shared value can affect business strategy at three mutually reinforcing levels: (a) Meeting social needs through new products; (b) Increasing productivity across the value chain; and (c) improving the business environment. From an investor perspective, instead of considering ESG metrics either as a measure of higher returns or as a tool to reduce regulatory or reputation risks, investors should seek out "... companies that achieve excellent economic performance by innovating to meet important societal needs." Porter, Michael E., George Serafeim, and Mark Kramer. "Where ESG Fails." Institutional Investor (October 16, 2019).

Electronic copy available at: https://ssrn.com/abstract=3696718

possibly meet this challenge for thousands of covered entities without resorting to empirical shortcuts and dodgy algorithms, and relying on a few available data inputs and proxies that may only be associated with a small fraction of the targeted ESG behavior? What about the causality and identification problems? Or ratings that rely on other raters' outputs for their own inputs? Things get even more challenging when efforts are made (as they should be) to incorporate inter-firm linkages through supply chains or input-output models to derive ESG summary scores. To what extent do even good faith but flawed metrics lose meaning when they are cooked down to aggregate scores that are inherently backward looking?

No less important is walking the talk. Firms that set themselves up as ESG performance paragons can be in for a rude awakening when they are found to have missed announced conduct targets. Just as the market can be unforgiving of missed earnings estimates, to the extent that ESG factors are priced into share values the consequences of ESG implementation failures may be similarly amplified. Markets may be ruthless in punishing firms that engage in ESG deception. The presumption is that ESG performance follows the rhetoric, and that investors really do care when they decide what to do with their financial assets.

To summarize, we concur that an ESG focus – in the context of a vigorous market-driven economy – is compatible with a social constraint platform that it looks to for alignment with the public interest. It is broadly consistent with the concerns of investors, the fiduciary obligations of asset managers, and a constructive framework for interaction with other (often overlapping and conflicted) stakeholders. As always, the devil is in the details.

We posit in this paper that, in a complex and often contradictory stakeholder environment, the conduct of any business enterprise tends to be calibrated against two sets of benchmarks: (1) Its performance in the competitive marketplace with respect to a well-established set of performance metrics; and (2) Its performance against a set of conduct standards that are partly exogenous and partly endogenous. These standards comprise the ESG world and extend from criteria anchored in regulation and law to fuzzy, weak-signal behavioral norms purported to define appropriate business and professional conduct.

Management straddles both sets of benchmarks. If it strays too far in the direction of meeting ESG pressures it runs the risk of poor performance in the marketplace, adverse shareholder reaction, leadership turnover, and possibly change of control. If it strays toward unrestrained market performance and sails too close to the wind, it can trigger regulatory or reputational pressure on the firm, its managers, its employees, and other stakeholders, which can ultimately find resonance with shareholders and is reflected in enterprise value.

There is constant tension between market performance and the social control platform outlined in this paper. Sometimes firms win battles (and

Electronic copy available at: https://ssrn.com/abstract=3696718

even wars), leading to periods of liberalization and paring away of constraints. Sometimes it is possible to convince the public that self-regulation or the adverse reputation effects of misconduct are powerful enough to obviate the need for external control. Sometimes the regulators can be convinced, one way or another, to go easy.

Within this heuristic, sensible ESG considerations can be a useful in diagnosing relevant issues, broadening the discourse, anticipating developments in the outer realms of the social control platform, and encouraging timely adjustment and engagement on the part of firms. To the extent that this increases revenues, lowers costs, and, especially, reduces idiosyncratic risk of the kind that is most difficult to calibrate, it can also enhance a firm's value as a going concern.

On the other hand, except as a way of flagging important social costs and signaling ways of dealing with them, ESG scoring and its misuse can be a convenient shortcut for stakeholders averse to hard work and critical thinking. Its key weaknesses run from identification and causality problems, self-reporting and greenwashing, quality of the primary data collection process, indicator mismatch in the use of secondary data, factor aggregation and weighting, setting scoring breaks, and reporting formats. This is a formidable list, with key dimensions that can frustrate transparency, replicability, and therefore credibility. Given the current state of play, it follows that there are a number of issues that must be addressed to reinforce ESG assessments and make them more defensible in capital allocation and enterprise management.

First, the identification problem must be resolved – what are the normative outcomes that are to be achieved, and how well anchored are they in consensus and public legitimacy? Inter-community and interpersonal utility differences hang over this issue and cannot be fully resolved, so solutions have to be framed within the theory of second-best.

Second, data inputs must be mapped specifically onto the normative outcomes that motivate the ESG exercise. What inputs can be captured, and are they internally coherent?[37] So far, buckshot is more descriptive than precision rifle fire in the use of whatever data happen to be available. But "good enough" is not good enough when serious capital-allocation decisions and fiduciary obligations are involved. So, plausibility and notional sufficiency – and the massive practical chore of generating and collating hundreds of inputs on thousands of rated entities by armies of analysts and statisticians of unknown competence – remain to be overcome.

Third, the weighting process used in creating composite ESG indexes is inherently vulnerable, since there are few objective sources of weights (unlike market capitalization or price volatilities in stock market indexes) in the world of ESG. Thus, weights too have to be judgmental, subjective, and subject to

[37]See Black and Yurtoglu (2017) on the concept of "construct validity."

Electronic copy available at: https://ssrn.com/abstract=3696718

challenge. This is unlikely to change in the short term, although cumulative data on stock-price sensitivity to ESG events may in the future make it possible to generate empirically defensible approaches to factor weighting.

Fourth, the unobserved variable issue needs to be addressed. This "blow-by" problem is commonly encountered in areas like sovereign risk ratings. Information that is not captured in data that is insufficiently granular – or dropped from weightings altogether – can turn out to be critical in performance outcomes. Vigorous debates within rating committees can help, but may not be feasible given the ESG rating caseload.

Fifth, transparency is critical. Data, ESG target specification, assumptions, adjustments for missing data, quantification of qualitative information, factor weighting, and index construction should all be available in the public domain. It should be possible for researchers to replicate ESG calibration and subject it to stress testing using alternative inputs and assumptions. Since ESG rating follows a subscription (user-pays) business model based on proprietary information and rating models, product differentiation may be marketed as the "secret sauce." So, transparency may be a long time coming.

Sixth, maturity and commonality are credibility factors that will come only with time. Credit ratings derived in the same way by different CRAs may not be identical, but they will tend not to show extreme differences. Rating convergence has long been common among the three leading credit rating agencies, and any rating differences and watch lists can be examined for remaining differences of view among the respective rating committees. The ESG rating business may well follow the same path, which implies a global oligopolistic market structure. But that will take time and creates a wholly different set of problems.

Seventh is the need for performance measurement and reporting on the part of investment funds marketed as ESG-driven. Investors are entitled to forensic audits confirming the extent to which progress toward the ESG goals has actually been achieved, alongside the conventional financial fund reporting. This will not be easy and is dependent on progress on the issue of defensible ESG performance metrics, noted above.

Lastly, there is the issue of certification. The CRA path offers a model in terms of the recognition by the SEC in 1975 of Nationally Recognized Securities Rating Organizations. Recognition by a serious, politically empowered regulator is important, as are the standards applied by that regulator. It is doubtful that NGO rules like the UN PRI will suffice. Regulatory oversight carries both costs and benefits. The costs of aligning with regulatory mandates are clear, both in obtaining operating licenses and ongoing compliance. The benefits derive from certification that is awarded in the public interest. The closer the ESG rating industry comes to the NRSRO model, the greater its credibility and impact – but at the same time the higher the likelihood of oligopoly.

Electronic copy available at: https://ssrn.com/abstract=3696718

These are among the key issues as the ESG ratings industry evolves. Investor and corporate behavior will place greater weight on substance, and some of the posturing by investors, asset managers, and executives will dissipate. No doubt the ESG rating industry will shake out over time. There will be new entrants, and incumbent players will consolidate and link to more substantial rating services with a lot to lose – notably the CRAs, stock exchanges, and data vendors. Improved and more granular metrics will develop. Decomposition of ESG into its constituent domains and narrowing its focus will ease some of ESG rating industry's "aggregation confusion" problems, improve the meaningfulness of information capture, and make calibration more credible, simpler, and easier to interpret.

References

Baum, A., C. Hackney, P. Medas, and M. Sy. 2019. "Governance and SOEs: How costly is corruption?" IMF Working Paper. International Monetary Fund. Washington, D.C., September 2019.

Berg, F., J. K. Julian, and R. Rigobon. 2019. "Aggregate confusion: The divergence of ESG ratings". MIT Sloan Research Paper No. 5822-19, August 17, 2019. Available at SSRN: URL: https://ssrn.com/abstract=3438533 or http://dx.doi.org/10.2139/ssrn.3438533.

Black, B. and B. B. Yurtoglu. 2017. "Corporate governance indices and construct validity". *Corporate Governance: An International Review*. 25: 397–410. URL: https://onlinelibrary.wiley.com/doi/10.1111/corg.12215.

Bradford, H. 2020. "ESG data market poised to hit $1 billion in 2021". Report, March 2020. Available at: URL: https://www.pionline.com/esg/esg-data-market-poised-hit-1-billion-2021-report.

Cornell, B. 2020. "The ESG concept has been overhyped and oversold". Open Parket Insight July 15, 2020. Financial Times. Available at: URL: https://www.ft.com/content/719e6253-4d07-402a-9f36-d461671657a1.

Doyle, T. M. 2018. "Ratings that don't rate: The subjective world of ESG rating agencies". American Council for Capital Formation, July 2018. Available at: URL: https://accfcorpgov.org/wp-content/uploads/2018/07/ACCF_RatingsESGReport.pdf.

Dunn, J., S. Fitzgibbons, and L. Pomorski. 2018. "Assessing risk through environmental, social and governance exposures". *Journal of Investment Management*. 16(1): 4–17.

Friede, G., T. Busch, and A. Bassen. 2015. "ESG and financial performance: Aggregated evidence from more than 2000 empirical studies". *Journal of Sustainable Finance & Investment*. 5(4): 210–233. DOI: 10.1080/20430795.2015.1118917.

Electronic copy available at: https://ssrn.com/abstract=3696718

Garvey, G. T., J. Kazdin, J. Nash, R. LaFond, and H. Safa. 2016. "A pitfall in ethical investing: ESG disclosures reveal vulnerabilities, not virtues". September 19, 2016. Available at SSRN: URL: https://ssrn.com/abstract=2840629.

Goodman, P. S. 2020. "A vow by big business proves too hard to keep". The New York Times, April 14, 2020.

Huber, B. M. and M. Comstock. 2017. "ESG reports and ratings: What they are, why they matter". Harvard Law School Forum on Corporate Governance.

Karp, E. 2019. "Investors, fasten your seat belts". Cornerstone Capital, July 2019. Available at: URL: https://www.corporatesecretary.com/articles/esg/31461/considering-esg-materiality-part-fiduciary-duty-says-cornerstone-capital-ceo/?utm_source=CS141218%5C&utm_medium=email%5C&utm_campaign=esg_forum.

Khan, M. 2019. "Corporate governance, ESG, and stock returns around the world". *Financial Analysts Journal*. 75(4): 103–123.

Kim, S. and A. Yoon. 2020. "Analyzing active managers' commitment to ESG: Evidence from United Nations principles for responsible investment". URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3555984.

Kotsantonis, S. and G. Serafeim. 2019. "Four things no one will tell you about ESG data". *Journal of Applied Corporate Finance*. 31(2). Spring 2019.

Lipsey, R. G. and K. Lancaster. 1956. "The general theory of second best". *Review of Economic Studies*. 24(1): 11–32.

Polk, D. 2020. "ESG disclosure frameworks – Recent updates". Client Memorandum. July 15, 2020. Available at: URL: https://ft.com/content/719e6253-4d07-403a-9f36-d461671657a1.

Serafeim, G. and A. Sikochi. 2019. *Why is corporate virtue in the eye of the beholder? The case of ESG ratings*. November 2019. Available at: URL: http://www.israeli-corporate-governance.org/files/content/banner/Sikochi_ESGRatings.pdf.

Smith, R. C. and I. Walter. 2006. *Governing the Modern Corporation: Capital Markets, Corporate Control and Economic Performance*. London: Oxford University Press.

Electronic copy available at: https://ssrn.com/abstract=3696718