GLOBAL ECONOMY & DEVELOPMENT
WORKING PAPER 103 | MAY 2017

Global Economy and Development at BROOKINGS

COMBINING GOOD BUSINESS AND GOOD DEVELOPMENT

EVIDENCE FROM IFC OPERATIONS

Raj M. Desai
Homi Kharas
Magdi Amin

B | Global Economy and Development
at BROOKINGS

Raj M. Desai is a visiting fellow in Global Economy and Development program at the Brookings Institution, and Associate Professor of International Development in the Walsh School of Foreign Service at Georgetown University.

Homi Kharas is a senior fellow and co-director in the Global Economy and Development program at the Brookings Institution.

Magdi Amin is the manager of Corporate Strategy and Partnership at the International Finance Commission.

Acknowledgments:

The authors are grateful to Brian Casabianca, Mozzafar Khan, Edmond Mjekiqi, Jane Nelson, and Elizabeth White for comments on earlier drafts. Lorenz Noe provided outstanding research assistance. All errors and omissions are the solely the authors' responsibility. Moreover, the findings, interpretations, and conclusions expressed are entirely those of the authors and do not reflect the views of the International Finance Corporation, the World Bank, their affiliated organizations, their executive directors, or the governments they represent.

Abstract

Most studies of the links between corporate environmental and socially sustainable actions find a positive link with financial performance. Most of these studies, however, have analyzed firms only in developed countries, and their identifying assumptions do not typically allow causal inferences. We are able to study firm performance in developing countries by using data from investments made by the International Finance Corporation (IFC) across over 1,000 projects between 2005 and 2014 in close to 100 middle- and low-income countries. The IFC provides information on the financial performance of their investments as well as ratings of environmental, social, and governance (ESG) outcomes. It is well established that such data may be affected by random and systematic error and that there is unavoidable endogeneity between ESG and financial performance since profitable firms can afford to behave more sustainably. Using an instrumental variables approach, we find that the relationship between a firm's sustainability behavior and profits disappears. However, there is some evidence that both sustainability and profits can jointly impact broader private sector development. Our results have implications for how best to blend public and private finance in the cause of sustainable development.

COMBINING GOOD BUSINESS AND GOOD DEVELOPMENT

EVIDENCE FROM IFC OPERATIONS

Raj M. Desai
Homi Kharas
Magdi Amin

INTRODUCTION

When businesses invest, they are expected to maximize financial profits for their owners. Investors look at firms' financial performance to guide capital allocations. Increasingly, they also look to material environmental, social, and governance (ESG) indicators as these have been shown to be leading indicators of future financial performance in developed countries (Khan et al., 2016). There is now a growing movement to integrate sustainability into the practices of all firms, including those in developing countries (Business and Sustainable Development Commission, 2017). But some developing countries have expressed concerns that high costs and new conditionalities for sustainability could undermine their growth (OECD, 2012). This paper analyzes the causal ESG/financial performance relationship in developing countries to address these concerns.

To understand the complementarity between financial performance and ESG performance, we analyzed investments carried out by the International Finance Corporation (IFC), the private sector arm of the World Bank Group in close to 100 lower- and middle-income countries between 2005 and 2014. During this period, the IFC invested in 2,475 firms in IFC-eligible, developing countries. This is a large and unique data set with which to explore the relationship between financial performance and developmental outcomes.

Current thinking on the relationship between ESG reporting and financial performance in developing countries is that there is indeed a positive correlation:

> *[A] convincing correlation exists between those investments that do well on a financial yardstick and those that show strong development results; moreover, integrating ESG criteria into the investment process appears to enhance financial performance (Wilson, 2013).*

This carefully worded view underscores the considerable causal ambiguity in interpreting the ESG/financial performance relationship. Better sustainability may trigger better financial performance, but the reverse can also be true. Better financial returns can permit more attention and resources to be dedicated to sustainability practices. Equally, when financial performance is poor, firms might cut sustainability activities first.

In fact, both sustainability and solid financial performance might result from an omitted variable: good management. In this case, it would not be fair to claim that ESG performance itself caused better financial performance, but only that the quality of management jointly determines both ESG and profits. The policy prescription would be to improve management, not to introduce regulations on ESG reporting. There is also reason to worry that subjective metrics like ESG performance that are constructed as judgments made by environmental and social specialists might be biased, again affecting empirical results, even though these judgments might be based in part on measurable indicators.

We address these identification issues by using instrumental variables (IV) estimation. We use two country-level variables as sources of exogenous variation in the ESG rating: (i) a measure quantifying a country's environmental performance; and (ii) the presence of the country as a member on the United Nations Security Council. We show that both these variables are highly significant determinants of project level ESG ratings—satisfying the relevance criterion—and argue that neither would be plausibly correlated with firm-level profitability—satisfying the exclusion condition.

The importance of correcting for endogeneity and measurement error is borne out by the empirical analysis. A simple ordinary least squares (OLS) shows a significant positive correlation between ESG and profitability, but this finding is not robust to our IV strategy.

We find no adverse causal link from ESG to lower profits (the principal concern of developing countries mentioned above). Conversely, however, there is no evidence that better ESG performance at the firm level causes improved financial performance. Moreover, there is limited evidence that both ESG and profitability independently contribute toward broader private sector development in the industry as a whole. One caveat: our project-level data usually only extend for two or three years, so we cannot explore how time lags may impact the ESG/profitability relationship.[1] Furthermore, we based our data set on clients of the International Finance Corporation, who have gone through several rounds of due diligence processes, and therefore our sample may not be representative of all firms in a country.

We draw two policy implications for the nature of public-private cooperation in developing countries from this analysis. First, there is evidence that individual firm activity can make markets in developing countries work better through improved competition, demonstration effects of a profitable business model that may not have been tried before in that context, or by spurring sectorwide policy, legal, or regulatory reforms. This suggests that commercial business activities can have a developmental impact beyond firm-level profits, providing a potential rationale for public intervention and the blending of public and private funds.

Second, we also find that ESG performance, by itself, can encourage firms to undertake activities that contribute to sustainable development, without noticeable harm to individual firms. Because there does not appear to be a financial incentive for firms to undertake ESG improving activities, governments may need to tilt incentives through regulatory requirements, or condition funding, technical assistance, or capacity building programs on ESG performance.

[1] By contrast, studies that focus on the ESG/stock price relationship implicitly factor in long-term impacts, as the stock price represents the discounted sum of all future profits. The short- to medium-term nature of our analysis is an inevitable limitation.

We find, additionally, no evidence to suggest that financial returns depend on geography or on the income level of the country. This finding has implications for portfolio adjustments (e.g., toward low-income countries), suggesting that such changes may be possible without foregoing financial returns. This conclusion, however, does not take into account any difference in origination costs that a company may incur in different regions. It is possible that these differ markedly across regions and in terms of size of the investment made. Finally, we find domestic financial sector investments tend to be more profitable than real sector investments, but real sector investments tend to have a greater impact on broad private sector development. If both profitability and sustainable development are objectives, then a balanced portfolio of financial and real sector investments will be needed.

THE PRIVATE SECTOR AND DEVELOPMENT IMPACT

Private sector growth is widely acknowledged to be an essential component in providing economic opportunity. Private sector activities can contribute to employment, wealth creation and poverty alleviation, taxes, and even provide essential services ranging from housing to heath to education to infrastructure. In recent years, the private sector has become more central to economic development. The 2015 Addis Ababa Action Agenda, which served as the financing framework for the U.N. Sustainable Development Goals, acknowledged that:

> *Private business activity, investment, and innovation are major drivers of productivity, inclusive economic growth, and job creation… We call on all businesses to apply their creativity and innovation to solving sustainable development challenges. We invite them to engage as partners in the development process, to invest in areas critical to sustainable development, and to shift to more sustainable consumption and production patterns (United Nations, 2015).*

The private sector, however, has a mixed track record with respect to managing inequality, environmental effects, human rights, or social conflicts. Hence a lively debate has ensued regarding the responsibility of businesses not simply as given by law, but to society at large.[2] Today, this debate is again taking place in boardrooms across the world. The Global Impact Investing Network (2016) identified $77 billion in assets under management by 156 impact investors. There are 1,500 signatories of the Principles for Responsible Investment with $60 trillion under management that have agreed as their first principle that "we will incorporate ESG issues into investment analysis and decisionmaking processes."

Mainstreaming ESG into core business models and taking it out of separate "corporate social responsibility" departments is a recent phenomenon. The contrary view was clearly and famously spelled out by Milton Friedman: "[T]here is one and only one social responsibility of business—to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game" (1962: 133). Four decades later, only 4 percent of respondents surveyed by *Business Week* supported that statement (Bernstein, 2000). More recently, 84 percent of a thousand global CEOs surveyed agreed that business "should lead efforts to define and deliver new goals on global priority issues" (Accenture and U.N. Global Compact, 2014). Partly, a new narrative that good sustainability practices can improve the reputation of a company—as well as increase sales, enhance employee loyalty, and attract better personnel—has driven this shift (Pfau et al., 2008). An added advantage for public companies is that sustainability activities may help them gain a possible listing in the *FTSE4Good Index,* the *Dow Jones Sustainability Index*, or other similar indices, potentially boosting stock prices (Economist Intelligence Unit, 2008). More recently, companies have emphasized the effect of sustainability as a driver of innovation.

[2] In the late 19th century, company towns, model villages, and, later, industrial colonies provided examples of what today would be called "corporate social responsibility." In the 1930s Adolf A. Berle argued that public policy should define a strict fiduciary duty for corporate management, while E. M. Dodd advocated for corporations to assume responsibility for their local communities, via public regulation if necessary. By the 1950s, Berle himself had become an advocate for progressive corporate law warning that, should businesses fail to voluntarily assume greater community responsibilities, the government would likely intervene in a less efficient way (see, e.g., Weiner, 1964; Berle, 1954; Page and Katz, 2011).

The balance of evidence, too, appears to tilt in favor of sustainability. One assessment finds that the pooled internal rate of return for investments made by "impact" private equity funds exceeds the small-cap market index (Gray et al., 2015). Private equity and venture capital funds with "impact missions" produce higher or equivalent returns as traditional funds (Cambridge Associates and Global Impact Investing Network, 2015). Meta-analyses show that social and (to a lesser extent) environmental responsibility pays off (Orlitzky, Schmidt, and Rynes, 2003). A review examining some 159 empirical analyses finds that the majority show a positive relationship between sustainability and financial performance (63 percent), while 15 percent of studies report a negative relationship, and 22 percent report a neutral or mixed relationship (Peloza and Yanchin, 2008). A meta-analysis of over 2,000 studies shows a strong business case for ESG investments, with 90 percent of these studies finding a positive relationship between environmental sustainability and financial performance (Friede, Busch, and Bassen, 2015).

For our purposes, however, three factors limit the validity of the conclusions that sustainability behavior and financial performance are positively related. First, the analyses mentioned above almost exclusively focus on publicly listed companies or investment funds within the United States, or within high-income OECD countries. There is precious little information on the relationship between sustainability and performance in low- and middle-income countries.[3] Moreover, as capital markets in developing countries are often small and underdeveloped, the reputational effects described as being the key channels from ESG to financial performance may be less important in developing countries.

Second, the definitions of what it means for business practices to be "sustainable" or "socially responsible" are inconsistent and widely diverse. In some instances, companies that provide favorable workplace environments for their employees are deemed "responsible." In other cases, the emphasis is on ecology and environmental stewardship. Moreover, some of the indicators used rely on business self-reporting rather than outside verification. Additionally, some companies have mixed historical records, further complicating how they should be assessed temporally:

> *Does Monsanto's support for genetically modified agriculture make it responsible or irresponsible? Should Dow Chemical be excluded from a sustainability-screened portfolio because of past pollution or included because of its recent leadership in tracking and reducing its environmental impact and that of its products? What about a fossil-fuel company that has recently expanded its renewable-energy investment portfolio? In the real world, a company's ethical behavior is both constantly changing and often mixed (Vogel, 2016).*

The salient indicators, therefore, are likely subject to measurement error due to critical information gaps, non-standard reporting and assessment (both across firms and across time periods), and rater bias.

Third, endogeneity presents a daunting problem in any examination of the relationship between sustainable practices and financial performance. There is the obvious possibility that causality is reversed: better-performing, more profitable companies may shift their portfolios or change their behaviors as or

[3] In the meta-analysis of 142 empirical studies undertaken by Hang, Geyer-Klingeberg, Rathgeber, and Stockl (2016) only 6.9 percent of studies have data from developing countries, while 54 percent of studies have data only from the U.S. or EU countries. This study suggests the effect between corporate environmental and financial performance in developing countries is positive, disappears in emerging economies, and reappears as positive in developed countries.

after they reach certain financial performance targets.[4] Alternatively, both sustainable practices and favorable business outcomes stem from good management, incentives, or supervision, creating a virtuous cycle whereby actions to increase sustainability can trigger good financial performance, and positive returns can permit additional investments in sustainability (Gaspar, 2013). There is also concern that sustainability can have an effect on the "metrics" used to measure financial performance or *vice versa* (see, e.g., Esty and Cort, 2016).

[4] This is similar to the problem of examining the wealth effects of "post-industrial" individual actions: do environmentally conscious citizens prosper, or do people value conservation and become more environmentally aware after they achieve a certain level of economic status (See Ott and Soretz, 2016)?

DATA

We rely principally on firm-level data from the International Finance Corporation,[5] IFC financial investments include direct and syndicated loans, equity, trade finance, and structured and securitized finance, as well as a number of other risk and liquidity instruments.

Since 2005, the IFC has used the Development Outcome Tracking System (DOTS) to measure the performance and development outcomes of its investment and advisory services (International Finance Corporation, 2011). DOTS data is collected at the client level, which may not correspond perfectly with the outcomes of the IFC projects. Within the DOTS are various measures of financial and economic performance, as well as indicators of environmental and social sustainability, chosen for relevance in eight different subsectors, indicators for corporate governance, and other indicators for broader impact on private sector development, such as on competition or influence on sectoral policy or regulations. Each IFC project is annually scored based on a judgment of company performance against a number of quantitative and qualitative indicators, identified at the start of the project. As the project matures, indicators are monitored on a yearly basis until the project closes.[6] Companies are given an overall development outcome rating that addresses a project's or company's contribution to the host country's development. This rating is based on several underlying performance areas such as demonstration effects of the investment, improvements to governance, improvements to the investment climate, and positive externalities for suppliers and customers. Details regarding the construction of individual scores may be found in the Appendix.

At the initial stage, a baseline, targets, and expected values for the indicators are developed. Some indicators are common across investments, some common across a sector, and some may be specific to a single investment. Following the approval of a project, the investment team must provide every year the actual value of each indicator (such as the actual number of employees). Each year, the indicator is rated on a four-point scale (surpassed, achieved, partly achieved, and not achieved) depending on the extent to which the actual value surpasses, matches, or does not meet the expected value.

DOTS ratings are "owned" by the investment departments that originate and supervise projects, but a central Development Impact Department analyzes and reviews the ratings to ensure the quality and consistency of the information contained in DOTS. In addition to ongoing quality control, DOTS ratings are comprehensively reviewed once a year during the Portfolio Review process to assess poor quality/insufficient data in support of ratings, inconsistencies between ratings and rationales, and outdated data. Finally, quality control includes assurance by external auditors on an annual basis.

[5] The International Finance Corporation was established in 1956 to further economic development by encouraging the growth of productive private enterprise in developing countries. It does this through financing private enterprise without public guarantees, where sufficient private capital is available on reasonable terms, and by creating the conditions conducive to the flow of private capital to the private sector. IFC's theory of change rests on demonstration effects. IFC seeks to identify underserved sectors or regions and support viable projects that serve as a signal to others whose further investments expand and complete markets. In 2003 a group of commercial banks launched the Equator Principles, modeled after the IFC's performance standards. IFC's new long-term strategy (IFC 3.0) expands on the positive externalities from private investment activity through market creation, better integrating its operations with efforts by the World Bank and others to address policy and regulatory constraints to markets, and increasing mobilization of third party capital. As market failures are addressed, IFC views its role in demonstrating sustainable investment with high development outcomes as essential to creating market solutions to development challenges.

[6] Further details on the DOTS scoring system are in the Appendix.

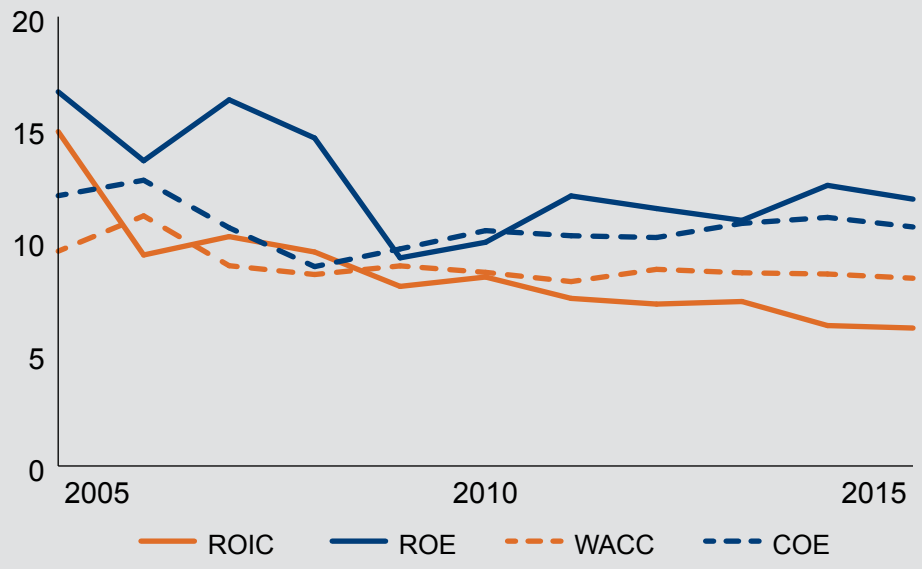
Financial Performance

The basic principle for rating the financial performance of a project or company is to assess whether the returns are sufficient to compensate the financiers for the risks they are taking. DOTS applies return on equity (ROE) for financial sector projects, and both annual and lifetime return on invested capital (ROIC) for real sector projects including infrastructure and natural resources. ROIC is computed as the ratio of net operating profits from core operations less adjusted taxes divided by total invested capital. ROE is the financial returns to equity holders (both dividend and capital gains) calculated as net income for the most recent year divided by average equity (average between the most recent year and the previous year). For each project, IFC also computes a weighted average cost of capital (WACC) for real-sector firms, average cost of equity (COE) for financial firms, that reflect the risk associated in the sector and country. Trends for both sets of return and cost indicators between, 2005 and 2015, are show in Figure 1. In our estimations of financial performance, we include the costs measures WACC or COE as independent variables.

As constructed, "financial performance" is a continuous variable, but it may be the case that ESG affects the probability of a firm being profitable rather than the actual level of profitability. Accordingly, we also construct and use a binary variable coded 1 if ROIC > WACC or ROE > COE, zero otherwise. Initially we combine both real and financial sector observations, but in subsequent specifications we split the sample between real and financial sector investments.

DOTS also provides a measure of economic returns defined as financial returns adjusted for taxes paid to the government, subsidies, externalities, import protection, consumer or producer surpluses, and worker benefits

Figure 1. Mean financial returns, 2005–2015



Source: Author's calculations

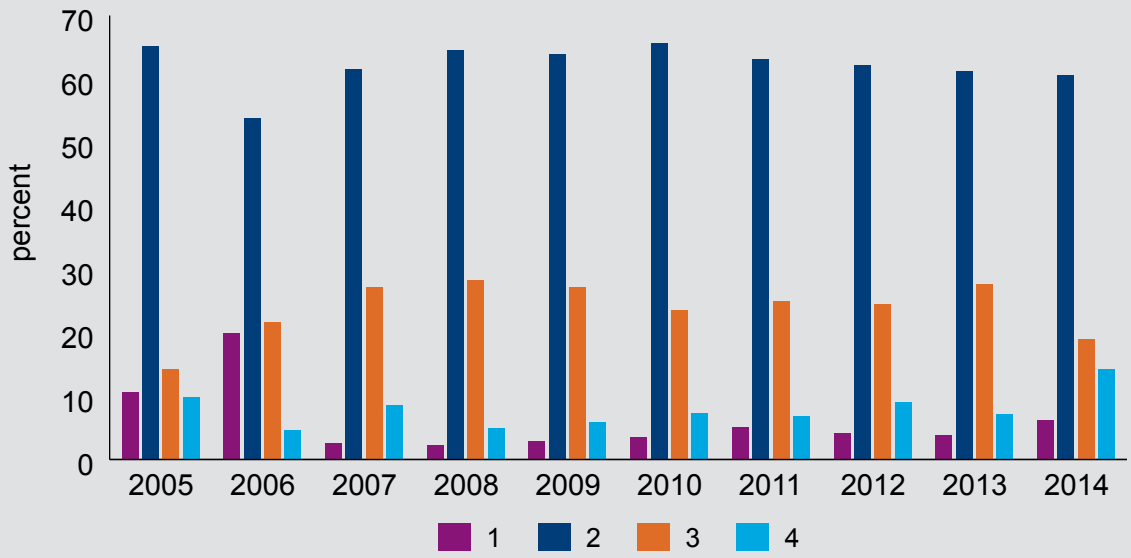
above their opportunity costs. We perform the same tests using economic profitability as the dependent variable.

ESG Variables

The degree to which a project meets IFC's environmental and social performance standards at approval, as well as the standards that would apply if the project were appraised today, provides the basis for rating a firm's environmental and social performance (IFC, 2012).[7] There are eight performance standards and 20 indicators that measure, *inter alia*, energy efficiency, pollution control, maintenance of health and safety standards, outreach to underserved markets, access to key services, gender issues, and other measures of community development.[8] The ESG score ranges from 1 ("excellent") to 4 ("unsatisfactory"). We rescale this from 0 (lowest) to 1 (highest).

Figure 2 graphs the distribution of raw scores for the ESG rating over time, and shows that over two-thirds of projects are consistently rated as "2." Of course, as with any indicator in which subjective assessment is a major component, there is a possibility of measurement error and bias. Measurement error can arise for many reasons, including the limited ability of project

Figure 2. ESG score distributions



Source: Author's calculations

[7] The World Bank Group Environmental, Health, and Safety Guidelines (EHS guidelines) are technical reference documents with general and industry-specific examples of good international industry practice. IFC uses the EHS guidelines as a technical source of information during project appraisal. The General EHS Guideline contains information on assessment and management of environmental and social risks and impacts; labor and working conditions; resource efficiency and pollution prevention; community health, safety, and security; land acquisition and involuntary resettlement; biodiversity conservation and sustainable management of living natural resources; indigenous peoples; and cultural heritage (IFC, 2012). The IFC also provides assistance to firms to meet these standards.

[8] IFC does not include governance indicators directly in this measure, but given the common usage, we refer below to the composite score as an "ESG" rating, even though it only covers environmental and social issues.

managers/evaluators to collect accurate information and the deviation between the variables specified in principle and data collected in practice. If an explanatory variable is measured with additive random errors, then the coefficient on that variable in an ordinary least squares regression will be biased toward zero the larger the sample; the higher the proportion of variability that is due to errors, the greater the bias (see, e.g., Angrist and Krueger, 2001; Murray, 2006).

Moreover, bias can also be systematic, based on personal perceptions that influence how any given rater evaluates a project. In this case the indices may not be independent of the error term. Specifically, firm performance and growth might influence judgments regarding sustainability.[9]

Instrumental Variables

We treat the project-level ESG rating as an endogenous regressor, and so need to select valid variables that may be used to instrument ESG. An instrument that is uncorrelated with errors in the ESG variable and the equation error from the model with the correctly measured data but correlated with the correctly measured variable, can provide a consistent estimate even in the presence of measurement error (Angrist and Krueger, 2001). We select two country-level instruments that satisfy these conditions.

First, we use an indicator of whether, in any given year, the country in which the project is being implemented is serving as a member of the United Nations Security Council (UNSC); if yes, the variable is coded 1, zero otherwise. While five members of the UNSC—China, France, Russia, the United States, and the United Kingdom—are permanent members, 10 other U.N. member-states serve two-year terms as elected members, having first been nominated by their regional caucus and having been approved by a two-thirds majority in the General Assembly. UNSC decisions require a majority of nine votes; each member-state has one vote, but the permanent members have veto power.

UNSC membership is considered a measure of the political "salience" of a country. There is considerable evidence that low- and middle-income countries serving on the UNSC benefit in a number of ways (Vreeland and Dreher, 2014). They receive increased bilateral aid from the U.S. and other major official donors (Kuziemko and Werker, 2006). They are required to comply with fewer conditions attached to loans from the International Monetary Fund (Dreher et al., 2006). More importantly, there is evidence of lower quality thresholds for World Bank projects in countries serving on the UNSC—projects that are rated lower than those in non-UNSC countries (Dreher et al., 2013).[10]

Our second instrumental variable is the country-level Yale-Columbia Environmental Performance Index (EPI), a measure of a state's policies regarding ecosystem protection, resource management, and prevention against environmental harm based on targets set forth in the U.N. Millennium Development Goals. The EPI was preceded by the Environmental

[9] This particular "poor is bad" bias has received attention in the literature on corruption indices (see, e.g., Søreide, 2006).

[10] Dreher et al. (2013) show that countries receiving World Bank projects as UNSC members are more likely to receive negative evaluations than those received by non-members, concluding that there may be political imperatives at work in extending projects that are lower-quality to UNSC members. They do not adjust for measurement error or bias.

Sustainability Index (ESI), published between 1999 and 2005.[11] The ESI was developed to evaluate environmental sustainability relative to the paths of other countries. The EPI uses outcome-oriented indicators, constructed as a benchmark index. Further details on the EPI may be found in the Appendix.

Both instruments—being country-level aggregates—cannot directly affect firm performance except through some effect on the sustainability of the project and are therefore expected to be orthogonal to the outcomes of interest. If it were the case that projects in politically important countries or projects in countries with better environmental records were rated better than other countries' projects, this would suggest a strongly relevant instrument in the form of UNSC membership or the EPI score. Additionally, both measures can also correct for a selection effect, wherein certain countries' firm-level investments are of higher quality at the outset (and therefore have a greater likelihood of sustainability). Finally, whether or not a country has a seat on the UNSC or whether a country's environmental policies are rated well or poorly, both being country-level processes, can be seen as quasi-random from the perspective of individual firms. Hence, these variables plausibly satisfy the exclusion criterion.

Control Variables

We include in our estimation, a number of project and country-level control variables. We include the total IFC commitment (either as a loan or in equity) in constant U.S. dollars, in natural logs, to account for project size. We also include GDP per capita also in U.S. dollars in natural logs, to control for overall country income and wealth effects. Given the effect of the overall legal-institutional environment on a firm's behavior, we include the "Rule of Law" indicator from the World Bank's *Worldwide Governance Indicators*. We include change in international reserves (in months of imports), on the assumption that firm performance can be affected by shifts in international liquidity and resulting pressures on local currency. We include a dummy coded "1" if the country is on the OECD "harmonized" list of fragile states and zero otherwise, to control for different risks and costs of businesses in states affected by conflict or instability. Finally, we include sector, geographic region, and time dummies.

[11] The ESI/EPI was developed jointly by the Center for Environmental Law and Policy (Yale University) and the Center for International Earth Science Information (Columbia University) in collaboration with the World Economic Forum and the Joint Research Centre of the European Commission.

METHODS AND RESULTS

We focus principally on the performance of firm *i* in country *c* at time *t* ($\pi_{i,c,t}$) and estimate the following model of the determinants of firm-level financial performance pooled across firms:

$$\pi_{i,c,t} = \beta_0 + \beta_1 s_{i,c,t} + \beta_2 \boldsymbol{x}'_{i,c,t} + \beta_3 \boldsymbol{t} + \varepsilon_{i,c,t} \quad (1)$$

where *s* is the (potentially endogenous) regressor measuring ESG, and $\boldsymbol{x}'$ is a vector of exogenous firm and country controls described above. Where π is measured as continuous, adjusted returns, we also include either the weighted average cost of capital (real sector), or the cost of equity (financial sector) in the vector of controls. Where profits are binary outcomes, we exclude costs as a control but use the linear probability model. Moreover, since we stack together real and financial sector observations in our preliminary estimations, we also include a "real sector" dummy variable. Finally, $\boldsymbol{t}$ is a vector of time dummies and ε is a random disturbance clustered at the country level.[12] Descriptive statistics for all variables used in our estimations can be found in the Appendix.

Table 1 shows our estimates from our basic, pooled OLS regression. Supporting earlier empirical analyses, these results show that higher sustainability ratings are consistently associated with greater financial and economic returns and a greater likelihood of profit for individual firms. A simple stochastic simulation of these firm outcomes, setting all other control variables at their sample means, suggests that a firm with the lowest-ranked ESG score (4) has a 45 percent chance of turning a profit, and an average return on capital/equity of 8 percent, while firms scoring the highest on the DOTS scale (1) have a 57 percent chance of showing profit that year, and an adjusted return of 11 percent.[13] Thus every incremental, single-point improvement in the ESG indicator raises the likelihood of profitability by 4 percentage points, and raises the return on capital/equity by 1 percentage point.

These initial regressions, however, may yield biased coefficients on ESG due to the endogeneity problem with the ESG rating. Accordingly, we shift to an instrumental variables (IV) regression, in which equation (1) becomes the second stage while in the first stage we estimate:

$$s_{i,c,t} = \alpha_0 + \alpha_1 \mathrm{UNSC}_{c,t} + \alpha_2 \mathrm{EPI}_{c,t} + \alpha_3 \boldsymbol{x}'_{i,c,t} + \alpha_4 \boldsymbol{t} + \mu_{i,c,t}, \quad (2)$$

where *s* is the firm-level ESG rating, UNSC and EPI are the two instruments described above, $\boldsymbol{x}'$ is again a vector of controls and μ is the country-clustered error term. We use the two-stage least squares (2SLS) estimator.

Instrument Validity

In Table 2 we present our first-stage results, that is, the coefficients from equation (2). The first stage relationship between both the UNSC membership indicator and EPI (our instruments) and the firm's ESG sustainability score is positive and significant. Investment projects in countries with seats on the UNSC, or in countries with higher EPIs are rated more highly in terms of environmental and social sustainability. This suggests that cluster-robust IV-2SLS estimates are not biased toward ordinary least squares (OLS) estimates. The Kleibergen-Paap/Wald rank statistic test rejects the null hypothesis that the equation is under-identified. The Sargan-Hansen test (Hansen's *J* statistic) for over-identification, on the other hand, does not reject the null suggesting that the instruments are

[12] Note that, although the maximum period is 2005 to 2015, the average number of years of data for each firm is less than 3 years.

[13] Sample means for profit likelihood and returns to capital/equity are 52 and 9.3 percent, respectively.

Table 1: Financial and economic performance, pooled OLS

VARIABLES	(1) Financial Profitability	(2) Economic Profitability	(3) Financial Returns	(4) Economic Returns
ESG rating	0.122*** (0.043)	0.142*** (0.043)	0.026** (0.011)	0.043*** (0.014)
Cost			0.142 (0.140)	0.408* (0.221)
Rule of law	0.136 (0.164)	0.049 (0.189)	-0.021 (0.048)	-0.036 (0.071)
GDP per capita (Ln)	-0.008 (0.019)	-0.008 (0.020)	-0.002 (0.004)	-0.003 (0.008)
Δ Reserves	-0.626 (0.494)	-0.634 (0.484)	-0.215 (0.165)	-0.137 (0.229)
Real sector	-0.146*** (0.031)	-0.144*** (0.032)	-0.029*** (0.009)	-0.038*** (0.014)
Fragile state	-0.057 (0.045)	-0.037 (0.047)	-0.015 (0.011)	-0.007 (0.015)
Total commitment (Ln)	0.055*** (0.011)	0.054*** (0.010)	0.010*** (0.002)	0.011*** (0.003)
Trend	-0.008 (0.006)	-0.006 (0.005)	-0.005*** (0.002)	-0.007*** (0.002)
N	2,969	2,951	2,886	2,873
k	102	102	102	102
R^2	0.049	0.048	0.035	0.040
(*p* > *F*)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Estimation is by ordinary least squares (OLS), pooled across investments, with standard errors robust to *k*-country clusters in parentheses. Constants are estimated but not reported. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

valid instruments, i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation.[14] In terms of other significant covariates, larger projects and projects in the real sector have higher ESG scores than projects with smaller commitments or financial-sector projects.

As an additional check of instrument validity, we specify a "false experiment" in which UNSC membership and the EPI in the following year—conditional on our vector of exogenous variables—are included as explanatory variables in estimating the firm-level ESG score. Since "future" UNSC membership or EPI should be orthogonal to today's ESG score, this constitutes a partial test of the identification assumption that any firm-level outcome would not have been observed in the absence of contemporaneous or precedent instrumental variable effects. Table 3 presents the

[14] While *J* will be identically zero for any exactly identified equation, it will be positive for an over-identified equation. If it is "too large," doubt is cast on the satisfaction of the exclusion conditions underlying 2SLS, and thus rejection of the null ($J = 0$) would call into question the validity of the instruments.

Table 2: Sustainability and profitability, first-stage IV results

VARIABLES	(1)	(2)	(3)	(4)
EPI Score	0.245***	0.242***	0.248***	0.237***
	(0.079)	(0.079)	(0.084)	(0.083)
UNSC membership	0.027**	0.030**	0.033***	0.036***
	(0.012)	(0.013)	(0.013)	(0.013)
Cost			0.390*	0.391*
			(0.220)	(0.205)
Rule of law	0.040	0.014	0.035	0.022
	(0.055)	(0.057)	(0.057)	(0.058)
GDP per capita (Ln)	-0.014	-0.013	-0.014	-0.012
	(0.012)	(0.011)	(0.011)	(0.011)
Δ Reserves	0.063	0.085	0.248	0.313*
	(0.182)	(0.182)	(0.170)	(0.171)
Real sector	0.030**	0.031**	0.041***	0.041***
	(0.014)	(0.014)	(0.015)	(0.014)
Fragile state	-0.007	-0.009	-0.013	-0.014
	(0.017)	(0.017)	(0.018)	(0.018)
Total commitment (Ln)	0.010**	0.011**	0.011**	0.011**
	(0.005)	(0.005)	(0.005)	(0.005)
Trend	-0.003	-0.001	-0.003	-0.002
	(0.003)	(0.002)	(0.003)	(0.002)
Kleibergen-Paap LM	6.024	6.153	6.596	6.532
($p > \chi^2$)	(0.049)	(0.046)	(0.037)	(0.038)
Hansen's *J*	0.660	0.155	0.054	0.082
($p > \chi^2$)	(0.417)	(0.694)	(0.816)	(0.775)
N	2,426	2,407	2,344	2,331
k	98	98	98	98
R^2	0.018	0.019	0.024	0.024
($p > F$)	(0.001)	(0.000)	(0.000)	(0.000)

Notes: Estimation by two-stage least squares (2SLS) with standard errors robust to *k*-country clusters in parentheses. First stage dependent variable is ESG rating, with EPI score and UNSC membership as instrumental variables excluded in the second stage. The null hypothesis of the Kleibergen-Paap Lagrange-multiplier test is that the structural equation is underidentified (i.e., the rank condition fails). For Hansen's *J* statistic, the joint null hypothesis is that the instruments are uncorrelated with the error term, and that the excluded instruments are correctly excluded from the second stage. Constants are estimated but not reported. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

coefficients from this experiment, where we compare the t, $t - 1$, and $t + 1$ values of instrumental variables on ESG scores in simple regressions. Lagged and contemporaneous values are positive and significant correlates of ESG; subsequent values are not, affirming the orthogonality of these country-level instruments to firm-level sustainability ratings.

The first-stage is presented graphically in the partial residual plots in Figure 3. We regress ESG on our exogenous

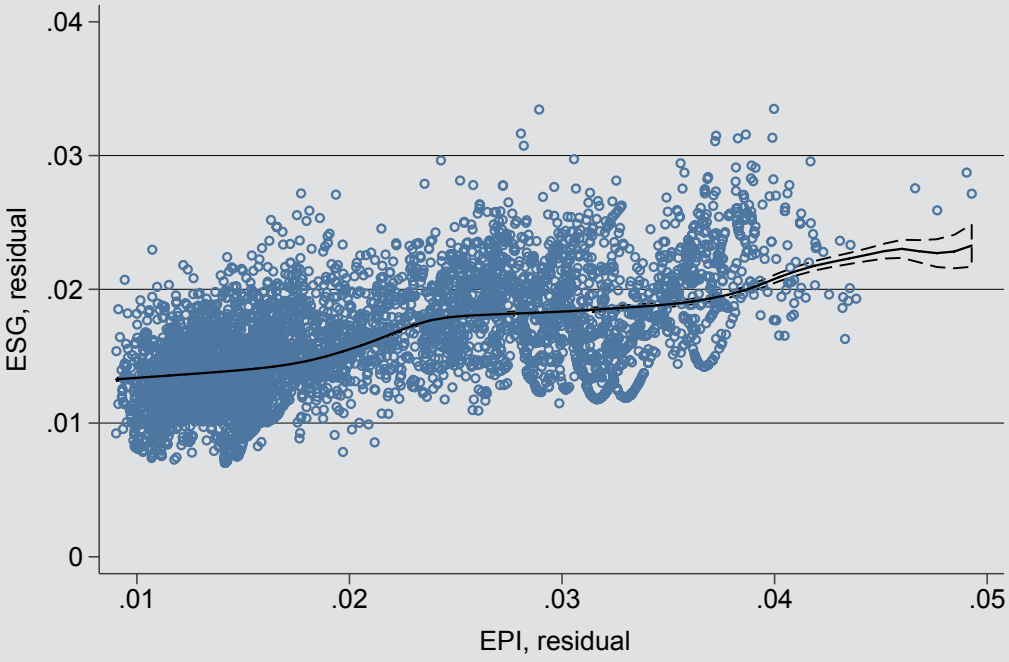
Table 3: Sustainability ratings: false-experiment estimates

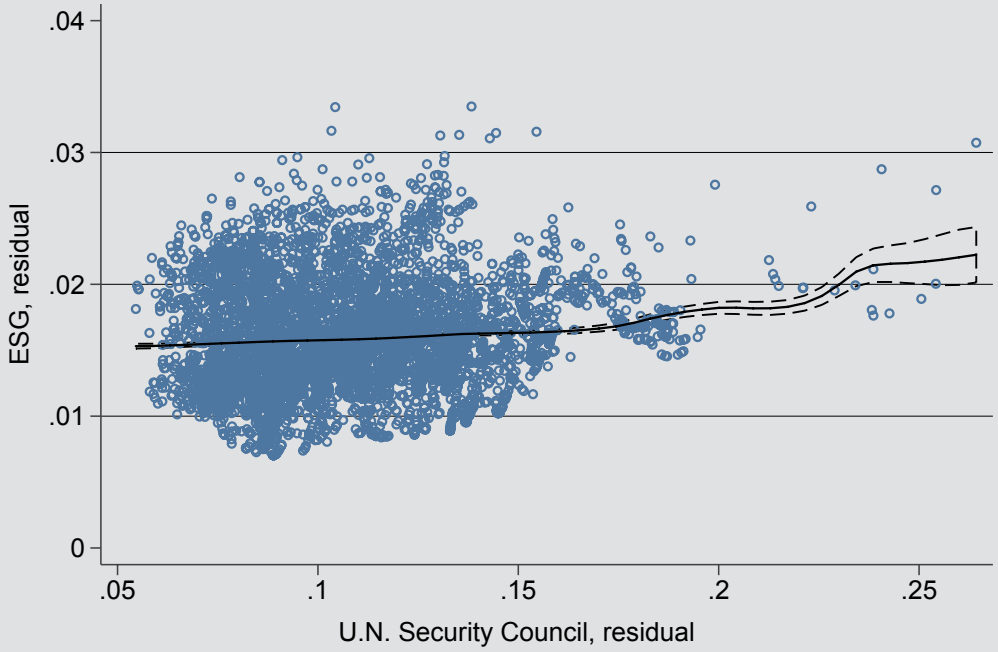
	(1)	(2)	(3)
EPI score$_t$	0.242*** (0.079)		
EPI score$_{t-1}$		0.191** (0.085)	
EPI score$_{t+1}$			0.149 (0.093)
UNSC membership$_t$	0.030** (0.013)		
UNSC membership$_{t-1}$		0.026** (0.013)	
UNSC membership$_{t+1}$			-0.010 (0.014)
Rule of law	0.014 (0.057)	0.050 (0.058)	-0.044 (0.065)
GDP per capita (Ln)	-0.013 (0.011)	-0.015 (0.013)	-0.002 (0.014)
Δ Reserves	0.085 (0.182)	-0.072 (0.221)	-0.004 (0.253)
Real sector	0.031** (0.014)	0.037** (0.016)	0.031* (0.017)
Fragile state	-0.009 (0.017)	-0.033* (0.020)	-0.021 (0.022)
Total commitment (Ln)	0.011** (0.005)	0.010* (0.006)	0.008 (0.006)
Trend	-0.001 (0.002)	-0.002 (0.003)	-0.001 (0.005)
N	2,407	1,924	1,362
k	98	96	90
R^2	0.019	0.023	0.015
(*p* > *F*)	(0.000)	(0.002)	(0.003)

Notes: Dependent variable is ESG rating. Estimation is by ordinary least squares (OLS), pooled across investments, with standard errors robust to *k*-country clusters in parentheses. Constants are estimated but not reported. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

variables and the time dummies, excluding the instrumental variables. We then regress our instrumental variables, in separate estimations, on the identical set of exogenous variables and examine the relationship between the two sets of residuals using non-parametric, locally weighted polynomial regressions. The resulting partial-residual plots show the significantly positive, approximately linear, relationship between the country-level instruments and the firm-level ESG scores (higher-order polynomial instrument terms are not significantly related to ESG).

Figure 3. Country-level instruments and firm-level environmental sustainability





Notes: Trendlines are smoothed values estimated with kernel-weighted local polynomial regressions with Epanechnikov (parabolic) kernels. Confidence intervals are estimated using Silverman bandwiths.

Source: Author's calculations

Main Results

We examine the second-stage to re-estimate the impact of sustainability on financial performance, correcting now for problems of measurement and endogeneity. We rely on the IV-2SLS estimation even though two of our four dependent variables (profitability and economic profitability) are binary dependent variables in the second stage. Since the use of a non-linear second stage requires assumptions regarding the precise functional form in order to be accurate, the linear IV-2SLS estimator is preferred (see, e.g., Angrist and Krueger, 2001; Wooldridge, 2002). Nevertheless, we also present estimates from a non-linear, two-stage procedure in which standard errors are corrected for second-stage binary variables in the Appendix, and note that the first- and second-stage results are similar to our IV-2SLS results.

Table 4 presents second-stage results. The significant, positive relationship between ESG ratings and firm profitability, seen in Table 1, disappears. In no estimation

Table 4: Sustainability and profitability, second-stage IV results

	(1) Financial Profitability	(2) Economic Profitability	(3) Financial Returns	(4) Economic Returns
ESG rating	-0.476 (0.860)	-0.850 (0.947)	0.170 (0.256)	-0.115 (0.342)
Cost			0.128 (0.181)	0.523* (0.296)
Rule of law	0.243 (0.176)	0.157 (0.197)	-0.019 (0.053)	-0.021 (0.076)
GDP per capita (Ln)	-0.004 (0.021)	0.001 (0.026)	-0.004 (0.006)	0.000 (0.010)
Δ Reserves	-0.548 (0.450)	-0.566 (0.421)	-0.243 (0.169)	-0.081 (0.222)
Real sector	-0.119*** (0.038)	-0.105** (0.044)	-0.029* (0.015)	-0.025 (0.023)
Fragile state	-0.075 (0.050)	-0.051 (0.056)	-0.019* (0.011)	-0.013 (0.018)
Total commitment (Ln)	0.059*** (0.013)	0.063*** (0.016)	0.008** (0.004)	0.012** (0.005)
Trend	-0.011 (0.007)	-0.009 (0.006)	-0.005*** (0.002)	-0.008*** (0.002)
N	2,426	2,407	2,344	2,331
k	98	98	98	98
RMSE	0.505	0.538	0.136	0.171
R^2	-0.022	-0.159	-0.013	-0.006
(*p* > *F*)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Estimation by two-stage least squares (2SLS) with standard errors robust to *k*-country clusters in parentheses. Second stage dependent variables are financial and economic profitability and returns. R^2 is centered. Constants are estimated but not reported. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Financial and economic performance, reduced-form estimates

	(1) Financial Profitability	(2) Economic Profitability	(3) Financial Returns	(4) Economic Returns
EPI score	-0.160 (0.224)	-0.219 (0.253)	0.052 (0.076)	-0.030 (0.100)
UNSC membership	0.000 (0.024)	-0.019 (0.027)	0.001 (0.010)	-0.003 (0.013)
Cost			0.044 (0.187)	0.353 (0.228)
Rule of law	0.200 (0.160)	0.105 (0.191)	-0.035 (0.050)	-0.054 (0.075)
GDP per capita (Ln)	0.004 (0.030)	0.013 (0.036)	-0.007 (0.009)	0.003 (0.014)
Δ Reserves	-0.512 (0.492)	-0.563 (0.447)	-0.177 (0.167)	-0.025 (0.239)
Real sector	-0.140*** (0.030)	-0.133*** (0.031)	-0.025** (0.010)	-0.032** (0.015)
Fragile state	-0.076 (0.048)	-0.047 (0.053)	-0.028** (0.014)	-0.015 (0.018)
Total commitment (Ln)	0.059*** (0.010)	0.059*** (0.009)	0.010*** (0.003)	0.010*** (0.004)
Trend	-0.009 (0.005)	-0.008 (0.005)	-0.006*** (0.002)	-0.008*** (0.002)
N	2,577	2,557	2,490	2,476
k	100	100	100	100
R^2	0.051	0.048	0.030	0.030
(*p* > *F*)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Dependent variable is financial and economic profitability and returns. Estimation is by ordinary least squares (OLS), pooled across investments, with standard errors robust to *k*-country clusters in parentheses. Constants are estimated but not reported. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

of the four of the measures of firm financial performance does sustainability have an effect—either on profitability or adjusted returns—in which we can have confidence. Although the joint significance tests reject the null of zero slopes, negative (centered) R-squared values (which are calculated from the sum of the squared residuals corresponding to the instruments rather than the original regressors, and can thus be larger than the total sum of squares) also indicate poor fit in the second-stage. Reduced-form estimates are in Table 5. The weakness of the second-stage relationships, indeed, removes any effects of the instrumental variables in the reduced-form specifications.[15]

[15] The weakness of the second-stage relationships, indeed, removes any effects of the instrumental variables in the reduced-form estimations.

Other explanatory variables behave as expected. Larger projects are more profitable/show higher returns. Real sector projects are less profitable, reducing the likelihood of profitability by between 11 and 12 percent, and reducing returns by 3 percent of capital/equity. Investments in financial sector projects tend to be more profitable for several reasons. Financial sector investments are more often in clients that are regulated by central banks according to increasingly global standards for capital adequacy and risk management. Regulatory requirements, minimum capital requirements, and high technology costs may also serve as barriers to entry. Even when outside the purview of central bank regulation, the portfolios of investee companies are often diversified across sectors and industries and less subject to economic volatility that may impact one industry or sector. Fragile state projects reduce financial returns, in our second stage results by 2 percent, but this effect is not consistent across outcomes. Finally, the time trend shows declining performance across our measures.

In Tables 6 and 7 we divide the sample between real- and financial-sector observations. In Table 6 we estimate financial and economic profitability as binary outcomes, whereas in Table 7 we use the continuous measures as dependent variables. Both tables show only the coefficients for the ESG scores in the second-stage equations, as well as for the instruments in the first stage. Here our instrumental variables are generally more robust in the real-sector regressions than in the financial sector regressions. But in no estimation is the relationship between sustainability and financial performance significantly different than zero. There are several possible explanations for this. Note that the IFC's portfolio has become increasingly concentrated in the financial sector; by the end of our sample period, over half of the IFC's portfolio was in financial intermediaries where the ESG rating is less obvious because it may not capture the end-user behavior. For example, the World Bank's Internal Evaluation Group concluded:

> *[M]easuring development results of financial projects at the sub-borrower level (or the level of end beneficiaries) is inherently difficult, and IFC has gaps in information. IFC has no direct relationship with, access to, or often even knowledge of the companies or microenterprises that are borrowing from the financial institutions. In practice, DOTS tracking for indicators such as number of small and medium-size enterprise (SME) borrowers is based on "proxy" figures from the financial institutions' portfolio: number of loans below a maximum, the total portfolio of the targeted business segment (such as housing, energy efficiency), and the credit quality of that portfolio (such as number of nonperforming loans) (World Bank, 2013).*

Moreover, these indicators do not necessarily assess the intermediary's record of extending credit to the most productive companies or their impact.[16] Consequently, one might expect larger, random measurement error with respect to ESG ratings in the financial sector, which would bias the effect of the instrumental variables toward zero.

[16] The World Bank's internal evaluation further notes that, although nonperforming loans of specific segments of financial intermediary clients are used as a proxy for the business performance of the sub-borrowers, the consideration of collaterals and other risk mitigations can hide the poor business performance of the financial institution clients. Often, IFC sets targets of extending loans to groups of previously unbanked micro-entrepreneurs, yet the extent to which borrowers had been previously unbanked cannot be confirmed because of a lack of data. Given that the client does not collect and report on whether new sub-borrowers had previously borrowed from the formal financial sector, it is not possible to assess reliably whether borrowers had been previously "unbanked." Assumption of new clients as fresh entrants to formal financial institutions needs to be questioned.

Table 6: Profitability by sector, IV results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ESG	Profit ROIC	ESG	Profit EROIC	ESG	Profit ROE	ESG	Profit EROE
ESG ratings		-0.402 (0.498)		-0.903 (0.563)		1.457 (0.936)		1.559* (0.856)
EPI score	0.448*** (0.099)		0.418*** (0.100)		-0.027 (0.099)		-0.011 (0.099)	
UNSC membership	0.033** (0.014)		0.035** (0.014)		0.044*** (0.017)		0.049*** (0.017)	
Rule of law	0.082 (0.066)	0.092 (0.166)	0.048 (0.066)	0.103 (0.171)	-0.034 (0.082)	0.544** (0.230)	-0.017 (0.082)	0.495** (0.220)
GDP per capita (Ln)	-0.037*** (0.013)	0.015 (0.018)	-0.033** (0.013)	0.029 (0.019)	0.013 (0.014)	-0.055* (0.030)	0.007 (0.014)	-0.061** (0.026)
Δ Reserves	0.434 (0.356)	-0.857 (0.815)	0.539 (0.360)	-0.241 (0.906)	-0.056 (0.329)	1.114 (0.979)	-0.062 (0.324)	0.256 (0.964)
Real sector	0.090 (0.179)	0.139 (0.314)	-0.127*** (0.016)	-0.606*** (0.078)	-0.209 (0.161)	-0.282 (0.300)	0.117*** (0.014)	-0.813*** (0.091)
Fragile state	0.000 (0.020)	-0.069 (0.046)	-0.002 (0.020)	-0.007 (0.049)	-0.041* (0.022)	-0.043 (0.071)	-0.044* (0.022)	-0.043 (0.070)
Total commitment (Ln)	0.027*** (0.006)	0.047*** (0.018)	0.027*** (0.006)	0.067*** (0.020)	-0.006 (0.005)	0.075*** (0.013)	-0.006 (0.005)	0.072*** (0.013)
Trend	-0.007** (0.003)	-0.025*** (0.008)	-0.006* (0.003)	-0.030*** (0.008)	0.002 (0.003)	-0.017* (0.009)	0.002 (0.003)	-0.012 (0.009)
Kleibergen-Paap LM ($p > \chi^2$)		20.080 (0.000)		17.760 (0.000)		8.529 (0.014)		9.699 (0.008)
Hansen's *J* ($p > \chi^2$)		0.134 (0.715)		1.742 (0.187)		1.825 (0.177)		1.817 (0.178)
N	1,405	1,405	1,386	1,386	936	936	933	933
R^2	0.038	-0.040	0.037	-0.225	0.028	-0.257	0.024	-0.345
($p > F$)	(0.000)	(0.000)	-	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)

Notes: Estimation by two-stage least squares (2SLS) with robust standard errors in parentheses. First stage dependent variable is ESG rating, with EPI score and UNSC membership as instrumental variables excluded in the second stage. Second stage dependent variables are financial and economic profitability (binary). The null hypothesis of the Kleibergen-Paap Lagrange-multiplier test is that the structural equation is underidentified (i.e., the rank condition fails). For Hansen's *J* statistic, the joint null hypothesis is that the instruments are uncorrelated with the error term, and that the excluded instruments are correctly excluded from the second stage. Constants are estimated but not reported. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Returns by sector, IV results

	(1) ESG	(2) ROIC	(3) ESG	(4) EROIC	(5) ESG	(6) ROE	(7) ESG	(8) EROE
ESG rating		0.153		-0.223		0.145		0.559
		(0.108)		(0.158)		(0.315)		(0.432)
EPI score	0.457***		0.428***		-0.032		-0.018	
	(0.100)		(0.101)		(0.098)		(0.098)	
UNSC membership	0.034**		0.037***		0.048***		0.051***	
	(0.014)		(0.014)		(0.017)		(0.017)	
Cost	0.375	0.483***	0.328	0.796***	0.495*	-0.537*	0.507**	-0.223
	(0.245)	(0.121)	(0.246)	(0.174)	(0.264)	(0.279)	(0.236)	(0.376)
Rule of law	0.091	-0.033	0.062	-0.014	-0.016	-0.003	0.009	0.038
	(0.066)	(0.040)	(0.066)	(0.051)	(0.083)	(0.070)	(0.084)	(0.093)
GDP per capita (Ln)	-0.037***	0.002	-0.034***	0.014**	0.017	-0.016*	0.010	-0.033***
	(0.013)	(0.004)	(0.013)	(0.005)	(0.014)	(0.010)	(0.014)	(0.013)
Δ Reserves	0.411	-0.315*	0.532	0.015	-0.043	-0.115	-0.055	-0.040
	(0.355)	(0.164)	(0.358)	(0.247)	(0.329)	(0.335)	(0.323)	(0.448)
Real sector	0.086	-0.018	-0.120***	-0.091***	-0.108	-0.089**	0.102***	-0.162***
	(0.168)	(0.033)	(0.017)	(0.030)	(0.120)	(0.038)	(0.017)	(0.047)
Fragile state	0.001	-0.020**	-0.001	0.006	-0.045**	-0.012	-0.048**	-0.005
	(0.020)	(0.010)	(0.020)	(0.014)	(0.022)	(0.020)	(0.022)	(0.029)
Total commitment (Ln)	0.028***	0.003	0.028***	0.015***	-0.005	0.013***	-0.005	0.014***
	(0.006)	(0.004)	(0.006)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
Trend	-0.007**	-0.005***	-0.005	-0.009***	0.001	-0.004	0.001	-0.006
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
Kleibergen-Paap LM ($p > \chi^2$)		20.681		18.743		9.968		10.650
		(0.000)		(0.000)		(0.007)		(0.005)
Hansen's *J* ($p > \chi^2$)		1.187		1.504		0.155		0.042
		(0.276)		(0.220)		(0.694)		(0.838)
N	1,410	1,410	1,398	1,398	939	939	936	936
R^2	0.040	-0.002	0.039	-0.146	0.030	0.008	0.029	-0.197
($p > F$)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)

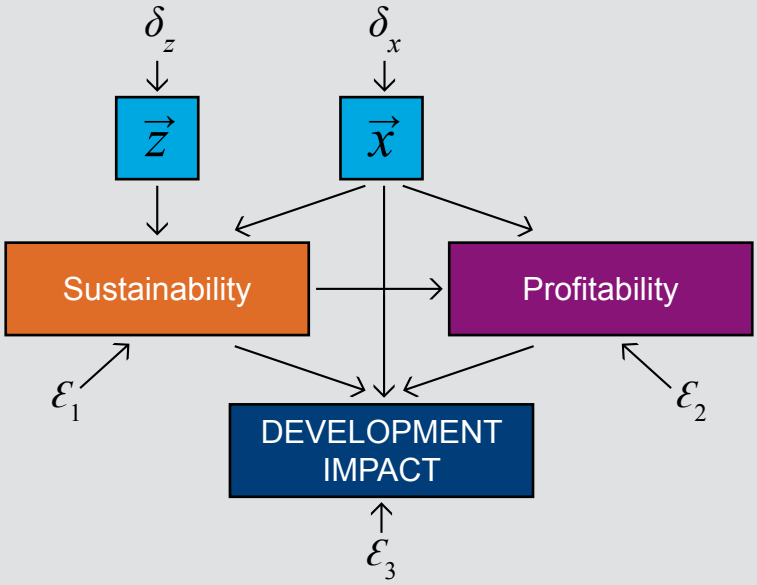
Notes: Estimation by two-stage least squares (2SLS) with robust standard errors in parentheses. First stage dependent variable is ESG rating, with EPI score and UNSC membership as instrumental variables excluded in the second stage. Second stage dependent variables are financial and economic (continuous) returns (ROIC and economic ROIC are real sector observations, ROE and economic ROE are financial sector). The null hypothesis of the Kleibergen-Paap Lagrange-multiplier test is that the structural equation is underidentified (i.e., the rank condition fails). For Hansen's *J* statistic, the joint null hypothesis is that the instruments are uncorrelated with the error term, and that the excluded instruments are correctly excluded from the second stage. Constants are estimated but not reported.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Development Impact

A final question concerns the broader, longer-term developmental impact of a firm's financial and economic activities. In addition to rating the project's environmental and social outcomes, the IFC also assesses its broader, private-sector development impact by looking at wider outcomes beyond the firm such as demonstration effects (where the project sends a signal to the market, other companies follow the example and offer a similar product, or make upgrades) or supplier upgrades (where a project that helps its suppliers upgrade the quality of their goods and services benefits other local companies that can now source local goods and services of higher quality), as well as effects such as improved competition, innovation, or sectoral transformations that spread the benefits of the growth of productive, private enterprises.

We examined the effects of firm performance on private-sector development outcomes through simple path analysis, in order to provide estimates of the magnitude and significance of hypothesized causal connections between sets of variables. The path diagram in Figure 4 depicts the various expected relationships. Our vector of exogenous variables (***X***) affects both firm sustainability and profitability, while our excluded country-level instruments (***Z***) affect profitability only through the rating of a firm's sustainability activities, δ represents random disturbances associated with exogenous variables. This is the IV set-up examined above. In addition, we now allow sustainability and profitability to be considered exogenous variables with direct effects on private-sector developmental impact. We model this as a simple path model—a special case of a structural equation model containing only observed variables. The advantage of the path model is that it takes into account the presence of other equations by recognizing that there will be a (contemporaneous) covariance structure between the error terms. In our model, by allowing the error terms ε_1 and ε_2 to co-vary,

Figure 4. Path diagram of firm development impact



we can generalize the 2SLS method to take account of simultaneous correlation between the first and second stage equations (in the same way, for example, that seemingly unrelated regression generalizes OLS). Additionally, given that the same exogenous factors may jointly affect the data-generating processes for "sustainability" and "development impact," we also allow ε_1 and ε_3 to co-vary.

Results from the path model are in Table 8 and Table 9. In Table 8, we use the binary "profitability" measure, whereas in Table 9 we rely on the continuous measure of returns to capital or equity. In both tables, our relationship identifying ESG remains valid, with UNSC membership and the EPI entering the system with significantly positive coefficients. As with 2SLS, the simultaneous setup also confirms the absence of any correlation between sustainability and profitability that is significantly different than zero. However, this path analysis does indicate that sustainable behaviors and profits have a positive influence on longer-term development impact on the local private sector. In columns (4) – (6) we eliminate all country-level exogenous variables with the exception of our instrumental variables. The previous results are robust to the exclusion of these variables form the path model. The instrumental variable effects of the 2SLS estimations remain intact, but once again, the effect of sustainability on performance in the second stage disappears. Both factors continue to exert a significantly positive effect on broader private sector development. For all path analyses, diagnostic test statistics reveal properly-fitted models. We cannot reject the null that our specified models fit more poorly than "saturated" models—a model that perfectly reproduces variances, covariance, and means of the observed variables—indicating that our path models fit well. Additionally, root mean squared errors and an incremental fit index—the Tucker-Lewis index comparing the fit of the target model to a null/independence model specifying that all measured variables are uncorrelated—are below and above benchmarks for goodness of fit, respectively.

Table 8: Sustainability, profitability, and development impact: path model results

	(1) ESG	(2) Profitability	(3) Development Impact	(4) ESG	(5) Profitability	(6) Development Impact
Profitability			0.175*** (0.064)			0.118*** (0.040)
ESG rating		-0.309 (0.601)	-0.354 (0.294)		0.289 (0.476)	-0.238 (0.233)
EPI score	0.249*** (0.070)			0.180*** (0.043)		
UNSC membership	0.028*** (0.010)			0.025*** (0.009)		
Rule of law	0.037 (0.053)	0.252** (0.123)	-0.050 (0.063)			
GDP per capita (Ln)	-0.013 (0.009)	-0.010 (0.015)	0.002 (0.008)			
Δ Reserves	0.013 (0.256)	-0.584 (0.583)	0.495* (0.294)			
Real sector	0.034*** (0.009)	-0.127*** (0.029)	0.067*** (0.018)	0.031*** (0.009)	-0.139*** (0.025)	0.050*** (0.013)
Fragile state	-0.004 (0.015)	-0.076** (0.033)	-0.010 (0.017)			
Total commitment (Ln)	0.011*** (0.004)	0.060*** (0.010)	0.025*** (0.006)	0.008** (0.004)	0.052*** (0.010)	0.027*** (0.005)
Trend	-0.003 (0.002)	-0.010* (0.005)	-0.019*** (0.003)	-0.003 (0.002)	-0.008 (0.005)	-0.020*** (0.002)
N	2,343			2,415		
Likelihood ratio (model vs. saturated)	0.225			0.213		
($p > \chi^2$)	(0.894)			(0.899)		
RMSEA	0.000			0.000		
Tucker-Lewis index	1.051			1.031		

Notes: Estimates are maximum likelihood from path analyses of figure 4, with robust standard errors in parentheses. Likelihood ratio p-values compare fit of specified model to saturated model (perfectly reproducing variances, covariance, and means of observed variables). Root Mean Squared Error of Approximation (RMSEA) indicates degree to which the path model, with unknown but optimally chosen parameter estimates (values above 0.10 indicate poor fit). The Tucker Lewis index is a relative fit index that compares a chi-square for the model tested against one from a baseline/null model that specifies that all measured variables are uncorrelated (TLI > 0.95 indicates good fit). *** p<0.01, ** p<0.05, * p<0.10.

Table 9: Sustainability, investment returns, and development impact: path model results

	(1) ESG	(2) Returns	(3) Development Impact	(4) ESG	(5) Returns	(6) Development Impact
Returns			0.327 (0.246)			0.461*** (0.176)
Cost		0.188* (0.104)	-0.351* (0.180)		0.193* (0.102)	-0.377** (0.171)
ESG rating		0.167 (0.161)	-0.412 (0.283)		0.076 (0.132)	-0.353 (0.239)
EPI score	0.261*** (0.068)			0.183*** (0.042)		
UNSC membership	0.028** (0.011)			0.024** (0.009)		
Rule of law	0.013 (0.053)	-0.010 (0.034)	-0.026 (0.062)			
GDP per capita (Ln)	-0.015* (0.009)	-0.005 (0.004)	-0.004 (0.008)			
Δ Reserves	0.179 (0.262)	-0.212 (0.165)	0.706** (0.308)			
Real sector	0.040*** (0.009)	-0.029*** (0.009)	0.058*** (0.015)	0.037** (0.009)	-0.026*** (0.007)	0.054*** (0.014)
Fragile state	-0.009 (0.014)	-0.021** (0.009)	-0.027 (0.018)			
Total commitment (Ln)	0.011*** (0.004)	0.008*** (0.003)	0.031*** (0.006)	0.008** (0.004)	0.009*** (0.003)	0.028*** (0.005)
Trend	-0.003 (0.002)	-0.005*** (0.002)	-0.021*** (0.003)	-0.003 (0.002)	-0.006*** (0.001)	-0.020*** (0.003)
N	2,264			2,334		
Likelihood ratio (model vs. saturated)	3.542			2.408		
($p > \chi^2$)	(0.315)			(0.492)		
RMSEA	0.009			0.000		
Tucker-Lewis index	0.988			1.008		

Notes: Estimates are maximum likelihood from path analyses of figure 4, with robust standard errors in parentheses. Likelihood ratio p-values compare fit of specified model to saturated model (perfectly reproducing variances, covariance, and means of observed variables). Root Mean Squared Error of Approximation (RMSEA) indicates degree to which the path model, with unknown but optimally chosen parameter estimates (values above 0.10 indicate poor fit). The Tucker Lewis index is a relative fit index that compares a chi-square for the model tested against one from a baseline/null model that specifies that all measured variables are uncorrelated (TLI > 0.95 indicates good fit). *** p<0.01, ** p<0.05, * p<0.1.

CONCLUSION

The evidence from firm-level interventions by IFC do indicate the presence of spillovers into better development of the market in which the firms operate, thus supporting the core theory of change underlying IFC's mandate, as well as that of other multilateral development institutions that invest in productive private enterprise.

The analysis also supports a policy rationale to induce firms to improve ESG performance because of the externality on broader development outcomes. However, it may be unlikely that firms in developing countries undertake ESG activities as a commercial decision, because the channels from ESG to profitability do not appear strong in the short run. That said, the key obstacle to firms to improve ESG activities are administrative costs, and in some countries the quality and relevance of environmental and social metrics, while the presence of externalities affecting longer term, private sector development suggests that there are broader benefits.

How companies address ESG issues, and whether they have risk management approaches in place, matters for development. Firms can be encouraged to address ESG issues through standardization in measurement and greater transparency in reporting. Lowering the administrative cost of reporting, and increasing the benefits associated with such reporting may therefore help increase take-up. In the latter respect, the lack of standardized ESG measurements and performance frameworks hinders strategic selection of companies, cross-comparison between different investment alternatives and long-term value definition. Increasing the standardization of such reporting may lower the administrative costs, more quickly build capacity of reporting firms, and increase their use by external stakeholders. There is a strong case for IFC and other multilateral development finance institutions with private sector operations to continue with, and perhaps refine and standardize, ESG and private sector developmental ratings frameworks, perhaps building on standards disseminated by such entities as the Sustainability Accounting Standards Board and other international organizations.

The analysis suggests important differences between the impact of investments into financial intermediaries, which are more profitable but have relatively less development impact, and investments in real sector firms, which are on average less profitable but have broader development impacts.

Investments in fragile states appear less profitable for an external investor, especially given relatively higher costs of origination and structuring to identify and mitigate the higher risks, and the relatively smaller project sizes on which to support these higher origination costs. This implies that instruments to mitigate risks of such investments could be useful. The 18th replenishment of the World Bank Group's facility for concessional finance for lower income and fragile states, the International Development Association (IDA), includes for the first time a private sector window to help mitigate risks of private investment in IDA countries.

Finally, there is solid evidence that stronger commercial performance can also contribute to broader development outcomes. Additionally, higher profits lead to a healthier private sector environment. Thus, the essentially commercial nature of private investment is consistent with economic and social benefits to society.

REFERENCES

Accenture Sustainability Services, United Nations. (2014). *The UN Global Compact-Accenture CEO Study on Sustainability*. Available at <https://www.accenture.com/us-en/~/media/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Industries_8/Accenture-The-UN-Global-Compact-Acn-CEO-Study-on-Sustainability-Metals-and-Mining.PDF>.

Angrist, J. D., & Krueger, A. B. (2001). Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *The Journal of Economic Perspectives, 15*(4), 69-85.

Berle, A. A. (1954). *The 20th Century Capitalist Revolution*. New York: Hartcourt, Brace & World.

Bernstein, A. (2000). "Too Much Corporate Power?" *Business Week* 3698 (September 11): 144-158.

Business & Sustainable Development Commission. (2017). *Better Business Better World – The report of the Business & Sustainable Development Commission*. Available at http://report.business-commission.org/.

Cambridge Associates & Global Impact Investing Network. (2015). *Introducing the Impact Investing Benchmark*. Available at <http://40926u2gov-f9kuqen1ndit018su.wpengine.netdna-cdn.com/wp-content/uploads/2015/06/Introducing-the-Impact-Investing-Benchmark.pdf>.

Dreher, A., Klasen, S., Vreeland, J. R., & Werker, E. (2013). The Costs of Favoritism: Is Politically Driven Aid Less Effective? *Economic Development and Cultural Change, 62*(1), 157-191.

Dreher, A., Sturm, J. E., & Vreeland, J. R. (2006). Does membership on the UN Security Council influence IMF decisions? Evidence from panel data.

Edmans, A., & Vogel, D. J. (2016). Does Socially Responsible Investing Make Financial Sense? *Wall Street Journal*. Available at <https://www.wsj.com/articles/does-socially-responsible-investing-make-financial-sense-1456715888>.

Esty, D. & Cort, T. (2016). "Corporate Sustainability Metrics: What Investors Need and Don't Get," Alliance for Research on Corporate Sustainability, 8th Annual ARCS Research Conference, Leeds School of Business, Colorado available at <http://corporate-sustainability.org/wp-content/uploads/Corporate-Sustainability-Metrics.pdf>.

Friede, G., Busch, T. & Bassen, A. (2015). ESG and Financial Performance: Aggregated Evidence from More Than 2000 Empirical Studies. *Journal of Sustainable Finance & Investment* 5(4), 210-233.

Friedman, M. (1962). *Capitalism and Freedom*. Chicago, IL: University of Chicago Press.

Gaspar, R. (2013). *Sustainability and financial performance: The chicken-egg dilemma*. GreenBiz. Available at <https://www.greenbiz.com/blog/2013/03/06/ sustainability-financial-performance-chicken-egg>.

Global Impact Investing Network. (2016). "Latest Data on Impact Investing Shows Diverse and Dynamic Global Market, According to 2016 Survey." Available at <https://thegiin.org/assets/GIIN_Annual%20Survey_Press%20Release%20FINAL_for%20web.pdf>.

Gray, J., Ashburn, N., Douglas, H. V., Jeffers, J. S., Musto, D. K., & Geczy, C. (2015). *Great Expectations: Mission Preservation and Financial Performance in Impact Investing*. Available at <https://socialimpact.wharton.upenn.edu/wp-content/uploads/2016/09/Great-Expectations-Mission-Preservation-and-Financial-Performance-in-Impact-Investing.pdf>.

Hang, M., Geyer-Klingeberg, J., Rathgeber, A., & Stoeckl, S., (2016). Economic Development Matters: A Meta-Regression Analysis on the Relation between Environmental Management and Financial Performance. *Journal of Industrial Ecology, forthcoming*. Available at SSRN: https://ssrn.com/abstract=2879330.

Hatashima, H. (2013). *Biennial report on operations evaluation: assessing the monitoring and evaluation systems of IFC and MIGA*. Washington, DC: World Bank.

International Finance Corporation. (2011). *Development Outcome Tracking System*. Available at <http://www.ifc.org/wps/wcm/connect/5fd0a7004a57bb0bb346bf8969adcc27DOTS+Handout+2011.pdf?MOD=AJPERES>.

_____. (2012). *Performance Standards on Environmental and Social Sustainability*. Washington, DC: IFC.

Khan, M., Serafeim, G., & Yoon, A. (2016). "Corporate sustainability: First evidence on materiality." *The Accounting Review, 91*(6), 1697-1724.

Kuziemko, I., & Werker, E. (2006). "How Much is a Seat on the Security Council Worth? Foreign Aid and Bribery at the United Nations." *Journal of Political Economy, 114*(5), 905-930.

Murray, M. P. (2006). Avoiding Invalid Instruments and Coping with Weak Instruments. *Journal of Economic Perspectives, 20*(4), 111-132.

Organization for Economic Cooperation and Development. (OECD). (2012). *Green Growth and Developing Countries: A Summary for Policy Makers*. Paris: OECD.

Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. *Organization Studies, 24*(3), 403-441.

Ott, I., & Soretz, S. (2015). Green attitude and economic growth. *Environmental and Resource Economics*, 1-23.

Page, A., & Katz, R. A. (2010). Is Social Enterprise the New Corporate Social Responsibility? *Seattle University Law Review* 34(1351).

Peloza, J., & Yachnin, R. (2008). Valuing business sustainability: A systematic review. *Research Network for Business Sustainability Knowledge Project Series*. London, Ontario: Research Network for Business Sustainability.

Pfau, M., Haigh, M. M., Sims, J., & Wigley, S. (2008). The influence of corporate social responsibility campaigns on public opinion. *Corporate Reputation Review, 11*(2), 145-154.

Søreide, T. (2006). *Is it wrong to rank? A critical assessment of corruption indices*. Chr. Michelsen Institute.

Unit, E. I. (2008). Doing good: Business and the sustainability challenge. *The Economist*, (January).

United Nations. (2015). *Addis Ababa Action Agenda of the Third International Conference on Financing for Development (Addis Ababa Action Agenda)*. Available at <http://www.un.org/esa/ffd/wp-content/uploads/2015/08/AAAA _Outcome.pdf>.

_____. *Principles for Responsible Investment. The Six Principles*. Available at <https://www.unpri.org/about/the-six-principles>.

Vreeland, J. R., & Dreher, A. (2014). *The political economy of the United Nations Security Council: money and influence*. Cambridge, UK: Cambridge University Press.

Weiner, Joseph L. (1964). The Berle-Dodd dialogue on the concept of the corporation. *Columbia Law Review, 64*(8), 1458-1467.

Wilson, G. E. R. (2013). Enhancing Financial Returns by Targeting Social Impact. In World Economic Forum Investors Industries, *From Ideas to Practice, Pilots to Strategy Practical Solutions and Actionable Insights on How to Do Impact Investing*. Geneva: World Economic Forum.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Boston, Massachusetts: MIT Press.

APPENDIX

Table A1: Variable Definitions

Variable	Measurement	Source
Environmental, Social, and Governance (ESG) rating	Degree of compliance with IFC's performance standards based on the following scoring: 1. Excellent = Project materially improved company's ESHS (Environmental, Social, Health, and Safety) performance and is in material compliance with IFC's ESHS requirements; 2. Satisfactory = Project is in material compliance with IFC's performance standard requirements, or is on track to meet the agreed environmental and social action plan; 3. Partly Unsatisfactory = Project is not in material compliance, but corrective action is underway or implementation of the action plan is proceeding with significant delays or shortcomings; 4. Unsatisfactory = Previous egregious non-compliance (e.g. leaving significant environmental or social harm even if performance shortfalls have since been addressed) or material non-compliance without realistic short-term prospects of corrective action. Rating is rescaled as (4 – *ESG*)/3	International Finance Corporation/Development Outcomes Tracking System
Environmental Performance Index (EPI) Score	The Environmental Performance Index (EPI) is constructed through the calculation and aggregation of more than 20 indicators reflecting national-level environmental data. These indicators are combined into nine issue categories, each of which fit under one of two overarching objectives. Source: www.epi.yale.edu.	Center for Environmental Law and Policy (Yale University) and the Center for International Earth Science Information (Columbia University)
UN Security Council membership	Dummy variable coded 1 if project is located in a country that is, in the project-year, a member of the United Nations Security Council, zero otherwise.	United Nations
Rule of law	Rule of law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.	World Bank, World Governance Indicators
GDP per capita (Ln)	GDP per capita is gross domestic product divided by midyear population in constant 2010 US dollars (in natural logs)	World Bank, World Development Indicators

Variable	Measurement	Source
Δ Reserves	Annual change in total reserves/months of imports.	World Bank, World Development Indicators
Total commitment (Ln)	Company IFC Disbursement Balance or original IFC Guarantee Commitment Balance in constant 2010 US dollars (in natural logs).	International Finance Corporation
Real Sector	Dummy variable coded 1 if project is a real sector project, zero otherwise.	International Finance Corporation
Fragile State	Dummy variable coded 1 if a project is, in that project-year, located in a country listed on the OECD harmonized list of fragile and conflict-affected states, zero otherwise.	Organization for Economic Cooperation and Development
Returns	Calculated as either return on invested capital (ROIC) or return on equity (ROE) for real and financial sectors, respectively: π = net operating profit, t = marginal tax rate, D = total debt, E = total equity	International Finance Corporation
Economic returns	Calculated as either economic return on invested capital (EROIC) or economic return on equity (EROE) for real and financial sectors, respectively: B_S = benefits to society, C_S = costs to society A = financial adjustments to net income, t = marginal tax rate, S = subsidies	International Finance Corporation
Cost	Calculated as either weighted average cost of capital (WACC) or cost of equity (COE) for real and financial sectors, respectively: D = total debt, r_D = cost of debt, t = marginal tax rate, E = total equity, r_E = cost of equity R_f = risk-free rate (5-year US Treasury rate), *Risk Premium* = IFC macro spread for the country + IFC project risk spread + 5%	International Finance Corporation
Profitability	Binary variable that takes on value of 1 if returns (ROIC or ROE) are greater than costs, zero otherwise:	International Finance Corporation
Economic profitability	Binary variable that takes on value of 1 if returns (EROIC or EROE) are greater than costs, zero otherwise:	International Finance Corporation

Variable	Measurement	Source
Development Impact	Degree to which project created conditions "conducive to the flow of private capital into productive investment." Scoring guidelines focus on, *inter alia*, changes in the legal and regulatory framework, demonstration effects (e.g., corporate governance), linkages, knowledge or know-how transfer, and increased competition or improvement in services, rated according to the following scale: 1. Excellent = Project substantially improved the enabling environment or contributed to the efficiency of markets; 2. Satisfactory = Project had some demonstrable positive impacts; 3. Partly Unsatisfactory = Project had some negative impacts that, however, are expected to be short-lived; 4. Unsatisfactory = Substantial negative effects on private sector development. Rating is rescaled as (4 – *Development Impact*)/3	International Finance Corporation/Development Outcomes Tracking System

Table A2: Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max	*N*	*n*	*T* (ave.)
ESG rating	0.562	0.217	0	1	2,969	1,022	2.905
Development impact	0.638	0.242	0	1	2,881	1,009	2.855
EPI score	0.466	0.105	0.178	0.820	2,426	976	2.486
UN Security Council membership	0.265	0.442	0	1	2,969	1,022	2.905
Rule of law	0.417	0.092	0	1	2,969	1,022	2.905
GDP per capita (Ln)	8.235	0.926	5.380	9.954	2,969	1,022	2.905
Δ Reserves	-0.001	0.016	-0.072	0.065	2,969	1,022	2.905
Total commitment (Ln)	16.982	1.276	13.816	20.977	2,969	1,022	2.905
Real sector	0.611	0.488	0	1	2,969	1,022	2.905
Fragile state	0.160	0.367	0	1	2,969	1,022	2.905
Profitability	0.520	0.450	0	1	2,969	1,022	2.905
Economic Profitability	0.603	0.489	0	1	2,919	1,005	2.905
Returns	0.093	0.132	-0.850	1.980	2,969	1,022	2.905
Economic Returns	0.126	0.167	-0.903	2.150	2,919	1,005	2.905
Costs	0.093	0.030	0.000	0.477	2,886	996	2.898
Real Sector:							
ROIC	0.080	0.109	-0.819	0.920	1,819	630	2.887
Economic ROIC	0.107	0.135	-0.869	1.011	1,782	616	2.893
WACC	0.086	0.030	0	0.477	1,738	604	2.878
Financial Sector:							
ROE	0.113	0.160	-0.850	1.980	1,156	394	2.934
Economic ROE	0.154	0.205	-0.903	2.150	1,141	390	2.923
COE	0.103	0.027	0	0.228	1,146	392	2.923

Table A3: Financial and economic performance, IV probit results

	(1) ESG	(2) Profitability	(3) ESG	(4) Economic Profitability
ESG rating		-1.268		-2.042
		(2.063)		(1.767)
EPI score	0.251***		0.249***	
	(0.077)		(0.075)	
UNSC membership	0.025*		0.028**	
	(0.013)		(0.013)	
Rule of law	0.039	0.608	0.013	0.368
	(0.055)	(0.414)	(0.057)	(0.466)
GDP per capita (Ln)	-0.014	-0.009	-0.014	0.001
	(0.011)	(0.055)	(0.011)	(0.061)
Δ Reserves	0.053	-1.350	0.077	-1.320
	(0.181)	(1.170)	(0.182)	(1.034)
Real sector	0.030**	-0.290**	0.031**	-0.243
	(0.014)	(0.133)	(0.014)	(0.151)
Fragile state	-0.007	-0.184	-0.010	-0.117
	(0.017)	(0.129)	(0.017)	(0.127)
Total commitment (Ln)	0.010**	0.146***	0.011**	0.147***
	(0.005)	(0.029)	(0.005)	(0.028)
Trend	-0.003	-0.026*	-0.001	-0.021
	(0.003)	(0.015)	(0.002)	(0.014)
Wald test		0.54		1.40
($p > \chi^2$)		(0.461)		(0.237)
k	98	98	98	98
N	2,426	2,426	2,407	2,407

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.

BROOKINGS
1775 Massachusetts Avenue, NW
Washington, DC 20036
202-797-6000
www.brookings.edu/global