Advancing Social Impact Investments through Measurement
Federal Reserve Board | Washington, DC
Steve Lydenberg | Panel #1

1

STEVE LYDENBERG

Okay, well, thank you, and I appreciate very much the opportunity to appear here on this panel on data collection and use. This is an extremely important topic here in the United States really and around the world, and getting more so. I come to it from a discipline where I've been looking at publicly disclosed data relating to large publicly-traded companies and how that data might be used to channel them toward the public interest, to improve their behavior in social and environmental ways. There's a lot of this data out there. It's a little different from the kind of impact investment data that is largely the focus of this, but I think there are some interesting lessons and ways of looking at that data that might be useful as we look at the impact question also, and as this whole field moves forward. I will be reminding you just first, briefly, of the importance of this data. It does have an impact. I will be making a distinction here that might be useful, between government-mandated data and voluntary data. I'll look a little bit at the challenges of both these kinds of data, possible solutions to those challenges and then open for discussion after the panel. First, I want to remind you of the variety of different kinds of data that are already out there. Again, this is for large, publicly-traded companies. But there is the toxic release inventory data on toxic chemicals; there's a home mortgage disclosure act data on lending by banks; there is the OSHA compliance data on safety records of US companies; there's the NLRB data; the National Labor Relation Boards data on union relations. There's a host of SEC

data on issues as varied as CEO compensation and the SEC will be requiring

corporations to explain what they're doing on diversity on boards of directors

coming up next year. There's the EEOC data on women and minority

employment. There is, then, the whole range of data that's being disclosed in

corporate CSR reports now, driven in part by the Global Reporting Initiative,

which is a worldwide standard for global reporting. There are the B-Lab and

Gears [?] rating systems which are also setting standards for reporting and

these are aimed a little bit more at the private equity space, small and medium-

sized enterprises. There's the CARS [?] data, the NCIF data, the CDFI data on

CDFIs that you'll be hearing more about today. My point here is that there is a

lot of data out here and it serves a lot of different purposes. It empowers

communities -- the TRI data and the Home Mortgage Disclosure Act data. It

informs investors; particularly in Europe investors are using these kind of data,

especially from the CSR reports, in their investments. It influences consumer

choice. We have the data for energy efficiency on appliances. We have the

nutrition data that's on the labels on foods, and, I believe it improves corporate

management. Believe it or not, corporations are not aware of what they are

doing in a lot of these areas until they are required to disclose this data. By

simply requiring it to disclose this data, you have made them more aware of

that and it is true that what gets measured gets managed. This data is also

useful in the setting of international standards for corporate conduct. The

other point I want to make about this is that this data is so varied and used for

Advancing Social Impact Investments through Measurement
Federal Reserve Board | Washington, DC
Steve Lydenberg | Panel #1

3

so many different purposes that it is in some way, when you look at it in the

aggregate, different from financial disclosure. Financial disclosure is used

primarily by the financial community and has a relatively discrete use, and this

data creates positive externalities, as Ben pointed out earlier today. So with

that by way of background, let me make a distinction among all these different

kind of datas. This is a slightly artificial distinction. It's one that comes from

historical accident, really, more than anything else, but it is real. And I think it

is useful in thinking about these, to make a distinction between data that it

has been required through legislation and data that is getting out there

through voluntary initiative. So the TRI, the Home Mortgage Disclosure Act

data, OSHA, NLRB, nutritional data, this has all come out essentially through

the regulatory and legislative processes, whereas the CSR reports, the Global

Reporting initiatives, Ecolabels essentially are voluntary initiatives. And this

distinction points out a little bit of the political nature of the legislative process,

and I think that's an important thing to keep in mind, on disclosure, that if

you're going to have disclosure that's broadly useful to wide varieties of people,

the political process produces that kind of data, whereas the voluntary process

tends to be aimed more at investors and consumers. Now let me characterize

these two different kinds of data, because they look and feel different. The

legislative data, the government-mandated data, is really aimed at empowering

citizens and empowering them locally. The EPA is quite specific about the toxic

release inventory. That information is there to empower local communities,

Advancing Social Impact Investments through Measurement
Federal Reserve Board | Washington, DC
Steve Lydenberg | Panel #1

4

local citizens. The Home Mortgage Disclosure Act was there to increase access

at a community level. And it is, therefore, detailed and narrowly targeted. The

TRI data is reported on a plant-by-plant level. You have to report it on a plant-

by-plant level if it's going to serve the local people in those communities.

Similarly with the Home Mortgage Disclosure Act Data, that has to be targeted

to different neighborhoods, different regions. So it has to be targeted narrowly.

It has a clear public purpose, and again this is tied to the fact that it comes

through the legislative and political process, and it is rich and comprehensive

and standardized. And again, it can be because it is mandated, and you can

ask for as much as you want and get it in exactly the form that you want. As a

consequence, these databases tend to be extremely rich. I'll get into some of

the problems that causes in a moment. And finally, it tends to be disclosed in

response to crises. The voluntary data is primarily targeted to those that are

trying to make an on-off decision. It's aggregated, company-level data.

Investors don't need facility-level data. They need company-level data.

Consumers want to know what the record of a company is; they're buying

multiple products, they want to know what the record of a company is, in

general. So that data tends to be aggregated, and it has a private interest.

Again, this distinction between public and private that I' m making is not

entirely clear, but it is aimed more at the market. And there is an implied

indirect public interest in having that information out there. It covers a broad

number of issues, because it's developed by consensus, and I'll get into that in

Advancing Social Impact Investments through Measurement
Federal Reserve Board | Washington, DC
Steve Lydenberg | Panel #1

5

a moment, it can cover all the pertinent issues for a corporation, and, increasingly it's expected of everybody. It's not required, but it's expected. So with that by way of a distinction as to the strengths of those, let me say a little bit about the challenges of each, because each has its own separate set of challenges. Because the legislation and regulation tends to be catalyzed by crisis, even though you can see areas where data will be necessary, it's hard to anticipate the crisis and get that information out there prior to the crisis. So just for example, it's fairly clear that one of the great difficult issues of the 21st Century is going to be water and water use. The Carbon Disclosure Project, a voluntary initiative, is trying to gather data on water. It's not mandated yet because the crisis hasn't hit. It's going to be very hard to get it mandated until the crisis does hit. So that data tends to be backward-looking in a sense, only in response to a crisis. Because the number of crises is limited and the number of specific issues that are covered tends to be a limited set of issues, it needs interpretive intermediaries. What I mean by this is the datasets are so huge and hard to work with, quite often, that you need people to process this data before it can be used. The HMDA data has something like 10 to 15 million data points in it each year. The TRI data has hundreds of thousands of data points in it. This data has to be processed before it can be used, and, like the voluntary data, it needs analysis.   The voluntary data comes through consensus process and can be built up through consensus, but it lacks consistent participation because it is, by nature, voluntary. It needs

Advancing Social Impact Investments through Measurement
Federal Reserve Board | Washington, DC
Steve Lydenberg | Panel #1

6

consistency and standardization, again, because it is voluntary. And, like the other, it needs analysis. I can't overstress the need for analysis on both these types of data, these datasets. There's a tendency to think that once you have the data out there, the problem is solved. I view data availability as the starting line, not the finish line. So, let me linger here for a moment on what government can do in this area to facilitate more data. These data that are out there create a tremendous amount of value, have made a tremendous amount of process. We need more data as more issues arise, and we need better analysis of the data. So what is it, roughly speaking, that government can do? One, it can facilitate the analysis of the data simply by having it reported out in forms that are easier to use, as, for example, aggregation of data to a company level. The problems with the toxic release inventory data and the HMDA data is that it is very hard to tell what the ultimate parent is. With the data that's going to start to reported under the Greenhouse Gas Registry next year, we'll have tags to the parent company, so that we'll make that easier. So making it easier to aggregate this data to the company level is an important thing. Industry-level benchmarks and comparisons with other databases are important here. The EEOC does aggregate at an industry level. EEOC establishes the aggregate level of employment of women and minorities at different levels so you can have an industry benchmark, but comparing also, for example, HMDA data with Census Tract data is a kind of additional step that government can do. So creating industry benchmarks and comparing

with other databases is something it can do, and identifying key data points.

These databases are so rich that it's very hard to tell what is meaningful.

Second, support for analysis. As I say, analysis is very difficult. Research is

expensive. It's very hard to fund research, and knowing that there is a market

out there is not the same as saying that market has been fully developed. So

markets for research, for those that are actually analyzing these kinds of datas

are really underdeveloped at the moment, and they need support, whether the

support comes from government directly, treating these datas as kind of pure

research and therefore needing subsidy from the government is one approach.

Another approach would be getting the corporations for which this data is most

relevant under one mechanism or another to actually pay for the research by

independent analysis.

Unidentified Male: I hate to interrupt, but any chance you can summarize?

Steve Lydenberg: Yes. This is a bit of the last line. My apologies. Finally,

mandating disclosure of aggregated data, the aggregated data has mostly been

disclosed at a voluntary level. Governments in France, Sweden, South Africa,

are tending toward mandating data disclosure at an aggregated level and

identifying key performance indicators. So that's my final point. Move on.

Thank you. [Applause]