The Governance of Clean Development
Working Paper Series

Exploring the measurement of sustainable development in the Clean Development Mechanism (CDM)

Rachel Killick

University of East Anglia

The Governance of Clean Development
Working Paper 019 – July 2012

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

About *The Governance of Clean Development* Working Paper Series

The Governance of Clean Development is an innovative research programme, which explores the actors, institutions, and policy-making processes at the national and international level involved in promoting clean development, particularly in the area of energy. This research programme is funded through the ESRC Climate Change Leadership Fellowship scheme and is based in the School of International Development (DEV) at the University of East Anglia (UEA).

The working paper series, which is open to all academics and policy practitioners, show cases the latest research and policy thinking on critical issues related to the governance of clean development.

Please note that the *Governance of Clean Development* working papers are "work in progress". Whilst they are subject to a light-touch review process, they have not been subject to a full peer review. The accuracy of this work and the conclusions reached are the responsibility of the author(s) alone and not *The Governance of Clean Development* Project.

For further information on the Working Paper Series, please contact:

Jon Phillips
International Development,
University of East Anglia,
Norwich NR4 7TJ
e: jon.phillips@uea.ac.uk
w: www.clean-development.com

For further information on DEV and International Development UEA, please contact:

School of International Development
University of East Anglia
Norwich NR4 7TJ, UK
T: +44 (0) 1603 592807
F: +44 (0) 1603 451999
E: dev.general@uea.ac.uk
W: www.uea.ac.uk/dev

Exploring the measurement of sustainable development in the Clean Development Mechanism (CDM)

Rachel Killick

Abstract

This paper explores the consistency of a proposed method of measuring sustainable development in the Clean Development Mechanism (CDM), with particular reference to the Gold Standard, a premium CDM label. The research applies a sustainability matrix to 40 CDM and 40 Gold Standard projects to explore the relative influence of different drivers of sustainable development. The study finds that Gold Standard accreditation and project host country are the primary drivers of whether CDM projects have potential to promote a range of sustainable development benefits in this sample. Other postulated drivers including number of CERs (Certified Emissions Reductions) generated and project type are found to be statistically inconsistent. The paper concludes that a set of quantitative measures for sustainable development could improve transparency in assessments of sustainable development potential.

Key words: Clean Development Mechanism (CDM); Gold Standard; sustainable development; multi-criteria evaluation (MCA)

About the author: Rachel Killick was a student at the University of East Anglia, Norwich, UK. This paper reports the findings of a Masters dissertation in the School of International Development. e: rlkillick@gmail.com

This publication should be cited as:
Killick, R. (2012) Exploring the measurement of sustainable development in the Clean Development Mechanism (CDM). Working Paper 019, *The Governance of Clean Development Working Paper Series.* School of International Development, University of East Anglia UK.

Abbreviations and acronyms

AGF	High-Level Advisory Group on Climate Change Financing
BaU	Business as Usual
CDM	Clean Development Mechanism
CER	Certified Emissions Reduction
ECO1	Regional economy
ECO2	Microeconomic efficiency
ECO3	Employment generation
ECO4	Sustainable technology transfer
ENV1	Fossil energy resources
ENV2	Air quality
ENV3	Water quality
ENV3	Land resource pressure
MATA-CDM	Multi-Attributive Assessment of CDM
MCA	Multi-Criteria Analysis
PDD	Project Design Document
SD	Sustainable development
SOC1	Stakeholder participation
SOC2	Improved service availability
SOC3	Equal distribution
SOC4	Capacity development

1. Introduction

1.1 The CDM

The potential for major climatic changes to negatively impact those in developing countries has been widely discussed in a variety of scientific, economic and political arenas (Giddens, 2009). In particular, the United Nations Framework Convention on Climate Change provides a forum for developed and developing nations to meet and discuss ways to mitigate and adapt to the impacts of climate change. Nations have "common but differentiated responsibilities" (United Nations Framework Convention on Climate Change, 1992:2) to respond to climate change, particularly for reducing levels of greenhouse gases[1] emitted into the atmosphere.

This forum led to the creation of the Kyoto Protocol. Ratified in 2005, this legally binds 37 developed countries to reduce emissions to set levels from 2008-12. To enable countries to meet their targets cost effectively, three market-based mechanisms were created (United Nations Framework Convention on Climate Change, 1997). Of these, the CDM in particular has seen significant growth; to date, 4370 projects have been registered (United Nations Framework Convention on Climate Change, 2012). The CDM aims to reduce emissions whilst contributing to sustainable development (SD) in non-Annex I countries (i.e. socially, environmentally, or economically improving the situation where the project is based; United Nations Framework Convention on Climate Change, 1997). It allows a developed country to claim and sell CERs for an emissions reduction project they finance in a developing country.

However, the extent to which the CDM achieves these goals has been questioned (Olsen, 2007; Paulsson, 2009). Criticism of the CDM's contribution to SD is most pronounced in development literature (e.g. Sutter & Parreño, 2007). However, on the whole, perhaps more attention is paid to CDM performance in emission reduction terms. In a bid to preserve national sovereignty, the country hosting a CDM project has final say whether that project contributes to their SD priorities (United Nations Framework Convention on Climate Change, 2006). The sustainable development benefits of projects are not monitored or rewarded, so in some cases benefits to the local community are minimal (e.g. Gilberertson & Reyes, 2009); consequently it has been asserted that SD in the CDM has been sidelined (Newell & Paterson, 2010)

1.2 The Gold Standard

The Gold Standard was created by the World Wildlife Fund to address observed shortcomings of the CDM (Gold Standard, 2009a). With stricter guidelines on which projects could qualify to be certified with the Gold Standard label, it was hoped these premium projects would achieve genuine SD and emissions reductions.

Three main 'screens' are employed by the Gold Standard to encourage best practice CDM projects (Drupp, 2011). First, only renewable energy and energy efficiency projects qualify for registration with the Gold Standard, on the basis that these projects are associated with greater SD benefits (Michaelowa, 2005). Second, the Gold Standard carries out its own, more conservative, assessment of a project's additionality. Because additionality involves a complex set of calculations largely independent of the proposed sustainable development benefits of a project, this will not be explicitly examined in this paper. Finally, every project applying for registration with the Gold Standard is required to submit a 'sustainability matrix'. This checklist approach requires project developers to state what impact their project has on a range of environmental, social, and economic indicators (Bumpus & Cole, 2010). Negative scores are not permitted on any SD indicator, and a net positive score must be attained to achieve Gold Standard certification. Sustainability claims are then monitored and verified (Gold Standard, 2011a).

1.3 Gold Standard Screens: a CDM blueprint post-2012?

The Wuppertal Institute (2009) suggest that elements of the Gold Standard's approach could be mainstreamed to all CDM projects once the Kyoto Protocol expires in 2012. In particular, the Wuppertal Institute (2009) and Boyd *et al.,* (2009) suggest that the Gold Standard's 'sustainability matrix' approach could be mainstreamed. Segregating SD into indicators or introducing minimum standards would require developers to outline predicted project impacts, and enables others to critique and compare 'development dividends' (Cosbey *et al.,* 2006). Secondly, many have suggested restricting further project types viable for registration with the CDM (e.g. Olsen & Fenhann, 2008). Because the Gold Standard restrict project types, the impact that doing so has on SD benefits can therefore be examined.

1.4 Is the Gold Standard a genuine improvement on its non-premium counterpart?

Given these suggestions, determining the relative promise of Gold Standard procedures is vital. To this end, research to date has taken three main directions. First, research has assessed the potential that a premium label has to engender greater SD (e.g. Godfrey-Wood, 2011). This is not explored here because the benefits of having a premium label will always necessarily be felt by a minority; people pay more for a premium label in part because it demonstrates exclusivity. This paper aims to examine ways to lift the SD baseline of CDM projects. Second, researchers have considered whether the Gold Standard's approach could be mainstreamed (Wuppertal Institute, 2009).

Third, research has explored whether Gold Standard projects have greater SD potential than non-labelled projects (Nussbaumer, 2009; Drupp, 2011). Nussbaumer (2009) compares Project Design Documents (PDDs)[2] of Gold Standard and non-labelled CDM projects. Using a sustainability matrix, Nussbaumer suggests that Gold Standard projects may marginally outperform non-labelled CDM projects in terms of the sustainable development benefits that they promise. However, Nussbaumer's small sample (due to the limited number of Gold Standard projects at the time of analysis) limits the certainty of his conclusions. Additionally, PDDs state the intent of CDM projects, not their outcomes; the extent to which these are similar has been questioned (e.g. Gaia Foundation *et al.,* 2011).

Drupp (2011) extends Nussbaumer's (2009) research. He attempts to capture SD differences between typical non-labelled CDM projects and Gold Standard projects. Using a 'representative portfolio', Drupp suggests that project type may be more important than project accreditation for SD potential (again, Drupp uses PDDs and so conclusions can only be made for project potential). If Gold Standard projects propose greater SD benefits then the literature to date suggests that the reasons are not well understood, and hence that the most appropriate methods to mainstream SD are still equivocal.

1.5 Gaps in current knowledge

It remains politically and methodologically unclear whether parts of the Gold Standard's procedures could be used as a blueprint for the CDM after 2012. In particular, the extent to which sustainability matrices accurately capture and encourage SD needs to be assessed. Drupp's (2011) conclusions in this regard are associated with an element of uncertainty. Because Drupp used only a small sample, it could be that his findings are not generalisable. Second, in order to compare CDM and Gold Standard projects directly, Drupp restricted his analysis to PDDs. However, there is more information available in the Gold Standard to those registering projects. This approach may therefore not prove as realistic as assessing all literature available to those registering a project.

Third, although Drupp (2011) suggests that project type may be a more significant driver of SD profiles than Gold Standard accreditation, this was not examined systematically. It may also be that other project attributes not assessed by Drupp could drive SD profiles. For example, Drupp's sample contained CDM projects that varied in their year of registration, but Drupp did not differentiate according to this. The CDM methodology has changed significantly over time, and so it could be that SD differences may be driven by project registration date rather than the stringency of SD indicators (Schneider, 2007). Finally, the Gold Standard created a new methodology in 2009 to improve flexibility and efficiency of their procedures (Gold Standard, 2009a); it could be that SD benefits have since changed. Godfrey-Wood (2011) examines this, but the methodology is mainly qualitative interviews with stakeholders; a systematic quantitative investigation could offer a different perspective.

1.6 Research aims and questions

Based on current literature and concerns, this research therefore aims to understand how assessing and capturing SD may work post-2012 by examining best practice through the Gold Standard. It focuses on the most fundamental of assumptions that need to be investigated:

1. Do refinements of Drupp's (2011) methodology mean that Gold Standard projects are consistently found to engender greater SD potential than non-labelled projects?
2. Can other drivers better explain differences in sustainable development potential between Gold Standard and non-labelled projects?
3. Can the Gold Standard methodology for assessing SD do so effectively for mainstream CDM projects?

Section 2 outlines the research methodology used to explore the SD potential of CDM projects. After outlining the approach to data collection, the method used to assess SD follows, along with sample selection methodology. An outline of data analysis concludes, with justifications and limitations of the approach discussed. Section 3 examines and discusses SD profiles of projects. Section 4 compares the relative importance of different drivers of SD, and Section 5 explores the potential of the Gold Standard methodology for a post-2012 CDM framework. Finally, Section 6 concludes by discussing the need for future research in this area.

2. Methodology

2.1 Data collection

2.1.1 Document selection

Data for analysis was collected from all publicly available documents required for registration with the CDM and Gold Standard, because SD benefits would need to be highlighted in these documents to be rewarded within the UNFCCC and Gold Standard framework. Only those available for the whole sample were used to enable valid comparisons. Gold Standard projects were analysed using PDDs and either the Gold Standard Annex (Gold Standard, 2006) or Passport[3] (Gold Standard, 2009b) depending upon Gold Standard registration version. Non-labelled CDM projects were assessed using PDDs.

Examining both PDDs and Gold Standard Passports/ Annexes is considered here to be a more accurate method than previous approaches that consider PDDs alone. For each Gold Standard project analysed, scores on the MCA matrix were logged when just the PDD was used, then re-rated using the Gold Standard PDD and Annex to determine whether comparisons could be made with past research. With the exception of indicators ECO1 (regional economy) and ECO3 (employment generation), all associations between MATA-CDM scores and documents used to give these scores (PDD vs. PDD and Gold Standard Passport/ Annex) were not statistically significant. Although regional economy and employment generation results should be treated with caution, subsequent analysis used ratings with both PDD and Gold Standard Annex/ Passport input to utilise the full range of available information.

It should be noted however that project potential does not always match project outcome, which has implications for conclusions to be made (eg. Subbarao & Lloyd, 2011). Nonetheless, PDD analysis represents standard practice in the literature (Nussbaumer, 2009), and creates the opportunity for comparisons with other studies. Finally, the reliability of the comparative study presented here is improved by examining rating relative to others; efforts were made to keep ratings consistent (Drupp, 2011). It does however need to be emphasized that this restricts conclusions to be drawn from this research; conclusions can only be made for project SD potential, not outcome.

2.1.2 Multi-Criteria Analysis (MCA)

Sutter (2003) identifies four main approaches for assessing SD potential of CDM projects: guidelines, checklists, negotiated targets, and MCA. The Wuppertal Institute (2009) have proposed MCA as a suitable way for rewarding SD dividends in a post-2012 framework. MCA appears to be the least prescriptive in what constitutes SD, and so was used for this analysis to avoid normative assessment as far as possible.

Within MCA, methods mainly differ in weighting of indicators, and how they assess SD (qualitatively or quantitatively). Sutter's (2003) Multi-Attributive Assessment of CDM (MATA-CDM) is judged to offer the most appropriate combination of flexibility and validity for the analysis undertaken here. This was the same approach taken by Drupp (2011) and Nussbaumer (2009), and so facilitates comparisons with past research.

When assessing multiple indicators, SD indicators should not be equated; different indicators will be prioritized as stakeholders see fit (Sutter, 2003). To weight indicator importance in the present study was judged inappropriate, as value judgments would be assumed for local host communities, and their SD priorities may not match those of their host country (Sutter 2003). One aggregate value could imply that negative scores could be compensated for by positive SD scores on different indicators. Consequently, although examining aggregate SD potential is standard practice in the literature (Aleexew *et al.,* 2010), in this instance results would be less objective and oversimplify findings.

MATA-CDM appropriately offers a range of SD indicators rather than one aggregate, and has been created after analyzing PDDs, thus enabling differentiation between projects. Appendix A presents details of the scoring matrix employed. Twelve indicators of SD were used: four environmental, four economic, and four social (Table 1). These indicators were scored using a combination of qualitative and quantitative methods. Each project was given a score from -1 to 1 for each indicator; the higher the score, the greater the SD potential of the project compared to Business as Usual. Each Gold

Standard project was rated using just the PDD, then re-assessed using all publicly available Gold Standard documents, to determine whether comparisons could be made with past research.

Acronym	Indicator
SOC1	Stakeholder participation
SOC2	Improved service availability
SOC3	Equal distribution
SOC4	Capacity development
ENV1	Fossil energy resources
ENV2	Air quality
ENV3	Water quality
ENV4	Land resource pressure
ECO1	Regional economy
ECO2	Microeconomic efficiency
ECO3	Employment generation
ECO4	Sustainable technology transfer

Table 1: MCA scoring indicators employed.
Source: Drupp (2011)

2.1.3 Sample selection

The Gold Standard Project Registry (1 June 2011; Gold Standard, 2011a) was used to select all 40 Gold Standard projects that are either registered or issued with credits. Projects still in the process of applying for registration with the Gold Standard could still be rejected and conclusions made using these projects could therefore be invalid (Schneider, 2007). Although using only projects with 'issued' status would facilitate comparisons between potential and outcome, this restricted sample would be too small for thorough analysis (Hinton, 2004).

The 40 projects forming the non-labelled CDM sample were randomly selected from a statistically representative pool of non-labelled registered CDM projects, thereby allowing for some of the inherent variability of the CDM portfolio to be retained (Fenhann, 2011). Project attributes used to create the proportionally representative sample were those available from the Risoe CDM Pipeline as of 1st June 2011 (Fenhann, 2011.): methodology, registration date, project submission date, type, size, scale, host country, and developer, Designated Operating Entity[4], PDD author, and number of CERs generated.

2.2 Data analysis

2.2.1 Statistical analysis

In total then, 80 projects were rated (40 Gold Standard; 40 non-labelled). The statistical significance of the differences between the datasets has been assessed. Conclusions presented here could have greater weight than Drupp (2011) and Nussbaumer's (2009), not least because a greater number of projects have been included in the sample.

The most appropriate statistical methodology here would be an ordinal regression; this could explain how much dataset variance could be accounted for by hypothesized predictor variables (Norusis, 2011). However, the sample was not large enough to facilitate valid and useful comparisons using this method (Hinton, 2004). Although categories could have been aggregated to increase expected frequencies (and therefore mean an ordinal regression could be used), this would have made it difficult to make any differential judgements with the data.

Therefore a weaker statistical test has been used, which is valid for categorical data. Fisher's Exact Test is used to determine the probability that associations between projects examined are due to chance. The strength of association between two variables was also assessed using Cramer's *V.* Measured between zero and one, the higher Cramer's *V*, the stronger the association between two variables (Field, 2009). This value could then be compared for various tests of association. Although many postulated drivers of SD are interrelated (Cosbey *et al.,* 2006), comparisons can still result in significant insights.

2.2.2 Statistical test procedure

For the main body of the research, separate analyses were undertaken for eight hypothesized drivers of associations between SD profiles which would best answer the research questions set. Fisher's Exact Test and Cramer's *V* were computed for each of the 12 indicators and postulated drivers (Annex B). Radar graphs or clustered bar charts showing sample SD benefits were created. Finally, association strengths were visualized, allowing conclusions regarding relative importance of SD drivers to be made.

3. SD profiles of postulated drivers: multi-criteria analysis (MCA) results

To answer research questions 1 and 2, SD sample profiles grouped by postulated drivers are examined; consistency of associations are discussed in light of existing research. Examining differences between drivers according to sample mean SD values offers immediate insight into differential SD profiles. Radar graphs show sample mean variation across the main hypothesized drivers of SD and illustrate relative differences within drivers. However, mean values do not take into account different sample sizes, and cannot explain strengths of these differences. Tests of association are therefore applied and discussed below.

3.1 Gold Standard Version

As outlined in Section 1, to determine whether Gold Standard projects have greater SD potential than non-labelled projects, it is necessary to examine whether differing methodologies of the Gold Standard engender different SD impacts. However associations between methodology and every indicator were found to be non-significant. This may be because the number of projects registered under Version 2 of the Gold Standard is currently very small (Gold Standard, 2011a). Version 1 projects score marginally higher on average on more indicators than Version 2 projects (Figure 1). The Gold Standard Foundation created Version 2 in order to increase efficiency (Gold Standard, 2009a); it could be that 'SD corners' are being cut to save time. The current research is however unable to offer robust conclusions here.

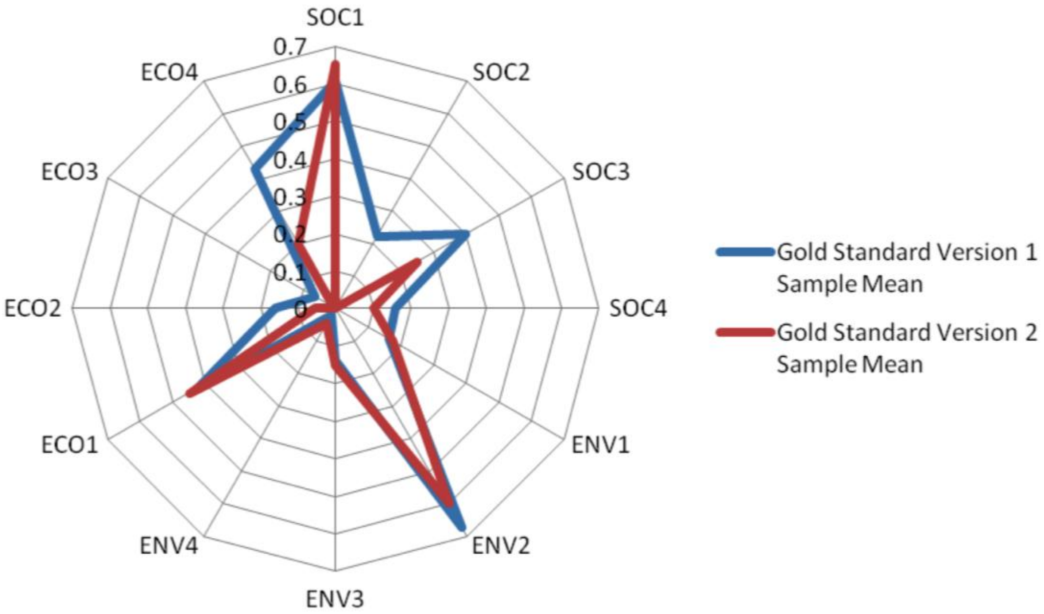

Figure 1: Radar graph showing SD profiles of sample means of Gold Standard projects according to registration version.

3.2 Project Accreditation

The Gold Standard sample is found to have higher average scores on every indicator except fossil energy resources (ENV1) (Figure 2). This could be due to Gold Standard's conservative additionality assessments (Gold Standard, 2009a); exaggerated claims of additionality would serve to increase ENV1 scores, but would be less likely to meet Gold Standard accreditation standards.

Results may suggest that Gold Standard projects emphasise environmental and economic SD over social, because Cramer's *V* values are not as strong for social indicator. This would counter much of the Gold Standard literature however (Gold Standard, 2011b); instead the rating scale or information base used could be ineffective at differentiating with social SD. This is discussed further in Section 5.

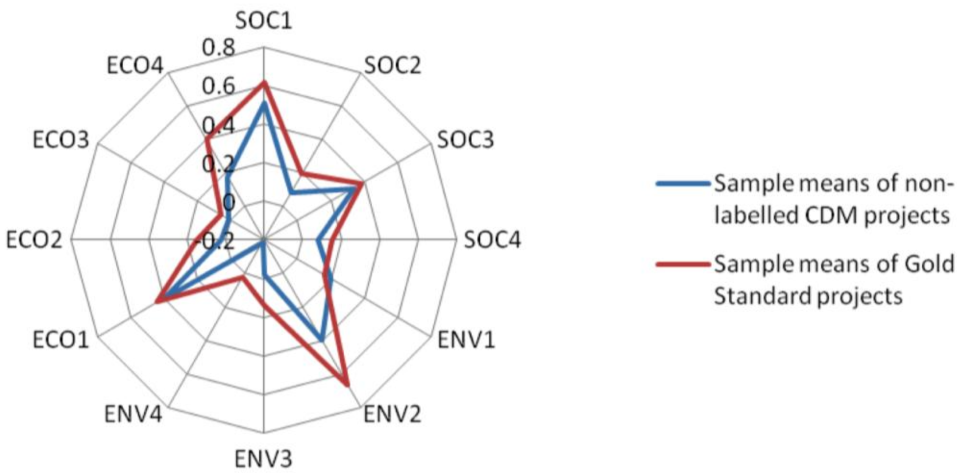

Figure 2: Radar graph of sample means clustered by project accreditation across twelve SD indicators.

All other drivers below present Gold Standard and non-labelled projects combined.

3.3 Association strengths between postulated SD drivers and MATA-CDM scores

Figure 3 shows significant Cramer's *V* values for all postulated drivers that use the same sample. Results presented in Figure 3 and sample mean MATA-CDM scores shall now be discussed to answer research questions set. Each of the charts that follow is associated with a significance test that is displayed in Figure 3.

The higher Cramer's *V* in Figure 3, the stronger the association is between that driver and SD indicator. So for example, in Figure 3, host country has a Cramer's *V* score of 0.7 for SOC1, the strongest association on that SD indicator. This suggests that the association between host country and SOC1 is statistically stronger than any other potential driver examined. If no Cramer's *V* value indicated on Figure 3, any Cramer's *V* value generated was not statistically significant, so results may have been due to chance.

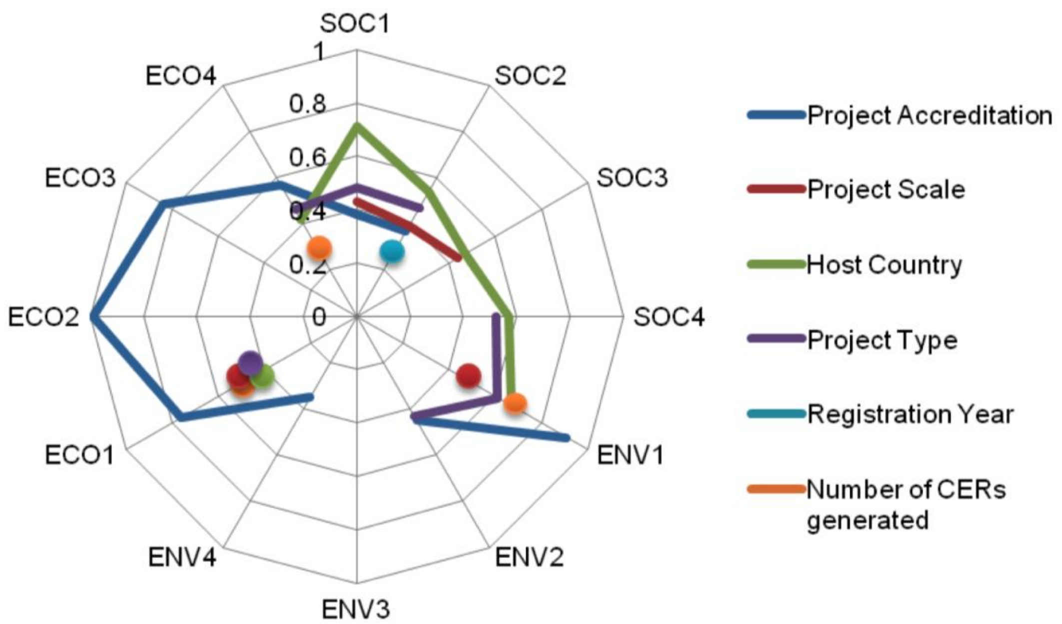

Figure 3: Radar graph of Cramer's *V* values for significant associations between indicator score and postulated drivers.

3.4 Project Scale

There is currently debate whether small-scale projects have greater SD potential than large-scale projects (Cosbey *et al.,* 2006; Olsen & Fenhann, 2008). Figure 4 shows distributions of sample mean scores by project scale. On average small-scale projects score higher than large-scale projects against all indicators other than ENV1 (fossil energy resources), for which the relationship is reversed. ENV1 (fossil energy resources) has a more direct relationship with CERs, and so is rewarded within the existing CDM framework regardless of scale.

The Project Scale driver had significant associations on mainly social indicators (Figure 3); these were assessed in terms of presence or absence of promised social outcomes, rather than the scale of change anticipated. Consequently project scale may be less important in driving SD than indicated by Figure 4 alone.

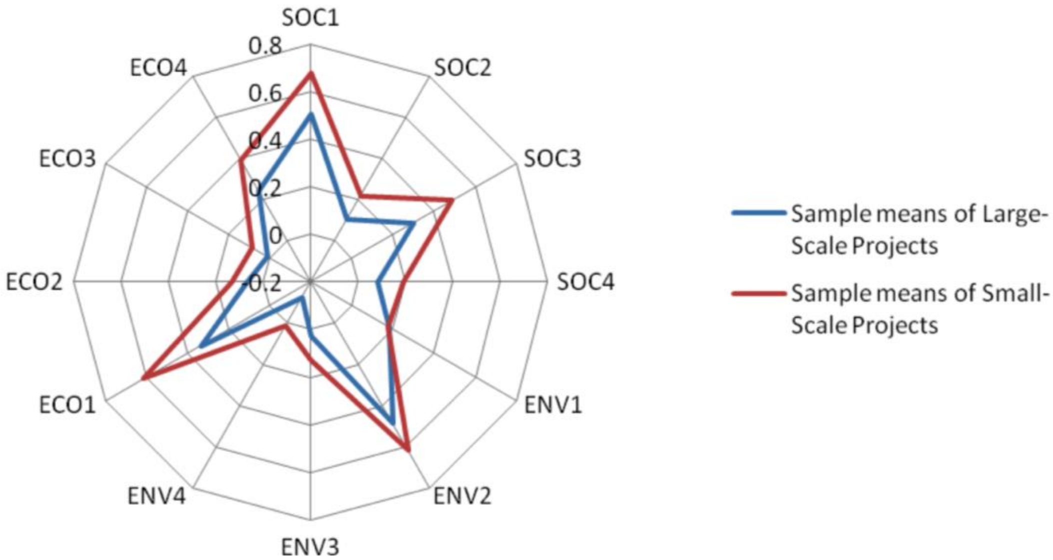

Figure 4: Radar graph of indicator-means of large-scale and small-scale projects.

3.5 Project host country

Because SD priorities are determined by project host country there could be a significant association between a project's host country and its SD profile. As Gold Standard accreditation is not country specific, associations may be blurred. Figure 5 outlines project SD impacts for different host countries. South Africa scores highly on a relatively consistent basis and Brazil scores lower on the majority of indicators, particularly when examining significant associations. Associations were significant for all social indicators, ENV1 (fossil energy resources), and ECO4 (sustainable technology transfer) (Figure 3).

Disch's (2010) research supports the present analysis; South Africa generally scores higher on SD measures than Brazil, except with environmental indicators. Although the present results do not show this reversed trend, differences between SD scores on environmental indicators are much lower than other indicators (Figure 6). Disch shows that host country SD priorities do not seem to influence SD profiles; again, this is reflected in the current results.nIt also appears that the most developed countries in the sample (Brazil, Mexico, India, and China) have a lower SD profile when compared to countries with a smaller share in the CDM portfolio (Fenhann, 2011).

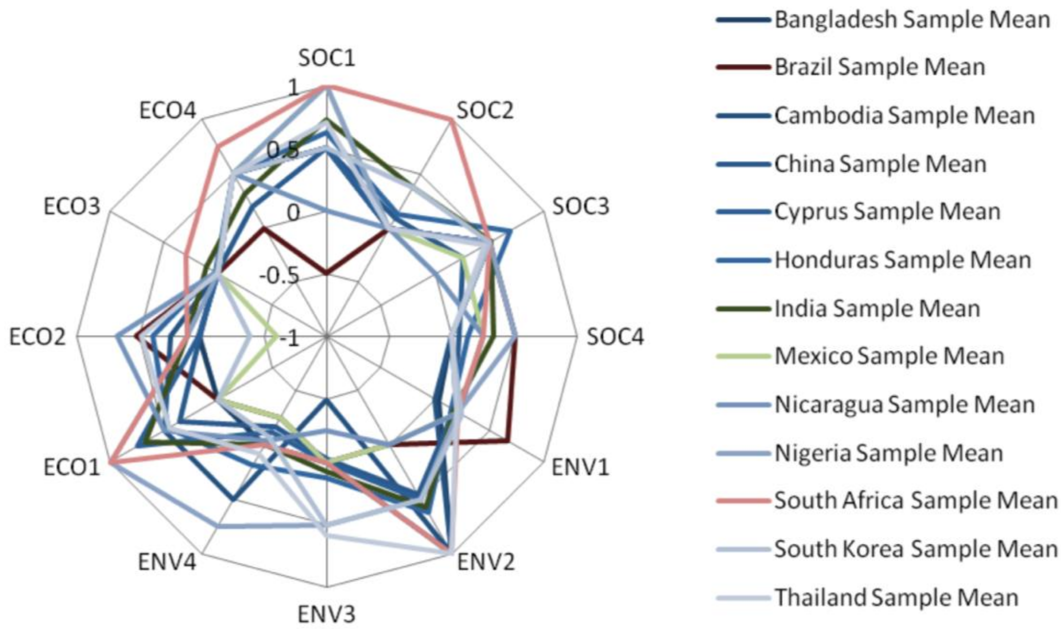

Figure 5: Radar graph of indicator-means of each project host country for twelve SD indicators.

3.6 Project Type

As outlined in Section 1, project type and SD are thought to be associated. However, the present analysis revealed no consistent results to support this (Figure 6); only around half of the associations were statistically significant, which had no bias towards an indicator type (Figure 3). Reasons for these results could therefore be due to indicator measurement, sample sizes, or project type. This shall be discussed in Section 5.

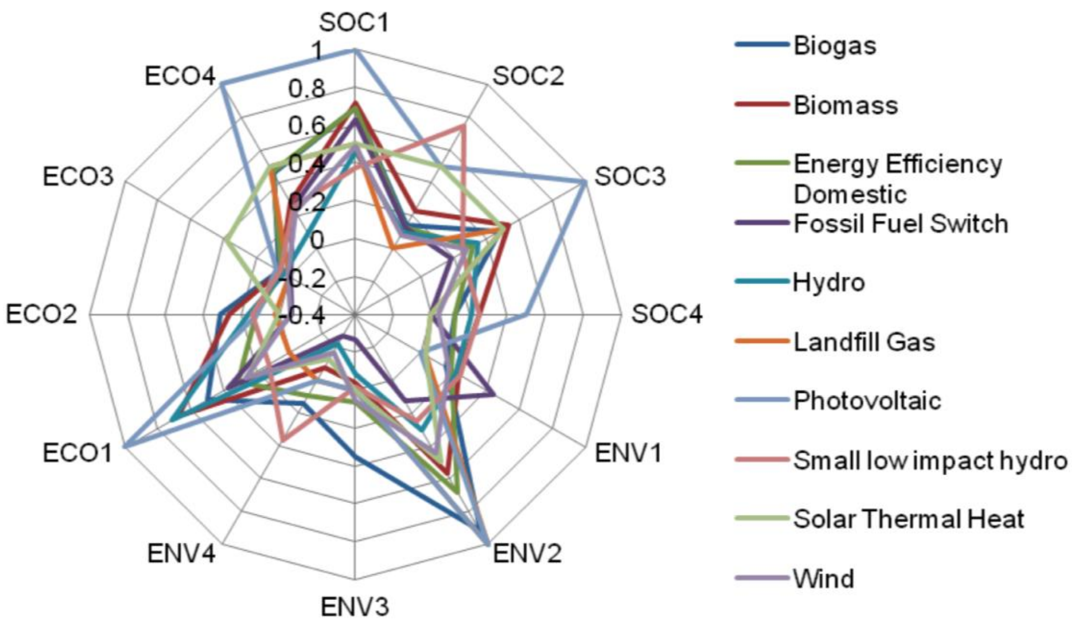

Figure 6: Radar graph of 12 SD indicator scores according to project type sample means.

3.7 Registration Year

As discussed above, SD profiles may vary according to year of registration. Figure 7 visualizes sample mean scores on 12 SD indicators according to project registration year. The trend seems to be that the most recent the project registration year, the lower the average SD score.

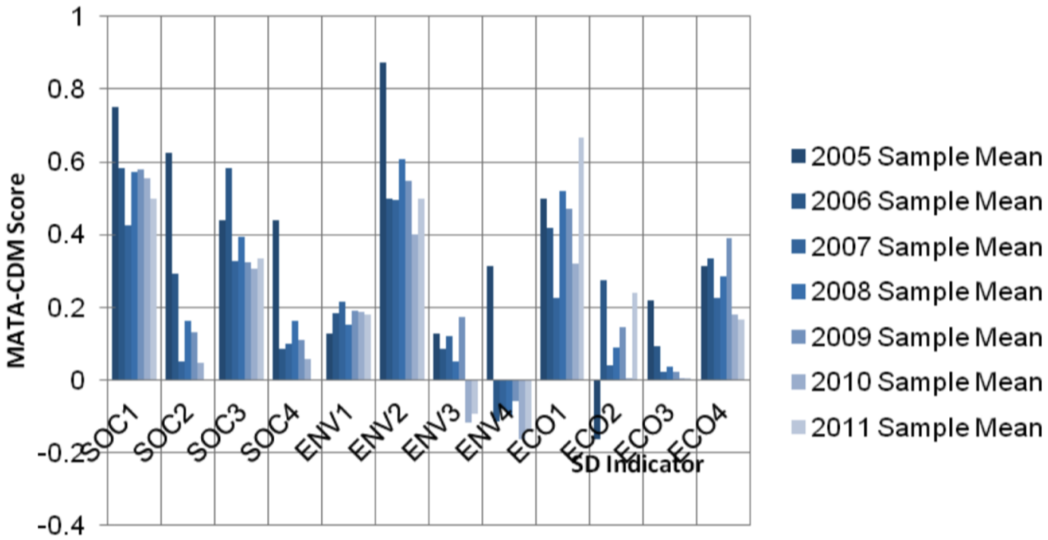

Figure 7: Bar chart of SD profiles grouped by registration year.

This could perhaps be because SD indicators have become stricter over time. Alternatively, now project developers have become practiced in implementing CDM projects, it may be that they no longer feel SD impacts have to be emphasized in PDDs; earlier in the process there might have been more uncertainty as to how much weight would be placed by host countries on SD claims.

However, only SOC2 (improved service availability) had a statistically significant association between registration year and SD profiles (Figure 3). Despite fairly convincing trends shown in Figure 8 then, these are not robust and consistent enough to rule out the possibility that they may be due to chance.

3.8 Number of CERs claimed

Aleexew *et al.* (2010) hypothesise an inverse relationship between number of CERs generated and the associated SD benefits possible. Figure 8 visualises this relationship in the present study.

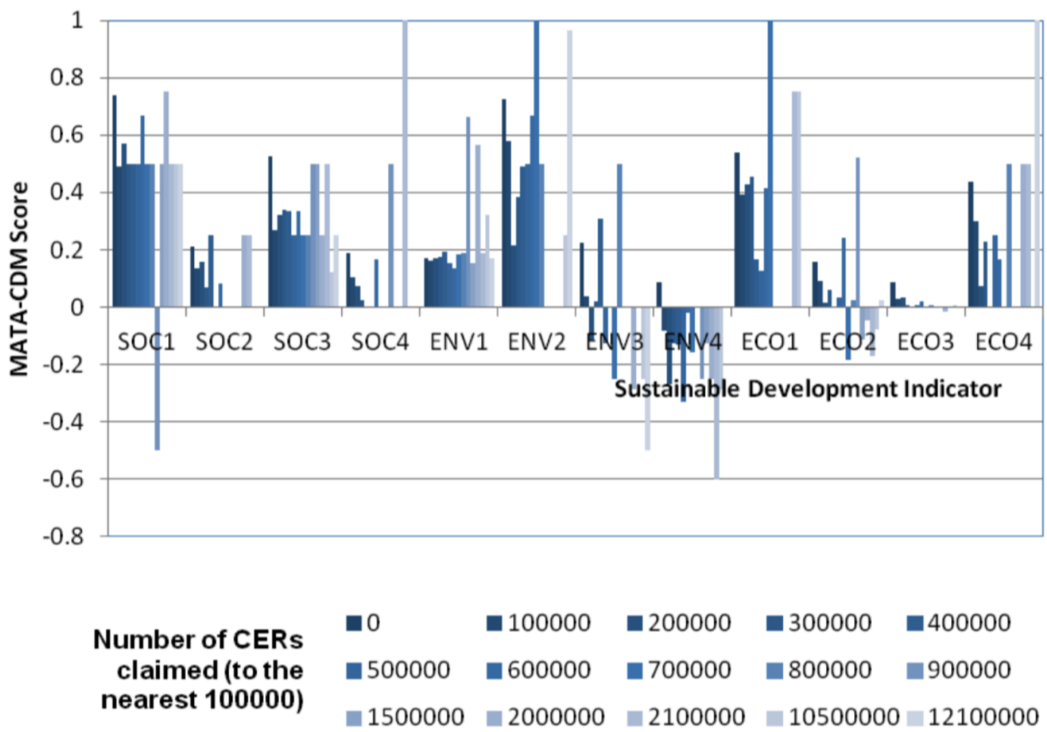

Figure 8: Bar chart of SD profiles clustered by number of CERs claimed (to the nearest 100,000)

As hypothesized, the lower the number of CERs generated, the higher the score. However this was only on social indicators, ENV3 (water quality), ENV4 (land resource pressure), and ECO3 (employment generation). An exception on some indicators was for the project with the highest number of CERs claimed; an energy efficiency project (Annex B; Fenhann, 2011), which scored relatively well here.

Associations were only significant for three indicators (Figure 3). Of these, there was generally a positive relationship between SD score and CERs generated, which was not in the expected direction. However, very little past research in this area (e.g. Olsen & Fenhann, 2008) assesses the likelihood of their results being due to chance. Literature may therefore not be robust enough to warrant conclusions made. Nonetheless, the large number of existing studies (albeit non-statistical in nature) suggesting a trade-off does suggest that the current research may be in some way flawed.

4. Relative strengths of association between postulated SD drivers and MATA-CDM scores

Having discussed SD profiles for postulated drivers individually, it is necessary to explore their relative implications when examined together. Figure 3 is again used to demonstrate significant associations between a postulated driver of sustainable development profile and indicator scores. This can generate insights into which drivers provide a valid base for mainstreaming SD in the CDM.

Sample-mean comparisons demonstrate a variety of potential SD drivers. It would appear that indicator differences are often consistent in the expected direction, with the exception of fossil energy resources. As discussed previously, this may be because the Gold Standard assess fossil energy resource savings differently to the CDM and this could be influencing results. However to ensure valid conclusions, it is vital that sample size is taken into account; sample means could be misleading. This is why one needs to remember there is a relatively small number of significant associations shown in Figure 3, so one should not just look at sample means, as some might be due to chance. This section compares drivers with one another to determine which account for more significant variation in MATA-CDM scores.

Section 2 explains why the measure of the strength of association between two variables is used to weigh the relative importance of different drivers in explaining SD profiles. Because not every association is statistically significant (Figure 3), comparing the relative importance of different drivers is a little complex. Because the Gold Standard combines various postulated drivers of SD, one would expect project accreditation to be the most consistent driver of SD profiles. This combination hypothetically means that the Gold Standard captures the highest performing projects in sustainable development terms. So for example, only certain project types are permitted, the Gold Standard portfolio has a higher proportion of small-scale projects, which in turn means they are more likely to claim smaller numbers of CERs; all of these have been previously demonstrated to be associated at least in part with greater SD benefits.

The main conclusions drawn here are therefore in line with existing assumptions of the relationship between sustainable development benefits and a variety of associated drivers. Although conclusions must be made with caution (particularly given the large number of non-significant associations), it appears that, as anticipated, project accreditation accounts for the most consistent differences in SD profiles. Surprisingly, host country results in stronger associations on half the indicators where project accreditation and host country are significant, but these occur across fewer drivers.

However, project accreditation is not found to unequivocally demonstrate the strongest association between driver and SD profile. For some indicators, there was no significant association between accreditation and indicator, but there were associations with other drivers. For other indicators, association between accreditation and indicator was significant, but other drivers were stronger. Indicators with associations not supporting this hypothesis were: stakeholder participation, improved service availability, equal distribution, capacity development, and water quality (although the latter had no significant associations for any postulated driver; Figure 3).

It seems that where associations are significant, host country has a stronger association with social SD indicators, whereas project accreditation has significant associations more consistently, and is particularly strong for economic SD indicators. The reasons for this must consequently either be differential priorities of Gold Standard or host countries, or the way in which information was gathered.

The reasons for this could be either that other drivers (particularly host country) have greater disparities in whether they emphasise that social indicators should be included in SD priorities, or that the indicators used do not accurately capture differences in social development benefits between Gold Standard and non-labelled projects.

As predicted and demonstrated in Section 3.3, project accreditation accounts for the most consistent differences in sustainable development profiles. Host country has stronger associations on half these criteria where project accreditation and host country are significant. However project accreditation is significant across more drivers. It is surprising that some associations are stronger than project accreditation; these are mainly on social criteria.

Registration year only had one statistically significant association, and when compared with other significant associations, was weakest. Associations were significant between regional economy and all other drivers except registration year. However, no large differences were found between association strengths on regional economy, except project accreditation, which had a stronger association than other drivers. Interestingly, significant associations tend to be clustered on certain indicators, particularly regional economy (although project accreditation remained the strongest association). Associations were significant between regional economy and all other drivers (except registration year), but there were no large differences between drivers on this indicators except project accreditation, which had a stronger association than other drivers.

Associations were not significant between accreditation and any social SD indicator (Figure 3); this is perhaps unexpected. Gold Standard literature does appear to emphasise the importance of social development (Gold Standard, 2011b). This would suggest that perhaps the tool used was not appropriate for capturing social SD as the Gold Standard frame it, but it was useful in terms of disaggregating by host country. Given that the Gold Standard generally favour a quantitative approach to measuring SD (such as requiring that the number of jobs created are listed), whereas host countries as a rule have more qualitative statement-like indicators, this could explain why these results were so. This is discussed in relation to a future research questions in the next section.

4.1 Summary

It can therefore be concluded that project accreditation, as predicted, has the strongest association with SD potential. There are important exceptions however for social indicators. The reasons for this could be either that other drivers (particularly host country) have greater disparities in whether they emphasise that social indicators should be included in SD priorities, or that the indicators used do not accurately capture differences in social development benefits between Gold Standard and non-labelled projects.

The issues raised by this section are clearly complex, and highlight some major issues with measuring SD in the CDM. However, it can be surmised that examining the strength of association between drivers does offer great insight into the complexities of capturing SD, and the dangers of omitting certain project characteristics like project types for project registration because they are assumed to not result in SD benefits (as, for example, the Gold Standard does; Gold Standard 2009b).

It appears that project accreditation is associated with significantly greater SD benefits. Relationships between other drivers are more complex but can be disaggregated. In terms of mainstreaming sustainable development benefits, it would appear that calls to focus on certain project types in particular may be misplaced.

The issues highlighted by these results and discussion will now be examined with relation to the methodology used by the present research, and how useful this would be in terms of practically assessing SD potential of projects in order to answer other research questions.

5. Using MATA-CDM to capture and differentiate by SD potential

Having presented and discussed the results generated when using MATA-CDM to explore consistency of variations in SD profiles of CDM projects, this section addresses the final research question, whether the methodology used to assess SD can do so effectively in a mainstream CDM environment. The main issues raised in the present study when using MATA-CDM to capture and differentiate by SD impacts of projects are discussed, with particular relation to whether this methodology would be feasible in a post-2012 framework, as has been proposed (Godfrey-Wood, 2011). Indicators are used as illustrative examples of the problems and questions that would need to be raised when examining in greater detail what SD assessment in a post-2012 framework would look like.

I argue below that there may be a trade-off between the transparency of methods used to assess SD potential, and the extent to which it is holistic and flexible. Given that SD in the CDM is already critiqued because it is too open to interpretation (e.g. Lohmann, 2006), I suggest that it may be preferable to have indicators which are externally reliable yet reductionist; these may omit important aspects of SD and prove difficult to discriminate between projects, but could be vital in restoring the credibility of the CDM process. This argument is developed by examining issues raised when assessing SD potential in the present research.

5.1 Information base

Firstly, assessing SD potential using PDDs and Gold Standard Annexes/ Passports is a problematic process. Often not enough information is included in these documents to enable one to accurately capture and differentiate by scores on certain indicators used in the analysis.

For example, SOC3 (equal distribution) proved very difficult to assess because information is not provided transparently in PDDs. Some countries (e.g. China) have tax levies on certain project types (Yang *et al.,* 2010). Because this information was not stated in the PDD, it could not be included in the score given; this indicator in some instances may not therefore have been an accurate reflection of SD potential.

This has clear implications post-2012; if SD were to be mandated, PDDs would need to be modified significantly. In addition to being politically difficult and time-consuming, this would mean that more resources would have to be used training project developers.

5.2 Indicator selection

MATA-CDM has been praised because it uses a range of SD indicators; this is seen as less normative (Drupp, 2011). However even when commonly accepted indicators of SD have been identified, deciding the best way to assess progress on these indicators is problematic. In the present research, it was found that subjective assumptions were often made in documentation as to what constitutes SD.

For example, Drupp (2011) created a new rating methodology to assess ECO1 (regional economy). Although also found by the present study to enable easily replicable project differentiation (indicated by the number of significant associations between this indicator and drivers in Figure 3), his rating can be seen as inflexible, with implications for national sovereignty over priority setting. Drupp suggests that the further away a project is from a community, the lower the SD score should be. However, in the case of windfarms for example, project proximity could be a disadvantage (Gamboa & Munda, 2007). If mainstreamed, this indicator may disadvantage certain project types.

5.3 Impact levels

This research argues that in order for SD benefits to be captured and differentiated effectively and transparently, indicators should specify involvement levels, as opposed to binary assessments of presence or absence. For example, SOC1 (stakeholder participation) has a mandatory inclusion in PDDs (Boyd *et al.,* 2007). The Gold Standard recognize the need for improvement on this particular indicator and mandate two stakeholder consultations compared to the CDM's one (Gold Standard, 2009b). However, increasing consultation numbers may not improve stakeholder involvement levels (Arnstein, 1969).

The scoring rationale used for this indicator did examine clear participation levels; it is suggested that this could be feasible for future use, particularly because several significant associations were found (Figure 3).

5.4 The SD Business as Usual

For many projects, the SD Business as Usual was not stated in PDDs and so had to be assumed for scoring. For example, job creation was a commonly stated project benefit. However, if Business as Usual were to build a coal-fired power station, then jobs would be created regardless. However, determining Business as Usual could be very time-consuming. Additionally, stating a counterfactual situation in order to claim financial rewards has already been critiqued with respect to emissions reductions (Schneider, 2009).

5.5 Qualitative vs quantitative indicators

5.5.1 Qualitative SD indicators

Qualitative indicators are problematic because they are more difficult to monitor and verify and hence are more open to interpretation. SOC4 (capacity development) was assessed qualitatively; it is a multi-faceted, complex dimension. However the scoring rationale used (Annex A) had no explicit definition of what this would constitute. Financially rewarding this in the UNFCCC framework would imply some level of measurement. However, this does retain host country sovereignty. With no clear definitions and indicators set, the reliability of such indicators and their verification by third parties could be questionable., particularly because this could mean social benefits are easier to promise.

5.5.2 Quantitative indicators

However, quantitative assessments of SD also have their problems. ENV4 (land resource pressure) was modified for the present study because the process of assigning a rating was judged to be unsystematic. Despite this however, there was only one significant association on this indicator here (Figure 2). This may be because environmental impact assessments in PDDs were fairly formulaic, and so provided little means to disaggregate between projects. This has clear implications; by stating which SD impacts can be included and how they should be measured, there is a risk that all PDDs will state the same impacts and so limit the ways in which they can be financially rewarded, and limit innovation.

5.6 Political volatility of defining indicators

SD registration requirements may mean host countries lose their national sovereignty (Boyd *et al.,* 2009). This could mean significant political divisions regarding how to define SD. For example, ECO4 (sustainable technology transfer) is currently used in the CDM framework. As Das (2011) found, varying definitions of technology transfer are used in PDDs; this may lead to difficulties in terms of differential financial rewards.

5.7 Summary

I argue that although quantitative indicators may overlook important aspects of SD, for a mechanism known to be open to abuse, specifying precisely what constitutes SD is of primary importance at this stage.

Host countries can still maintain control by specifying their own benchmarks for scores on these indicators, and any additional requirements for registration. Given however that the CDM has previously seen limited engagement with SD by host countries, it is more probable that certain minimum mechanism-wide standards would have to be set for SD (perhaps a 'do no harm' approach). If SD were then financially rewarded, there would be a greater incentive to improve on this minimum standard.

Although the inclusion of some indicators is open to question, this doessuggest that MATA-CDM could be a way forward for the CDM post-2012, particularly as it is flexible enough to allow host countries some autonomy over what they see as SD.

6. Conclusions

6.1 Research summary

This research aimed to explore to what extent current efforts to measure and improve SD in the CDM have been successful. A potential way forward for assessing and capturing SD impacts of CDM projects once the Kyoto Protocol expires in 2012 could then be highlighted.

To do so, three main research questions were set: these aimed firstly to improve current research regarding relative potential of Gold Standard and non-labelled CDM projects. Having done so, consistency of proposed drivers could then be investigated with a view to exploring whether they could be encouraged post-2012. Throughout this research, the importance of the methodology used became clear; this is vital in capturing SD. Issues raised when using the present methodology were therefore explored.

Although many issues were raised with this exploratory research, some of the main conclusions supported existing literature. This analysis aimed to update and extend preliminary research undertaken by Nussbaumer (2009) and Drupp (2011) and has broadly supported their findings. Gold Standard projects generally capture greater SD benefits than non-labelled projects. However, the robustness of some of these associations is open to question; generally speaking, economic and social indicators were more easily differentiated than environmental indicators. Interestingly, differences between Gold Standard versions 1 and 2 were not significant on the whole; this may suggest that the Gold Standard has achieved its aim of improving efficiency without compromising SD.

It has been suggested that encouraging certain project attributes may engender greater SD in the CDM. However, the present research found that associations between many drivers were not as robust as previously thought, particularly in the case of project type. Despite current suggestions that host country influence may have a negligible impact on SD profiles (Schneider, 2007), the present study found that, after project accreditation, this had the strongest association with indicator scores, particularly for social indicators.

This highlights the second main issue highlighted by the present research; it is vital to see measurement and outcome as inextricably linked. Social SD indicators used here did provide ways for assessors to disaggregate between projects. However, these differences are argued to be subjective and difficult to replicate objectively; often they do not quantitatively assess the impact of a project, and if they do so, only look at the presence or absence of an indicator, rather than levels of impact. If MATA-CDM were to be used post-2012 to assess SD, these indicators would largely be open to interpretation by stakeholders, and would not be transparent.

This paper therefore argues that, if SD were to be assessed on a wider scale post-2012, MATA-CDM could be useful in doing so. However, contrary to other existing opinion (e.g. Olsen, 2007), it is argued that it is actually more beneficial in this regard to have minimum quantitative indicators that need to be met, rather than all-encompassing statements such as "a significant contribution to capacity development" which may not be externally verifiable. Although this means it may be more difficult to discriminate between projects, I argue this is preferable because it is more transparent.

Agreeing on minimum standards for SD indicators would obviously be very difficult politically, particularly if these are not as open to interpretation as they are currently. However, I argue that having politically feasible indicators precisely because they leave the mechanism open to interpretation would actually lend a false sense of legitimacy to the CDM process. Any asserted progress regarding SD in the CDM, without quantitative indicators, could be seen as redundant. A possible solution would be to have a 'do no harm' approach, with agreed quantitative indicators; any improvement upon this baseline situation could be left to host country national sovereignty.

6.2 Further research

This paper has examined and highlighted issues regarding the most appropriate way to capture SD in the future, if one accepts that mainstreaming sustainable development benefits in non-labelled CDM projects will require dealing with project developers that may not be as well-intentioned as those seeking Gold Standard accreditation, and that resources may not be as plentiful.

Explicitly examining distributions of scores along indicators and significant standardized residuals would be of value in examining why the Gold Standard scores better on the MATA-CDM matrix than non-labelled CDM projects: is this because they are consistently higher, or because there are a few exceptionally high scoring projects, for example? Due to the time-consuming nature of results gathered, this was beyond the scope of this paper. This research could also be updated, as the number of registered projects in the Gold Standard portfolio has increased substantially. This could mean that ordinal regression could be a viable research method.

Examining project outcome rather than potential is also vitally important. One could argue that if project outcomes were no different between Gold Standard and non-labelled CDM projects, then there would be little point in examining in more detail how to assess SD potential of CDM projects.

I argue that research would do well to consider the value of determining statistical significance of results (this could be done on existing research too); the present research does suggest that previous conclusions regarding SD differences may have placed too much weight on their data, particularly given small sample sizes.

This research provides an exploratory starting point in suggesting that SD indicators should be quantitative and easily replicable. Future research would need to examine how this would be received by stakeholders in practice, and additionally whether the experiences of the current research are replicated with larger samples. Particularly valuable would be comparing non-labelled CDM projects with greater sample sizes to determine the consistency of these indicators; in many instances sample sizes were too small to result in statistically significant conclusions.

6.3 Summary

This paper had an overall goal: to greater understand how assessing and capturing SD may work in a post-2012 framework, building on current practical and academic efforts examining best and most feasible practice.

The paper has suggested that there is something to be gained from understanding the way in which the Gold Standard operates, as it does appear to capture greater SD benefits than non-labelled CDM projects. Indeed, this does not seem to be exclusively due to encouraging project type, as Drupp (2011) suggests. The current research focused on examining SD potential, using a methodology previously suggested to be a valid way of assessing SD.

Results generated suggest that to capture and differentiate between CDM projects in SD terms, indicators along a variety of indicators are desirable in order to maintain the ability of host countries to prioritise their own SD needs. However, contrary to existing research, I suggest that indicators should be quantitative in nature. Although this will omit aspects of SD, this will be done in a more transparent manner, and will enable host countries to set qualitative indicators. Academic literature could focus on examining the consistency of these associations, as conclusions are often made without expressing how small sample sizes limit research; it is hoped that the present study has demonstrated this need.

Acknowledgements

I thank the University of East Anglia for allowing me the time and space to complete this research as part of my dissertation for my Masters at UEA. In particular, I thank Professor Peter Newell for his comments on drafts of this paper, and for his supervision throughout the dissertation process. For comments and suggestions to this working paper, I thank Jon Phillips.

---

[1] Atmosphereric gases which absorb infrared radiation (Harvey, 2000).
[2]'Emissions Reductions' within the United Nations Framework Convention on Climate Change and this paper exclusively refer to decreasing levels of anthropogenically emitted greenhouse gases (United Nations Framework Convention on Climate Change, 1997)
[3]United Nations Framework Convention on Climate Change project outlines required for registration with the CDM.
[4]DOEs are organisations approved for verifying or rejecting PDDs according to CDM guidelines (Boyd et al. 2007).

References

Aleexew, J., Bergset, L., Meyer, K., Petersen, J., Schneider, L., & Unger, C. (2010). An analysis of the relationship between the additionality of CDM projects and their contribution to sustainable development. *International Environmental Agreements,* 10(1): 233-248.

Arnstein, S.R. (1969) Ladder of Citizen Participation. *Journal of the American Institute of Planners* 35(4): 216-224.

Boyd, E., Hultman, N., Roberts, J.T., Corbera, E., Cole, J., Bozmoski, A., Ebeling, J., Tippman, R., Mann, P., Brown, K., & Liverman, D.M., (2009) Reforming the CDM for sustainable development: lessons learned and policy futures. *Environmental Science & Policy* 12(7): 820-831.

Boyd, E., Hultman, N.E., Roberts, T., Corbera, E., Ebeling, J., Liverman, D.M., Brown, K., Tippmann, R., Cole, J., Mann, P., Kaiser, M., Robbins, M., Bumpus, A., Shaw, A., Ferreira, E., Bozmoski, A., Villiers, C., & Avis, J. (2007) *The Clean Development Mechanism: An assessment of current practice and future approaches for policy.* [Online]. Available at: www.scribd.com/doc/5283244/The-Clean-Development-Mechanism-An-assessment-of-current-practice-and-future-approaches-for-policy [accessed 5 June 2011]

Bumpus, A.G., & Cole, J.C., 2010. How can the current CDM deliver sustainable development? *Wiley Interdisciplinary Reviews: Climate Change,* 1(4): 541-547.

Cosbey, A., Murphy, D., Drexhage, J., & Balint, J. (2006) *Making Development Work in the CDM: Phase II of the Development Dividend Project. Pre-Publication Version.* [Online]. Available at: www.iisd.org/pdf/2005/climate_realizing_dividend.pdf [accessed 14 June 2011]

Das, K. (2011) *Technology Transfer under the Clean Development Mechanism: an empirical study of 1000 CDM projects.* [Online]. Available at: www.uea.ac.uk/dev/gcd/Das+2011 [accessed 10 August 2011]

Disch, D. (2010) A comparative analysis of the 'development dividend' of Clean Development Mechanism projects in six host countries. *Climate and Development* 2: 50-64.

Drupp, M.A. (2011) Does the Gold Standard label hold its promise in delivering higher Sustainable Development benefits? A multi-criteria comparison of CDM projects. *Energy Policy* 39: 1213-1227.

Fenhann, J. (2011) UNEP Risoe CDM/JI Pipeline Analysis and Database, June 1st 2011. [Online]. Available at: http://cdmpipeline.org/ [accessed 12 June 2011]

Field, A. (2009) *Discovering Statistics Using SPSS.* (3rd Ed.). London: Sage Publishing.

Gaia Foundation, African Biodiversity Network, Carbon Trade Watch, Timberwatch, and Biofuelwatch (2011) *The CDM and Africa: Marketing a new land grab.* [Online]. Available at: http://www.gaiafoundation.org/sites/default/files/CDM_Briefing_Feb2011_lowres.pdf [accessed 27 August 2011]

Gamboa, G., & Munda, G. (2007) The problem of windfarm location: A social multi-criteria evaluation framework. *Energy Policy* 35(3): 1564-1583.

Giddens, A. (2009) *The Politics of Climate Change.* UK: Polity Press.

Gilbertson, T., & Reyes, O., 2009. *Carbon Trading: How it works and why it fails.* Sweden: Dag Hammarskjöld Foundation.

Godfrey-Wood, R. (2011) *Carbon Finance and Pro-Poor Co-Benefits: The Gold Standard and Climate, Community and Biodiversity Standards,* London: IIED.

Hinton, T. (2004) *Statistics Explained: A Guide for Social Science Students.* (2nd Ed.). London: Routledge.

Lohmann, L. (2006) *Carbon Trading: A Critical Conversation on Climate Change, Privatisation, and Power.* Sweden: Dag Hammarskjöld Foundation.

Michaelowa, A. (2005) *CDM: Current Status and Possibilities for Reform.* Hamburg: Hamburg Institute of International Economics.

Newell, N., & Paterson, M., 2010. *Climate Capitalism: Global Warming and the Transformation of the Global Economy.* Cambridge: Cambridge University Press.

Norusis, N. (2011) *IBM SPSS Statistics 19 Advanced Statistical Procedures Companion.* U.S.: Pearson.

Nussbaumer, P. (2009) On the contribution of labelled Certified Emission Reductions to sustainable development: A multi-criteria evaluation of CDM projects. *Energy Policy* 37: 91-101.

Olsen, K.H., & Fenhann, J. (2008) Sustainable development benefits of clean development mechanism projects: A new methodology for sustainability assessment based on text analysis of the project design documents submitted for validation. *Energy Policy* 36(8): 2819-2830.

Olsen, K.H. (2007) The clean development mechanism's contribution to sustainable development: a review of the literature. *Climatic Change* 84: 59-73.

Paulsson, E. (2009) A review of the CDM literature: from fine-tuning to critical scrutiny? *International Environmental Agreements – Politics, Law and Economics* 9(1): 63-80.

Schneider, L. (2007) *Is the CDM fulfilling its environmental and sustainable development objectives? An evaluation of the CDM and options for improvements.* [Online]. Available at: www.oeko.de/oekodoc/622/2007-162-en.pdf [accessed 5 June 2011]

Schneider, L. (2009) Assessing the additionality of CDM projects: practical experiences and lessons learned. *Climate Policy* 9: 242-254.

Subbarao, S., & Lloyd, B. (2011) Can the Clean Development Mechanism (CDM) deliver? *Energy Policy* 39: 1600-1611.

Sutter, C., & Parreno, J.C. (2007) Does the current Clean Development Mechanism (CDM) deliver its sustainable development claim? An analysis of officially registered CDM projects. *Climatic Change* 84: 75-90.

Sutter, C. (2003) *Sustainability Check-Up for CDM Projects. How to assess the sustainability of international projects under the Kyoto Protocol.* [Online]. Available at: www.up.ethz.ch/publications/documents/Sutter_2003_Sustainability_Check-Up_for_CDM_Projects__e-book_.pdf [accessed 10 August 2011]

The Gold Standard (2006) *The Gold Standard Manual for CDM Project Developers.* [Online]. Available at: wwf.panda.org/what_we_do/how_we_work/businesses/climate/offsetting/gold_standard/ [accessed 14 June 2011]

The Gold Standard (2009a) *The Gold Standard Requirements Version 2.1 effective June 2009.* [Online]. Available at: www.cdmgoldstandard.org/fileadmin/editors/files/6_GS_technical_docs/GSv2.1/GSv2.1_Requirements.pdf [accessed 15 September 2011]

The Gold Standard (2009b) *The Gold Standard Rewards Excellence in Carbon Markets: 2009 Annual Report.* [Online]. Available at: www.cdmgoldstandard.org/fileadmin/editors/files/1_communication/Annual_Report_Gold_Standard_2009_FINAL.pdf [accessed 20 February 2011]

The Gold Standard (2011a) *Gold Standard CDM/JI Projects.* [Online]. Available at: https://gs1.apx.com/myModule/rpt/myrpt.asp?r=113 [accessed 12 June 2011]

The Gold Standard (2011b). *The Gold Standard Brochure 2011.* [Online]. Available at: www.cdmgoldstandard.org/fileadmin/editors/files/1_communication/brochure/May_2011_-_GS_Brochure_2011.pdf [accessed 15 September 2011]

UNFCCC (1992) *United Nations Framework Convention on Climate Change.* [Online]. Available at: http://unfccc.int/resource/docs/convkp/conveng.pdf [accessed 5 June 2011]

UNFCCC (1997) *Kyoto Protocol to the United Nations Framework Convention on Climate Change*. [Online]. Available at: http://unfccc.int/resource/docs/convkp/kpeng.pdf [accessed 25 September 2011]

UNFCCC (2006) *Report of the Conference of the Parties serving as the meeting of the Parties to the Kyoto Protocol on its first session, held at Montreal from 28 November to 10 December 2005.* [Online]. Available at: http://cdm.unfccc.int/Reference/COPMOP/08a01.pdf#page=43 [accessed 5 September 2011]

UNFCCC (2012) *Registered CDM Project Activities.* [Online]. Available at: http://cdm.unfccc.int/Projects/registered.html [accessed 23 July 2012].

Wuppertal Institute for Climate, Environment and Energy (2009) *Further Development of the Project-Based Mechanisms in a Post-2012 Regime: Final Report of the Project Commissioned by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety,* Berlin: German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety.

Yang, M., Nguyen, F., De T'Serclaes, P., & Buchner, B. (2010) Wind farm investment risks under uncertain CDM benefit in China. *Energy Policy* 38(3): 1436-1447.

Annex A: MCA scoring criteria and rationale employed

Acronym	Indicator	Score	Scoring Rationale
SOC1	Stakeholder participation	1	If stakeholders can participate actively in the decision process
		.5	If stakeholders are invited to give inputs and raise concerns
		0	If stakeholders are informed
		-.5	If stakeholders are only informed upon request
		-1	If stakeholders are not involved at all; no access to data is possible
SOC2	Improved service availability	*± .25*	*For each service access is gained or lost compared to BaU; this includes compensation for loss of land.*
SOC3	Equal distribution	1	If the largest fraction of the profits from CER revenues flows to the poorer 50% of the host country population (e.g. the project owner is a small producer, local association)
		.5	If the largest fraction of the profits from CER revenues flows to the host country population (e.g. the project owner is a corporation of the host country, a host country owned entity)
		0	If the largest fraction of the profits from CER revenues flows to people outside the host country (e.g. the project owner is an internationally held corporation)
		-.5	If the project activities reduce revenues of the host country.
		-1	If the project activities reduce revenues of the poorer 50% of the host country population
SOC4	Capacity development	1	If significant increase in opportunities for capacity development
		.5	If moderate increase in opportunities for capacity development
		0	If no change compared to baseline
		-.5	If moderate decrease in opportunities for capacity development
		-1	If significant decrease in opportunities for capacity development
ENV1	Fossil energy resources	$\frac{tC_B - tC_P}{CER * 0.744}$	Where: tC_B : tons of coal equivalent per year used by baseline; tC_P : tons of coal equivalent per year used by project
ENV2	Air quality	*1*	*If directly improving air quality in local*

			households
		±.5	*If increase or decrease in gases (other than those CERs have been claimed for)*
			If increase or decrease in odour
		0	If no change compared to baseline
ENV3	Water quality	1	Significant decrease of pressure on the water supply
		.5	*Water quality stated as 'improved' compared to BaU; moderate decrease of pressure on the water supply*
		0	No change compared to baseline
		-.25	*If water sprayed to mitigate dust pollution during project operation*
		-.5	Moderate increase of pressure on the water supply
		-1	Significant increase of pressure on the water supply
ENV4	Land resource pressure	*±.25*	*For each land resource pressure that changes compared to BaU*
		0	If no change compared to baseline
		-.035	*If project is located on infertile or barren land*
		-.25	*If project is located on occupied land*
ECO1	Regional economy	1	If located in poor rural community or if otherwise stated that location is economically disadvantaged
		.75	If located in village/ rural area *(if located in factory in rural area also)*
		.5	If located in town/ low income suburb
		.25	If located in suburb
		0	If located in city, industrial zone, factory, or oil field offshore
ECO2	Microeconomic efficiency	1	If Internal Rate of Return (IRR) ≥ 20%
		(IRR/10) - 1	If 0% < IRR > 20%
		-1	If IRR ≤ 0
ECO3	Employment generation	$\frac{J_P - J_B}{CER * 0.22}$	where: J_P : man-months created by the project; J_B : man-months created by the baseline case. CER: average annual CERs. *Following Drupp's example PDD analysis, If stated that "indirect opportunities" would be created, value is divided by seven. If number of jobs created not quantified, a minimal nominal value of one or two jobs assumed.*

ECO4	Sustainable technology transfer	*A flexible definition of technology transfer was adopted; although the UNFCCC specify technology transfer should be international, no standardized definition of the term exists in PDDs (Das, 2011) and so the same definition was used as contained in the PDD.*	
		1	If the technology is innovative and the capacity exists locally to maintain and manage it
		.5	If the technology is innovative but external assistance is required to develop local skills
		0	If there is no technological transfer or the innovative technology requires durable external assistance
		-.5	If external skills must be imported and the project creates dependence
		-1	If the new technology cannot be maintained and managed in the long-term

Source: Adapted from Drupp, 2011 (adaptations in *italics*)

Annex B: Fisher's Exact Test and Cramer's *V* values for associations between 12 SD indicators and six postulated drivers

Criteria	Driver	Fisher's Exact Value	Cramer's *V*	Significance Level
SOC1	Project Accreditation	11.19	.38	$p \geq .05$
SOC2		10.23	.37	$p \geq .05$
SOC3		8.95	.35	$p \leq .05$
SOC4		3.98	.22	$p \leq .05$
ENV1		60.87	.91	$p \geq .05$
ENV2		15.83	.45	$p \geq .05$
ENV3		9.37	.35	$p \leq .05$
ENV4		33.35	.67	$p \geq .05$
ECO1		46.28	.76	$p \geq .05$
ECO2		73.80	.99	$p \geq .05$
ECO3		55.02	.84	$p \geq .05$
ECO4		26.89	.57	$p \geq .05$
SOC1	Project Scale	12.51	.43	$p \leq .05$
SOC2		11.29	.39	$p \leq .05$
SOC3		13.89	.44	$p \leq .05$
SOC4		5.84	.26	$p \geq .05$
ENV1		15.05	.43	$p \leq .05$
ENV2		4.10	.24	$p \geq .05$
ENV3		8.39	.34	$p \geq .05$
ENV4		20.69	.55	$p \geq .05$
ECO1		19.58	.51	$p \leq .05$
ECO2		60.14	.90	$p \geq .05$
ECO3		40.50	.73	$p \geq .05$
ECO4		8.69	.33	$p \geq .05$
SOC1	Host country	138.36	.71	$p \leq .05$
SOC2		86.19	.54	$p \leq .05$
SOC3		125.96	.47	$p \leq .05$
SOC4		70.99	.57	$p \leq .05$
ENV1		146.02	.67	$p \leq .05$
ENV2		135.68	.33	$p \geq .05$
ENV3		Fisher's Exact Test timed out		
ENV4		Fisher's Exact Test timed out		
ECO1		60.25	.48	$p \leq .05$
ECO2		1607.12	.73	$p \geq .05$
ECO3		979.90	.60	$p \geq .05$
ECO4		144.62	.42	$p \leq .05$
SOC1	Project Type	124.38	.48	$p \leq .05$
SOC2		109.29	.47	$p \leq .05$
SOC3		Fisher's Exact Test timed out		
SOC4		60.76	.52	$p \leq .05$
ENV1		141.47	.61	$p \leq .05$
ENV2		143.29	.44	$p \leq .05$
ENV3		Fisher's Exact Test timed out		
ENV4		Fisher's Exact Test timed out		
ECO1		76.36	.54	$p \leq .05$
ECO2		Fisher's Exact Test timed out		
ECO3		Fisher's Exact Test timed out		
ECO4		138.94	.46	$p \leq .05$
SOC1	Registration Year	30.11	.24	$p \geq .05$
SOC2		38.73	.35	$p \leq .05$
SOC3		33.87	.28	$p \geq .05$
SOC4		17.12	.29	$p \geq .05$
ENV1		Fisher's Exact Test timed out		

ENV2		Fisher's Exact Test timed out		
ENV3		Fisher's Exact Test timed out		
ENV4		Fisher's Exact Test timed out		
ECO1		29.56	.33	$p \geq .05$
ECO2		Fisher's Exact Test timed out		
ECO3		Fisher's Exact Test timed out		
ECO4		Fisher's Exact Test timed out		
SOC1	Number of CERs generated	130.01	.58	$p \geq .05$
SOC2		110.86	.41	$p \geq .05$
SOC3		118.27	.57	$p \leq .05$
SOC4		54.30	.49	$p \geq .05$
ENV1		130.64	.74	$p \leq .05$
ENV2		170.68	.65	$p \leq .05$
ENV3		Fisher's Exact Test timed out		
ENV4		Fisher's Exact Test timed out		
ECO1		70.31	.48	$p \leq .05$
ECO2		Fisher's Exact Test timed out		
ECO3		Fisher's Exact Test timed out		
ECO4		159.56	.36	$p \leq .05$