THE HONG KONG
POLYTECHNIC UNIVERSITY
香港理工大學
Pao Yue-kong Library
包玉剛圖書館

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

Pao Yue-kong Library, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

http://www.lib.polyu.edu.hk

ON SUSTAINABILITY

HELEN SHANYIN CHEN

PhD

The Hong Kong Polytechnic University

2021

The Hong Kong Polytechnic University
Department of Logistics and Maritime Studies

On Sustainability

Helen Shanyin Chen

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

May 2020

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgment has been made in the text.

Helen Shanyin Chen

To those ordinary people who see the world as it is,

and still love it deeply

ACKNOWLEDGEMENT

I am grateful to my chief advisor, Prof. Edwin Cheng, for his support and guidance. I sent Prof. Cheng my first research proposal a decade ago but took five years to bid farewell to my corporate career and fully commit to this intellectual journey. Prof. Cheng has always been supportive and believed in me, even and especially when I had much to prove, nothing to show, and still chose a path that is steep and risky. He is exceptionally sharp, demands hard work and excellence, and expects nothing in return but his students' professional development and growth. I feel indebted to Prof. Cheng and look up to him as an academic role model.

I appreciate the opportunity to have spent two and a half years at INSEAD and had Prof. Luk Van Wassenhove as my host supervisor. I am grateful for the valuable lessons learned from Prof. Van Wassenhove, especially those on practice-based impactful research.

Many of you have inspired and encouraged me in one way or another. To you all, my gratitude. I am particularly thankful to Liam Barker for proofreading this thesis and for always being there.

My deepest gratitude belongs to my parents, as they have shaped me into who I am today. My mother has taught me the traditional Chinese virtues of sympathy, generosity, and selflessness. My late father had taught me discipline, hard work, and the meaning of time and space in life. After an odyssey through peaks and valleys around the world searching for a place where I could find and free myself, I have now come back to where I started feeling whole and mellow, with only one wish that I had returned earlier before my father's departure. I find solace in the thought that perhaps father can see me all the time now watching over me from above.

This thesis is written at the time of a pandemic when the world is going through much unspoken pain and suffering. Death has never felt so real and life, so fragile. Yet, it is through death that we learn to better appreciate life. Despite – or rather, because of – the acute awareness that at some indefinite time, gravity will bring us to the ground after all our upward struggles, I still believe that the way upward is the only way forward as at the end of the day, to have truly lived and lived well is to not surrender to the heaviness of void but have exerted to elevate and transcend oneself through love, faith, and hope. I hope you will join us in this journey.

Now, let the adventure begin, and may the most faithful win.

ABSTRACT

This thesis studies sustainability at the organizational level in the field of business and management research. It explores different aspects of the subject with two complementary essays: the first essay develops a theoretical model to explicate the interactions on the environmental-economic intersection of sustainability in the private sector; the second essay designs a strategic supply chain framework to enable sustainable practices for social value creation in the humanitarian sector. The purpose that motivates both essays is to ground the sustainable development goals (SDGs) in contexts and develop practically relevant knowledge in sustainability research that is readily translatable into actionable solutions.

Essay 1, *Green Competitiveness*, studies the environmental-economic intersection of corporate sustainability. It develops an intermediate theory of green competitiveness (GC) on the nexus' context-contingent interactions. The rationale underpinning the study is that a pragmatic approach aligning environmental protection with business' economic logic can be more effective in driving and sustaining corporate green practices. The project's mixed-methods design takes three phases: (1). Phase 1 – systematic literature review retrieves and identifies 171 relevant and quality academic articles on the topic of GC; (2). Phase 2 – quasi meta-analysis codes and categorizes the dataset by research characteristics to quantify empirical evidence for data-driven theorizing; (3). Phase 3 – abductive reasoning develops progressively six moderated-mediated propositions through a rigorous and replicable process of hypothesis generation, refinement, evaluation, and acceptance. Also proposed in conjunction with the preliminary GC theoretical model are a new green typology, an intermediate construct of GC, and four concepts of key contextual contingencies. Together, they elevate the GC research stream from a nascent to an intermediate theory. During the theory development process, the essay also makes three peripheral contributions that address the following research gaps in the existing literature: (1). The G-C logic chain designed to support abductive reasoning can serve as a coherent framework to integrate fragmented studies on different subdimensional environmental-economic (G-C) relations; (2). The extensive list of measures for green and competitiveness can be used to improve consistency and validity of construct operationalization in future GC studies; and (3). The holistic organizational performance framework (HOPF) designed for hypothesis evaluation can be refined into a theoretical model for sustainability research. Academics interested in furthering GC research may consider the essay's recommendations to (1). Go beyond the construct level to examine specific G-C subdimensional links with strong theoretical matching and proximal causal connection; (2). Limit the study to a single sector or region if the key contextual contingencies cannot be measured and controlled; (3). Collect longitudinal data, use nonlinear models, and vary lag lengths between focal constructs to capture the time-variant dynamics of G-C interactions. Practitioners may refer to the GC five-forces framework

for corporate environmental strategy planning, the roadmap towards the green market for implications on building a context conducive to corporate green practices, and the hierarchy of green forces for an institutional view on the structure and functions of different green motivations – all developed by applying the proposed GC theoretical model in different contexts. Future research may empirically test the proposed hypotheses, develop measurement items and scales for the intermediate construct of GC, or design an interdisciplinary framework for environmental performance measurement which is fundamental to harnessing GC as a force for sustainability.

Essay 2, *Sustainable Humanitarian Supply Chains*, designs a strategic supply chain framework to enable context-contingent sustainable humanitarian operations. The sustainable humanitarian supply chain (SHSC) strategic framework is developed through a four-phase design process: observing the contexts, framing the problem, eliciting the design imperatives, and developing the solution. The solution imperatives are elicited by specifying the performance objective of sustainable humanitarian operations, identifying the required supply chain capabilities, and distilling the practical SHSC tactics. The framework is then formulated based on the specified design imperatives. It proposes four context-contingent SHSC strategies: Operational Leanness, Scalable Readiness, Proximal Readiness, and Collaborative Preparedness and Response. Humanitarian practitioners may use the framework to identify contextually congruent strategies to improve the sustainability performance of HSCs. A roadmap for phased implementation is also proposed based on humanitarian organizations' mission profile and capability portfolio. The framework fills the research gap of a strategic supply chain framework with sustainability as the performance objective in the humanitarian context. Future research may design an interdisciplinary framework for social performance measurement or experiment to elicit humanitarian practitioners' mental decision models in emergency response situations.

Tables of Content

List of Figures

List of Tables

Chapter 1 Introduction

This thesis studies sustainability at the organizational level in the field of business and management research. It explores different aspects of sustainable development with two complementary essays: Essay 1 (Chapter 2) develops a theoretical model to explicate the interactions on the environmental-economic intersection of sustainability in the private sector; Essay 2 (Chapter 3) designs a strategic supply chain framework to enable sustainable practices for social value creation in the humanitarian sector. The essays are driven by the rationale to ground SDGs in specific contexts and develop practically relevant knowledge in sustainability research that is readily translatable into actionable solutions. Towards this end, both projects apply novel yet rigorous mixed-methods to explore the research topics.

Above and beyond the specific findings from each essay, three lessons can be synthesized from both studies. First, sustainability research needs to take an interdisciplinary approach. Broadly speaking, sustainability is a multifaceted subject concerning every discipline in academia and every sector in society. An interdisciplinary approach can avoid tunnel vision and bring together divergent perspectives from different domains to enable and enrich coordinated solution design towards SDGs. Second, supply chain management can be an ideal anchor point for interdisciplinary sustainability research. As a field of study in management systems engineering, it can provide a variety of concepts and methods to bridge the gap between theory and practice in sustainable development. Lastly, given its intrinsic normative nature, sustainability research as a scientific field of inquiry shall not and cannot be entirely value-free. It is perhaps necessary to adopt a pragmatic approach applying positivist methodology for normative application in future sustainability research. The rest of the thesis demonstrates in vivid detail how these lessons are learned. A brief overview is presented in the following executive summaries. Figure 1.1 illustrates the research positioning of the thesis.

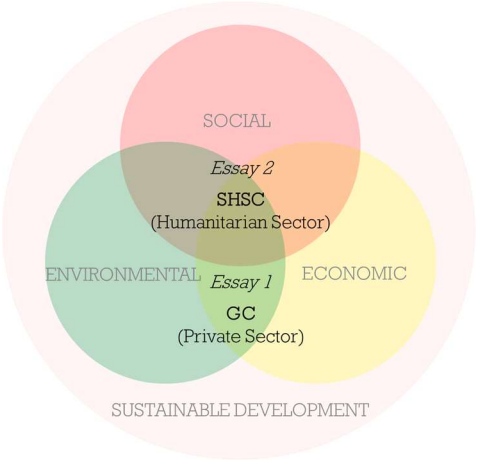

Figure 1.1 Research Positioning of the Thesis

1.1 Executive Summary of Essay 1 – Green Competitiveness

- Research Objectives

This essay explores the interactions on the environmental-economic (G-C) intersection of corporate sustainability. Motivating the study is the belief that a pragmatic approach aligning environmental protection with business' economic logic can be more effective in driving and sustaining corporate green practices. Research interest in this area that emerged several decades ago has continued to grow, but there is still no conclusive consensus regarding "*whether it pays to be green*". A close examination of seven GC systematic reviews published between 2009 and 2018 finds that they all identified a positive but weak link between the dyad. Yet, it is unclear whether the empirical evidence can be directly cumulated for research synthesis as the articles they reviewed are highly heterogeneous in terms of construct operationalization, methodological artifacts, and study contexts. Built upon the frontier of GC research, this essay intends (a) to identify the key contextual contingencies on the G-C nexus, and (b) to develop a preliminary theoretical model that explicates the interactions of the dyad. These research objectives address the questions of "*when/why it pays to be green*".

- Methodology

This study conceptualizes the relationship between environmental management practices (EM) and business performance (BP) as a neural network, mediated by GC and moderated by contextual contingencies. It adopts a multi-phased mixed-methods design that includes systematic literature review (SLR) in Phase 1, quasi meta-analysis (QMA) in Phase 2, and abductive reasoning (AR) in Phase 3.

SLR in Phase 1 takes three steps: database search, relevance screening, and quality screening. It searches and collects the comprehensive dataset of academic GC articles as *raw materials* in preparation for data-driven theorizing. The process recalls 1,847 potentially relevant articles from the Web of Science (WoS) database with a replicable search protocol, identifies 230 publications after screening for relevancy and retains 171 articles after screening for quality.

QMA in Phase 2 analyzes the collection of GC publications as *work-in-progress* for evidence cumulation, pattern detection, and inductive generalization. It codes and classifies the dataset in the QSR Nvivo software from the dimensions of construct operationalization, methodological artifacts, and contextual contingencies. The results from QMA are a set of generalized observations that provide initial descriptive answers to the question of "*when it pays to be green*".

AR in Phase 3 further processes the QMA results into the *final product* of a preliminary GC theoretical model by *reasoning from effects to causes*. The theorizing takes four steps: hypothesis generation, hypothesis refinement, hypothesis evaluation, and hypothesis acceptance. The abductive inference process goes like this: (1). D is a collection of data from SLR; (2). F is the QMA results from D; (3). H explains F, and among all the hypotheses, no others can explain F as well as H does – therefore, H, the proposed GC theoretical model, is probably true. Before accepting the theoretical model as a tentative answer to "*why it pays to be green*", its validity is evaluated through a holistic organizational performance framework (HOPF) and is reconciled with existing management theories relevant to GC research.

- Results and Contributions

The primary contribution of this essay is a preliminary theoretical model that explicates the G-C interface of corporate sustainability. The model, together with its supplementary theoretical components – a new green typology, an intermediate construct of GC, and four concepts of key contextual contingencies – elevate GC research from a nascent to an intermediate theory. Based on the theoretical model, the essay also develops a GC five-forces framework for corporate environmental strategy planning, a roadmap towards the green market for implications on building a context inducive to corporate green practices, and a hierarchy of green forces with an institutional perspective to depict the structure and functions of different green motivations.

The essay's secondary contribution is that it also addresses three research gaps in the process of developing the GC theoretical model. First, the G-C logic chain, developed as the analytical AR framework, closes the gap that there previously had not existed a unifying framework to integrate fragmented studies on different aspects of the G-C dyad. Second, it collects, compiles, and classifies an extensive list of green and competitiveness measures that can be used to improve the validity and consistency of construct operationalization in future GC research. Third, HOPF designed for hypothesis evaluation has the potential to be further developed into a theoretical model for sustainability research.

- Limitations

The essay's fundamental limitations are associated with its secondary data source of academic articles and its methodological choice of theorizing by abduction. Using academic articles as the data source for theory building might expose the project to potential publication bias, *the drawer effect*, where only positive results are reported and published. The risk of publication bias has been mitigated by assigning a higher weight to negative research findings in the data analysis and hypothesis development process. Such a dataset might also be afflicted by the observational bias of

the streetlight effect in that answers are only sought where data are available. The risk of observational bias is reduced by the theorizing method of abduction that goes beyond immediate evidence to infer to best explanations. Furthermore, theorizing through abduction might present threats to reliability. Uncertainty is inherent in abductively developed hypotheses – they are neither strictly induced from generalized observations nor precisely deducted from well-established theories, but also make use of relevant suppositions, intuition, and disciplined imagination that cannot be fully articulated, codified, and replicated. The study makes the following efforts to reduce the threats to reliability: (1). It assures falsifiability of the proposed hypotheses by methodically documenting the reasoning processes and making transparent where the logical inference flow might go adrift; (2). It enables critical assessment of the confidence level in the theoretical model by presenting empirical observations alongside the inferred hypotheses and highlighting the areas where evidence might be lacking or gaps might exist. Nonetheless, although the threats to validity and reliability have been minimized through these efforts, the proposed GC model is but a preliminary theory that needs to be empirically validated in future research.

- Future Research

Future research in GC may focus on three areas: (1). Empirically test the moderated-mediated hypotheses proposed by the preliminary GC theoretical model; (2). Develop measurement items and scales for the intermediate construct of GC; (3). Design an interdisciplinary framework for environmental performance measurement which is fundamental to harnessing the force of GC for sustainability.

1.2 Executive Summary of Essay 2 – Sustainable Humanitarian Supply Chains

- Research Objectives

This essay designs a strategic supply chain framework to enable sustainable practices in humanitarian operations. Its objective is to develop a practical tool to integrate sustainability considerations into operational strategy decision-making in the humanitarian context. The research is positioned at the intersection of humanitarian operations, supply chain management, and operations strategy within the broader field of sustainability research.

- Methodology

The project conceptualizes the research problem based on the strategy-structure-capabilities-performance framework (SSCP). Adapted to this study, SSCP implies that for a humanitarian organization to achieve its sustainability performance objective, it needs to develop context-congruent supply chain capabilities to enable its intended supply chain strategy and structure.

The SHSC strategic framework is designed through a four-phase process: observing the contexts, framing the problem, eliciting the design imperatives, and developing the solution. The design imperatives are elicited by (a) specifying the performance objective of sustainable humanitarian operations (net sustainability value), (b) identifying the required enabling capabilities (responsiveness, efficiency, flexibility, reliability, and agility), and (c) distilling practical SHSC tactics (localization, collaboration, and dematerialization). The framework is then formulated based on the specified design imperatives

- Results and Contributions

The SHSC strategic framework proposes four context-contingent strategies: (a) Operational Leanness in low urgency/low unpredictability contexts, (b) Scalable Readiness in low urgency/high unpredictability contexts, (c) Proximal Readiness in high urgency/low unpredictability contexts, and (d) Collaborative Preparedness and Response in high urgency/high unpredictability contexts. Humanitarian practitioners may use the framework to identify contextually congruent HSC strategies to improve sustainability performance. The framework fills the research gap of a supply chain strategic framework with sustainability as its performance objective in the humanitarian context. A roadmap for phased implementation is then proposed taking into consideration the resource and capability constraints of humanitarian organizations. Finally, the essay discusses the potential of taking supply chain management as an anchor point for sustainability studies – the systematic, universal, and practical nature of the field makes it a natural fit for the multidimensional, interdisciplinary, and conceptually abstract yet practically relevant subject of sustainability research.

- Limitations

The key limitation of this essay is that the proposed SHSC strategic framework is not implemented in practice for design testing and refinement. Its applied research method of design thinking is not the full, but only a partial, version of the design science approach. However, the framework's pragmatic validity can still be defended since it is developed through a transparent and falsifiable process grounded on the author's intensive field experience. The validity of the design concept is also proved by enfolding relevant theories to elaborate on its plausibility. Nevertheless, the framework needs to be empirically tested in future studies through focus group interviews, experimentation, or pilot implementation by collaborating with humanitarian organizations.

- Future Research

Future research in this line of inquiry may focus on two timely and interesting areas: (1). Develop an interdisciplinary framework for standardized social performance measurement at the organizational level; (2). Experiment to elicit humanitarian practitioners' mental decision models in

emergency response situations. Decision-making in emergency response is particularly challenging as it involves human suffering and life loss. Scenario-based experimentation can explicate the rationales and biases behind practitioners' trade-off and prioritization decisions. Findings from such behavioral research can help design foolproof SHSCs with layered default choices that lighten practitioners' cognitive load and reduce reliance on their value preferences in high-stake emergency response situations.

Chapter 2 Green Competitiveness

1. Introduction

1.1 Background

From the Amazon to Australia, the world is on fire. The private sector has a critical role to play in the environmental crisis. First, the sector's natural resource consumption and pollution generation are responsible for most of the environmental damage caused by economic activities. For example, a report by the Principles for Responsible Investment (2010) found that the top 3,000 public companies together caused US$ 2.15 trillion of environmental costs in 2008, over one-third (35%) of annual global externalities. Second, business can be a powerful force for good as the largest and most influential global institution in terms of employment size and technological capabilities (Hoffman, 2018). Yet, proactive corporate environmentalists are rather rare since "*it is not easy being green*" (Walley and Whitehead, 1994). So the assertion that "*the business of business is business*" (Friedman, 1970), although not socially desirable, is indeed closer to reality.

A normative approach to promoting voluntary green practices has proven ineffective (Buhr and Gray, 2012). Environmentalism, like other social movements, often has some of its central actors motivated by deeply ingrained beliefs that drive pronounced involvement and commitment (Gore, 1992, Etzion, 2007). It is unrealistic to expect such an ideology to become mainstream. Therefore, a pragmatic approach is necessary to align environmental practices with business' economic logic. This is the motivation behind the study of green competitiveness (GC), the synergy on the environmental-economic intersection (G-C) of corporate sustainability.

1.2 Business and Society

Sustainable development, or sustainability, is "*development that meets the needs of the present without compromising the ability of future generations to meet their own needs*" (The Brundtland Report, 1987, p. 37). In the private sector, corporate sustainability can be illustrated with a triple bottom line incorporating the economic, social, and environmental dimensions (Elkington, 1994). It is also closely associated with the concept of Corporate Social Responsibility (CSR), although they have different origins and focuses. The idea that business has societal obligations was evident as early as the 19th century (Smith, 2003) and research interests in CSR started to emerge in the US in the 1950s, traditionally focused on the social aspect such as worker welfare and philanthropy (Bowen, 1953). In contrast, discourses on sustainability originated in Europe, focusing on the environmental aspect initially (Brudtland Report, 1987). Over time, CSR and corporate sustainability have converged

into a comprehensive triple-dimensional concept and are now often referred to interchangeably, although the term CSR still has a nuanced normative connotation in contrast to the more positive posture of sustainability. This essay belongs to the camp of corporate sustainability. As shown in Figure 2.1, it is positioned at the environmental-economic intersection of corporate sustainability in the broader context of sustainable development.

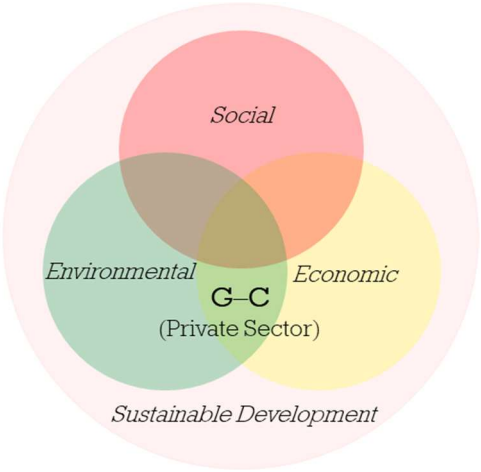

Figure 2.1 Research Positioning of Essay 2

The two opposing views on the business-society relation are represented respectively by Friedman and Freeman, their pioneering proponents. The Friedman view asserts that "*the social responsibility of business is to increase its profits*" (Friedman, 1970). This shareholder-centric perspective posits that pursuing a CSR agenda increases costs and contradicts corporations' obligation to create profits for shareholders. On the other hand, the Freeman view argues that a corporation has relationships with many constituent groups and/or individuals who can affect or be affected by its achievements (Freeman, 1984).

A stream of research adopts the instrumental stakeholder view (Donaldson and Preston, 1995) to study the effect of CSR performance (CSP) on corporate financial performance (FP). CSP has been defined as "*a business organization's configuration of principles of social responsibility, processes of social responsiveness, and policies, programs, and observable outcomes as they relate to the firm's societal relationships*" (Wood 1991, p. 693). Such a conceptualization of CSP as a meta-construct has been criticized by some scholars (Peloza, 2009), questioning it as "*(a) mishmash of uncorrelated variables*" (Rowley and Berman, 2000, p. 405). Without intending to step into the debate, research on CSP-FP will not be frequently referenced in this study.

1.3 Business and the Natural Environment

On the interface between business and the natural environment, a critical question that has been studied for decades is "*whether it pays to be green*" (Hoffman and Georg, 2012). Broadly

speaking, there are also two opposing views on this topic: the Traditionist perspective and the Revisionist perspective. The Traditionists assert that "*talk is cheap, environmental initiatives are not*" (Walley and Whitehead, 1994, p. 46). According to this trade-off view, stringent environmental regulations increase corporate costs and force companies to sacrifice their profits to be green. Conversely, the Revisionists posit that win-win opportunities exist on the environmental-economic interface. Notable leaders in this narrative shift include politician environmentalist Gore (Gore, 1992) and strategy scholar Porter (Porter, 1991a, Porter and van der Linde, 1995). The Porter Hypothesis postulates the concept of environment-competitiveness, proposing that stricter environmental regulations can motivate innovations that eventually overcompensate for investment costs and result in net benefits (Porter and van der Linde, 1995). It has stimulated and spurred the research stream on "*whether it pays to be green*". However, although this question "*has probably generated more research pages than any other single question, the answer remains (largely) unresolved*" (Hoffman & Bansal, 2012, p. 14) to the present day.

1.4 Review of Reviews

There are a number of studies that recount literature on the interface of business and the natural environment. For example, Etzion (2007) provided an overview of research on organizations and nature at organizational, sectoral, and societal levels. Ambec and Lanoie (2008) reviewed the conceptual arguments and presented supporting evidence of the channels where environmental performance (EP) can lead to better FP. Endrikat (2016) analyzed stock market reactions to corporate environmental events and found that the stock market generally reacted positively to positive events and negatively to negative events. Geng et al. (2017) reviewed studies on the effects of green supply chain management on different dimensions of corporate performance. These research syntheses differ in their review scope and construct conceptualization, hence do not provide a clear collective overview of the business-nature relation.

We review the systematic reviews that focus specifically on the environmental-economic nexus of corporate sustainability, including seven studies conducted between 2009 and 2018. The review identifies the knowledge front and research gaps in GC research as follows: (1). The common theme of these reviews is centered on "*whether it pays to be green*"; (2). They all find a positive, albeit weak, connection between EP and FP; (3). The empirical articles they covered are highly heterogeneous regarding construct operationalization, methodological artifacts, and study settings; and (4). The findings on moderating effects are inconclusive or even contradictory. The details from the review of reviews are reported in Table 2.1. Learnings from the review inform the study's objectives and research design.

Author(s) (Year)	Synthesis Method	Timespan (No. of Study)	Main Findings
Molina-Azorín et al. (2009)	Systematic Review	1995 - 2008 (32)	(1). Results mostly positive (21 out of 32); (2). Construct measures vary (14 used EM, 12 EP, and 6 both); (3). Regression the dominant analysis method (50%); (4). Study settings vary in terms of sectors & regions
Horváthová (2010)	Meta-Regression	Up to Feb 2009 (37)	(1). Positive results: qualitative > quantitative EP measures & common law (US/CA/UK) > civil law countries (EU); (2). Negative results: correlation analysis > econometric methods; (3). No difference between different FP measures
Dixon-Fowler et al. (2013)	Meta-Analysis	1970 - 2009 (39)	(1). EP-FP positive & significant (r=.06); (2). Moderation effects: small > big firms, US > Non-US firms, market-based FP measures > other FP measures; (3). No moderation effects: public vs. private firms, high-polluting vs. other sectors, reactive vs. proactive vs. hybrid environmental strategies, TRI vs. non-TRI EP data, self-reported vs. objective EP measures, concurrent vs. 1-year lag
Albertini (2013)	Meta-Analysis	1975 - 2011 (52)	(1). EP-FP positive & significant (r=.09); (2). Stronger effects with accounting-based (.24) than market-based (.03) FP measures; (3). Moderation effects: Non-US/Non-EU (r=.12) > US/CA (r=.08) > EU (r=.03), Cross-sectional (r=.13) > longitudinal (r=.07); (3). No moderation found with sector or timing of study (pre- vs. post-1995)
Golicic & Smith (2013)	Meta-Analysis	1990 - 2011 (31)	(1). GSCM-FP positive & significant; (2). No differences between market-based, operational, accounting-based FP measures; (3). Moderating effects: significant for region, sector, timing of study (pre-2003, 2003-2008, post-2008), no effect from firm size
Endrikat et al. (2014)	Meta-Analysis	1990 - 2013 (149)	(1). EP-FP positive & significant (r =.102); (2). Process-based EP: significant on accounting-based but no effect on market-based FP; (3). Outcome-based EP: significant on both accounting- & market-based FP (r=.096); (4). Reverse causality: accounting-based FP on process-based EP positive & significant, but no effect from market-based FP on process-based EP, and no effects from either accounting- or market-based FP on outcome-based EP; (5). Moderation effects: controlling for financial risk & possible endogeneity, single- or multiple sectors; (6). No moderation effects from firm size, industry, timing of study, or R&D/advertising/capital intensities
Hang et al. (2018)	Meta-Analysis	Up to 2015 (142)	(1). FP-EP: positive effect in 1 year, but the effect disappears after 1 year; (2). EP-FP: no short-term effect, but positive & significant effect after 1 year; (3). Bidirectionality and strength of causality contingent on study timeframe

Table 2.1 Review of Reviews on GC Research

1.5 Research Objectives

Based on the finding from the review of reviews that *it pays to be green in certain contexts*, this essay intends to push GC research forward by addressing the questions of *when/why it pays to be green*. When there is empirical evidence on the economic values of green practices, the next step following the contingency view is to understand in what contexts such synergies exist (Sousa and Voss, 2008). The question of "*when it pays to be green*" has been addressed with anecdotal data (Orsato, 2006). However, this research area up to date is still in its infancy (Albertini, 2013, Lucas and Noordewier, 2016). As for the explanatory question of "*why*", it remains a territory largely uncharted. Motivated by these questions, this essay's research objectives are as follows:

Research objective 1 (RO1): Identify the key contextual contingencies on the G-C nexus
Research objective 2 (RO2): Develop a preliminary theoretical model on GC

1.6 Definitions and Terminology

Environmental aspect (EA) is "*(the) element of an organization's activities or products or services that can interact with the environment*" (ISO 14001: 2015). **Environmental impact (EI)** is "(*the) change to the environment, whether adverse or beneficial… resulting from an organization's environmental aspects*" (ISO 14001: 2015). We define **environmental management (EM)** as *an organization's management on its EAs with the objective for EI reduction* and **environmental performance (EP)** as *the performance result of EM in terms of EI reduction*. Broadly speaking, the environment interacts with all organizations in a similar way: it provides them with valuable resource inputs ("source" on **EIA**, the **environmental input aspect**) and bears undesirable by-product outputs from economic activities ("sink" on **EOA**, the **environmental output aspect**). Specifically, although within-sector EAs are largely similar, cross-sector EAs can vary considerably (e.g., EIA of resource extraction vs. EOA of chemical manufacturing).

Business performance (BP) is conceptualized as *a company's interim or eventual achievements on the economic dimension that affects its financial results*. Such a broad conceptualization of BP enables us to capture the environmental-economic interactions at different stages along a company's performance causal chain. **Green competitiveness (GC)** is defined as *the superior competitive position a company obtains from its EM that enables it to achieve better BP relative to its rivals*.

The term, *GC*, may be used in two self-explanatory contexts, referring specifically to the intangible economic value created by EM or generally to academic research on the environmental-economic dyad. *G-C* refers to environmental-economic interactions at a more granular

subdimensional level. *EI* refers only to negative EI, considering that economic activities' primary effects on nature are unfavorable. *Hypothesis* and *proposition* are used interchangeably in this study.

1.7 Conceptual Model

We adapt the structure of the Elman neural network (Elman, 1990) to conceptualize the environmental-economic interactions of corporate sustainability. GC, by definition, is the intermediary between the EM of a company (the *Input Layer*) and its materialized BP (the *Output Layer*). A company's competitive advantages are mainly intangible, GC is therefore the *Hidden Layer* in the neural network. The context a company operates in affects the magnitude of its output given a certain level of input, GC is hence moderated by the *Context Layer* in the neural network[1]. Since EM and BP are both multidimensional meta-constructs, it is plausible that the connections between them are at the subdimensional level – it can then be inferred accordingly that the latent construct of GC also has multiple dimensions as the mediators between the EM and BP subdimensions. Figure 2.2 illustrates such a moderated-mediated G-C neural network as the conceptual model of this study: (1). RO1, *Identify the key contextual contingencies on the G-C nexus*, aims to delineate the *Context Layer*; and (2). RO2, *Develop a preliminary theoretical model on GC*, aims to explicate the *Hidden Layer* and elucidate the entire G-C neural network.

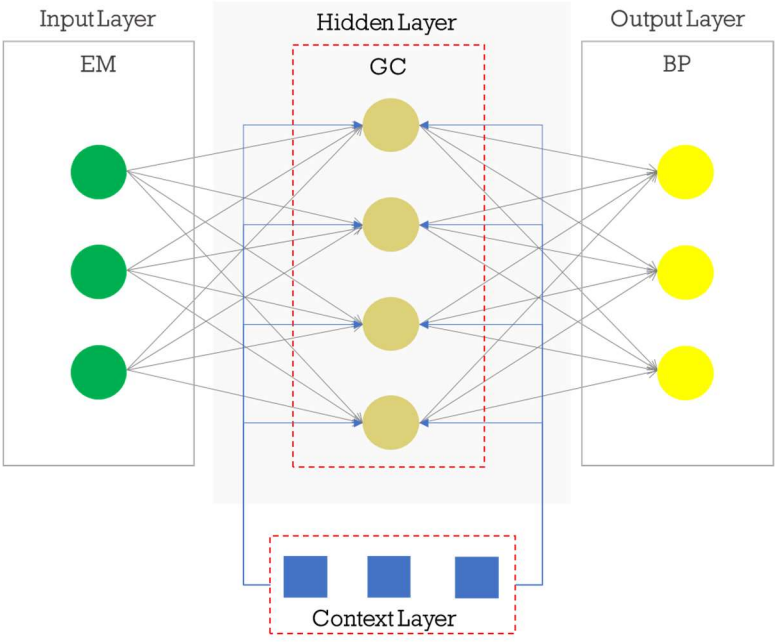

Figure 2.2 Conceptual Model – The G-C Neural Network

[1] The Context Layer in the neural network is not conceptualized as recurrent with bidirectional interactions but unidirectional from the contexts to an organization, since the reverse influence flow from one single average company to its external context can be rather insignificant, hence neglectable, from the individual company's perspective. Either way, such potential reverse flow is beyond the scope of this study.

2. Methodology

2.1 Mixed-Methods Design

- **Intermediate Theory Building**

Research in a field typically develops from description to explanation to testing (Meredith, 1993). Following such a trajectory, management theories evolve through three stages from nascent to intermediate to mature: (1). Nascent theories observe new phenomena and develop new concepts; (2). Intermediate theories propose provisional explanations of the new phenomena, often introducing new constructs and linking them with established ones; (3). Mature theories provide well-defined constructs and models that have been studied over time with broad agreements among researchers (Edmondson and McManus, 2007).

For nascent theories, research questions are exploratory and open-ended, qualitative data are collected, and content analysis is used to code data for evidence identification and sensemaking (Edmondson and McManus, 2007). For intermediate theories, research questions are about the relationship between a proposed new concept and established constructs, hybrid qualitative and quantitative data are collected, and the purpose of data analysis is preliminary building or exploratory testing of new propositions (Edmondson and McManus, 2007). For mature theories, hypotheses are generated, and inferential statistics applied to test the hypotheses within specified confidence levels to decide on their rejection or acceptance (Flynn et al., 1990).

Although the GC phenomenon has been examined in various specific settings for decades, as a collective body of work at the knowledge cumulation level, it is still a nascent theory considering that currently there are still no shared definitions and common measurements for the focal constructs (Meredith, 1993, Trumpp et al., 2015). This essay aims to develop GC research from a nascent to an intermediate theory.

The foundations for a theory-building study are not well-specified hypotheses, but rather some assumptions or tentative suppositions (Flynn et al., 1990); and data are used not for theory validation and refinement but to ground and strengthen the preliminary theory in formulation (Glaser and Strauss, 1967, Yin, 1994). The essay's underpinning supposition is the finding from the review of reviews that *it pays to be green in certain contexts* – which can be viewed from the theoretical lens of the Porter Hypothesis (Porter and van der Linde, 1995) and the contingency theory (Donaldson, 2001). The research questions that flow naturally from the supposition and the research objectives specified accordingly are (a) the *exploratory* question, *when it pays to be green*, addressed by RO1, *Identify*

the key contextual contingencies on the G-C nexus; and (b) the *explanatory* question, *why it pays to be green*, addressed by RO2, *Develop a preliminary theoretical model on GC*.

- **Mixed-Methods Design**

Mixed-methods design is necessary for this essay as it seeks to answer both exploratory and explanatory questions. Additionally, such a design is also more powerful – by allowing for versatility while assuring methodological rigor, it enables incremental learning and synthesis that can lead to a richer and deeper understanding of the complex phenomenon under study. Given that the multidimensional G-C interactions may vary in different contexts, it is unlikely that a narrowly defined dataset for a specific group of companies can cover the breadth and richness of the topic. Therefore, the population of published academic GC articles will be used as the data source for this study. Figure 2.3 and Table 2.2 summarize the multi-phased mixed-methods design. The details of each phase are delineated as follows.

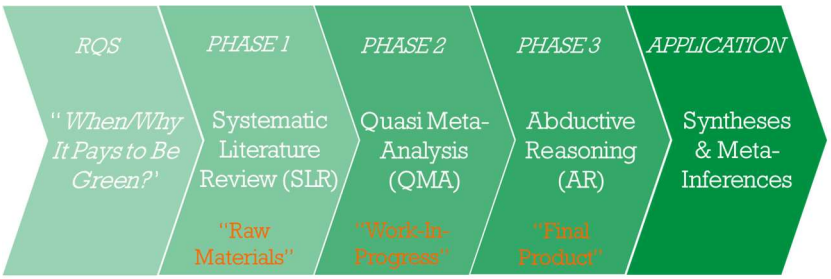

Figure 2.3 Multi-phased Mixed-methods Design

Theoretical Foundation	Porter Hypothesis and Contingency Theory; Supposition: It pays to be green in certain contexts
Research Objectives	1. Identify key contextual contingencies on the G-C nexus; 2. Develop a preliminary GC theoretical model
Research Design	Multi-phased mixed-methods design
Phase 1_Data Collection	Systematic Literature Review on published academic articles on the G-C nexus
Phase 2_Data Analysis	Quasi Meta-Analysis adapting meta-analysis & content analysis techniques
Phase 3_Theory Building	Abductive reasoning from descriptions to preliminary explanatory propositions

Table 2.2 Multi-phased Mixed-methods Design

2.2 Phase 1 – Systematic Literature Review for Data Collection

Systematic literature review (SLR) in Phase 1 identifies and collects the comprehensive dataset of the population of academic GC studies. This phase extracts *raw materials* for the grounding and development of an evidence-based GC theoretical model. The method of SLR differs from traditional narrative literature review in that it uses a structured process to increase transparency, ensure replicability, and reduce biases through exhaustive literature search and rigorous literature selection

in a field of study (Tranfield et al., 2003). Phase 1 takes eight steps in three stages: (a) database search, (b) relevance screening, and (c) quality screening.

2.2.1 Database Search

Web of Science (WoS) is chosen as the database for literature search for its multidisciplinary coverage of over 12,000 academic journals (Reuters, 2017). Database search strings are developed with the ''ancestry'' approach and the ''snowballing'' technique, described as follows. The guideline here is to develop search strings that balance both recall and precision in information retrieval (Van Rijsbergen, 1986) – higher recall identifies results as comprehensive as possible, while higher precision reduces false positives as far as necessary in order for the searched results to stay focused and relevant.

1) **Identify keywords**: By thoroughly combing through the search strings and cited references of the systematic reviews relevant to GC (cited in Section 1.4), we identified and retrieved two categories of tentative keywords (noun words and phrases, i.e., *terms*) relevant to *green* (EA, EM, EI, and EP) and *competitiveness* (BP and GC).
2) **Create search strings**: We then created another category for *linkage* terms such as ''affect'', ''contribute'', and ''impact'' (developed extensively based on ancestry citations and thesauruses). The scope for the category of *linkage* was set as ''Topic'' in order to also search in abstracts and keywords beyond article titles, and the scope for the categories of *green* and *competitiveness* was set as ''Title''. The search strings were created by connecting first (a) within-category terms with the Boolean operator ''OR'', and then (b) between-category terms with ''AND'' (i.e., ((green term 1 OR green term 2 OR…) in ''Title'') AND ((linkage term 1 OR linkage term 2 OR…) in ''Topic'') AND ((competitiveness term 1 OR competitiveness term 2…) in ''Title'')).
3) **Refine search strings:** After that, we refined the terms in the search strings using the wildcard of asterisk (''*'') (e.g., ''eco-efficiency'' and ''eco-innovation'' were both replaced by ''eco*''), reducing redundant terms and making the strings precise and concise.
4) **Check for robustness**: We then checked the robustness of the search strings by going through all the cited references in ''*The Oxford Handbook of Business and the Natural Environment*'' (Bansal and Hoffman, 2012). When coming across a relevant citation, the search strings were applied virtually to test the recall. None of the relevant studies had slipped through the virtual recall test, the search strings were hence considered sufficiently robust to be finalized.
5) **Search in database**: We conducted an ''Advanced Search'' in WoS with the search strings and other relevant settings (including published articles in peer-reviewed academic journals in

English from 1970 till January 2019, as detailed in Table 2.3)[2]. After eliminating the recalled results outside the scope of business and management by filtering through the WoS categories, the database search identified 1,847 articles that were potentially relevant to GC research.

ITEM	SEARCH SETTING
Database	Web of Science Core Collection (1970-present)
Search strings	Green terms: (eco* OR environment* OR green* OR "ISO 1400*" OR carbon OR emission* OR pollut* OR social OR sustainabl* OR responsib*) in Title,
	Linkage terms: (affect* OR associat* OR barrier* OR benefit* OR "business case" OR causal* OR connect* OR consequence* OR conserv* OR contribut* OR correlat* OR determin* OR driv* OR effect* OR evidence* OR impl* OR impact* OR improve* OR influence* OR interact* OR link* OR outcome* OR outperform* OR overcom* OR prevent* OR promot* OR reduc* OR relat* OR return* OR reward*) in Topic, AND
	Competitiveness terms: (advantage* OR *competitive* OR *effective* OR *efficien* OR cost* OR earning* OR financ* OR growth OR investor* OR market OR outcome OR pay* OR performance OR productivit* OR profitab* OR quality OR return OR revenue* OR risk OR share* OR stock* OR success*
Timespan	1970 - Present (Jan 2019)
Research areas	Business and management
Language	English
Document type	Articles

Table 2.3 SLR Search Setting

2.2.2 Relevancy Screening

The recalled and remained articles were then screened for relevancy and conceptual consistency. An article needs to meet two criteria to be included: (1). It studies the G-C nexus, conceptualizing it as either a correlative or causal relation, regardless of the co-variation direction (excluding CSR-BP studies); (2). Its unit of analysis needs to be set at the organizational level (excluding sectoral or regional level studies).

6) **Screen articles for relevancy**: Following these relevancy criteria, we screened the titles and abstracts, as well as full texts when needed, of the 1,847 articles. Those irrelevant but recalled articles were eliminated (e.g., articles using the term ''environment'' to refer to companies' business contexts). Articles that fulfill the inclusion criteria were saved in the ''Marked List'' in WoS.
7) **Download articles**: After all the articles went through the process, those saved in the ''Marked List'' were then reviewed again to ensure the reliability of relevancy screening. This resulted in a list of 230 studies that evidently conform to the relevancy inclusion criteria. These articles were then downloaded and saved as PDF files.

[2] Some articles identified via the ''snowballing'' method had to be manually collected from Google Scholar as they were published earlier than WoS indexing those journals, including articles in BSE prior to 2009, CSR&EM prior to 2008, EE prior to 1993, JBE prior to 1982, and JCP prior to 2002).

2.2.3 Quality Screening

We referred to the Academic Journal Guide 2018 of the Chartered Association of Business Schools (Chartered Association of Business Schools, 2018) to screen articles for quality assurance. Other systematic literature reviews have also used the Guide for quality control purposes (Geng et al., 2017, Wilding et al., 2012).

8) **Screen articles for quality**: The quality of the downloaded articles was evaluated according to the Guide's ratings of the journals. Those publications in journals rated less than "2" were excluded, and the remaining 171 articles published in journals rated from "2" to "4*" were retained and included in the study. The results of SLR are summarized in Figure 2.4.

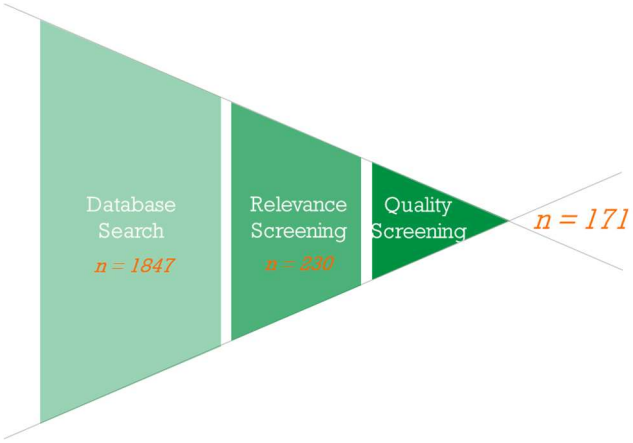

Figure 2.4 Results of Systematic Literature Review

2.3 Phase 2 – Quasi Meta-Analysis for Data Analysis

2.3.1 Design Rationale

Phase 2 addresses the research question, *when it pays to be green*. It codes and analyzes cumulated evidence from the collection of GC studies for pattern detection and inductive generalization. This phase supplies *work-in-progress* to Phase 3 for further processing into the *final product* of a preliminary GC theoretical model. This is achieved by developing and applying a new method, quasi meta-analysis (QMA), with the adaptation and integration of traditional meta-analysis and content analysis techniques. The design rationale for QMA is elaborated as follows.

The review of reviews in Section 1.4 finds that GC studies are highly heterogenous in construct operationalization, methodological design, and study settings. Such heterogeneity in research characteristics can moderate research outcomes through methodological or contextual factors. We

hence hypothesize that a significant portion of the variance in the empirical GC research findings can be attributed to three categories of research characteristics:

1) **Construct operationalization**: How the multifaceted meta-constructs of EM and BP are operationalized can affect the findings on the existence and direction of G-C co-variations. Since different G-C subdimensional relations may require different theoretical explanations, the collection of GC research needs to be examined separately to disentangle and distinguish the potentially contradictory value generation mechanisms;
2) **Methodological artifacts**[3]: The methodological factors (such as data types and time-lags between focal constructs) may exert *artificial* moderating effects on the G-C relations. Such *artificial* methodological moderators need to be controlled to enable the identification of *authentic* moderating effects from contextual contingencies;
3) **Contextual contingencies**: The study contexts (such as sectors and regions) can affect the research findings and reveal *authentic*, i.e., naturally present and real, moderating effects on the G-C nexus. Such *authentic* moderators can answer *when it pays to be green* and interest RO1 in this study.

Meta-analysis as a research synthesis method can identify moderating effects from study characteristics that may not be obvious in individual empirical studies (Lipsey and Wilson, 2001). Therefore, it may help unravel the *authentic* contextual moderators. However, traditional meta-analysis is built on the assumption of study homogeneity – it therefore needs to be adapted to account for the heterogeneity in GC research. The adaptation is illustrated in the QMA conceptual framework in Figure 2.5.

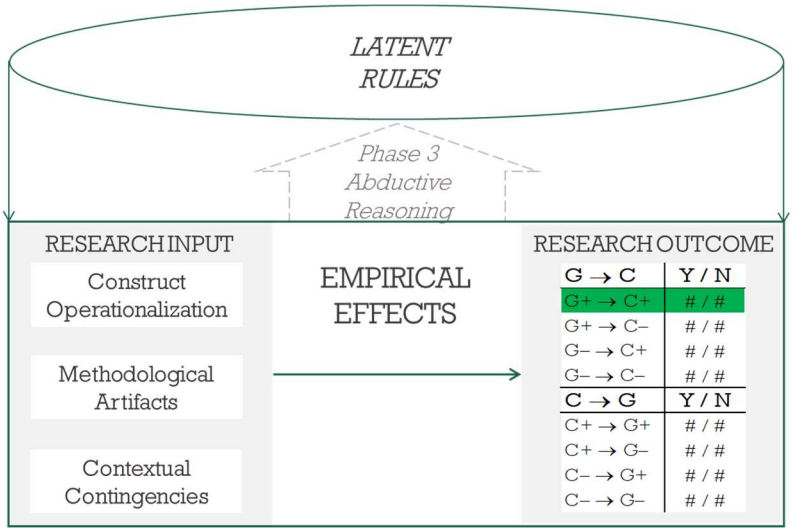

Figure 2.5 QMA Conceptual Framework

[3] An artifact is "*something observed in a scientific investigation or experiment that is not naturally present but occurs as a result of the preparative or investigative procedure*" (Oxford Dictionary).

The QMA conceptual framework posits that there are *latent rules* behind the manifest *empirical effects* in the collection of GC research. These latent rules are the *theoretical causes* that will be abductively inferred in Phase 3. The empirical effects are the associations between *research inputs* and *research outcomes*. Research inputs include three categories of study characteristics: (a) construct operationalization, (b) methodological artifacts, and (c) contextual contingencies. Research outcomes are the findings on the existence and direction of G-C co-variations, including by enumerating eight types of co-variations with either positive, negative, or neutral findings. The primary interest of this study is the positive causal G-led/C-lagged relation (as highlighted in the above framework); the other types are included to enrich the understanding of the topic and serve as counterfactuals in the theory formulation process.

Given the level of study heterogeneity, these latent *theoretical causes* behind the manifest *empirical effects* are not directly discernable from the collection of GC research en masse. But they may be inferred if the body of work is segmented into more homogeneous subgroups to enable evidence cumulation and pattern detection – this is the rationale of QMA. In sum, QMA is tasked to code and categorize GC studies by *research inputs* and *research outcomes* as *work-in-progress* to enable reasoning from *empirical effects* to *theoretical causes* through abduction in the next phase.

2.3.2 Design Elaboration

- **Meta-Analysis and its Fit for this Study**

Meta-analysis "*deals with the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings*" (Hartung et al., 2011, p. xiii). A meta-analysis aims to synthesize and quantify the intervention effects by following a systematic process to identify, select, collect, correct for sampling errors when applicable, and pool and analyze the weighted mean and distribution of the effect sizes (Lipsey and Wilson, 2001). A well-conducted meta-analysis can provide more accurate parameter estimates since it is evidence-based, disciplined, and transparent – with all these features combined to reduce errors and biases that might exist in individual studies (Littell et al., 2008).

However, traditional meta-analysis needs to be adapted before it is applied in this study. The primary reason is that it is generally risky to reduce an individual empirical study into a single statistic and pool such values from multiple studies together into a weighted average: (1). A lot of valuable information from the original studies is removed as if irrelevant; (2). The exchangeability assumption that the study subjects and contexts are similar enough for the results to be aggregated may not be justified (Walker et al., 2008, Cleophas and Zwinderman, 2017).

Such risks are particularly pertinent to GC research. First, with its roots in medicine, traditional meta-analysis is built on the assumption that the intervention effects can be precisely measured, such as that in randomized controlled trials – this is the rationale behind the singular significance placed on a study's effect size. However, such an assumption is often violated in business and management research. Management studies are usually conducted in noisy contexts beyond the researchers' control, and the ''dose-response'' effect is more a metaphor than a reality: (1). The intervention effect of management practices (in this case, the marginal change in GC and BP caused by the implementation of EM) usually cannot be precisely measured and quantified as an effect size; and (2). It is often a challenge to rule out omitting unobserved variables. Therefore, traditional meta-analysis in its original form is only applicable to a limited subset of management studies – e.g., those conducted with lab experiments – and strictly forcing it upon other areas may lead to invalid findings.

Second, the collection of GC studies violates the assumption of exchangeability. GC research is highly heterogeneous involving meta-constructs operationalized in different measures and investigated in various settings. From a theoretical standpoint, this research area is still at its nascency without commonly accepted focal construct definitions and measures. Furthermore, it is difficult, if at all possible, at this stage to define a standard effect size, i.e., the quantitative economic impact of EM, applicable to all the GC studies. It will most likely raise validity concerns if the artificial ''effect sizes'' from GC studies are pooled together for statistical analysis and synthesis despite the research context of study heterogeneity and theoretical nascency. Hence, traditional meta-analysis needs to be adapted before its application to fit the features of GC research.

- The Adaptation of Meta-Analysis

The primary adaptation of meta-analysis is to analyze the *effects* instead of the *effect sizes* in GC studies. The task of Phase 2 is to code and categorize GC research for the detection of hidden patterns behind inconsistent research effects. As the first question of this study is *when it pays to be green*, the primary classifier to categorize the dataset is the direction (*does it pay to be green*?), not the magnitude (*how much does it pay to be green*?), of G-C co-variation in the research findings. Hence, the research outcome of interest is not a precise *effect size* (the quantitative magnitude) but rather the *effect* (the qualitative direction) of G-C co-variation. Using *effect size* would not be appropriate here as pooling positive and negative results altogether would just let them offset each other and confound the aggregated findings. The ''vote-counting'' method in meta-analysis – referring to the collection and cumulation of positive, negative, or non-significant results (Hunter and Schmidt, 1990) – also focuses on *effect* instead of *effect size* (QMA in this study is more granular and powerful than this method since it also applies content analysis techniques). The ''vote-counting'' approach has been criticized for reasons such as ignoring sample size differences, lacking a point estimate of pooled effect size, and compromising statistical power with increased number of studies included

(Hunter and Schmidt, 1990, Orlitzky et al., 2003). However, such criticisms are justifiable only if the measurement of *effect size* is not an issue – if the raw data in management studies are already crude to start with, it probably makes more sense to focus on clarity rather than precision in the analysis process.

The secondary adaptation of meta-analysis is to focus more on quality control in the data collection and cleansing processes before pooling and analyzing the end results. As a piece of general advice for meta-analysis, rather than concerning about correcting for sampling errors and pooling weighted effect sizes which is "*almost a breeze*" if the foundational work has been done well, the focus should instead be placed on "*the hard work in meta-analysis… (i.e.) the validly collecting, digesting, and rating of often very heterogeneous studies*" (Barendregt and Doi, 2016, p. 7). Such advice is particularly relevant to GC research given its theoretical nascency and study heterogeneity. Therefore, SLR is first conducted in Phase 1 for rigorous data collection and cleansing before the techniques of content analysis – a research method used to analyze textual materials in order to make valid and replicable inferences (Franzosi, 2008) – are brought in to integrate into adapted meta-analysis for data analysis in Phase 2. From the perspective of content analysis, QMA in Phase 2 is similar to data coding for a multi-case study, with 171 *empirical cases*, in that each individual study is taken as an empirical *case* and coded for its key study characteristics. The coding results can then be categorized for within- and cross-group "vote-counting" meta-analysis in order to surface the latent patterns behind the inconsistent research effects.

2.3.3 Design Implementation

The QSR NVivo 12 Plus software was used for QMA. Beyond traditional code-and-retrieve functions, the software also supports versatile code shuffling and combination – which is instrumental for pattern detection – and flexible customization of report templates (QSR International, 2020). From WoS, the 171 GC articles saved in the "Marked List" were downloaded in PDF format and imported into NVivo. Each article was then coded according to the coding scheme shown in Table 2.4.

CODE	CODE SCALE
Research Outcome	
Hypothesis	G→C (++, +-, -+, --), C→G (++, +-, -+, --), Others (code as is)
Result	Significant, Insignificant
Construct Operationalization	
G Measures	(code as is first & categorize/organize with QMA framework along the process)
Data Type	Physical, financial, evaluative, communicative, perceptual, anecdotal
C Measures	(code as is first & categorize/organize with QMA framework along the process)
Data Type	Financial, evaluative, perceptual, anecdotal
Contextual Contingencies	
Sector	High-polluting sectors, manufacturing, manufacturing & services, services
Region	*(code as is first and categorize & organize along the process)*
Size	Public, large, SMEs, varied
Methodological Artifacts	
Data Timeframe	Longitudinal, cross-sectional
Data Timelag	Concurrent, 1Y, 2Y, 3Y+, event window (code as is)
Data Age	(Code as is if reported, otherwise code the latest citation year)

Table 2.4 QMA Coding Scheme

For each independent GC study, only one *effect* – from the most conceptually relevant hypothesis – and its related research characteristics were included for analysis to avoid non-independent sample bias. Therefore, (1). When an article included a main and several more granular hypotheses, only the *effect* from the main hypothesis was included; (2). When a structured model included several hypotheses, the most theoretically relevant path was included; and (3). When mix-methods were used in an article, only the *effect* and the characteristics of the quantitative method were included.

Coding was done twice in reverse chronological order: first actual coding, then cross-checking for accuracy and completeness. Some minor modifications were made to the results in cross-checking, mostly to re-organize the construct measures into a more intuitive and coherent structure following the logical flows of business processes and performance results. After that, we re-organized the codes (''Nodes'' in NVivo) into a one-layer flat structure, redesigned the template for ''Coding summary by File Report'' in ''Report Designer'', ran the report, and exported the coding results into an Excel file. The retrieved data were first cleansed and then analyzed using Pivot tables, with different data slicing by the key variables in construct operationalization, methodological artifacts, and contextual contingencies. The QMA results are reported in Section 4.

2.4 Phase 3 – Abductive Reasoning for Theory Building

2.4.1 Design Rationale

Phase 3 addresses the research question, *why it pays to be green*. It applies abductive reasoning (AR) to develop a preliminary GC theoretical model based on the generalized descriptive

findings from the QMA results. Abduction, or inference to the best explanation, is a method of theory formulation that goes from data describing something to hypotheses that best explain the findings (Josephson and Josephson, 1996). Abduction is chosen as the theorizing method for the following reasons:

1) First, it fits well methodologically. Phase 3 aims to identify the *theoretical causes* that explain the *empirical effects* of inconsistencies in GC research, and abductive theorizing ''reason(s) from effects to causes'' – the method fits well the purpose;
2) Second, it is a rigorous theorizing method. AR follows a structured and well-documented process of hypothesis generation, hypothesis evaluation, and hypothesis acceptance. By making explicit the places where inferences might go wrong, abductively developed theories can be verifiable, replicable, and falsifiable;
3) Third, abduction can develop more generalizable propositions. The primary power of abduction lies in its broad coverage of candidate propositions (Josephson and Josephson, 1996). If such propositions are built on (a) a systematically collected and rigorously analyzed dataset and (b) a holistic conceptual foundation, the derived propositions can be much more generalizable compared with theories developed from anecdotal ''success stories'' or ''best practices'' (Flynn et al., 1990). This again explains the rationales of the study's mixed-methods design: SLR aims for exhaustive data collection and QMA for rigorous data preparation – they are designed to build a broad and solid foundation to enable data-driven theorizing of generalizable propositions through abduction.

Furthermore, as a counterfactual argument, AR is likely the only feasible option for GC theorizing. Induction alone with the grounded theory approach (Glaser and Strauss, 1967) cannot develop generalizable propositions for this topic, since (1). It is unrealistic to collect primary data that fully cover different aspects of the multifaceted construct of GC; and (2). The secondary data of existing literature most likely have not yet explored all dimensions of the G-C nexus in different empirical settings. Deduction is not suitable either, because (1). GC theory needs to be built first before hypotheses can be developed and tested; and (2). Existing management theories have mostly been ''virtually silent'' on the idiosyncratic features of the organization-natural environment dyad (Starik and Kanashiro, 2013). This leaves abduction the only viable option to theorize the G-C relation.

2.4.2 Design Elaboration

The force of abduction is limited by its coverage of candidate hypotheses, which must be broad enough to include the true one (Josephson and Josephson, 1996). Given that the G-C subdimensional links are potentially moderated by a variety of contextual contingencies, enumerating all the combinations would lead to a set of candidate hypotheses that are exhaustive but too broad to

manage. However, enumerating all the possible combinations is neither practical nor necessary since not all G-C subdimensional connections have probable or strong theoretical underpinnings: e.g., it is rather implausible that compliance-driven pollution reduction would lead to price premium in the product market. We hence adapt the logic model (Ebrahim and Rangan, 2014) into a G-C logic chain (Figure 2.6) and use it as the abduction framework to support the generation of a sufficiently broad yet reasonably succinct set of candidate hypotheses.

The G-C logic chain first breaks down on the vertical axis EM and its BP implications into cost and revenue flows, then organizes them along the horizontal axis into Input (what goes in), Process (what happens), Output (immediate results), Outcome (mid- to long-term results), and Impact (long-term effects on root causes). Conceptually, it covers the EMs and their BP effects at different stages throughout an organization's performance causal chain, thereby enabling the inference and generation of extensive yet parsimonious candidate propositions. The G-C logic chain can be used to map QMA results to ground hypothesis generation, and to guild the derivation of theoretically plausible propositions when empirical evidence is absent (the existing literature might not have covered all relevant aspects of the topic). As a result, AR guided by the G-C logic chain can minimize the threat of publication bias – in that only positive and statistically significant results are reported and published while those with null or negative results are left in file drawers (Dickersin, 2005) – that might present in the dataset of GC research.

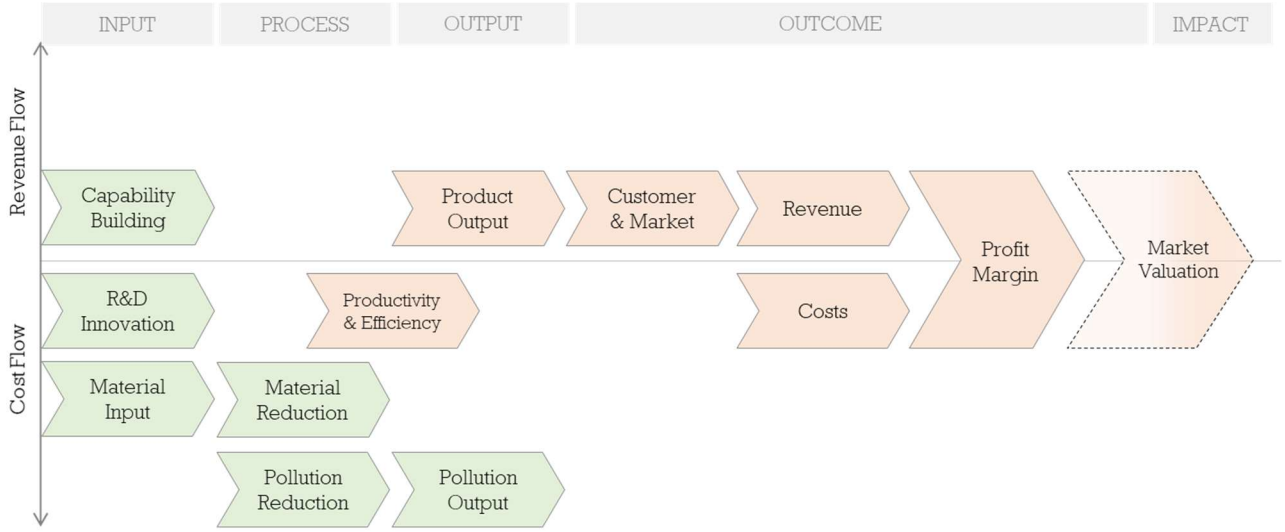

Figure 2.6 Abduction with the G-C Logic Chain

2.4.3 Design Implementation

The AR process goes like this: (1). D is a collection of data; (2). F is the findings from D; (3). H explains F, and among all the hypotheses, no others can explain F as well as H does – therefore, H is probably true (Josephson and Josephson, 1996). D (data) in this study is the *raw materials* collected

via SLR, F (findings) the *work-in-progress* from QMA, and H (hypotheses) the *final product* of a preliminary GC theoretical model, developed through a four-step process as follows:

1) **Hypothesis Generation** (Section 5.1): This step developed the main hypotheses and unveiled translucently the *Hidden Layer* in the G-C neural network. We first mapped the key QMA results on the G-C logic chain to enable a structured assessment of EM's BP effects. When there was *strong* empirical evidence on *specific* measures with *clear* causal links, the latent rules behind the manifest empirical effects were developed as tentative propositions. Where findings were based on high-level composite measures, plausible propositions were generated and evaluated, and the best possible explanations were retained while implausible ones were eliminated.
2) **Hypothesis Refinement** (Section 5.2): This step further refined the main propositions into moderated-mediated hypotheses to unravel the *Context Layer* in the G-C neural network. We delved into the QMA results to distill four essential contextual moderators. The tentative propositions were then refined by adding in empirically evidenced and theoretically plausible contextual contingencies. These refined hypotheses were synthesized into a preliminary GC theoretical model following the structure of the G-C neural network.
3) **Hypothesis Evaluation** (Section 5.3): We then developed a holistic organizational performance framework (HOPF) to evaluate the plausibility of the GC theoretical model. We also contrasted and reconciled it with existing relevant theories to prove its validity.
4) **Hypothesis Acceptance** (Section 5.4): Finally, the proposed GC theoretical model was accepted as a preliminary answer to *why it pays to be green*. The confidence level in the model is relatively high, although the hypotheses are still provisional and need to be empirically tested in further research.

The key processes in the mixed-methods design are summarized in Figure 2.7:

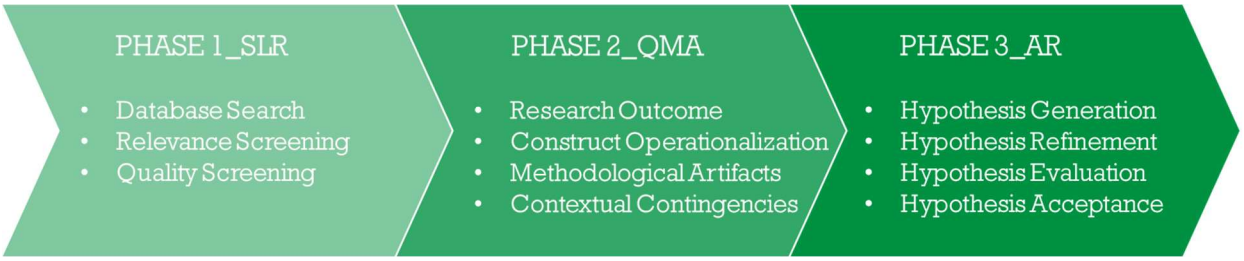

Figure 2.7 Research Process in the Mixed-methods Design

2.5 Methodological Rigor

The following practices assure the methodological rigor of this study:

1) **Reliability**: Both SLR and QMA were done twice following structured and well-documented protocols to ensure verifiability, replicability, and inter-coder reliability. The process of abductive theorizing in Phase 3 was also methodically documented, with empirical evidence and logical inferences presented alongside the propositions-in-formulation to enable falsifiability and replicability.
2) **Construct validity**: The focal constructs relevant to green are defined according to the ISO 14001 standards for clarity and consistency with clearly specified scopes and boundaries.
3) **Internal validity**: The establishment of causal relationships between the focal constructs is underpinned by both empirical evidence and theoretical inferences.
4) **External validity**: The theoretical model is grounded on the population of academic GC articles covering different industrial sectors and management fields, so the confidence level in the inductively developed hypotheses is comparable to generalizing from inferential statistical analysis based on large-sized samples. The threat to generalizability from potential publication bias is mitigated by abduction that reasons to best possible propositions beyond immediately observable evidence.

3. Systematic Literature Review Results

3.1 By Year and Journal

Table 2.5 presents the summary of SLR results by publication year and journal. It shows that research interest in GC has been growing, especially in the past decade. Academic journals that cover the topic spread across different disciplines, including management science (MS), operations management (JOM), strategic management (AMJ & SMJ), accounting and finance (AR), engineering (IEEE), ecological economics (EE), and interdisciplinary journals dedicated to sustainability research (BSE, CSR&EM, and O&E). A detailed list of the 171 GC articles and their research characteristics is available in Appendix I.

		BSE	JBE	AMJ	IJOPM	IJPE	CSR&EM	JBR	JCP	JOM	MD	O&E	AR	EE	EJOR	IEEE	JAMS	MS	OMEGA	SMJ	TM	Other (33)
Total	*171*	*35*	*16*	*6*	*6*	*6*	*5*	*5*	*5*	*5*	*5*	*5*	*4*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*44*
2019	2															1					1	
2018	13	5	2				2	1			1											2
2017	18	9	3			1					2						1					2
2016	12	2	1		1	3	1				1	1			1							1
2015	19	2	4			1		1	1		1	3									2	4
2014	11	4	1		1				1				1									3
2013	13	3		1	1							1					1					6
2012	11	2						1						1					1			6
2011	10		1		1				2					1								5
2010	5	1	1						1	1					1							
2009	4	2	1																			1
2008	7		1			1	1													1		3
2007	3	1								1						1						
2006	5		1					1								1						2
2005	7				2			1									1		1			2
2004	4	1		1						1			1									
2003	2									1												1
2002	2						1											1				
2001	4	1													1				1			1
2000	5			1						1			1					1				1
1999	2			1																		1
1998	3																			1		2
1997	2	1		1																		
1996	2	1																1				
1995	1																			1		
1993	1													1								
1980	1												1									
1978	1																					1
1975	1			1																		

Table 2.5 By Year and Journal

3.2 By Hypothesis and Finding

Table 2.6 summarizes the SLR results by hypothesis and research outcome. A majority (76%, 130 out of 171) of studies on the G-C nexus hypothesized a positive G-led/C-lagged causal relation (those articles that only hypothesized the causal link without specifying the direction of co-variation ex-ante were included based on ex-post findings). Within this category, the overwhelming majority (90%, 117 out of 130) found a positive G-C causal link. This may suggest that *it indeed pays to be green*, at least as reported by the published GC articles in those specific circumstances.

HYPOTHESIS	RESEARCH OUTCOME		SUB-TOTAL
	Y	N	
G → C (G-led/C-lagged)			
G+ → C+ (4A_Green Reward)	117	13	130
G+ → C– (3A_Green Liability)	15	-	15
G– → C– (2A_Grey Penalty)	7	-	7
G– → C+ (1A_Grey Exploitation)	2	-	2
C → G (C-led/G-lagged)			
C+ → G+ (4B_Slack Resources)**	4	-	4
C– → G+ (3B_Resource Diversion)**	-	-	-
C– → G– (2B_(Resource Constraints)**	-	-	-
C+ → G– (1B_Irresponsible Growth)**	1	-	1
Others_Bi-Directional			
1A & 2A *	1	-	1
2A & 4A *	6	-	6
3A & 4A *	1	-	1
4A & 4B **	1	-	1
Others_Association			
0_GC_Positive Association **	1	-	1
0_GC_Neutral **	2	-	2
TOTAL	**158**	**13**	**171**

Table 2.6 By Hypothesis and Finding

Since this study is primarily interested in the G-led/C-lagged causal relation (1A to 4A), the articles with C-led/G-lagged causal (1B to 4B) or associative (indicated by "**" in Table 2.6) relations are out-of-scope and excluded from further analysis. Those articles with bidirectional findings (indicated by *) are grouped into 1A, 2A, and 3A to offset the potential publication bias. Figure 2.8 visualizes the cleansed results of the 162 articles in a G-C matrix. The four quadrants in the matrix are not balanced in size, with the category of 4A – the positive G-led/C-lagged causal relation – accounts for 80% of the studies (this reflects the research interest in finding positive G-C causation, although the other quadrants might represent the majority in reality).

Figure 2.8 SLR Results By G-C Matrix

4. Quasi Meta-Analysis Results

4.1 By Construct Operationalization

Table 2.7 summarizes the QMA results by key construct measures. The results bring to light the inconsistencies of construct operationalization in GC research.

			C1	C2	C3	C4	C5	C6	C7	C8	C9
Construct Operationalization	Sub-Total	%	Productivity & Efficiency	Customer & Market	Costs	Revenue	Profitability	Market Valuation	Composite Accounting	Composite Accounting & Market	Overall & Others (NEC)
Sub-Total			10	6	4	2	27	27	10	24	52
%			*6%*	*4%*	*2%*	*1%*	*17%*	*17%*	*6%*	*15%*	*32%*
G1_Green Investments	5	*3%*	1				2	1		1	
G2_Green Innovation	11	*7%*	1				3	1			6
G3_Product Design	1	*1%*		1							
G4_Pollution Reduction Practices	3	*2%*						3			
G5_Pollution Output	30	*19%*	1				7	5	1	13	3
G6_Environmental Certification	11	*7%*		1	1		1	1	2	1	4
G7_Green SCM	15	*9%*	2		2		1	1	1		8
G8_Incidents & Penalties	2	*1%*					1	1			
G9_Rating, Ranking & Awards	8	*5%*					3	3		2	
G10_Overall & Others (NEC)	76	*47%*	5	4	1	2	9	11	6	7	31

Table 2.7 By Construct Operationalization (No. of Studies)

As shown in the table, EM has been operationalized with ten different subdimensional measurement groups and BP, nine groups (the detailed descriptions of the measures are available in Appendix II). The collection of 162 articles examines in total 43 distinct G-C subdimensional relations. These findings prove that traditional meta-analysis is not directly applicable to synthesize empirical GC research outcomes for a generalized answer to *whether it pays to be green* – these EM measures are highly inconsistent and can have different, or even contradictory, economic effects that offset each other and obscure the cumulated results. It is necessary to develop and apply the new method of QMA to disentangle these heterogeneous studies into homogeneous subgroups for more granular analysis and meaningful interpretation.

The results in Table 2.7 show that approximately half (47%) of the articles used high-level composite measures of EM that are not elsewhere classified (G10_Overall & Others), while only a few articles used more specific EM measures (G1 to G4, 13% in total). The operationalization of BP shows a similar pattern, with about one-third (32%) of studies used composite measures that are not elsewhere classified (C9_Overall & Others), yet only a few used specific measures on the *front-end* of the performance logic chain (C1 to C4, 14% in total). E.g., there are few studies in EM's impact on C3_Costs (2%) and C4_Revenue (1%), yet there are many on C5_Profitability (17%), which is the composite measure combining C3 and C4 – it appears that although it is not uncommon to hypothesize EM's economic impacts through the cost and revenue channels, it is indeed uncommon to measure such effects on either the cost or the revenue channel at a more granular level.

The common practice of using composite and *back-end* measures in GC research makes it difficult to ascertain the mechanisms through which that EM affects BP – due to the likelihood of omitted variables and confounding effects involved between the focal constructs. This makes abduction a fit for GC theorizing as the grounded approach is limited in interpreting the dataset which is fraught with noises mixing up with and weakening valid signals (e.g., the 31 articles that applied the G10-C9 composite measures could imply anything or nothing regarding *why it pays to be green*),

Table 2.8 further breaks down the QMA results into three categories by research outcome: (1). 4A (Y) includes studies that hypothesized and found positive G-led/C-lagged causal relations; (2). 4A (N) includes those hypothesized but did not find such causal relations; and (3). 1A-3A combines the rest of non-4A studies.

Construct Operationalization	Sub-Total	C1 Productivity & Efficiency	C2 Customer & Market	C3 Costs	C4 Revenue	C5 Profitability	C6 Market Valuation	C7 Composite Accounting	C8 Composite Accounting & Market	C9 Overall & Others (NEC)
4A (Y)	*72%*	*5%*	*3%*	*2%*	*1%*	*9%*	*8%*	*5%*	*10%*	*28%*
G1_Green Investments	2%	1%				1%			1%	
G2_Green Innovation	5%	1%				1%	1%			3%
G3_Product Design	1%		1%							
G4_Pollution Reduction Practices	1%						1%			
G5_Pollution Output	12%	1%				1%	2%	1%	5%	2%
G6_Environmental Certification	6%		1%	1%		1%		1%	1%	2%
G7_Green SCM	7%	1%		1%						4%
G9_Rating, Ranking & Awards	4%					2%	1%		1%	
G10_Overall & Others (NEC)	35%	2%	2%	1%	1%	4%	3%	3%	2%	17%
4A (N)	*8%*	*1%*	*1%*			*1%*	*1%*	*1%*	*1%*	*2%*
G2_Green Innovation	1%					1%				
G4_Pollution Reduction Practices	1%						1%			
G7_Green SCM	2%					1%		1%		1%
G10_Overall & Others (NEC)	5%	1%	1%				1%	1%	1%	2%
1A-3A	*20%*	*1%*				*6%*	*7%*		*4%*	*1%*
G1_Green Investments	1%					1%	1%			
G2_Green Innovation	1%					1%				1%
G4_Pollution Reduction Practices	1%						1%			
G5_Pollution Output	7%					3%	1%		3%	
G6_Environmental Certification	1%						1%			
G7_Green SCM	1%						1%			
G8_Incidents & Penalties	1%					1%	1%			
G9_Rating, Ranking & Awards	1%						1%			
G10_Overall & Others (NEC)	7%	1%				1%	3%		1%	1%

Table 2.8 By Construct Operationalization (%)

The following patterns emerge from comparing within- and between-group frequency of occurrence in the research findings: (1). When BP is measured by C9_Overall & Others, positive findings outnumber negative findings overwhelmingly (28% vs. 1%), but the difference disappears when BP is measured by C6_Market Valuation (8% vs. 7%); (2). Similarly, when EM is measured by G10_Overall & Others, positive findings largely outnumber negative findings (35% vs. 7%), but the gap narrows when EM is measured by G5_Pollution Output (12% vs. 7%). Findings from the metrics in C9 and G10 may provide more nuanced insights, but it is risky to draw inferences directly from such results as it is challenging to isolate EM's BP effects for examination using high-level composite measures. Conversely, G5_Pollution Output and C6_Market Valuation can lead to results that offer better clarity and validity since they are both specific objective measures.

In addition to construct operationalization, the data types[4] applied in these studies are also coded and reported in Table 2.9.

[4] Types of data: (a) Anecdotal (anecdotal evidence such as "best practices" or "success stories"), (b) Communicative (data from sources such as corporate websites or announcements and media coverage), (c) Evaluative (data based on evaluation by third-parties such as ESG rating agencies), (d) Financial (data collected from corporate annual reports or financial databases), (e) Perceptual (data collected from questionnaire surveys), (f) Physical (physical measurement in-natura, such as material input in metric tons)

		BP	Anecdotal	Evaluative	Financial	Perceptual
EM			*1%*	*1%*	*51%*	*47%*
Anecdotal	*1%*		*1%*			
4A (Y)	1%		1%			
Communicative	*11%*			*1%*	*10%*	*1%*
4A (Y)	6%				5%	1%
4A (N)	2%			1%	1%	
1A-3A	4%				4%	
Evaluative	*18%*				*17%*	*1%*
4A (Y)	14%				14%	
4A (N)	1%				1%	
1A-3A	3%				2%	1%
Financial	*3%*				*3%*	
4A (Y)	2%				2%	
1A-3A	1%				1%	
Perceptual	*51%*				*5%*	*46%*
4A (Y)	41%				2%	38%
4A (N)	5%				1%	4%
1A-3A	5%				1%	4%
Physical	*16%*			*1%*	*15%*	
4A (Y)	9%				9%	
4A (N)	1%				1%	
1A-3A	7%			1%	6%	

Table 2.9 By Data Type

Table 2.9 show that GC research relies heavily on Perceptual data – 51% of the articles used Perceptual data for EM, 47% for BP, and 46% for the EM-BP dyad. Perceptual data in general, and when applied to sustainability-related topics in particular, can be afflicted with social desirability bias and susceptible to threats to validity. Therefore, the predominantly positive findings from Perceptual data (38% positive vs. 4% negative) should be interpreted cautiously. On the other hand, fewer articles (16%) used Physical data for EM. Since Physical data usually are less subjective and more reliable when reported consistently, research findings from this category presumably are more valid and will be closely examined in the theorizing stage.

4.2 By Methodological Artifacts

Table 2.10 presents the QMA results by the methodological artifacts of data timeframe, time-lag, and data age.

	Sub-Total	% 4A (Y)	% 4A (N)	% 1A-3A
BY TIMEFRAME				
Cross-Sectional	94	81%	11%	9%
Longitudinal	68	60%	15%	35%
BY TIME-LAG				
Concurrent	112	76%	9%	15%
1 Year	18	83%		17%
2 Years	5	80%		20%
1-3 Years	12	50%	17%	33%
3 Years+	4	100%		
EW: 1-3 Day	*11*	*27%*	*9%*	*64%*
BY DATA AGE				
2018	1			
2017	2	100%		
2016	3	100%		
2015	3	100%		
2014	12	58%	25%	17%
2013	9	89%	11%	
2012	12	92%		8%
2011	8	88%	13%	
2010	14	71%		29%
2009	9	67%	11%	22%
2008	3	100%		
2007	8	75%	13%	13%
2006	14	79%	7%	14%
2005	7	43%	14%	43%
2004	9	78%	11%	11%
2003	8	100%		
2002	1			
2001	6	67%		33%
1999	3	100%		
1998	3	67%		33%
1996	4	50%	25%	25%
1995	6	50%		50%
1994	2	50%		50%
1993	1			
1992	5	40%		60%
1991	1			
1990	2		50%	50%
1989	2	50%		50%
1987	1			
1971	1			
1970	2	50%	50%	

Table 2.10 By Methodological Artifacts

Table 2.10 shows that negative results are more likely to be found in studies using longitudinal instead of cross-sectional data (35% vs. 9%, between-group comparison of within-group proportions)[5]. *Ceteris paribus*, longitudinal data can better capture EM's long-term BP effects, while cross-sectional data reflect only static time-specific snapshots. Hence, if the potential confounding factors within the timeframe under study had been controlled, then findings from longitudinal data compared with which from cross-sectional data would be more realistic – in other words, the probability of getting green rewards might not be as high as studies using cross-sectional data have suggested.

[5] In the analysis here, we look more closely into the negative findings in order to counter the potential publication bias in the dataset – the rationale being that if positive results are more likely to get reported and published, studies with negative findings need to outperform in other ways to survive in a harsh context, so these tried-and-tested survivors presumably are able to withstand scrutiny and offer more credible insights.

As for the time-lag between focal constructs, the length of 1-3 years (stepwise year by year) is more likely to lead to negative findings (33%) – compared with concurrent (15%), 1-year (17%), and 2-year (20%) lag-lengths. It is not immediately clear whether the results indicate diminishing marginal returns on EM investments, so the methodological factor of time-lag needs to be further examined to ascertain its moderating effects on GC research findings.

Regarding data age (as reported for cross-sectional data and round-up mean year for longitudinal data), there seems to be a weak but positive trend of GC associated with *data recency* (a potential *authentic* moderator). Such a pattern is not observable when the results are organized by publication year, ruling out the alternative explanation of a positive trend associated with *study recency* (a potential *artificial* moderator). If data recency indeed is associated with more positive findings despite study recency, then this may indicate the trend of a growing market force that rewards corporate green practices.

4.3 By Contextual Contingencies

Table 2.11 presents the results by the contextual factors of sector, region, and firm size.

	Sub-Total	% 4A (Y)	% 4A (N)	% 1A-3A	4A GC++(Y)
BY SECTOR					
HPS	24	46%	8%	46%	11
M&S	55	64%	9%	27%	35
MFG	68	84%	7%	9%	57
SVC	15	93%	7%	0%	14
BY REGION					
NA	53	58%	6%	36%	31
EU	41	73%	10%	17%	30
CN	19	79%	11%	11%	15
JP	6	100%	0%	0%	6
ANZ	7	86%	0%	14%	6
APO	14	100%	0%	0%	14
MENA	3	67%	33%	0%	2
SA	1				1
AF	1				
CR	17	71%	18%	12%	12
BY SIZE					
Public	66	61%	6%	33%	40
Large	12	67%	25%	8%	8
SMEs	10	80%	0%	20%	8
Mixed	74	82%	8%	9%	61

Table 2.11 By Contextual Contingencies

Table 2.11 shows that studies with samples from high-polluting sectors (HPS) are more likely to get negative results (46% compared with 27% for Manufacturing & Services and 9% for Manufacturing), while service sectors (SVC) are the least likely to get negative results (0%). In terms of regional differences, companies in North America (NA) – mainly from the US with very few data points from Canada – are the most likely to get negative results (36% compared with 17% for Europe, 11% China, 14% Australia and New Zealand, 12% Cross-Regional studies, and 0% for the rest of the world). As for firm size, public firms are more likely to get negative results (33%) compared with large

companies (Large, 8%), small and medium-sized enterprises (SMEs, 20%), and companies of varying sizes (Mixed, 9%).

4.4 Summary

The QMA results are summarized in Table 2.12 in the structure of the G-C neural network.

RESEARCH INPUT	RESERCH OUTCOME
Construct Operationalization	
EM/EP Measures	G10_Overall & Others (47%), G5_Pollution Output (19%), G7_Green SCM (9%); Others
EM/EP Data Type	Perceptual (51%), Evaluative (18%), Physical (16%); Others (15%)
BP Measures	C9_Overall & Others (32%), C5_Profitability (17%), C6_Market Valuation (17%); Others
BP Data Type	Financial (51%), Perceptual (47%), Anecdotal (1%), Evaluative (1%)
Methodological Artifacts	
Timeframe	Longitudinal vs. Cross-sectional: 35% vs. 9% negative findings
Time-lag	1-3Y vs. Concurrent/1Y/2Y/3Y+: 33% vs. 15% / 17% / 20% /0% negative findings
Data Age	Subtle positive trend in data recency
Contextual Contingencies	
Sector	HPS vs. M&S/MFG/SVC: 46% vs. 27% / 9% / 0% negative findings
Region	NA vs. Rest-of-World average: 36% vs. 8% negative findings
Firm size	Public vs. Large/SMEs/Mixed: 33% vs. 8% / 20% / 9% negative findings

Table 2.12 QMA Results (Summary)

In summary, the QMA finds that GC research as a collective body of work is highly heterogenous in construct operationalization – the collection of 162 articles study in total 43 distinct G-C subdimensional relations. The inconsistencies in construct operationalization require that the abduction process dives deeper into the empirical data to better understand the G-C co-variations at the subdimensional level as they may need different theoretical explanations. Furthermore, the number of studies that used more specific EM measures (G1 to G4, 13% in total) or *front-end* BP measures (C1 to C4, 14%) is limited. The reliance on high-level composite measures can entangle and obscure the mechanisms through which that EM affects BP – very much like observing the surface of the G-C neural network from a distance trying to figure out what is going on inside the black box. This again justifies using abduction to go beyond the immediate empirical data and reason to the best explanations on the G-C nexus.

Regarding methodological artifacts, Longitudinal data with a time-lag of 1-3 years are more likely to associate with negative findings than concurrent data, or data with 1- and 2-year time-lags. As for contextual contingencies: (1). Service sectors are the most, while high-polluting sectors the least, likely to get green rewards; and (2). Companies located in NA or traded in the stock market are less likely to get green rewards. The robustness of the results was tested through analyses with two-variable data slicing, combining data timeframe with the three contextual factors (sector, region, and firm size), respectively. Except for NA firms with cross-sectional data, the other combinations exhibit the same patterns as results from single-variable data slicing – largely ruling out the alternative

explanation that methodological artifacts may have exerted significant *artificial* moderating effects on the findings.

To sum up, the QMA results show that, in general, companies are more likely to get green rewards when they are in services but not high-polluting sectors, located outside of NA, or not publicly traded. These contextual contingencies unveil, translucently, the *Context Layer* in the G-C neural network and will be applied for hypothesis refinement in the theorizing process.

5. Theorizing GC

5.1 Hypothesis Generation

This subsection develops tentative hypotheses to unravel the *Hidden Layer* in the G-C neural network. The hypothesis generation process takes three steps: (a) detecting from the research outcomes within- and between-group patterns in the frequency of occurrence, (b) developing provisional explanations as *latent rules* for the *empirical effects* based on observed evidence and logical inferences, (c) evaluating and retaining the best possible explanations. Figure 2.9 visualizes on the G-C logic chain the QMA results from two main EM measures that together represent 66% of the dataset (19% from G5_Pollution Output and 47% from G10_Overall & Others). Table 2.13 supplements the details.

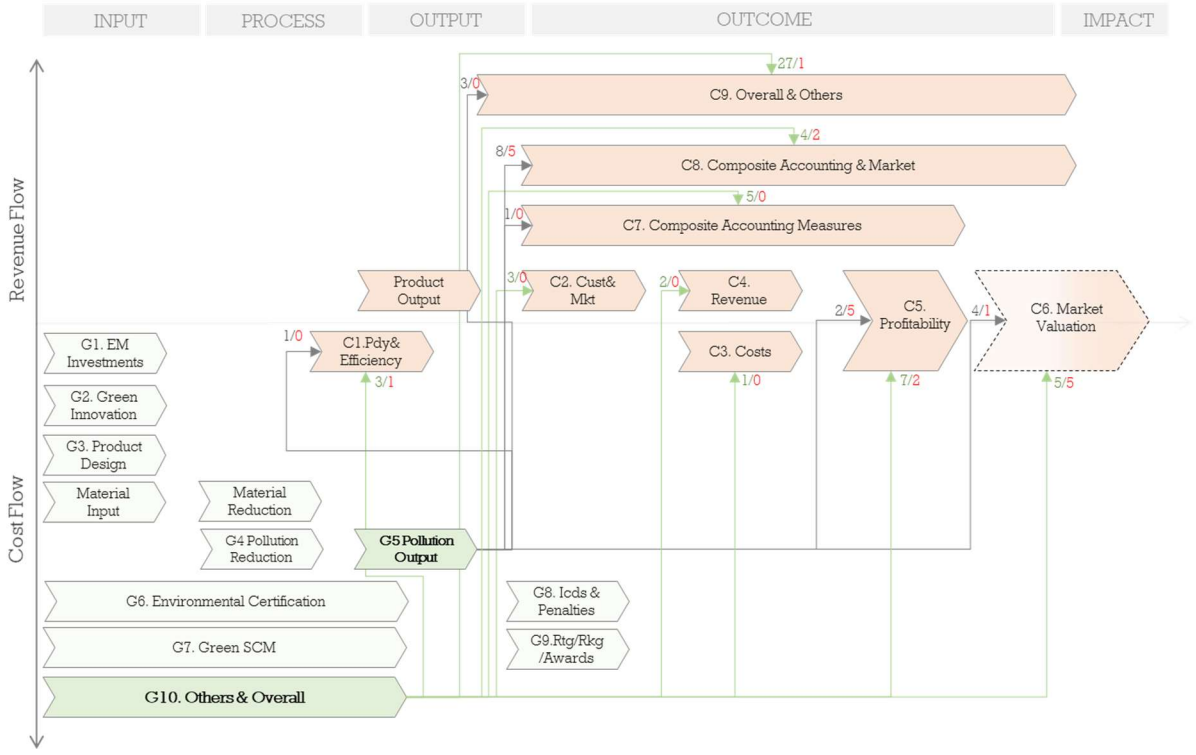

Figure 2.9 QMA Results (G5 and G10)

	Sub-Total	C1 Produtivity & Efficiency	C2 Customer & Market	C3 Costs	C4 Revenue	C5 Profitability	C6 Market Valuation	C7 Composite Accounting	C8 Composite Accounting & Market	C9 Overall & Others (NEC)
G5_Pollution Output	**30**	**1**				**7**	**5**	**1**	**13**	**3**
4A (Y)	19 (60%)	1				2 (30%)	4 (80%)	1	8 (60%)	3
1A-3A	11 (40%)					5 (70%)	1 (20%)		5 (40%)	
G10_Overall & Others (NEC)	**76**	**5**	**4**	**1**	**2**	**9**	**11**	**6**	**7**	**31**
4A (Y)	57 (80%)	3 (60%)	3	1	2	7 (80%)	5 (45%)	5 (80%)	4 (60%)	27 (90%)
4A (N)	8	1	1				1	1	1	3
1A-3A	11 (10%)	1 (20%)				2 (20%)	5 (45%)		2 (30%)	1 (3%)

Table 2.13 QMA Results (G5 and G10)

G5_Pollution Output is the pollution level of production activities, the result from pollution reduction practices (EM) or the lack thereof, with lower pollution output indicating better EP. Research findings based on G5 are chosen as the primary empirical evidence for hypothesis generation since most studies with this measure (73%, 22 out of 30) used Physical data from official databases (e.g., the Toxic Release Inventory of the US Environmental Protection Agency). All else being equal, findings from Physical data – compared with those from Perceptual or Evaluative data – are more valid and reliable, hence can be taken as a more solid foundation for abduction with a lower level of uncertainty.

Among all the BP measures that are linked with G5_Pollution Output, C5_Profitability is the most likely (70% within-group), and C6_Market Valuation the least likely (20%), to get negative results (C7, C8, and C9 are left out of analysis here as it is challenging to digest findings from composite measures due to confounding effects). In other words, the results suggest that lower pollution output generally reduces profitability but increases market valuation.

As profit margin is derived by subtracting costs from revenue and pollution reduction incurs abatement costs but does not lead to revenue directly, it may therefore be inferred from the results that *reducing pollution output decreases profitability due to increased abatement costs*.

Market valuation (C6) predominantly reflects shareholders' perception of a stock's returns and risks (another less plausible explanation that investors reward environmentally responsible practices out of altruism will be discussed later). From the perspective of returns, pollution output – as the undesirable by-product of production activities – is unlikely to have direct economic benefits or be perceived as such since it may be managed by reactive ''end-of-pipe'' pollution control technologies rather than proactive pollution prevention practices that can also improve efficiency. From the perspective of risks, high pollution output may imply intangible costs of potential penalties imposed by environmental regulations, and lower pollution output can reduce a company's exposure to such un-booked liability risks.

We therefore hypothesize that:

H1a Pollution reduction mitigates risks (1)

H1b Pollution reduction increases costs (1)

The category of G10_Overall & Others includes measures that cannot be neatly classified into other groups (e.g., environmental capabilities, staff training, and participation in environmental programs), and two-thirds of studies (67%, 57 out of 76) in this category used Perceptual data collected chiefly from questionnaire surveys. These features make interpreting the results from G10 less straightforward but more nuanced.

Table 2.13 shows that except C6_Market Valuation (45% positive and 45% negative), the research outcomes for all other BP measures linked with G10 are predominantly positive (80% on average). A closer look into the data types applied in these studies finds that the majority (73%, 8 out of 11) of the G10-C6 dyad used Communicative-Financial data (with Market Valuation data collected from financial databases such as Compustat), indicating that the G10-C6 dyad primarily measures financial market's reaction to environmental events. On the other hand, the rest of the studies in this category (G10 and non-C6) mostly used Perceptual-Perceptual data (71%, 46 out of 65), usually by surveying internal stakeholders such as managers with environmental responsibilities. In other words, the results indicate that EM and green projects are more likely to be perceived as economically beneficial by internal stakeholders (managers) but less so when communicated to external stakeholders (stock investors).

The predominately positive findings in the group of G10 studies (except when G10 is linked with C6_Market Valuation) may be explained by social desirability biases potentially present in the Perceptual data. However, in the absence of relevant evidence, another plausible explanation is that the managers surveyed in these studies, and overwhelmingly responded with positive evaluations, had first-hand insider experiences implementing green projects and witnessing their intangible economic benefits, although such positive effects might not be readily quantifiable or observable to outsiders. Such internally perceptible, yet externally unobservable, economic values of EM may be materialized through cost reduction (e.g., efficiency gain from process reengineering and material reduction) or revenue growth (e.g., price premium from green products with lower energy consumption). Although empirical evidence that precisely investigates these two mechanisms is limited in the rest of the dataset[6] for out-of-sample verification, such reasoning nevertheless is theoretically plausible.

[6] Except for one study that used specific EM measure (G3_Product Design) and front-end BP measure (C2_Customer & Market) with positive finding, the rest of the dataset do not directly provide corroboration with clarity as they either used high-level composite EM measures, or back-end composite BP measures, or both.

On the other hand, when such environmental events are communicated to the financial market, investors might evaluate the economic implications of these events positively or negatively due to information asymmetry and different value preferences. For example, when a company's investment in green production technology is announced, some investors might respond to the event negatively based on the perception that it incurs high present costs without promised future returns (Traditionist investors), while the others might respond to the same event positively, perceiving it as an economically profitable investment (Revisionist investors) or simply rewarding the company for its environmentally responsible practice (social investors).

We therefore hypothesize that:

H2 Material reduction reduces costs (1)

H3 Green products may lead to revenue growth (1)

H4 Shareholders may value environmental events positively or negatively (1)

The five tentative hypotheses (H1a, H1b, H2, H3, and H4) are developed by abductively reasoning to plausible *theoretical causes* from the observed *empirical effects* of research outcomes. The hypotheses are built upon 66% of the collection of GC articles (19% from G5_Pollution Output and 47% from G10_Overall & Others). A close examination of the remaining 34% of the dataset finds no further evidence that either corroborates or contradicts the proposed hypotheses, elaborated as follows (examined from the dimension of BP operationalization, excluding those linked with G5 and G10):

1) Results from the BP measure, **C6_Market Valuation**, are too dispersed for meaningful synthesis, with 11 studies (7% of the dataset) spreading across seven different dyads. If the results were to be cumulated – although inappropriate to do so as the EM measures differ significantly in nature – the pooled results would form no apparent patterns, with four positive, two neutral, and five negative findings. Hence, findings from this category neither strengthen nor invalidate the proposed hypotheses that financial markets *reward lower pollution output as a risk mitigation mechanism (H1a)* and *may respond to green events positively or negatively (H4)*.
2) The nine studies (6%) that applied the specific front-end BP measures from C1 to C4 (**C1_Productivity & Efficiency**, **C2_Customer & Market**, **C3_Costs**, and **C4_Revenue**) cannot be pooled for pattern detection as they represent different value creation mechanisms (and scatter over seven different EM-BP dyads).
3) The 11 studies (7%) that applied **C5_Profitability** include six positive, two neutral, and three negative findings. Since C5 is a composite measure and its associated EM measures differ in nature, insights with certainty cannot be distilled from such entangled results to either validate or refute the propositions that *pollution reduction increases tangible costs (H1b) and reduces*

intangible risks (H1a), *material reduction reduces costs (H2)*, and *green products may increase revenue (H3)*.

4) The 25 studies (15%) that applied the measures from C7 to C9 (**C7_Composite Accounting**, **C8_Composite Accounting & Market**, and **C9_Overall & Others**) are all based on back-end composite measures that offer no clarity on specific causal links, hence provide no further evidence for theorizing due to difficulties disentangling and isolating the effects under examination.

To conclude, the out-of-sample verification of the 34% of the dataset does not provide further evidence to corroborate or contradict the proposed hypotheses. Therefore, we consider the five propositions built upon the majority (66%) of the dataset representative of the body of work in GC research. These preliminary propositions are delineated on the G-C logic chain in Figure 2.10.

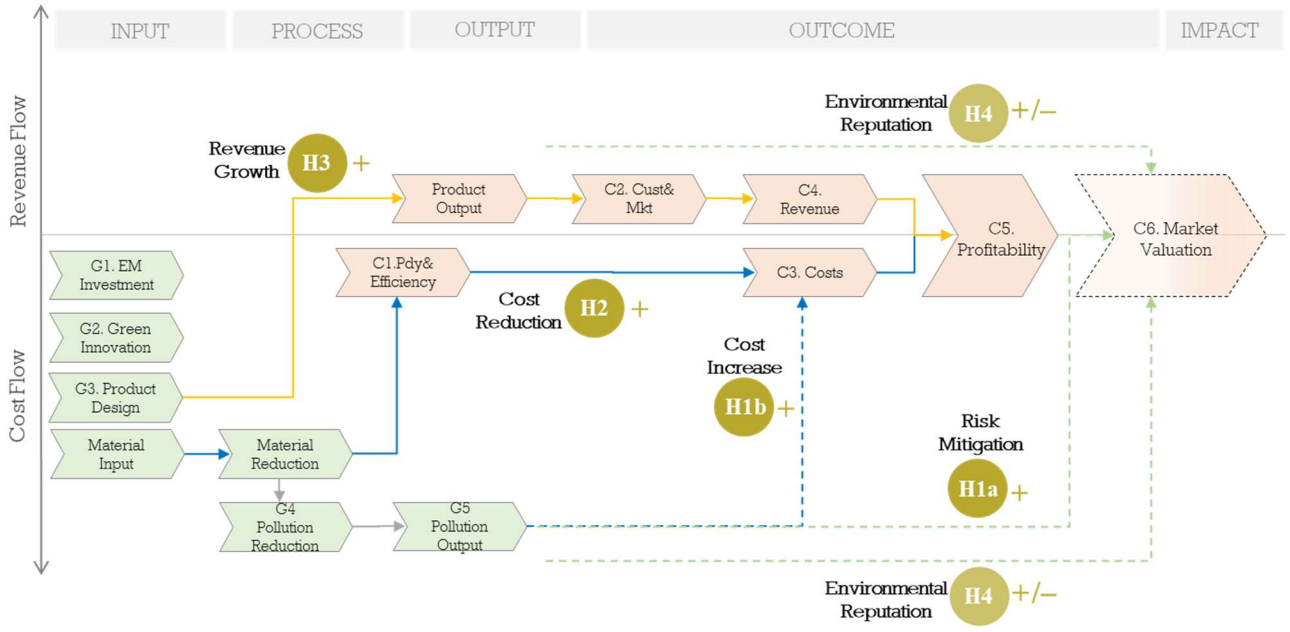

Figure 2.10 Provisional Propositions in the G-C Logic Chain

5.2 Hypothesis Refinement

5.2.1 The Context Layer – Contextual Contingencies

This subsection infers key contextual contingencies on the G-C nexus from the QMA results to refine the tentative hypothesis. The hypothesis refinement process takes three steps: (a) examining the patterns in the cumulated evidence of contextual contingencies, (b) proposing plausible contextual moderators that elucidate the observed empirical effects, (c) explicating how the contextual contingencies moderate the G-C nexus. Compared with the process of hypothesis generation, which is firmly grounded on empirical observations, hypothesis refinement draws on, wherever appropriate, relevant theories, experience, intuition, and *disciplined imagination*. Applying tacit knowledge that cannot be fully articulated is necessary and valuable in abduction, while it also

inevitably introduces uncertainty in abductively developed theories. Therefore, the refined propositions will have to go through a structured evaluation process before being accepted as preliminary hypotheses.

The QMA results on contextual contingencies suggest that (1). Among all sectoral groups, high-polluting sectors are the least likely, while service sectors the most likely, to see positive GC results; and (2). NA-based and publicly traded companies are less likely to get green rewards. These findings lead us to propose four contextual contingencies on the G-C nexus: (a) pollution saliency, (b) regulatory stringency, (c) material saliency, and (d) stakeholder eco-propensity.

5.2.1.1 Pollution Saliency

Pollution saliency is defined in this study as *the significance of pollution's negative EIs*. Basically, pollution on EOA includes all undesirable by-products from economic activities that are harmful to nature and/or human health (e.g., wastes, air emissions, nuisances of sound, light, and odor). Pollution saliency is a contextual factor at the sectoral level since within-sector EOAs are mostly similar even across geographical locations. It can be assessed from three dimensions: (a) pollution intensity – the level of pollution output per unit of product output of goods or services; (b) pollution toxicity – the degree of pollution's harmfulness to nature and/or human health, and (c) pollution longevity – the lifespan of the harmful substance's existence.

High pollution saliency sectors are exposed to greater risks of potential liabilities (e.g., fines and penalties), so the level of risk mitigation from pollution reduction in such sectors is commensurably higher. Hence, pollution saliency positively moderates the link between pollution reduction and risk mitigation. The link between pollution reduction and cost increase is also amplified by pollution saliency, as the abatement costs are generally higher in high polluting sectors.

Although risk mitigation from pollution reduction can reduce intangible liability costs, such economic benefits may not be recognized since they are not materialized transactions. On the other hand, the costs invested in pollution abatement are tangible as they are typically materialized and booked expenditures. As a result, pollution reduction may appear to incur net costs – it may indeed be the case, in narrow financial accounting terms, if organizations voluntarily implement such initiatives (on a side note, this is probably the base of the Traditionist argument that "*talk is cheap, environmental initiatives are not*" (Walley and Whitehead, 1994, p. 46)). However, if pollution reduction is mandated by environmental legislation or "soft regulations" in the market (or, from a normative standpoint, the moral obligation of sustainable development), then the intangible benefits from risk mitigation (and reduced negative environmental impacts) may compensate or overcompensate for the costs of pollution reduction.

H1a (1) and H1b (1) are refined as follows:

H1a Risk mitigation from mandatory pollution reduction is amplified by pollution saliency (2)

H1b Cost increase from voluntary pollution reduction is amplified by pollution saliency (2)

5.2.1.2 Regulatory Stringency

Regulatory stringency refers to *the strictness of environmental legislation and regulations*. Regulatory stringency can be assessed from three aspects: (a) scope of regulation (e.g., specific treatment of water drainage or chemical wastes), (b) standards of regulation (e.g., the capped level of GHG emissions or the use of toxic materials in products), and (c) severity of violation consequences (e.g., fines, penalties, or legal actions). Regulatory stringency is both a region- and sector-based contingency since environmental regulations are usually enacted within the broader regional legislative framework and applied to specific sectors operating from the region. Therefore, the same sector operates in different regions may be subject to varying degrees of regulatory stringency and competes in an unlevel playing field in the global market (hence the phenomenon of ''exporting pollution'' to less strictly regulated regions, often located in developing nations).

The QMA finds that companies based in NA (mostly in the US as there are very few data points from Canada) – are less likely to get green rewards. Interpreting the moderating effects of geographical location is less straightforward – location usually is not a *root cause* explanator but a proxy of various underlying factors. Among the geo-proxied PESTEL factors (political, economic, sociocultural, technological, environmental, and legal), the political-legal factor of stringent environmental regulations in the US seems to explain the results best considering the facts that (1). The annual cost of environmental compliance in the US accounts for about 2.1% of GDP in the 1990s, compared with only 1% of GDP in most developing countries (Jaffe et al., 1995); and (2). The number of US federal environmental laws and amendments has increased from about 70 in the 1980s to about 120 in the early 2000s (Allen and Shonnard, 2012).

The other PESTEL factors do not offer equally plausible explanations. In terms of economic development levels, there does not seem to be theories or empirically observed differences between the US and the other developed or developing nations that can explain the QMA results (the Environmental Kuznets Curve that assumes a reverse causality does not apply at the micro-level and is rather controversial itself). Sociocultural values may affect market behaviors and corporate performance results – however, there is currently no evidence to suggest that stakeholders in the US are generally less pro-environment than the rest of the world, to the best of our knowledge. The same argument applies to the technological (innovation, R&D) and environmental (terrains and natural resources) factors. Therefore, environmental regulatory stringency seems to be the most likely explanator of the QMA finding despite the uncertainty.

In stringently regulated regions or sectors, companies are exposed to higher liability risks as the probability and consequences of violation are both higher. Compliance with environmental regulations through pollution reduction mitigates such regulatory risks, and the higher the regulatory stringency, the higher the mitigated risks through pollution reduction.

H1a (2) is therefore refined as below:

H1a *Risk mitigation from mandatory pollution reduction is amplified by pollution saliency and regulatory stringency (3)*

5.2.1.3 Material Saliency

Material saliency is defined as *the intensity of physical assets and material inputs that enable production activities*. Material saliency covers three areas: (a) physical assets such as plant, property, and equipment; (b) primary material inputs to be processed as product outputs; and (c) secondary material inputs that facilitate and support production activities. Material saliency can be assessed from three dimensions: (a) material intensity – the volume of physical assets and material inputs in relation to product output of goods or services, (b) material scarcity – the remaining availability and substitutability of the materials, and (c) material renewability – the level of renewability and speed of replenishment during material depletion.

High material saliency sectors have physical assets that are ''stickier'' (taking longer to depreciate and are more costly to replace) and more ''rigid'' (harder to mold or modify) and/or use material inputs that are scarce, non-substitutable, or non-renewable. As proactive pollution reduction practices that also lead to efficiency gains require moving upstream from corrective ''end-of-pipe'' pollution control to preventive practices that may require upgrading production equipment or substituting input materials, implementing such EM practices would be less costly and more economically rewarding in low material saliency sectors – such as services sectors. Material saliency thus explains the QMA finding that services sectors are the most likely to get green rewards.

We therefore refine H2 (1) as follows:

H2 *Cost reduction from material reduction is attenuated by material saliency (2)*

5.2.1.4 Stakeholder Eco-Propensity

We define stakeholder eco-propensity as *stakeholders' readiness to take pro-environment actions, given the existence of other necessary conditions*. The antecedents of eco-propensity include (a) eco-awareness, i.e., knowledge about negative EIs caused by economic activities, and (b) eco-conscience, i.e., preference for the right actions to reduce economic activities' negative EIs. Built on

eco-awareness and eco-conscience, eco-propensity is the state of being ready to act on pro-environment knowledge and values.

The QMA finding that public firms are less likely to get green rewards can be explained by the level of stakeholder eco-propensity in the stock market. Stock investors react to an environmental event based on (a) their value preferences in terms of eco-propensity, and (b) their perception of the event's economic relevance. These share-trading stakeholders have either high or low eco-propensity. High eco-propensity stakeholders either take green events for granted, expecting those companies to assume their environmental responsibilities as an obligation, hence are more likely to penalize ''grey'' events instead of reward ''less grey'' ones; or are outweighed by low eco-propensity investors who perceive green to be an economic liability and cannot alter their collective force that responds to green (grey) events negatively (positively). As a result, public firms' green deeds are less likely to get rewarded economically due to low levels of collective eco-propensity in the financial market.

Stakeholder eco-propensity functions in the product market in a similar way, but to a lesser extent. Given the other necessary conditions (e.g., credible product information and customers' ability-to-pay), high eco-propensity customers are more inclined to buy – and may even be willing to pay a premium for – green products that may not benefit themselves economically. Therefore, revenue from green products without additional customer economic values is moderated by the level of eco-propensity in the product market. Although empirical evidence on this proposition is limited in the dataset of GC research, such reasoning is theoretically plausible and may also resonate with personal experience.

The hypotheses of H3a (1), H3b (1), and H4 (1) are refined accordingly as follows:

H3a Green products with additional customer economic value can lead to revenue growth (2)

H3b Revenue growth from green products without additional customer economic value is moderated by stakeholder eco-propensity (2)

H4 Market valuation of environmental events is moderated by stakeholder eco-propensity (2)

To sum up, four key contextual moderators are identified from the QMA results: pollution saliency, regulatory stringency, material saliency, and stakeholder eco-propensity. These contextual contingencies are applied to refine the main propositions generated in Section 5.1 into six moderated hypotheses. As summarized in Table 2.14, these refined hypotheses reveal the *Context Layer*, and make translucent the *Input Layer* and the *Hidden Layer*, of the G-C neural network.

INPUT LAYER		CONTEXT LAYER	HIDDERN LAYER	OUTPUT LAYER
H1a	*Pollution Reduction (Mandatory)*	Pollution Saliency Regulatory Stringency	*Risk Mitigation*	Cost Reduction
H1b	*Pollution Reduction (Voluntory)*	Pollution Saliency	-	Cost Increase
H2	*Material Reduction*	Material Saliency	*Efficiency*	Cost Reduction
H3a	*Green Products (w Customer Value)*	-	*Differentiation*	Revenue Growth
H3b	*Green Products (w/o Customer Value)*	Stakeholder Eco-Propensity	*Differentiation*	Revenue Growth
H4	*Overall Greenness*	Stakeholder Eco-propensity	*Reputation*	Market Valuation

Table 2.14 Refined Hypothesis in the G-C Neural Network

5.2.2 The Input Layer – A Green Typology

This subsection crystalizes the *Input Layer* of the G-C neural network into a green typology. Figure 2.11 shows that an organization's EM can be classified along the dimensions of EA and Stakeholder – the primary parties affected by the EM practices. Stakeholders include three groups: Producer, Customer, and Nature. EAs include three aspects: EIA, EOA, and External EA beyond the organizational boundary.

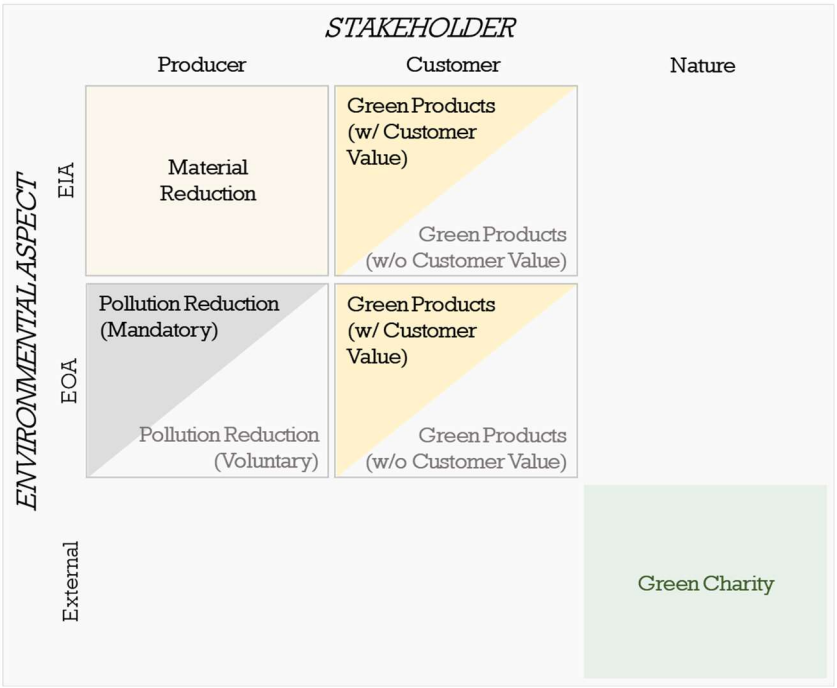

Figure 2.11 Green Typology

Material reduction is the EM on Producer's EIA and pollution reduction on its EOA. Material reduction decreases material saliency and increases green efficiency. Pollution reduction responding to present or projected environmental regulations mitigates liability risks and reduces intangible costs, and voluntary pollution reduction increases costs without necessarily yielding economic returns.

The EM practice interfacing with Customer is the development of green products. Green products can be categorized into two groups: (a) those with added customer economic value (w/

customer value), e.g., products with longer durability on EIA or higher recyclability on EOA; and (b) those without added customer economic value (w/o customer value), e.g., products made from recycled materials on EIA or with reduced packaging on EOA. Green products with added customer value can more readily grow sales and revenue through differentiation in a market, while those without have to leverage customers' stakeholder eco-propensity to penetrate the market.

On external EA beyond organizational boundaries, green charity[7] refers to altruistic EM practices, such as nature conservation (e.g., tree-planting), restoration (e.g., waterbody cleaning), giving to environmental causes, or sharing with competitors best green practices. Green charity does not directly benefit an organization economically, although it may improve its environmental reputation and corporate image.

In summary, the green typology suggests that there are four main categories of EM practices: material reduction, pollution reduction (mandatory or voluntary), green products (with or without customer value), and green charity.

5.2.3 The Hidden Layer – The New Construct of GC

This subsection crystalizes the *Hidden Layer* of the G-C neural network into a new construct of GC. GC is *the superior competitive position a company obtains from its EM that enables it to achieve better BP relative to its rivals*. As the value creation and apportion mechanism between EM and BP, GC is the answer to *why it pays to be green*. Figure 2.12 shows the four subdimensions of GC: risk mitigation, green efficiency, green differentiation, and green reputation.

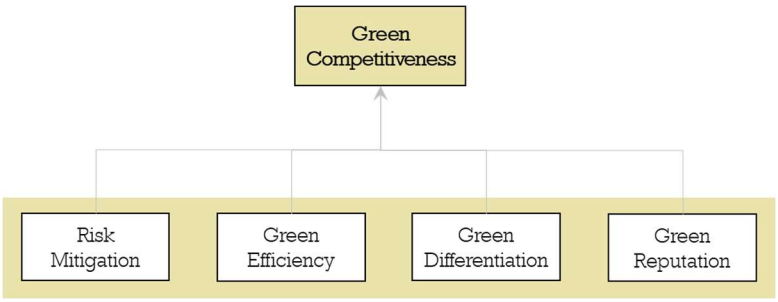

Figure 2.12 GC Subdimensions

- **Risk Mitigation**

Risk mitigation is the primary mechanism of economic value creation that links EM with BP.

[7] The term *green charity* instead of *environmental philanthropy* is chosen here as it is more appropriate in this context: *charity* usually refers to short-term giving and donations, while *philanthropy* refers to addressing root causes and long-term development programs.

Risk mitigation is realized primarily through pollution reduction since pollution is the area that currently gets the most attention from regulators and other key stakeholders presenting material or potential risks. Pollution reduction mitigates an organization's potential liability risks by reducing probabilities of violation and financial impacts of non-compliance. Risk mitigation through pollution reduction is not easy to measure with financial indicators since the mitigated liability costs are usually intangible and un-booked. Pollution reduction can also signal to investors an improved risk profile and positively affect market valuation. Such function has been examined using the event study methodology, although this approach captures only short-term effects rather than long-term firm value changes.

- Green Efficiency

Green efficiency is created through material reduction and dematerialization. It is the area where quick-win opportunities with ''low-hanging fruits'' exist, as it has a shorter payback period and does not actualize through external markets. Dematerialization reduces costs in material inputs and, to a lesser extent, non-production supporting activities. As cost reduction in these areas is captured in the financial metrics of costs-of-goods-sold and operating costs, these indicators after normalization may be used as proxy measures of green efficiency.

- Green Differentiation

Green differentiation is obtained from the superior market position created by green products. It is built upon green products that reduce EIs with or without added economic value to customers. Green differentiation requires a higher level of resource commitment since it is likely to involve substantial capital investment, higher return uncertainty, and a longer payback period. Green differentiation is reflected in market share and revenue, hence normalized financial indicators of revenue and sales growth may be used as its proxy measure.

- Green Reputation

Environmental reputation is the beliefs and opinions held by stakeholders about a company's environmental values, activities, and performance. The environmental reputation of a company may be positive ''green reputation'' or negative ''grey reputation'' (green reputation is of primary interest to this essay, although grey reputation is also discussed together to understand the dynamics better). External stakeholders form impressions about a company's environmental reputation through acquiring and processing relevant information from various unstructured data sources, including (a) compulsory corporate disclosure (e.g., TRI), (b) voluntary corporate disclosure (e.g., sustainability reports), (c) external exposure, either positive or negative, by media and/or activists (e.g., news reports of green awards or environmental incidents). As a result, a company's environmental

reputation is not only based on its EM and EP but also affected by a variety of other factors and noises. Additionally, due to information asymmetry, information quality, and information processors' cognitive biases and value preferences, a company's environmental reputation may significantly differ from reality.

Environmental reputation affects market valuation through a rather complex and dynamic mechanism. The short-term manifestation of this mechanism is filtered through two layers: (a) investors' eco-propensity, and (b) investors' perception of the environmental events' economic effects. From these two dimensions, stock investors can be categorized into three groups: (a) profit-driven rational investors with a Traditionist view that *green penalizes, grey pays*; (b) profit-driven rational investors with a Revisionist view that *green pays, grey penalizes*; and (c) principle-based social investors with the normative value that *green is virtue, grey is vice* – these investors base their capital allocation decisions on the social impacts, rather than the economic effects, of the environmental events. These three groups of investors presumably respond to the same type of environmental events differently: (1). Traditionists tend to respond negatively to green events (sell-off) and positively to grey events (buy-in); (2). Conversely, Revisionists and social investors tend to respond positively to green events (buy-in) and negatively to grey events (sell-off). Table 2.5 summarizes these hypothetical investor reactions to environmental events.

STAKEHOLDER ECO-PROPENSITY PERCEIVED VALUE RELEVANCE	TRADITIONIST "Green Penalizes, Grey Pays"	REVISIONIST "Green Pays, Grey Penalizes"	SOCIAL "Green is Virtue, Grey is Vice"
Green Events	*Sell-off*	Buy-in	Buy-in
Grey Events	*Buy-in*	Sell-off	Sell-off

Table 2.15 Hypothetical Investor Reactions to Environmental Events

As the stocks change hands, the collective buy-in (sell-off) decisions can pull up (push down) stock prices through the bid-and-ask trading mechanism. If the collective force of Traditionists outweighs that of Revisionists and social investors together, they can turn their *perception* into *reality* in the stock market – green events are punished with price decrease (as ''Green Liability'' in the G-C matrix in Figure 2.8), while grey events are rewarded with price increase (as ''Grey Exploitation'' in the G-C matrix). Because *what gets rewarded gets repeated*, these investors' Traditionist view will be reinforced in a vicious cycle where *green penalizes, grey pays*. Vice versa, if Traditionists are outweighed by Revisionists and social investors, their *presumption* becomes *misconception* when green is rewarded (''Green Rewards'' in the matrix) and grey penalized (''Grey Penalty'' in the matrix). If this happens frequently enough, realistic Traditionists will adjust their perception to fit the new reality – converted by ''enlightened'' self-interest into Revisionists – and join forces with social investors in creating a virtuous cycle where *green pays, grey penalizes*. Therefore, although rather simplistic still, this *seesaw* game between the two opposing forces implies that the key to scaling social investing is

to (1). Grow the number of social and Revisionist investors until it reaches a critical mass that can trigger a positive change; and (2). Sustain the momentum until it converts Traditionists into Revisionists and becomes a self-sustaining self-reinforcing virtuous force.

The discussion so far has been focused on short-term market reaction to environmental events, but the dynamics between the two opposing forces function in a similar way when environmental reputation is considered in long-term investment decisions. It is less straightforward to assess environmental reputation's effect on long-term market valuation – one reason being that the environmental dimension is only one aspect in the all-encompassing environmental, social, and governance (ESG) performance considered in social investment. Environmental reputation's short-term effect on market valuation is usually measured using event study with abnormal return as a proxy.

- Summary

To conclude, GC is the preliminary answer to *why it pays to be green*. This new construct has four subdimensions: risk mitigation, green efficiency, green differentiation, and green reputation. These subdimensions of GC are formative rather than reflective since they do not, for the most part, covariate but together form the higher-level construct. They are mutually exclusive and exhaustive, with each representing one aspect of, and all collectively and sufficiently represent, the construct of GC. However, the logical homogeneity among the four subdimensions is still rather limited. For example, these subdimensions currently need to be measured at different organizational levels: risk mitigation at site level, green efficiency at plant level, green differentiation at product-line or business-unit level, and green reputation at corporate level. Therefore, GC as an intermediate construct in formulation needs to be further refined in future research.

5.2.3 A Preliminary GC Theoretical Model

Figure 2.13 synthesizes the results and presents the deliverable of this project, a preliminary GC theoretical model. The model's overarching proposition is that *it pays to be green as GC converts values created by EM in different contexts into economic returns of BP*. It comprises six moderated-mediated hypotheses as follows:

H1a *Mandatory pollution reduction mitigates risks; pollution saliency and regulatory stringency amplify risk mitigation from mandatory pollution reduction*

H1b *Voluntary pollution reduction increases costs; pollution saliency amplifies cost increase from voluntary pollution reduction*

H2 *Material reduction creates green efficiency; material saliency attenuates green efficiency from material reduction*

H3a *Green products with added customer value lead to green differentiation*

H3b *Green products without added customer value may lead to green differentiation through stakeholder eco-propensity*

H4 *Green reputation leads to positive market valuation; stakeholder eco-propensity amplifies positive market valuation from green reputation*

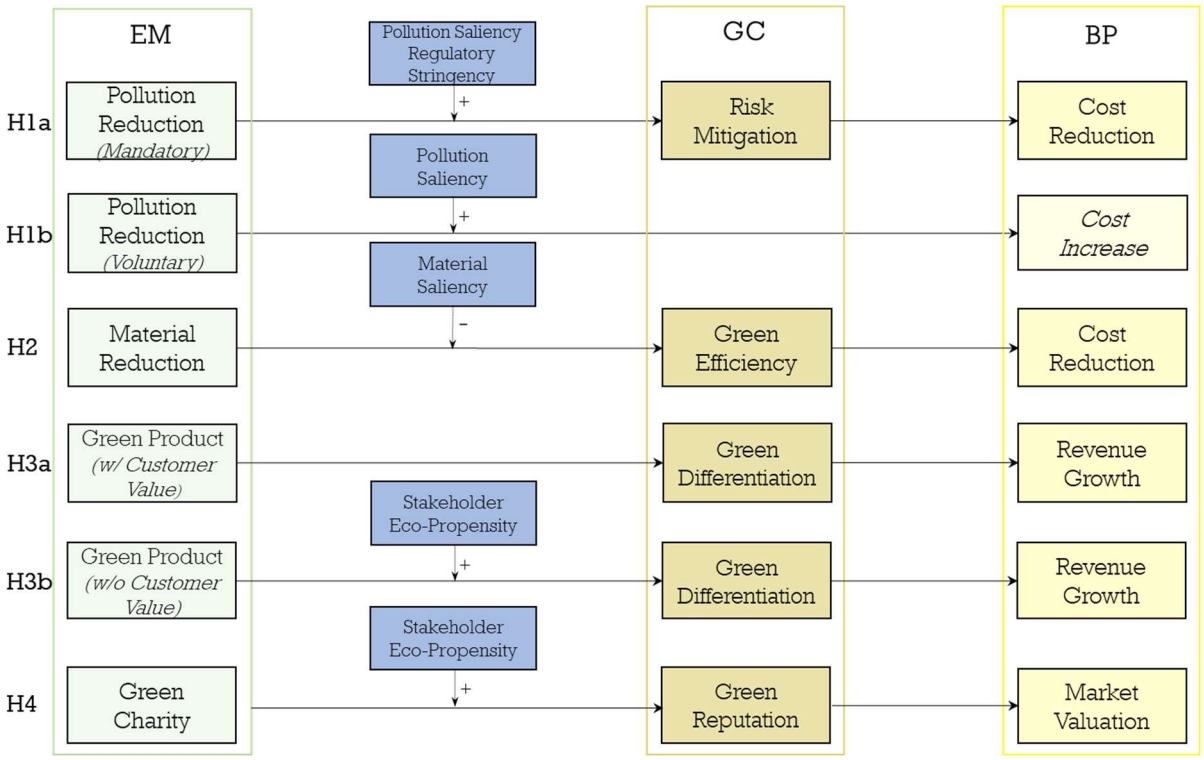

Figure 2.13 The Preliminary GC Theoretical Model

5.3 Hypothesis Evaluation

Before accepting an abductively generated theory, its explanatory power and theoretical plausibility need to be evaluated first. The GC theoretical model is developed through reasoning from empirical evidence to the best possible explanations, so its explanatory power has been tested and deemed satisfactory during hypothesis generation and refinement processes. This subsection is focused on evaluating its theoretical plausibility by juxtaposing the model alongside existing relevant theories (Eisenhardt, 1989). As shown in Figure 2.14, we develop a holistic organizational performance framework (HOPF), with the GC model mapped onto it, as a theoretical benchmark to approach this task.

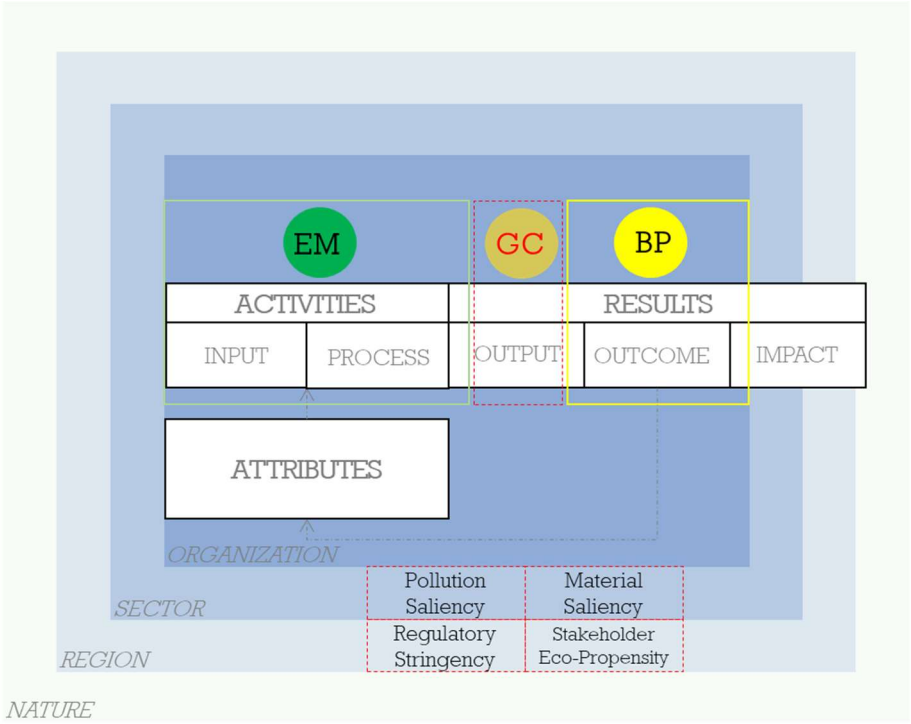

Figure 2.14 Holistic Organizational Performance Framework (HOPF)

HOPF is inspired by the value chain (Porter, 1985), the dynamic theory of strategy (Porter, 1991b), and the logic model (Ebrahim and Rangan, 2014). The overarching proposition of HOPF is that *an organization's performance is the results of its activities enabled and influenced by the attributes in its contexts*. In the short term along the cross-sectional causal chain, organizational performance is the Results of its Activities: (1). Activities on Inputs and Processes are primary value-creating practices; (2). Results include short-term Outputs, mid- to long-term Outcomes, and long-term Impacts with sustained significant effects on root causes; (3). Attributes of the organization's internal and external contexts, which are variables in the long run but may be fixed or sticky in the short term, enable and influence the organization's managerial choices of, and moderate the Results from, its Activities. Over the long run along the longitudinal causal chain, performance Results also feedback to affect the organization's internal Attributes and, to a lesser extent, its external contexts.

HOPF can be used as a benchmark to evaluate theories relevant to corporate sustainability due to its three distinctive features: (1). It deconstructs through an in-built temporal dimension the relationships between active organizational functions of Activities, stable contextual characteristics of Attributes, and time-variant Results of organizational performance; (2). It delineates the closed-loop performance causal chain of Attributes-Activities-Results-Attributes; and (3). It depicts along a spatial dimension the embedded structure of an organization within its sectoral, regional, and natural contexts.

As highlighted in Figure 2.14, the preliminary GC theoretical model proposes a performance causal chain from green Activities (EM) to economic Results (BP) through the mediation of the intangible GC. These Activities reduce negative environmental Impacts on Nature and create positive economic Outcomes for the Organization with moderation from Attributes in the contexts of Region and Sector. At a more granular level, the theoretical model proposes six GC-mediated Activities-Results links moderated by four contextual Attributes. Using HOPF as a benchmark, the model appears to be logically coherent and theoretically plausible.

We next juxtapose and reconcile the model with four prominent theories applied in GC research: the Porter Hypothesis, the resource-based view, the stakeholder theory, and the slack resource hypothesis.

- Porter Hypothesis

The Porter Hypothesis (Porter and van der Linde, 1995) has been a key driving force behind the research stream of *whether it pays to be green* – if it has not single-handedly shifted the environmental-economic discourse from ''Traditionist/trade-off'' to ''Revisionist/win-win'', it has at least largely accelerated the transition by providing a powerful and well-articulated argument. However, although not always made explicit, the Porter Hypothesis is not directly applicable at the organizational level as its unit of analysis is an industrial sector within a specific region, given that environmental regulations are usually enacted by government agencies in a region and imposed on specific sectors. Therefore, the ideal setting to empirically test the hypothesis is to assess the net costs/benefits experienced by the regulated sectors post-strengthening of environmental regulations in a region that is open to global competition – without external competition, there may be a *leakage effect* where the regulated sectors manage to pass the additional compliance costs to customers without being subjected to the expected *pressure effect* that is conjectured by the Porter Hypothesis to drive innovations and eventually deliver net benefits.

According to HOPF, the Porter Hypothesis proposes that changes in the Attributes (strengthened environmental regulations) of an organization's external contexts (Sector and Region) can propel Activities (innovations) that reduce negative environmental Impacts on Nature and deliver positive economic Outcomes (competitiveness) to Organizations. With HOPF as a benchmark, the hypothesis is theoretically plausible, although its focus is at a high level on the contextual antecedents (environmental regulations) to EM rather than its relations with BP. The GC theoretical model can be considered an extension of the Porter Hypothesis, in that it (1). Adapts the hypothesis' macro-level win-win concept of environment-competitiveness to study the phenomenon at the organizational level; (2). Moves forward from the antecedents to the consequences of EM and focuses on its short- to mid-term economic implications along the performance causal chain; (3). Goes beyond the surface and

dives into the G-C neural network to understand different aspects of the EM-BP nexus. To sum up, the GC theoretical model further develops the Porter Hypothesis by providing a richer and more precise language for the Revisionist discourse on the environmental-economic interface of corporate sustainability.

- Resource-Based View

The resource-based view of the firm is another theory frequently applied in GC research. It suggests that a firm's competitive advantages originate from its resources and capabilities that are valuable, rare, imperfectly imitable, and non-substitutable (Barney, 1991). The resource-based view has been extended to the context of the natural environment (Hart, 1995) and refined to account for external dynamics (Teece et al., 1997).

Mapped onto HOPF, the resource-based view can be considered proposing a causal link between an Organization's Attributes and its performance Results. Although this is in line with HOPF's Attributes-Activities-Results-Attributes causal chain, its primary focus on internal Attributes and downplaying of Activities make it rather elusive from an operationality perspective. Organizational attributes are relatively stable and passive characteristics that enable an organization to carry out its intended Activities by providing the required resources and capabilities. Suggesting they are the *causes* that *affect* performance Results would be making quite a leap of logic along the causal chain, bypassing the critical role of Activities that actually make things happen. Therefore, although a prominent theory in strategy, the resource-based view is not directly applicable to explicate the G-C nexus. Since the GC theoretical model is focused on the Activities-Results link at a more granular level, it does not overlap or contradict with the resource-based view.

- Stakeholder Theory

The stakeholder theory suggests that a corporation has relationships with many constituent groups and/or individuals that can affect or be affected by its activities (Freeman, 1984). The instrumental stakeholder theory suggests that the satisfaction of certain stakeholder groups can be instrumental for an organization to achieve its objectives (Donaldson and Preston, 1995). From this perspective, fulfilling the expectations of high eco-propensity stakeholders can benefit a company's economic performance (Busch and Hoffmann, 2011).

Using the language of HOPF, the stakeholder theory proposes that the Attributes (stakeholders) in an organization's contexts (Organization, Sector, and Region) can affect its performance Results. This is in alignment with the Attributes-Activities-Results causal chain of HOPF. However, the stakeholder theory is not as directly applicable to the G-C dyad as it is to the CSR-FP relation, since nature, for the most part, is a voiceless ''stakeholder'' that needs human agents to act on its behalf

(HOPF illustrates this through the indirect interface between Nature and Organization). Therefore, before applying it to the G-C nexus, the stakeholder theory needs to be adapted to reflect the human agents' function on the Organization-Nature relation. Such adaptation is captured in the contextual contingencies of stakeholder eco-propensity and regulatory stringency in the GC theoretical model. Stakeholder eco-propensity reflects the inclination of human agents in the private sector (as customers and investors) and civil society (as environmental activists and the general public) to act on nature's interests with their disposable power and influences. Regulatory stringency reflects the institutionalized interest in and power on environmental protection of the stakeholder group of government agencies, which might reflect the interest and influence of local constituents in democratic nations. To sum up, the stakeholder theory needs to be adapted before applied to the G-C nexus, and its adaptation and application are reflected in the contextual contingencies of stakeholder eco-propensity and regulatory stringency in the GC theoretical model.

- Slack Resource Hypothesis

The slack resource hypothesis posits a reverse causation from organizational slacks to green practices (Waddock and Graves, 1997). Organizational slack is the cushion of actual or potential resources that allows a firm to invest in areas necessary for long-term development but may not pay back immediately (Bourgeois, 1981). The hypothesis suggests that companies with slack resources can afford discretionary expenditures on socially responsible practices, regardless of the payback prospects of these investments.

With HOPF as a benchmark, the hypothesis proposes that organizational Attributes (slack resources) lead to its Activities of green practices. This is in congruence with HOPF's Attributes-Activities-Results-Attributes performance causal chain, although it addresses only the Attributes-Activities link. Therefore, the slack resource hypothesis complements rather than competes with the GC theoretical model that focuses on the Activities-Results link.

- Summary

This subsection evaluates the theoretical plausibility of the preliminary GC theoretical model. According to HOPF, the model proposes that the relations between Activities (EM) and Results (BP) are mediated via GC and moderated by Attributes of an Organization's contexts. The juxtaposition and reconciliation with existing relevant theories suggest that (1). The GC model can be considered the extension of the Porter Hypothesis at the micro and a more granular level; (2). It does not overlap or contradict with the resource-based view; (3). It adapts and applies the stakeholder theory through the contextual moderators of stakeholder eco-propensity and regulatory stringency; and (4). It

complements rather than competes with the slack resource hypothesis. To conclude, the hypothesis evaluation process proves that the preliminary GC model is theoretically plausible.

5.4 Hypothesis Acceptance

Finally, we accept the proposed GC theoretical model as a tentative answer to *why it pays to be green*. Grounded on a large dataset of empirical GC studies, the generation and refinement processes of the hypotheses have iteratively sharpened the model's explanatory power and pragmatic validity. Its theoretical validity has also been proved through juxtaposing and reconciling it with existing relevant theories. However, although developed through a data-driven methodical approach, uncertainty is inevitable in abductively generalized theories. Therefore, the preliminary GC theoretical model needs to be empirically tested for validation in future research.

6. Implications

6.1 Implications on Research

6.1.1 Theoretical Implications

The GC theoretical model illustrates that the environmental-economic interaction is multifaceted and context-contingent. As such, we recommend the following practices in future GC research:

1) **Construct operationalization**: Focus on specific subdimensions with strong theoretical matching and proximal causal connection. The G-C logic chain (Figure 2.6) may be used as a reference framework for this purpose. Applying composite measures for the multidimensional meta-constructs of EM and BP can involve many variables interacting at once and confound the effects under study (this point is illustrated in the study's hypothesis generation process where only two-thirds of the dataset can be used to ground the propositions). Using measures with weak theoretical matching or low causal proximity may lead to obscure results or even invalid findings.
2) **Contextual contingencies**: Design studies within a single sector (region) to limit potential interactions of contextual factors if the key sectoral (regional) contingencies cannot be controlled. Cross-sectoral (cross-regional) studies are only suitable within homogeneous settings.
3) **Methodological artifacts**: Use longitudinal data with various lag lengths between the focal constructs and non-linear models to capture the time-variant G-C co-variations. This study does not identify a generalizable EM payback period – it might be contingent on specific projects – hence experimenting with different lag lengths is recommended.

6.1.2 Methodological Implications

We also coded the research methods and theoretical perspectives of GC studies out of curiosity. The results regarding methodology are as follows: (1). Only four articles in the dataset (2%) used non-quantitative methods (Shrivastava, 1995, Orsato, 2006, Perotti et al., 2012, Primc and Cater, 2015); (2). The rest of the dataset (98%) all used quantitative methods – 99 articles (58%) used econometrics, 37 articles (22%) structural equation modeling, 14 articles (8%) event study, and 17 articles (10%) other quantitative methods such as portfolio study or difference-in-difference. As for theoretical perspectives, 112 articles (65%) were a-theoretical, 31 articles (18%) applied the resource-based view, and the rest of 28 studies were based on theories such as the Porter Hypothesis or the stakeholder theory. In general, the body of work in GC research relies disproportionately on quantitative methods, while as a nascent theory, its conceptual foundation is still weak and wobbly.

Although beyond the scope of this study, the theory-methodology mismatch in GC research seems rather concerning. We do not imply here that quantitative methods are not suitable for individual studies[8] – judging by the fact that they have been accepted for publication, they most likely have been conducted methodologically rigorously within specific research settings. However, at the aggregate level from a knowledge cumulation perspective, this pattern can be problematic – especially if it has persisted over a long period but goes unnoticed and left unchallenged. We believe that such prevalence of theory-methodology mismatch has been a critical issue hindering the progress of GC research despite several decades of ongoing efforts. This opinion is further elaborated based on Figure 2.15 as below.

[8] The critiques on existing methodologies are not an attack launched by arrogance, but distilled learnings from the essay's own rollercoaster journey. The project's original proposal included a systematic literature review, a traditional meta-analysis, an empirical study testing hypotheses with structured datasets, and a multi-case study. However, it got stuck for quite a while stumbling upon the discovery of the frail conceptual foundation and inconsistent data quality in this line of inquiry. Knowing well that an attempt at theory building would most likely be thankless for a junior researcher trying to earn her entry into the community, this unconventional essay is the fruition of confusion, curiosity, naïve bravery, and the gracious support from my advisors. So, in case of unintended offense, please take this as a sincere apology.

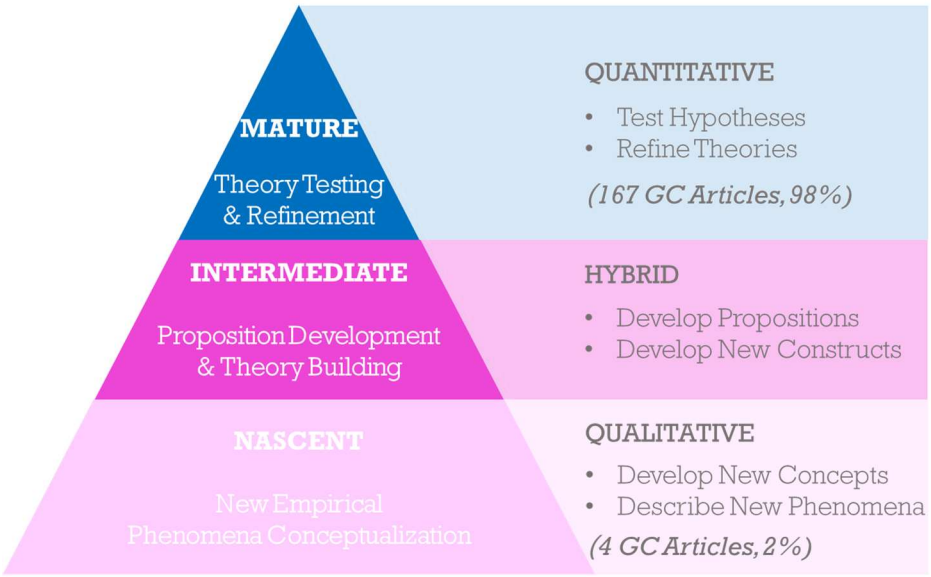

Figure 2.15 Theoretical Development Stages and Research Methods

Generally speaking, the methodological choice of a research area needs to be congruent with its theoretical development stage: (1). For nascent theories, qualitative methods are necessary to describe new phenomena and develop new concepts based on broad empirical and intellectual foundations; (2). Intermediate theories require hybrid qualitative and quantitative methods to develop new constructs, provisional propositions, and preliminary theories (this essay may be taken as an example); and (3). For mature theories, quantitative methods are appropriate to test hypotheses deduced from established theories (Flynn et al., 1990, Meredith, 1993, Edmondson and McManus, 2007). Following such a trajectory of theoretical development, the cumulated works in a healthy research field shall form into the shape of a pyramid, as illustrated in Figure 2.15. However, mapping the GC articles onto the structure would only get 2% of qualitative research at the base, carrying the weight of 98% of quantitative studies at the top with almost nothing in between.

As a collective body of knowledge, such a structure is rather unhealthy. The richness and complexity of the environmental-economic intersection of corporate sustainability are not yet fully understood and clearly described – bypassing the groundwork of building a solid conceptual foundation in favor of quantitative studies is no different from laying bricks for a treehouse on a trunk that is weak and shaky. From a knowledge cumulation perspective, if each published GC article is taken as a vote jointly cast by its author(s), editors, and reviewers, for several decades, the community has collectively overwhelmingly voted for quantitative methods. Imagine, for the lack of a better analogy, what it would be like if most employees in a highly autonomous company appoint themselves for the prestigious and neat position of financial controller and shun from the humble and messy work of administration and operations – obviously, such a hypothetical company would not be able to survive in a competitive market. However, such practices seem to have become the norm within, and perhaps far beyond, the academic community of GC research. Some changes perhaps are needed.

Hopefully, going forward, methodological rigor would not be reduced to numerical precision, and the value of conceptual rigor would be better recognized.

6.2 Implications on Practices

6.2.1 A Corporate Perspective – GC Five-Forces Framework

Figure 2.16 shows a GC five-forces framework for corporate environmental strategy planning. The framework is developed by applying the GC theoretical model through Porter's five-forces framework of competition (Porter, 1979). Managers may use it to scan their operating contexts and plan strategic EM priorities accordingly.

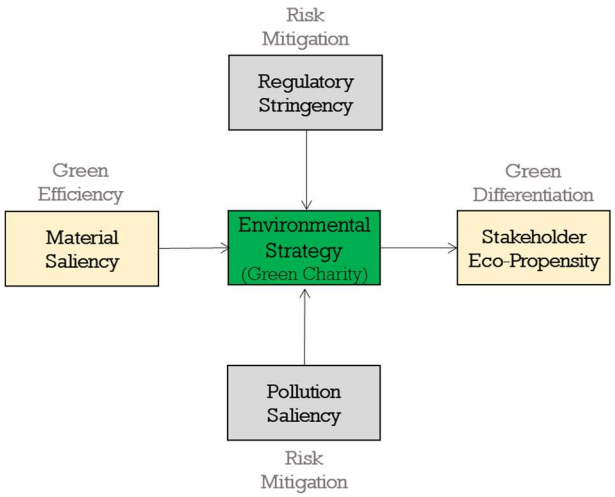

Figure 2.16 GC Five-Forces Framework

Risk mitigation should be the first priority in environmental strategy planning. This is done by analyzing the environmental liability risks facing a company at present or in the foreseeable future to determine the courses of action. The level of such risks is assessed through two contextual factors: pollution saliency and regulatory stringency. When pollution saliency is low, the firm's potential liability risks are likely to be low either due to low probabilities of violation or the absence of regulation. When pollution saliency is high, (1). If environmental regulations are stringent, pollution reduction shall be of the highest priority (H1a); (2). If regulations are not strict (do not exist) currently but may be tightened (enacted) in the foreseeable future, plans should be developed to reduce pollution output before the projected regulation strengthening (enactment); (3). Otherwise, it is managers' discretion to either invest in voluntary pollution reduction (H1b) or allocate the resources elsewhere when organizational slack is limited.

The next step in environmental strategy planning is to identify green efficiency opportunities. Material reduction is an area most companies can exploit for cost reduction – there are likely some

low-hanging fruits somewhere for quick-wins, such as better housekeeping and process improvement. However, when material saliency is high, material reduction may involve production equipment replacement, material substitution, or business model modification. Such EM practices may require longer planning horizons and higher capital commitments, hence may not lead to net benefits of cost reduction in the short term (H2).

After that, firms may want to evaluate the possibilities of green differentiation. Green differentiation is an area that fewer can afford to explore – the planning horizon is much longer, and the payback prospect is rather uncertain as it relies on external mechanisms for its realization. If a company has the resources and capabilities to design green products with added customer values, it will be better positioned to deploy the strategy more easily and reap the returns more readily (H3a). On the other hand, developing green products without added customer values can be risky since they will have to leverage customers' eco-propensity to penetrate a niche green market (H3b). Therefore, the environmental strategy of green differentiation requires innovation capabilities, sufficient resources, and an appetite for risks.

Finally, companies may want to engage in green charity. Together with other EM practices, green charity can improve a company's corporate image, green reputation, and market valuation (H4). However, if the involvement in green charity is just window dressing for positive publicity, while the interior aspects of EM are neglected, such inconsistencies may backfire at some point if they are exposed and perceived as greenwashing and hypocrisy. Therefore, a green reputation will be more solid if it is built up naturally over the long run as the by-product of a company's authentic identity, rather than the designed outcome of deliberate impression management through superficial ''green charity'' or selective disclosure about corporate sustainability.

To sum up, the GC five-forces framework suggests that corporate environmental strategy planning should prioritize risk mitigation, exploit green efficiency wherever possible, explore green differentiation when affordable, and let green reputation grow naturally from its authentic identity. The framework demonstrates that the GC theoretical model can be applied at the corporate level to guide environmental strategy planning and prioritization. Given that *there is nothing more practical than a good theory* in social science, it also serves to prove the pragmatic validity of the GC theoretical model.

6.2.2 A Societal Perspective – Roadmap towards Green Market

From society's perspective, the GC theoretical model is applied to chart the roadmap towards a green market. Figure 2.17 shows that there are three prerequisites to harness the force of GC for

sustainability: green performance measurement (GPM), green performance valuation (GPV), and green value apportion (GVA).

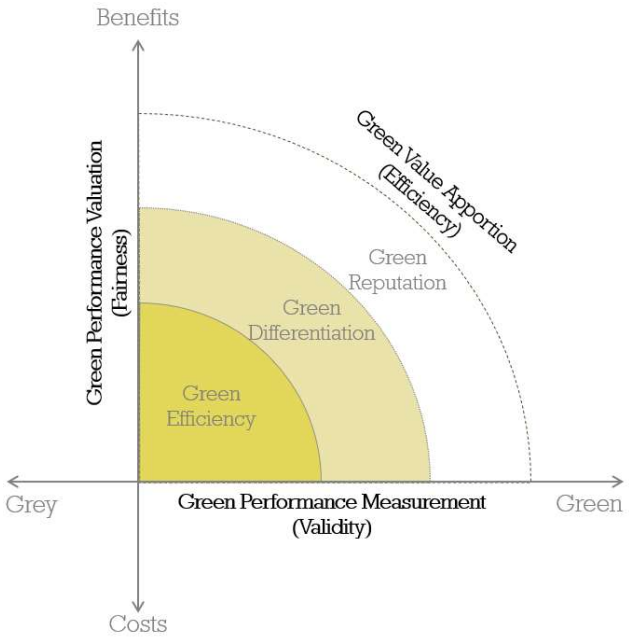

Figure 2.17 Roadmap towards Green Market

The first prerequisite for a green market is valid GPM. GPM is fundamental to the endeavor as *we cannot manage what we do not measure*. GPM is technically complex, but many efforts have been devoted to this area – one example is the UN's System of Environmental-Economic Accounting, a statistical system that measures at the macro level the interactions between the economy and the natural environment. The current challenge is not necessarily to develop new metrics but to reconcile these disparate developments into a standardized interdisciplinary GPM framework applicable at the organizational level.

GPV is another prerequisite for a green market. GPV converts EI caused by business activities into equivalent economic values, such as the monetization of GHG emission costs. It is a subject beyond management or environmental economics that involves ethical and philosophical debates – e.g., it would be controversial to monetize health impacts from air pollutions differently in developed and developing nations. Therefore, the fairness and equity principles need to be explicitly addressed besides the technical aspects of GPV, taking into account relevant factors in the contexts.

Finally, efficient GVA is another prerequisite to growing a green market. GVA refers to the apportion and allocation to responsible parties the commensurate economic values of environmental impacts caused by business activities. GVA is a multifaceted issue beyond the managerial domain involving sociopolitical and technological challenges. It may be accomplished through non-market regulatory mechanisms (such as environmental taxes and subsidies) or the market mechanism of GC

(such as green differentiation and market valuation). For a GC market that trades public goods (environmental values) and private gains (economic values) to function efficiently, the number of potential buyers and suppliers needs to be sufficiently large for the transaction costs to be affordable and the exchanges scalable, besides a governance system safeguarding the interests of both parties – perhaps a new social contract built on transparency, trust, and goodwill.

In summary, the roadmap towards a green market shows that despite the challenges, GC can be leveraged as a force for environmental sustainability through valid GPM, fair GPV, and efficient GVA. The roadmap also proves the GC theoretical model's pragmatic validity by demonstrating its application to build a context conducive to corporate sustainability.

6.2.3 An Institutional Perspective – Hierarchy of Green Forces

The hierarchy of green forces (Figure 2.18) puts GC in a broader institutional context and explicates the structure and functions of different green motivations. The hierarchy is built on the GC theoretical model, the hierarchy of needs (Maslow, 1943), and the two-factor motivation model (Herzberg, 1968). It suggests that there are four types of driving forces towards sustainability: normative, regulatory, rational, and moral.

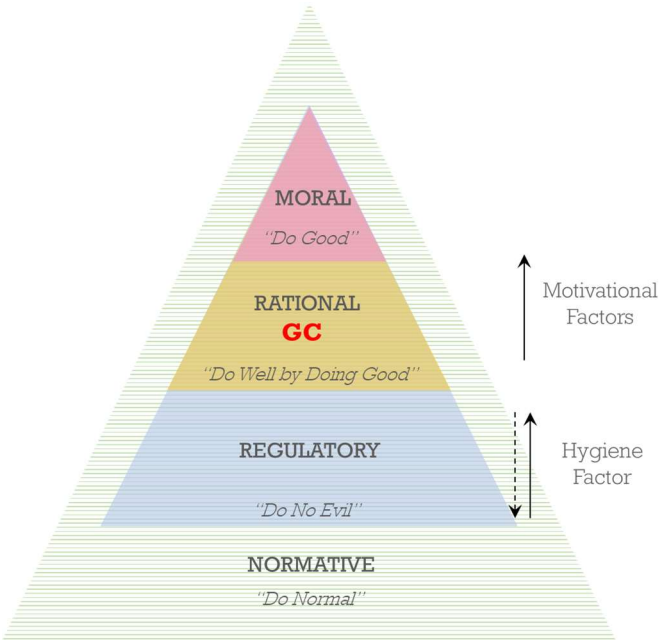

Figure 2.18 Hierarchy of Green Forces

The normative force towards green ("do normal") originates from civil society. It is a conforming force that normalizes and institutionalizes pro-environment values, beliefs, attitudes, and behaviors. Compared with other forces, its influence is more subtle and less straightforward – it works as a form of "informal regulation" that makes people, on their own will, want to do what they perceive society expects them to do, i.e., do what is "normal". Once formulated and institutionalized, the

normative force is ingrained in the collective mentality of the citizens and becomes part of the culture, exerting far-reaching and long-lasting influences through its manifestations in different institutions in society – e.g., it may manifest as the regulatory force where government agencies enact stringent environmental regulations in response to the expectations of high eco-propensity constituents. Hence, fostering and institutionalizing the normative force for green is the fundamental aspect to bring about substantial long-term changes for environmental sustainability. The normative force can be related to stakeholder eco-propensity and regulatory stringency in the GC theoretical model.

The regulatory force (''do no evil'') is primarily a repressive force with threatened disciplinary actions against violation of environmental regulations. The force has the enforcement power to *alleviate* environmental unsustainability by *punishing grey* through mechanisms such as fines, penalties, or lawsuits. It is not a motivational but a hygiene factor – it creates ''deficiency needs'' in the regulated when the level of compliance is inadequate (''dissatisfaction'', liabilities and risks); and adequate compliance reduces but does not increase motivation (''satisfaction'', economic returns). Similar to physiological needs, the regulatory force takes precedence and shall be prioritized over other higher-level needs. It can be related to regulatory stringency and mandatory pollution reduction in the GC theoretical model.

The rational force (''do well by doing good'') drives firms to carry out mutually beneficial green practices. It is a force that *elevates* environmental sustainability by *rewarding green* and a motivational factor for ''growth needs'' – meeting such needs increases motivation and leads to a self-reinforcing virtuous cycle. The rational force is essentially GC, the focus of this study.

The moral force (''do good'') comes from within and is the intrinsic motivation to protect the natural environment. The influence of the force is exerted through a company's key internal stakeholders, such as the board or executives. The moral force drives the firm to engage in green charity, e.g., giving up profitable opportunities with high EIs or sharing with competitors green best practices. Although it may not be intended, economic benefits may accrue to the company through its green reputation and stakeholder goodwill. The moral force is similar to the need for self-actualization and self-fulfillment in Maslow's needs hierarchy. It can be related to voluntary pollution reduction and green charity in the GC theoretical model.

In summary, the hierarchy shows that none of the forces alone is sufficient, but all of them are necessary, to build a context conducive to sustainable development: (1). The regulatory force can be a powerful impetus to generate a level of traction that exceeds resistance to change; (2). The normative and moral forces should be fostered for self-initiated motivations towards sustainability; (3). The rational force of GC needs to be harnessed to create self-reinforcing virtuous cycles of green rewards. The brief analysis of different green forces demonstrates the multidimensionality of

sustainability and suggests that research in this area needs to take an interdisciplinary approach. It also implies that sustainability research shall not and cannot be entirely value-free – not with its scientific criteria but its motivation and application, as this study strives to accomplish – given its normative character and inherent value dependency.

Finally, a revisit is probably warranted for the current narrative of sustainability. The commonly accepted definition of sustainability, or sustainable development, is "*development that meets the needs of the present without compromising the ability of future generations to meet their own needs*" (The Brundtland Report, p. 37). Put it bluntly in an unpolished way: *exploiting nature is fine so long as current exploitation does not compromise future opportunities in so doing*. The not-so-implicit instrumentality underpinning the narrative is that the value of nature is conditioned on its economic utilities to meet the needs of humans, not necessarily in and of itself. Such an anthropocentric conception of nature, together with capitalism and technological advancements, has forged an institutional setting where the human species possess enormous power that is essentially unregulated – either through morality internally or legislation externally – to use and abuse the natural environment. Hence, sustainable development probably needs a different narrative based on a reconceptualized human-nature relation. The new narrative appreciates nature unconditionally, respects its well-being as its conscious experience, and replaces egotistical ethics towards future generations with ecological ethics towards nature at present. Such an anthropomorphic perspective can probably better turn sustainability from an ideal into a reality.

7. Conclusions and Future Research

- Contributions

This essay addresses the questions of *when/why it pays to be green*. Its primary contribution is a preliminary GC theoretical model that explicates the environmental-economic interactions of corporate sustainability. The model comprises six moderated-mediated hypotheses and is built upon a green typology, a new construct of GC, and four concepts of key contextual contingencies – all developed in this study. The theoretical validity of the model is proved by contrasting and reconciling it with existing relevant theories. Together with its supplementary components, the preliminary theoretical model elevates GC research from a nascent to an intermediate theory.

We recommend the following practices in future GC research: (1). Focus on specific G-C subdimensions with strong theoretical matching and proximal causation to reduce confounding effects and improve validity; (2). Conduct studies within a single sector (region) if the key sectoral (regional) contingencies cannot be controlled; (3). Use longitudinal data with various lag lengths and

non-linear models to capture time-variant G-C co-variation. We also report the prevalence of theory-methodology mismatch in this body of work and propose balancing numerical precision with conceptual rigor in future GC research.

The GC theoretical model is then applied to design a GC five-forces framework to support corporate environmental strategy planning, a roadmap towards green market for implications from society's perspective on creating a context conducive to corporate green practices, and a hierarchy of green forces for an institutional view on the structure and functions of different green motivations. These applications prove the pragmatic validity of the GC theoretical model.

The secondary contribution of the essay is that it also addresses three research gaps in the process of developing the theoretical model. First, scholars have highlighted "*the complex nature of the link between environmental and financial performance* (Corbett and Klassen, 2006)" and advised that "*the key to fully addressing issues of business and the natural environment lies in treating its extant questions holistically*" (Hoffman and Bansal, 2012). However, prior to this study, a unifying conceptual framework did not incorporate the different G-C subdimensional relations (Trumpp and Guenther, 2017). The G-C logic chain closes the gap as a holistic framework that can integrate the different aspects of the "empirical elephant" of GC research into a coherent structure.

Second, scholars have also pointed out that the lack of consistency in construct operationalization hinders the progress of GC research (Busch and Hoffmann, 2011, Russo and Minto, 2012). This gap has been narrowed by the extensive green and competitiveness measures (Appendix II) compiled through this study's grounded approach. These measures can be used to support consistent construct operationalization in future GC research.

Third, HOPF developed to evaluate the GC model's plausibility can potentially address the theoretical gap in sustainability research. Starik and Kanashiro (2013) pointed out that "*sustainability management appears to require one or more dedicated theories because no other theories of management appear to have expressly included attention to human individuals, organizations, and societies and multiple other systems and their mutual embedding with the natural environment… traditional management theories are virtually silent on (these aspects)*" (p. 13). HOPF depicts and explicates the causal chain of corporate performance along the temporal dimension, and the embeddedness of organizations within sectoral, regional, and natural contexts along the spatial dimension. Therefore, it has the potential to be further refined into a fully-fledged theoretical model for corporate sustainability research.

- Limitations

The key limitation of this essay is that its proposed theory is built upon a secondary dataset of published academic articles. This might pose threats to validity, exposing it to the *drawer effect* of publication bias and the *streetlight effect* of observational bias. The potential publication bias is reduced by focusing more on negative research findings in the inference processes, and the potential observational bias is mitigated by using abduction that goes beyond observed empirical evidence. The theoretical validity of the proposed GC model has been proved through the hypothesis evaluation process and its pragmatic validity through its applications at the corporate and societal levels.

Furthermore, the methodological choice of abduction for theorizing might present threats to reliability. A certain level of uncertainty is inherent in abductively inferred hypotheses since they are not strictly based on observed effects or deducted from well-established theories, but also make use of relevant theories, intuition, and disciplined imagination that cannot be fully articulated and codified. This is not optimal, but inevitable, given that abduction is the best available method for theorizing the multidimensional G-C interactions – existing empirical studies do not cover all aspects of the rich and complex phenomenon for inductive generalization, and there are no established theories on this subject for deductive reasoning. This study has endeavored to improve the reliability of the proposed theory by (1) assuring falsifiability of the hypotheses by documenting the theorizing processes methodically and making explicit where logical inferences might go adrift, and (b) enabling critical assessment of confidence levels in the model by presenting empirical data alongside the hypotheses and highlighting the areas where evidence might be weak or gaps might exist. Nevertheless, although threats to reliability from abduction have been minimized with these efforts, the proposed GC theoretical model is but preliminary and needs to be empirically tested in future research.

- Future Research

Future GC research can focus on three areas that are practically relevant and intellectually stimulating. First, empirically test the propositions of the preliminary GC theoretical model – the findings would ideally not just stand alone within specific settings but could also feed into the holistic structure to refine the GC theoretical model. Second, develop measurement items and scales for the intermediate construct of GC. Third, design an interdisciplinary GPM framework as the foundation for the green market to harness the force of GC for sustainability.

Chapter 3 Sustainable Humanitarian Supply Chains

1. Introduction

- **Background and Motivation**

Four European countries recently set high temperature records in July 2019 (Climate Change Service, 2019). Climate change has led to more frequent and severe natural disasters (ReliefWeb, 2019). Pursuing the United Nations Sustainable Development Goals (SDGs) "*are not optional, they are about survival*" (Van Wassenhove, 2019). Sustainability concerns are particularly pertinent to humanitarian organizations as they operate in vulnerable regions (Pedraza-Martinez et al., 2011) with budget constraints (OCHA, 2018, Besiou and Van Wassenhove, 2020).

Yet, research on sustainable humanitarian operations is scarce. Supply chain management offers a systematic approach to operationalizing sustainability considerations in humanitarian operations (Starr and Van Wassenhove, 2014). It is also an area with significant cost impact in practice as more than 70% of funding is allocated to supply chain expenditures in disaster responses (HELP Logistics et al., 2018). However, current supply chain theories and methods are built in the commercial context and cannot be directly applied to the humanitarian sector since the latter differs significantly from the former (Van Wassenhove and Pedraza Martinez, 2012). This study addresses the gaps through a strategic supply chain framework for sustainable humanitarian operations.

- **Research Positioning**

As shown in Figure 3.1, the essay is positioned at the intersection of humanitarian operations, supply chain management, and operations strategy in the broader context of sustainable development. Specifically, this study develops a framework to support humanitarian supply chain strategy planning aligned with the sustainable development goals. It applies design thinking to create practically relevant, solution-oriented, generalizable knowledge in this nascent area of interdisciplinary research.

Figure 3.1 Research Positioning of Essay 2

- **Definitions and Terminology**

As the means to the end of sustainability, sustainable development is "*development that meets the needs of the present without compromising the ability of future generations to meet their own needs*" (The Brundtland Report, 1987, p. 37). We define sustainable humanitarian supply chain (HSC) management as *a systematic approach to designing, organizing, and managing key intra- and inter-organizational processes to reduce vulnerabilities of certain populations by fulfilling their immediate needs and supporting them in building long-term prosperity prospects with minimum possible economic and environmental resources.*

2. Review

2.1 Humanitarian Operations

Many humanitarian organizations have dual missions – disaster responses involving emergency relief and development programs with steady-state operations (Besiou et al., 2014). Such dual tasks may be referred to as "service mix" as the concept of "product mix" in business. Disaster responses are often rife with urgency, unpredictability, and complexity: urgency and unpredictability are contingent on the timing, type, location, scale, and severity of disasters; complexity results from urgency and unpredictability compounded by different external and organizational factors. Given the challenges posed to sustainable supply chain strategies, discussions on humanitarian operations hereafter will focus primarily on disaster responses.

- Urgency

High urgency is the dominating characteristic of most disaster responses, especially at the emergency relief stage. Speed often becomes the overriding operational objective, particularly in the first 72 hours right after the onset of disasters, in order to maximize life-saving opportunities for the affected populations (Tomasini and Van Wassenhove, 2009). Any delay in disaster responses could exponentially drive up deprivation costs – the proxy economic value of human suffering caused by a lack of essential life-sustaining goods or services (Holguín-Veras et al., 2013).

- Unpredictability

High unpredictability in disaster responses manifests on both the demand and supply sides of HSCs. On the demand side, the severity of damage and the demand for aid are mostly unknown until after initial post-disaster assessments (Balcik and Beamon, 2008). On the supply side, depending on the scale of disaster and media attention, a variety of organizations are often involved without effective coordination (Acimovic and Goentzel, 2016), making it challenging for any humanitarian organization to assess the needs for and potential effects of its response within the collective efforts. Improved forecasts can reduce unpredictability to some extent in natural disasters' occurrence and impact (e.g., tropical cyclones' overall frequency and intensity can be forecasted, and disaster profiles of different regions can be outlined). Unpredictability in post-disaster demand can be reduced by estimating the required quantities for essential items based on experience and expertise (Tomasini and Van Wassenhove, 2009). However, although reduced unpredictability can be helpful at the aggregate level (e.g., improving the estimation of regional prepositioned safety stocks), its usefulness at the initial emergency relief stage is rather limited when forecast errors are not pooled geographically.

- Complexity

Complexity is the result of urgency and unpredictability of humanitarian missions compounded by various contextual factors. First, unlike companies who can cherry-pick desirable markets, disaster responses often take place in harsh contexts involving scattered last-mile distribution points, underdeveloped or damaged infrastructure, and even turmoil and conflicts. Furthermore, supply chain planning problems encountered in different crises can vary significantly from one another (Starr and Van Wassenhove, 2014): responses to epidemics (e.g., Ebola) differ from that to earthquakes (e.g., the Sulawesi earthquake in 2018) or long-term conflicts (such as the Syria crisis). It is challenging to establish and manage new HSCs in response to different crises in different places (Van Wassenhove, 2019).

Second, from a sustainability perspective, humanitarian operations need to balance the divergent expectations of different stakeholders, including donors, beneficiaries, and nature. Without

considering potential geopolitical agendas (which is beyond this study), donors generally value economic efficiency without compromising the objectives of saving lives, easing pain, and improving life quality (Pedraza-Martinez and Van Wassenhove, 2016, Kretschmer et al., 2014). In disaster responses, beneficiaries value supply speed and sufficiency as human suffering at emergency relief could increase exponentially as time passes by without essential life-sustaining items (Holguín-Veras et al., 2013); in development programs, beneficiaries value self-reliance and prospects of long-term prosperity (Starr and Van Wassenhove, 2014). Nature values sustainable consumption of renewable resources, minimal consumption of non-renewable resources, and minimal pollution and wastes. The values of these key stakeholders only partially overlap and can pull humanitarian organizations in different directions.

Third, some organizational characteristics of humanitarian organizations can complicate their operations. Unlike businesses, humanitarian organizations do not have a profit-seeking agenda or a command-and-control structure. As a result, bureaucracy may impose excessive controls and slow down operations (Tomasini and Van Wassenhove, 2009). Furthermore, personnel turnover is often high, making organizational learning ineffective and knowledge cumulation difficult (Van Wassenhove and Pedraza Martinez, 2012).

In sum, humanitarian operations are challenging due to urgency, unpredictability, and complexity in the operating contexts. Supply chain management methods and models developed in the commercial context need to be adapted before being adopted in the humanitarian sector.

2.2 Sustainable Humanitarian Operations

SDGs "*are about survival*" (Van Wassenhove, 2019). Yet, it is challenging to incorporate sustainability considerations into humanitarian practices, and research on this topic is still a nascent area of study.

- Economic Sustainability

It is commonly accepted that in disaster responses, pursuing economic efficiency is subordinate to easing human pain and saving lives (Gralla et al., 2014, Haavisto and Kovács, 2019). Such perspective is rarely challenged, probably because of its moral backing in the "*Rule of Rescue*" which proposes that society has a moral obligation to rescue individuals facing avoidable peril, even if so doing involves high costs and inefficiencies (Jonsen, 1986). Although few would argue against applying the "*Rule of Rescue*" when resources are abundant, it is of little guidance when resources are scarce (Holguín-Veras et al., 2013) – which is often the case when considering in full the geographical scope and temporal horizon of humanitarian missions. In resource-constrained situations, the neglect of economic efficiency in some regions (at the early stage) can negatively affect

resource availability for other regions (at the later stage), leading to cross-regional (inter-temporal) opportunity costs and potential social inequity. Therefore, the design of SHSCs needs to adopt a holistic view taking into account a broad coverage both geographically and temporally.

- Environmental Sustainability

The negative environmental impacts of humanitarian operations can be significant (Pedraza-Martinez et al., 2011). Besides natural resource consumption and pollution generation, disaster responses often occur in areas that are already fragile. Neglecting these negative environmental impacts can lead to a vicious cycle in between, increasing the vulnerability of affected regions, and intensifying subsequent damage afterwards (Sodhi, 2016). Yet, research in green humanitarian operations is relatively scarce (Sarkis, 2012), with a few notable exceptions that investigate material convergence effects (Holguín-Veras et al., 2012a), beneficiary livelihood greening practices (Haavisto and Kovács, 2014), and the consequence of cumulated wastes (Humanitarian Research Group, 2011). Green products in humanitarian supplies are rarely addressed, while considerations of transportation emissions in humanitarian logistics just start to emerge.

In general, research on sustainable humanitarian operations is still at an early stage. Therefore, before designing the SHSC framework, this essay will first build an elementary theoretical base by conceptualizing sustainability in the humanitarian context and clarifying the interrelations among its subconstructs.

2.3 Sustainable Humanitarian Supply Chains

Scholars have advised taking a supply chain perspective as a critical first step to study the interface between sustainable development and operations management (Kleindorfer et al., 2005). Humanitarian practitioners have also recognized the essential role of supply chain management in delivering goods and services to beneficiaries (Besiou and Van Wassenhove, 2020). A supply chain approach provides a basis for solid analyses to replace well-intended intuition with a systematic perspective that accounts for all relevant aspects in humanitarian operations (Starr and Van Wassenhove, 2014). Such a systematic approach can detect partial solutions that displace rather than resolve the multifaceted sustainability issues, expose silo effects, pinpoint synergy areas, and identify optimal systemwide supply chain configurations. Therefore, supply chain management can serve as an anchoring point to integrate sustainability concerns in humanitarian operations.

Supply chain strategy is a critical starting point for the design of SHSCs since it deals with decisions with long-lasting effects on organizations (Simchi-Levi et al., 2004). Research on supply chain strategy is rich. Fisher (1997) proposed a framework for choosing the right supply chain strategies based on product types (''efficient'' for ''functional'' vs. ''responsive'' for ''innovative''). Lee

(2002) further developed the framework into a two-dimensional 2×2 matrix to align supply chain strategies with demand and supply uncertainties (''efficient'', ''responsive'', ''risk-hedging'', and ''agile''). Christopher and Towill (2001) recommended integrating lean and agile hybrid supply chain strategies based on the contexts. Simchi-Levi et al. (2014) advised managing unpredictable supply chain disruptions by identifying and mitigating low-probability and high-impact risks.

Despite the enlightening insights offered by these studies, they all focused on commercial supply chains. The supply chain strategies they proposed based on product- and profit-related characteristics are not directly applicable to humanitarian operations – humanitarian organizations do not make products or chase profits, and they compete with rivals in different ways. Furthermore, organizing completely new supply chains for disaster responses is more than just another application of existing theories since the contexts differ significantly from commercial operations (Van Wassenhove, 2019). Hence, sustainable HSC strategies need to fit the peculiar humanitarian operations contexts.

3. Method

3.1 Research Model

The study adopts the strategy-structure-capabilities-performance framework (SSCP) (Chen et al., 2009) as its conceptual foundation. SSCP combines the resource-based view with the strategy-structure-performance framework: the latter suggests that for an organization to perform well, its strategies need to fit its contexts, and its structures fit its strategies (Miles, 1978); the former proposes that an organization can obtain competitiveness by developing valuable, rare, and inimitable resources (Barney, 1991). Built on both, SSCP posits that for an organization to achieve its desired performance, it needs to develop capabilities that fit its contexts to enable its strategy and structure (Chen et al., 2009). In the context of this study, SSCP implies that for a humanitarian organization to achieve its sustainability performance objectives, it needs to develop context-congruent supply chain capabilities that enable its intended supply chain strategy and structure.

Figure 3.2 shows the study's research model based on SSCP and guided by design thinking. The research process takes four steps: Step 1 defines the performance objectives of SHSCs; Step 2 identifies context-contingent supply chain capabilities; Step 3 distills practical SHSC tactics; Step 4 formulates the strategic framework for SHSCs.

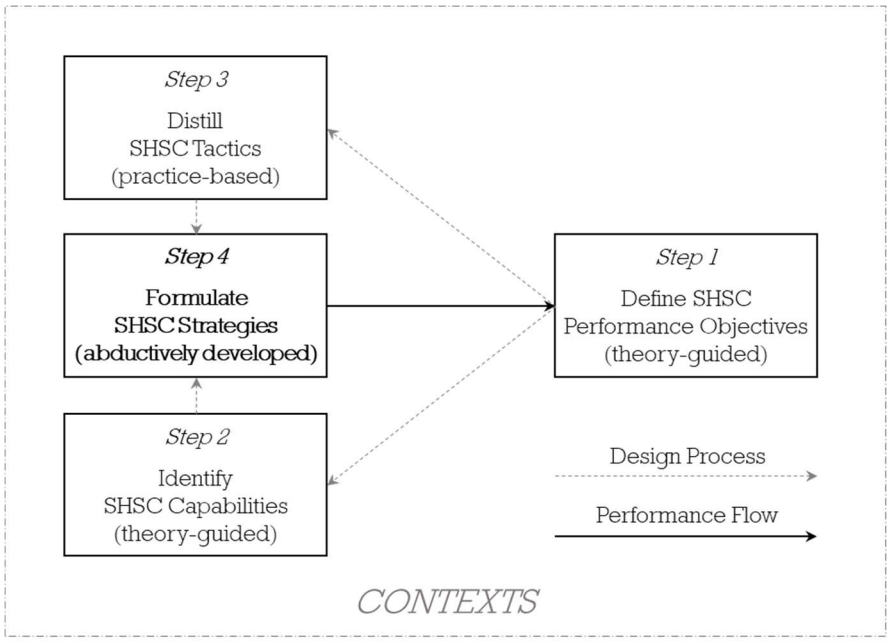

Figure 3.2 Research Model

3.2 Research Method

Design thinking is chosen as the method given the study's solution-oriented nature and the currently nascent theoretical development of this topic. Design thinking can be considered a reduced version of the design science method that develops a generic solution aimed at practical problem-solving (Van Aken et al., 2016) without real-life implementation. It may also be seen as a management engineering approach, adapting relevant theories and methods in innovative ways to design generalizable solutions for practical problems (Corbett and Van Wassenhove, 1993).

Beckman and Barry (2007) propose a four-phase process to develop solutions with design thinking: (1). Observing to gain a deep understanding of the problem's contexts; (2). Framing to identify what is missing and specify the problem that needs solving; (3). Deriving imperatives as guidelines of the intended solutions; (4). Designing solutions according to the imperatives. Observing takes place in Section 2 in the review of the contexts of humanitarian operations. Framing, also in Section 2, identifies the practical challenges and research gaps in sustainable humanitarian operations, and the need for an SHSC framework. The imperatives will be specified through three steps: Step 1 defines the performance objectives of SHSCs, Step 2 identifies the required supply chain capabilities, and Step 3 elicits practical tactics from sustainable humanitarian operations. Finally, these imperatives will guide the design of the solution, a strategic framework for SHSCs.

3.3 Research Process

The design imperatives are elicited through three steps. Step 1 defines the performance objectives of SHSC by adapting to the humanitarian context existing sustainability-related theories.

Step 2 infers from relevant theories deductively to identify context-contingent supply chain capabilities. This step makes visible the *enabling* layer underneath the solution to be designed, in the sense that capabilities enable strategies. Step 3 makes visible the *enabled* layer above, in that strategies enable tactics. Based on the author's practical experiences, it decomposes the structural components and operational elements of HSCs, enumerates alternative options for each operational element, evaluates each option's performance impact to map out the configuration with higher systemwide performance, and generalizes practical SHSC tactics from emerging patterns in the operational choices[9].

After that, the design prototype is unveiled by putting together the lower layer of enabling capabilities and the upper layer of enabled tactics – as illustrated in the research model in Figure 3.2. The formulation of the solution, an SHSC strategic framework, uses abduction to reason backwards and "*infer from effects to causes*": (1). The *causes* are the SHSC strategies; (2). The *effects* are the required SHSC capabilities, the enabled SHSC practical tactics, and the intended sustainability performance objective. This non-linear approach is similar to the emergent, as opposed to the deliberate, strategic planning process in ambiguous and dynamic situations (Mintzberg, 2000).

4. Design

4.1 Defining Performance Objectives

Corporate sustainability is often conceptualized as Triple Bottom Line (TBL), incorporating the economic, environmental, and social dimensions (Elkington, 1994). TBL has been effective in communicating the abstract concept of sustainability to practitioners using intuitive and relatable language. However, we argue that the TBL metaphor does not apply to humanitarian organizations – they have only one Single Bottom Line (SBL), and this shall be the performance objective of the designed SHSCs. This view is elaborated as below.

- **SBL with Economic Primacy in the Corporate Sector**

From the perspective of for-profit corporations, the three TBL dimensions are not homogenous elements of corporate sustainability performance. Profits and losses on environmental and social

[9] This process can be considered similar to lifecycle analysis in product design, where product engineers assess the aggregated environmental impacts of alternative design blueprints with different parts and subassemblies in order to identify optimal product configurations.

aspects are a means to the end of economic returns, since traditional financial accounting – the primary base for measuring corporate performance – books only environmental and social impacts that are internalized financially. The corporate SBL perspective, simplified yet realistic, differs from the conception of TBL sustainability performance with a normative connotation where externalities are considered but not practically internalized. In this sense, TBL in the corporate sector remains a concept, while SBL is a reality – an SBL with economic primacy.

- **SBL with Social Primacy in the Humanitarian Sector**

The TBL concept, however, does not apply to humanitarian organizations as they are non-profits. Humanitarianism, per se, is about *doing good*, implying that the normative performance objective for HSCs is to create and deliver social value. We define social value in the humanitarian context as *value created from humanitarian interventions through relieving suffering, saving lives, improving life quality, and increasing prosperity prospects of vulnerable populations.* Social value in the humanitarian setting can be analogized as gross revenue in the corporate context – it is therefore referred to as social revenue (SR) hereafter. Underneath this ''top line'' of a humanitarian organization's ''sustainability income statement'' are the ''middle lines'' of economic costs (EC_1) and environmental costs (EC_2) – and the bottom line is the balance after deducting EC_1 and EC_2 from SR, referring to as net sustainability value (NSV) hereafter.

NSV measures sustainability performance in humanitarian operations ($NSV = SR - EC_1 - EC_2$), and maximizing the value of NSV shall be the performance objective of an SHSC. As such, NSV is the SBL in the humanitarian context – an SBL with social primacy.

- **NSV as the Performance Objective of SHSCs**

Measuring NSV quantitatively is challenging since both SR and EC_2 are context-specific and hard to quantify. Yet, NSV as a qualitative indicator is still helpful for the design of SHSCs. First, as an all-encompassing composite measure that fully captures the dynamic relations among the three sustainability dimensions, it can avoid partial assessment of sustainability performance in humanitarian operations. E.g., NSV may indicate that the *firefighting* mode in emergency relief with an *as-much-as-possible-as-soon-as-possible* approach is unsustainable when the expended EC_1/EC_2 combined exceed the SR value created. Second, NSV provides a clear baseline to measure sustainability performance impacts of supply chain operational options qualitatively. Therefore, defining NSV as the performance objective is a critical first step in the design of SHSCs to enable the *as-much-as-necessary-as-fast-as-necessary* approach.

4.2 Identifying Required Capabilities

This step identifies the supply chain capabilities needed to enable sustainable humanitarian operations strategies. Capability is *the condition of having the capacity to do something*. The SSCP conceptual framework suggests that capabilities need to be congruent with contexts to enable intended strategies. The contexts of humanitarian operations may be examined from geopolitical, sociocultural, or organizational dimensions. This study analyzes the contexts based on the urgency and unpredictability (U2 hereafter) of humanitarian missions, as they are the key characteristics of disaster responses that pose challenges to the design of SHSCs.

- **Disaster Response Lifecycle and Critical Success Factors**

Figure 3.3 shows three typical stages in the lifecycle of disaster responses: Emergency Relief, Prolonged Relief, and Recovery & Rebuild. In the three stages, the level of U2 in the operating contexts varies significantly.

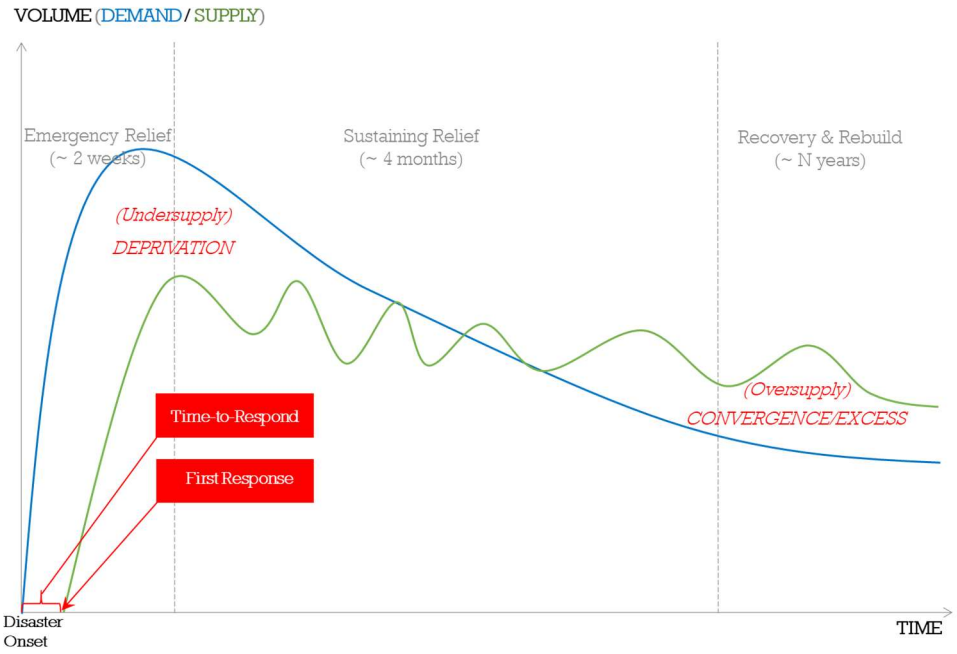

Figure 3.3 Lifecycle of Disaster Response

Emergency Relief is the initial stage of humanitarian response right after the onset of a disaster. It may take roughly two weeks on average (Besiou et al., 2014). The level of U2 is high at this stage (e.g., earthquake, flood, or sudden outbreak of civil conflicts). In this context, the critical supply chain capability is to deliver sufficient supply swiftly, as undersupply in Emergency Relief can cause suffering and causality, increase deprivation costs exponentially, and negatively affect NSV.

Prolonged Relief is the stage after Emergency Relief. It may take months on average (Besiou et al., 2014). At this stage, the level of U2 starts to decrease as the situation calms down and visibility improves. Speed of response becomes less critical as supplies gradually arrive from different sources and local capacities recover partially. The mismatch between supply and demand during Prolonged

Relief is likely to fluctuate with alternate periods of undersupply and oversupply (caused by, e.g., delayed response, unsuitable items from in-kind donations, or uncoordinated responses of different organizations). The supply-demand mismatch affects NSV negatively, as undersupply leads to unmet demand and losses in SR while oversupply wastes and increased costs. At this stage, the critical supply chain capability transitions from fast and sufficient supply to calibrating speed and sufficiency by situational dynamics.

After Prolonged Relief, the situation improves and the response mission slowly transitions into a more stable stage of Recovery & Rebuild. The level of U2 in the operating context decreases gradually and may eventually plateau when the relief mission becomes new years-long development programs. The supply chain capabilities needed at this stage shift to efficiency and productivity.

The brief analysis of the disaster response lifecycle surfaces two critical success factors (CSFs) for SHSCs[10]: supply speed and supply sufficiency. CSF in strategy literature is defined as *the crucial aspect that a strategy needs to address effectively in order to achieve its objective* (Boynton and Zmud, 1984). Speed and sufficiency together constitute supply availability (defined in this study as *the condition of having adequate suitable relief items at time of need*), a key determinant of NSV in disaster response. SHSC strategies need to effectively address the CSFs of supply speed and sufficiency dictated by U2 to maximize its performance objective of NSV. Therefore, the required capabilities for SHSCs are contingent on the level of urgency and unpredictability in the contexts.

- Context-Contingent SHSC Capabilities

The concept of supply chain capability is similar to manufacturing capability and competitive priority in the literature of operations strategy (Skinner, 1974, Hayes and Pisano, 1996, Ferdows and De Meyer, 1990). Built on the trade-off theory (Fisher, 1997, Skinner, 1974) and the cumulative capability model (Ferdows and De Meyer, 1990, Noble, 1995), we propose as follows (Figure 3.4) a context-contingent SHSC capability matrix.

[10] The concept of CSF in SHSC is similar to the seven ''rights'' in commercial logistics: right time, right quantity, right price, right product, right place, right people, and right condition.

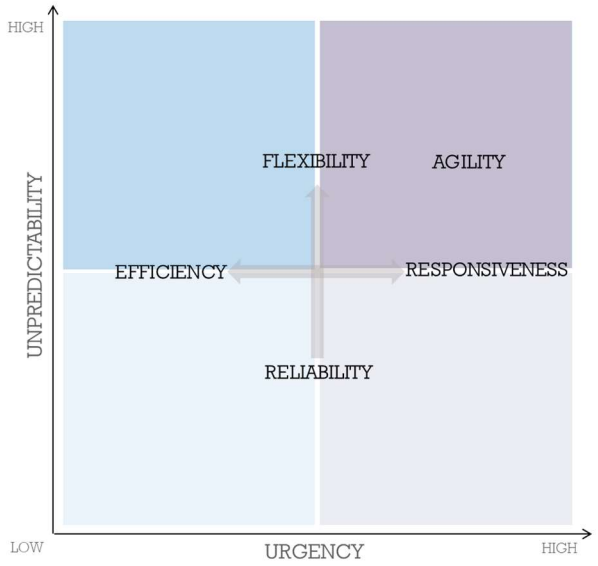

Figure 3.4 SHSC Capability Matrix

Supply chain capabilities required for sustainable response operations are contingent on the level of U2 in the contexts. The five key SHSC capabilities are responsiveness, efficiency, reliability, flexibility, and agility: (1). Responsiveness is the speed of response, the ability to configure and deploy resources to respond to humanitarian missions rapidly; (2). Efficiency is leanness, the ability to create more SR with less economic and environmental resources; (3). Reliability is dependability, the ability to deliver goods/services in different situations as expected with a predictable level of service; (4). Flexibility is adaptability, the ability to reconfigure and redeploy resources to serve different humanitarian missions; (5). Agility is a composite capability comprising flexibility and responsiveness – flexibility is the ability to adapt, responsiveness to act rapidly, and agility to adapt and respond rapidly to different humanitarian missions.

The level of urgency determines the priority of responsiveness or efficiency. High urgency situations (e.g., Emergency Relief of an earthquake) require responsiveness to rapidly activate and deploy resources and deliver emergency relief items. Responsiveness is not the priority in low urgency situations (e.g., Rebuild & Recovery after the earthquake), so the focus can be shifted to efficiency. Therefore, responsiveness and efficiency are *either/or* trade-off capabilities. Furthermore, the quest for efficiency is usually achieved through scaling and standardization – both tend to make the processes more established and harder to change, leading to increased rigidity and decreased responsiveness.

The level of unpredictability determines whether only reliability is or both reliability and flexibility are required. In low unpredictability situations (e.g., long-term recovery) with a longer horizon and sufficient time for preplanning, reliability is needed, but flexibility may not be necessary. In high unpredictability situations, flexibility is required to acquire additional supply without pre-

deploying physical resources with high uncertainty. Flexibility and reliability are *first-then* cumulative capabilities. Reliability is the necessary prerequisite for flexibility to ensure that the service level of HSCs is not compromised when adapting to other missions.

To sum up, SHSCs need a portfolio of capabilities to enable response operations that are fit with contextual dynamics, including responsiveness, efficiency, reliability, flexibility, and agility.

4.3 Distilling Practical Tactics

This step identifies the tactics that SHSC strategies need to enable to maximize NSV. Strategy is *a pattern in a stream of decisions* (Mintzberg, 1978), and tactics are *specific actions used when applying those strategic choices* (www.web-strategist.com). This step adopts a grounded approach to identify practical SHSC tactics: it first maps the structure of an HSC, then deconstructs it into operational decisions and assesses the options' NSV impacts to determine the optimal HSC configuration, and finally distills patterns in the operational choices to generalize them into practical SHSC tactics.

- **Mapping HSC Structure**

Figure 3.5 shows the typical structure of an HSC based on a generic supply chain framework proposed by Chen et al. (2019): on the upstream of the focal humanitarian organization, supply sources include prepositioned stocks, ad-hoc procurement, and in-kind donation; on the downstream, customers include beneficiaries directly served and implementing agents. As implies by Figure 5, evaluating NSV impacts of operational choices in the primary processes of Procure, Prepare, and Deliver can provide grounded insights to inform the design of Supply Chain Strategies of the focal organization.

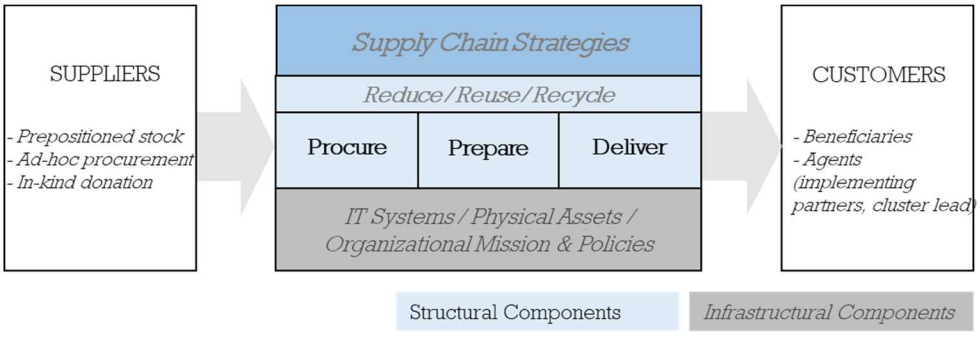

Figure 3.5 Humanitarian Supply Chain Structure

- **Identifying Optimal Configuration**

First, Procure, Prepare, and Deliver processes in HSCs are deconstructed into key operational decisions. The dichotomous options for each decision are then enumerated (simplifying the range of

potential options into dichotomous alternatives at both ends can increase clarity without compromising validity), with each option's NSV impact on SR, EC_1, EC_2 evaluated respectively through pairwise comparison. After that, those operational choices with more positive NSV impacts are identified and highlighted in Table 3.1 below.

PROCESS	OPERATIONAL DECISIONS	OPTIONS	NSV IMPACT		
			SR	*- EC_1*	*- EC_2*
P1_PROCURE	P1.1_Location of Supply Source	- Local	+	+	+
		- Global	-	-	-
	P1.2_No. of Supply Sources	- Single	-	-	=
		- Multiple	+	+	=
	P1.3_Type of Supply	- Generic Products	=	=	-
		- Green Products	=	=	+
	P1.4_Timing of Procurement	- Prepositioned Procurement	+	-	-
		- Postponed Procurement	-	+	+
P2_PREPARE	P2.1_Level of Inventory	- Low	-	+	+
		- High	+	-	-
	P2.2_Management of Inventory	- Self-Managed	+	-	-
		- Vendor-Managed	-	+	+
	P2.3_Level of Processing	- Low	-	+	+
		- High	+	-	-
	P2.4_Location of Inventory	- Local	+	+	+
		- Global	-	-	-
D3_DELIVER	D3.1_Mode of Transport	- Air	+	-	-
		- Land/Ocean	-	+	+

Table 3.1 Evaluating NSV Implications of HSC Operational Choices

The operational decisions in Procure (P1) include P1.1_Location of Supply Source, P1.2_No. of Supply Sources, P1.3_Type of Supply, and P1.4_Timing of Procurement. The alternative options for P1.1 are Local and Global. Local compared with Global has higher positive performance impacts on all three NSV dimensions (indicated by "+"): faster delivery (SR), lower transportation costs (– EC_1), and lower environmental impacts (– EC_2). The alternatives for P1.2 are Single and Multiple: Multiple outperforms Single on SR, as supply availability is better assured by different sources; it may also outperform on EC_1, given the potential economy of scale and the possibility to elicit in-kind donations without buying costs; but its advantage is uncertain on EC_2 (indicated by "=") – Single source of prepositioned stock and ad-hoc procurement can reduce environmental impacts, but Multiple sources with in-kind donations may use second-hand items or repurposed overstocks. Altogether, Multiple still outperforms Single in its overall NSV impacts. Following a similar pairwise comparison process, the rest of the evaluation is self-explanatory and not described at length here.

The SHSC configuration with higher NSV performance can be mapped after the evaluation. It comprises the following operational choices (highlighted in red in Table 3.1): (a) Procure: multiple local sources of green products with postponed procurement, (b) Prepare: low levels of inventory managed by vendors and stocked in locations proximal to origin or destination with a low level of further processing, and (c) Deliver: delivered by the transport mode of ocean or land.

- Distilling Practical Tactics

A closer examination of the SHSC operational choices detects three emergent patterns of sustainable tactics: localization, collaboration, and dematerialization. The column of Elaboration describes the inference process from sustainable operational choices to practical tactics.

PROCESS	OPERATIONAL CHOICES	ELABORATION	TACTICS
P1_PROCURE	P1.1_Local Supply Source	Location of supply sources proximal to origin of supply or destination of demand	Localization
	P1.2_Multiple Supply Sources	Collaboration with commercial suppliers and civil society institutions for alternative sources	Collaboration
	P1.3_Green Products	Reduced consumption of products with high pollutions or short lifespans	Dematerialization
	P1.4_Postponed Procurement	Delayed deployment of physical resources until uncertainty reduced	Dematerialization
P2_PREPARE	P2.1_Low Level of Inventory	Reduced tangible redundancy in the supply chain	Dematerialization
	P2.2_Vendor-Managed Inventor	Collboration with suppliers for third-party inventory management	Collaboration
	P2.3_Low Level of Processing	Elimination of non-essential further processing of purchased products	Dematerialization
	P2.4_Local Location of Inventory	Location of strategic stocks proximal to origin of supply or destination of demand	Localization
D3_DELIVER	D3.1_Land/Ocean Transport	Delivery of shipments via the modes of sea and land	Dematerialization

Table 3.2 Distilled SHSC Practical Tactics

4.4 Formulating Strategic Framework

The solution, an SHSC strategic framework, is developed according to the specified design imperatives: (1). Its objective is to maximize the performance of NSV; (2). The framework is built on five SHSC capabilities – responsiveness, efficiency, reliability, flexibility, and agility; (3). It informs the sustainable tactics of localization, dematerialization, and collaboration. Figure 3.6 presents the formulated SHSC strategic framework.

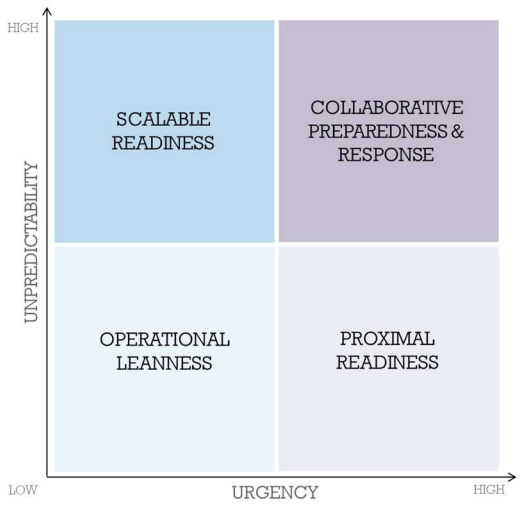

Figure 3.6 The SHSC Strategic Framework

The Framework proposes four context-contingent SHSC strategies: (a) Operational Leanness in low urgency/low unpredictability (L_{UG}/L_{UP}) contexts, (b) Scalable Readiness in low urgency/high unpredictability (L_{UG}/H_{UP}) contexts, (c) Proximal Readiness in high urgency/low unpredictability (H_{UG}/L_{UP}) contexts, and (d) Collaborative Preparedness and Response (CPR) in high urgency/high unpredictability (H_{UG}/H_{UP}) contexts.

- Operational Leanness

In L_{UG}/L_{UP} contexts, Operational Leanness should be the strategic focus of SHSCs. Such situations do not require high speed of response, and demand is relatively predictable for supply planning. Therefore, the tactic of dematerialization should be prioritized, and efficiency improvement practices (such as process standardization and utilization optimization) can be adopted for EC_1/EC_2 reduction without necessarily affecting SR negatively. Operational Leanness is where operational research techniques can contribute the most, given their strength in optimizing utilization from constrained resources (Besiou and Van Wassenhove, 2020).

- Scalable Readiness

The SHSC strategy in L_{UG}/H_{UP} contexts is Scalable Readiness. In such situations, supply planning and stock prepositioning are challenging as demand is unpredictable, but there might be sufficient time to organize flexible supplies as speed is not a priority. The collaboration tactic can be applied to enable scalable supply sufficiency and postponement of physical resource deployment, reducing EC_1/EC_2 without compromising SR. Collaboration in L_{UG}/H_{UP} contexts includes such practices as inventory pooling and resource sharing among agencies and NGOs (a shared resource pool), partnership with the private sector, and collaboration with media and civil society for in-kind donation elicitation.

- Proximal Readiness

In H_{UG}/L_{UP} contexts, Proximal Readiness is the SHSC strategy since although responsiveness is required in such situations, demand is relatively predictable. The localization tactic should be the priority to ensure proximity of preparedness. Localization includes such practices as pre-positioning vendor-managed inventory close to disaster-prone locations and developing long-term local partners to prepare for the readiness of local response.

- Collaborative Preparedness & Response

CPR is the SHSC strategy in H_{UG}/H_{UP} contexts where rapid response is critical and demand is unpredictable. The objective of CPR is to develop distributed preparedness by a virtual resource pool at scale ready for proximal response. The implementation of CPR involves all three SHSC tactics:

(a) collaboration with the private sector and civil society to build a committed virtual supply network, (b) localization by scaling of geographical coverage, (c) dematerialization by substituting intangible redundancy of committed supply for tangible redundancy of prepositioned stocks. Therefore, although CPR requires the more sophisticated capability of agility, it can achieve commensurable NSV performance in challenging H_{UC}/H_{UP} contexts.

- **Design Propositions**

The Framework's design propositions are summarized in Table 3.3 following the Context-Intervention-Mechanism-Outcome logic (Denyer et al., 2008):

CONTEXTS	STRATEGIES	OPERATIONAL MECHANISMS	OUTCOMES
L_{UC}/L_{UP}	Operational Leanness	Process standardization, efficiency improvement, utilization optimization	Cost reduction, dematerialization
L_{UC}/H_{UP}	Scalable Readiness	Inventory pooling/resource sharing with other agencies, partnership with suppliers, collaboration for in-kind donation	Cost reduction, flexible sufficiency
H_{UC}/L_{UP}	Proximal Readiness	Pre-positioning of vendor-managed stocks, developing local capacities for preparedness of priximal response	Local capacity building, dematerialization
H_{UC}/H_{UP}	Collaborative Preparedness & Response	Collaboration with private sector to build virtual supply network, scaling for broad geographical coverage, replacing prepositioned stocks with committed supply	Local capacity building, dematerialization

Table 3.3 Design Propositions of the SHSC Strategic Framework

The overarching proposition of the framework is that HSC strategies congruent with contexts can lead to better NSV performance. As many humanitarian organizations have a service mix of dual missions and the operating contexts can be volatile, a portfolio of SHSC capabilities is necessary to enable the fluid transition of strategic priorities in tune with contextual dynamics.

5. Proof of Concept

This essay designs the SHSC strategic framework with a four-phase process: context observing, problem framing, imperative derivation, to solution design. The elicitation of design imperatives takes three steps: Step 1 and Step 2 apply deductive inferences to identify the performance objective and enabling capabilities of SHSCs, and Phase 3 applies inductive inference and grounds on the author's field experiences to generalize practical tactics from sustainable SHSC operations. The design of the solution is based on the imperatives elicitated from these analyses. We hence argue that the proposed framework has in-built pragmatic validity, considering its practical

foundation and methodical formulation approach. This section further proves the framework's validity by enfolding relevant literature and elaborating on its plausibility.

5.1 SHSC Strategies and Performance Frontiers

We adapt the performance frontier theory (Schmenner and Swink, 1998) to demonstrate the plausibility of applying the framework to improve NSV performance in HSCs. The theory suggests that an organization's maximum performance is bound by its operating frontier immediately and its asset frontier ultimately (Schmenner and Swink, 1998). A supply chain is a social-technical system (Holguín-Veras et al., 2012b), whereby a social network of individuals organize a set of activities within structured operational processes enabled by infrastructural resources and capabilities. Therefore, an HSC's NSV performance can be considered bound first by its behavioral frontier (set by employees' daily activities), then by its structural frontier (the structural configuration of operational options), and ultimately by its infrastructural frontier[11]. As visualized in Figure 3.7, the proposed framework's objective is to improve the NSV performance of HSCs by flattening the structural frontier rightwards.

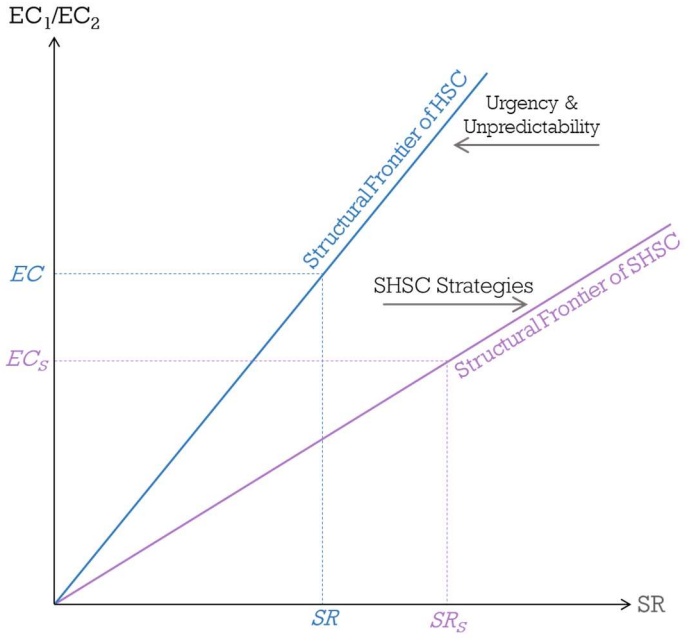

Figure 3.7 SHSC Strategies and NSV Performance Frontiers

[11] In manufacturing strategy literature, the term "structural" is usually used to refer to choices on physical assets and "infrastructural" on manufacturing systems. These terms are swapped here following the definitions of the Oxford Dictionary: "*structure – the arrangement of and relations between the parts or elements of something complex*", and "*infrastructure – the basic physical and organizational structures and facilities needed for the operation of a society or enterprise*".

The above figure illustrates that NSV performance on the structural frontier is determined by the SR and EC_1/EC_2 values. Without sustainability considerations, the theoretically achievable SR and the commensurable EC_1/EC_2 of a given HSC are generally positively correlated in the same context. When contextual U2 increases, the performance frontier will be pushed leftwards: either a higher level of EC_1/EC_2 is needed to maintain the same SR, or the theoretically achievable SR is reduced if EC_1/EC_2 remains unchanged. The practical tactics informed by the SHSC framework either address urgency through localization (Proximal Readiness), or unpredictability through collaboration (Scalable Readiness), or both through both (CPR) – all the operational options are designed to ensure the same level of SR (SRs) with minimal possible EC_1/EC_2 (ECs) resources. As a result, these contextually contingent SHSC strategies can counteract contextual U2 and push the structural frontier rightwards for improved NSV performance.

5.2 CPR vs. Prepositioning

To further prove the SHSC framework's theoretical plausibility, this subsection compares the CPR strategy's NSV performance implications with which of prepositioning, a prominent topic in humanitarian operations research. Figure 3.8 illustrates the strategic stock locations of CPR and prepositioning, respectively, on a generic HSC structure (Chen et al., 2019).

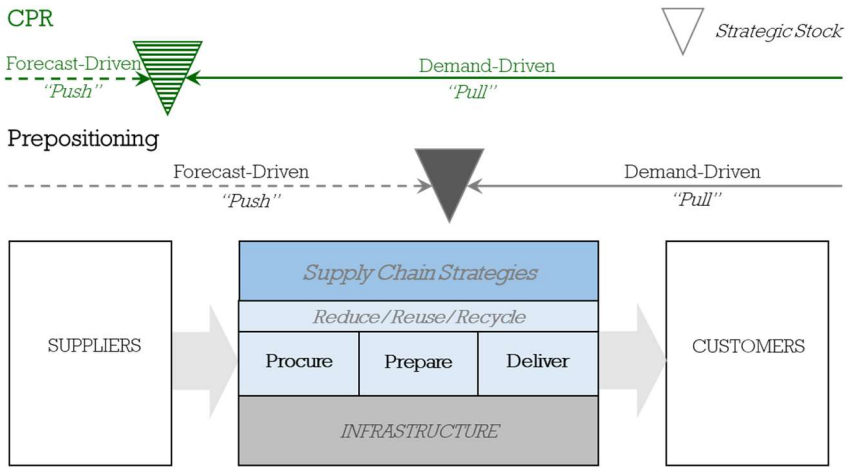

Figure 3.8 Strategic Stock Locations of CPR and Prepositioning

Prepositioning is the "*stockpiling of equipment and supplies at or near the point of planned use*" (Air Force Institute of Technology, 1981). Research in HSC prepositioning seeks to optimize responses by determining optimal facility locations and quantities of pre-acquired relief items (Acimovic and Goentzel, 2016). It is a type of Buy-to-Stock supply strategy, similar to the Make-to-Stock production strategy in manufacturing organizations. Figure 3.8 shows that the Push-Pull decoupling point of Prepositioning is close to the back-end of the Prepare process within the focal organization: supplies are pre-acquired, semi-processed, and stored in strategic locations virtually

ready for dispatch when needed. The effectiveness of prepositioning depends largely on forecast accuracy regarding the location, type, scope, and scale of disasters – in the same way that Make-to-Stock relies largely on forecast accuracy of demands in the market. Since not all disasters can be accurately forecasted, prepositioning is more effective in predictable situations – in this sense, it is similar to the SHSC framework's Proximal Readiness strategy in H_{UG}/L_{UP} contexts. However, prepositioned stocks are tangible redundancy built into HSCs – they come with prepaid buying costs (EC_1) and potential material deterioration and obsolescence (EC_2). Therefore, prepositioning is a suboptimal strategy from a sustainability perspective as it trades off EC_1/EC_2 for SR in H_{UG}/L_{UP} contexts.

Compared with prepositioning, the CPR strategy can achieve higher NSV performance. CPR aims to develop a virtual supply network of committed resources with broad geographical coverage for proximal responses. It can be likened to the Make-to-Order postponement production strategy. The Push-Pull decoupling point of CPR is positioned at the back-end of the Deliver process on the immediate upstream of the focal organization. The types and quantities of relief items can be pre-negotiated, semi-processed (plus pick & pack before Deliver), stored in facilities managed by suppliers – virtually ready for dispatch once a proximal region is hit by disasters. The committed supplies do not necessarily need to be safety stocks earmarked for disaster responses. Instead, they may be part of the cycle stocks in commercial supply chains that can be repurposed and prioritized for emergency relief in times of need. The CPR strategy can also be extended from goods to services, e.g., to logistics service providers for their capacity commitment during disaster responses. CPR replaces prepositioned tangible redundancy with prepared intangible redundancy to ensure the service level of SR without trading off EC_1/EC_2 in H_{UG}/H_{UP} contexts. Therefore, CPR compared with prepositioning is a better strategy from the perspective of sustainability.

6. Implications

6.1 Implications on Practices

Humanitarian practitioners may use the SHSC strategic framework for auditing internal supply chain capabilities and identifying improvement opportunities. A starting point is to review the organization's applied practices in recent response operations and compare them against the framework's implied tactics and strategies. The retrospective gap analysis can reveal the capabilities that the organization needs to build. The next step after that is to map the organization's existing supply network and identify, if any, those potentially more cost-effective sources not activated during the response – or, if absent, where additional sources need to be acquired to bolster the supply network. Findings from both can pinpoint the weak points in the current supply base and quick-win

opportunities to develop the required yet lacked HSC capabilities for more sustainable operations in future responses.

Implementing the framework's recommended SHSC strategies requires a long-range planning horizon for capability development. The first step to chart the roadmap for courses of action is to create the organization's current or prospective service mix profile and compare it with its existing capability portfolio. The gaps revealed are those areas that the organization needs to address in the medium to long term. Considering the practical challenges involved and the potential resource constraints, an easier place to start is to experiment with the Scalable Readiness strategy. Scalable Readiness aims to achieve supply sufficiency through collaboration without committing significant investments in tangible redundancy. It provides a low-stake context for building the capability of flexibility towards supply sufficiency in order to address high contextual unpredictability in the absence of pressure for supply speed from high contextual urgency. After that, the next step can be piloting Proximal Readiness at a small scale to test its time-constrained operationality in high urgency contexts. Finally, the CPR strategy is much more demanding to implement yet commensurably rewarding in improving NSV. Besides operational capabilities, CPR requires the infrastructural foundation of a digitalized platform and a community-based approach in collaboration with the private sector and civil society at large. The operationalization of CPR can be analogized as the sharing economy in the humanitarian sector, which we conceptualize as a type of social supply chain that can be established and dissolved fluidly on demand, governed by pre-negotiated framework agreements, and mediated by an implicit social contract based on reputation, goodwill, and trust.

6.2 Implications on Research

The key learning from this study is that supply chain management can be an instrumental anchor point for sustainability research. The systematic, universal, and practical characteristics of supply chain management make it a natural fit for sustainability research – a subject that is multidimensional, interdisciplinary, and abstract conceptually yet relevant practically.

First, supply chain management enables systems thinking to conceptualize the multidimensional multitemporal sustainability issues. Systems thinking is critical to sustainability research as it facilitates considerations of both proximal and distant, beneficial and harmful, short- and long-term effects of alternative options (Sterman, 2000). Supply chain management has its intellectual roots in industrial dynamics and value chain, and it can therefore be seen as a field of management systems engineering geared towards problem-solving. This is demonstrated in the study's elicitation process of design imperatives (Section 4.3): using a generic supply chain framework, it deconstructs humanitarian operations into integrated supply chain processes and evaluates the performance implications of alternative operational decisions – which involve trade-offs on different sustainability

dimensions – to identify the HSC configuration with higher systemwide NSV performance. Such a supply chain-based systematic method is necessary for the solution design, given the many variables involved in the evaluation process – nine operational decisions, two simplified alternative options each, with different performance implications on three sustainability dimensions, respectively. Furthermore, it probably is the only viable approach as narrative descriptions of sustainable practices and analytical methods focusing on a few limited aspects can lead to superficial or symptomatic problem-framing and partial solution. Hence, supply chain management can be an instrumental anchor point for multidimensional multitemporal sustainability research.

Second, as a field of management systems engineering, supply chain management offers various practical tools that can be applied for problem-solving in sustainability research. Sustainability as a high-level societal goal can be rather abstract to relate to in day-to-day activities. Supply chain management can be applied to ground the concept of sustainability and develop practical solutions for sustainable development since it is grounded on operations and remains close to reality. This argument is evidenced by the strategic framework designed in the study that can be used to identify sustainable practices in different humanitarian operations contexts.

Third, supply chain management can serve as a common ground to develop interdisciplinary sustainability research. As illustrated in the UN's Sustainable Development Goals, sustainability has increasingly been recognized as concerning every sector in society and every discipline in research. However, sustainability research in different fields tends to be fragmented and isolated from one another. Supply chain management traditionally focuses on manufacturing, but it is sector-neutral at its essence, and supply chain thinking can be universal – in a broad sense, even a peer-to-peer transaction of service between two individuals can be conceptualized as a supply chain. Therefore, supply chain management can provide a neutral language to integrate sustainability research in different domains for cross-learning and development. Such possibility has been demonstrated in a previous study where efforts were made to break down the wall between manufacturing and services supply chains (Chen et al., 2019). This study adds further evidence to this line of inquiry by building a bridge between commercial and humanitarian supply chains. There are rich opportunities in applying supply chain management for crossdisciplinary communication, transdisciplinary learning, multidisciplinary synthesis, and interdisciplinary enrichment in future sustainability research.

7. Conclusions and Future Research

- Contributions

This study proposes a strategic supply chain framework for sustainable humanitarian operations. The framework is developed through a four-phase design process: observing the contexts, framing the problem, eliciting the design imperatives, and designing the solution. The solution design imperatives are elicited by specifying NSV as the performance objective, identifying required supply chain capabilities, and distilling practical SHSC tactics. Humanitarian practitioners may use the framework to identify contextually congruent strategies to improve the sustainability performance of HSCs. The framework fills the research gap of a strategic supply chain framework with sustainability as the performance objective in the humanitarian context.

- Limitations

This key limitation of this essay is that it does not implement the framework in real life to test and refine the designed solution. Its applied methodology is design thinking, a partial but not the full version of the design science research strategy. The lack of practical implementation can obscure the framework's practical feasibility. However, its pragmatic validity may still be claimed since it is based on intensive field experiences and developed through a transparent and falsifiable process grounded on end-to-end HSC operations. Nonetheless, the designed framework needs to be empirically tested in future research via either focus group interviews, experimentation, or pilot implementation through collaboration with humanitarian organizations.

- Future Research

Future research in this line of inquiry may focus on two areas. First, develop an interdisciplinary framework for social performance measurement at the organizational level. Systematic supply chain thinking can be applied to integrate and refine the currently siloed developments on this topic in different fields, such as social entrepreneurship and impact investing, into a standardized measurement framework. Second, experiment to elicit humanitarian practitioners' mental decision models in emergency relief situations. Decision-making in emergency response is particularly challenging as it involves human suffering and life loss. Scenario-based experimentation can explicate the rationales and biases behind practitioners' trade-off and prioritization decisions. Findings from such behavioral research can help design foolproof SHSCs with layered default choices that lighten practitioners' cognitive load and reduce reliance on their value preferences in high-stake situations of emergency response.

With so many daunting challenges ahead, sustainability may seem like an unattainable goal. But to (mis)quote Lennon, ''*A dream we dream alone is only a dream. A dream we dream together is reality*.''

References

ACIMOVIC, J. & GOENTZEL, J. 2016. Models and metrics to assess humanitarian response capacity. *Journal of Operations Management,* 45, 11-29.

AIR FORCE INSTITUTE OF TECHNOLOGY 1981. Compendium of Authenticated Systems and Logistics Terms, Definitions, and Acronyms. Maxwell Air Force Base, Alabama: Department of the Air Force, Air University.

ALBERTINI, E. 2013. Does environmental management improve financial performance? A meta-analytical review. *Organization & Environment,* 26, 431-457.

ALLEN, D. T. & SHONNARD, D. 2012. *Sustainable Engineering: Concepts, Design, and Case Studies,* New Jersey, United States, Prentice Hall.

AMBEC, S. & LANOIE, P. 2008. Does it pay to be green? A systematic overview. *Academy of Management Perspectives,* 22, 45-62.

BALCIK, B. & BEAMON, B. M. 2008. Facility location in humanitarian relief. *International Journal of Logistics,* 11, 101-121.

BANSAL, P. & HOFFMAN, A. J. 2012. *The Oxford Handbook of Business and the Natural Environment,* Oxford, Oxford University Press.

BARENDREGT, J. J. & DOI, S. A. 2016. *MetaXL User Guide* [Online]. Available: http://www.epigear.com/ [Accessed 19 March 2019].

BARNEY, J. 1991. Firm resources and sustained competitive advantage. *Journal of Management,* 17, 99-120.

BECKMAN, S. L. & BARRY, M. 2007. Innovation as a learning process: Embedding design thinking. *California Management Review,* 50, 25-56.

BESIOU, M., PEDRAZA-MARTINEZ, A. J. & VAN WASSENHOVE, L. N. 2014. Vehicle supply chains in humanitarian operations: Decentralization, operational mix, and earmarked funding. *Production and Operations Management,* 23, 1950-1965.

BESIOU, M. & VAN WASSENHOVE, L. N. 2020. Humanitarian operations: A world of opportunity for relevant and impactful research. *Manufacturing & Service Operations Management,* 22, 135-145.

BOURGEOIS, L. J. 1981. On the measurement of organizational slack. *Academy of Management Review,* 6, 29-39.

BOWEN, H. R. 1953. *Social Responsibilities of the Businessman,* Iowa, United States, University of Iowa Press.

BUHR, N. & GRAY, R. 2012. Environmental management, measurement, and accounting: Information for decision and control? *In:* BANSAL, P. & HOFFMAN, A. J. (eds.) *The Oxford Handbook of Business and the Natural Environment.* Oxford: Oxford University Press.

BUSCH, T. & HOFFMANN, V. H. 2011. How hot is your bottom line? Linking carbon and financial performance. *Business & Society,* 50, 233-265.

CHARTERED ASSOCIATION OF BUSINESS SCHOOLS 2018. Academic Journal Guide 2018.

CHEN, H., DAUGHERTY, P. J. & LANDRY, T. D. 2009. Supply chain process integration: A theoretical framework. *Journal of Business Logistics,* 30, 27-46.

CHEN, H. S., LAI, K.-H. & CHENG, T. 2019. Sustainable service supply chains. *In:* SARKIS, J. (ed.) *Handbook on the Sustainable Supply Chain.* Cheltenham, United Kingdom: Edward Elgar Publishing.

CHRISTOPHER, M. & TOWILL, D. 2001. An integrated model for the design of agile supply chains. *International Journal of Physical Distribution & Logistics Management,* 31, 235-246.

CLEOPHAS, T. J. & ZWINDERMAN, A. H. 2017. *Modern Meta-Analysis: Review and Update of Methodologies,* Berlin, Germany, Springer.

CLIMATE CHANGE SERVICE, C. 2019. *Record-breaking temperatures for July* [Online]. Available: https://climate.copernicus.eu/ [Accessed 24 July 2019].

CORBETT, C. J. & KLASSEN, R. D. 2006. Extending the horizons: environmental excellence as key to improving operations. *Manufacturing & Service Operations Management,* 8, 5-22.

CORBETT, C. J. & VAN WASSENHOVE, L. N. 1993. The natural drift: What happened to operations research? *Operations Research,* 41, 625-640.

DENYER, D., TRANFIELD, D. & VAN AKEN, J. E. 2008. Developing design propositions through research synthesis. *Organization Studies,* 29, 393-413.
DICKERSIN, K. 2005. Publication bias: Recognizing the problem, understanding its origins and scope, and preventing harm. *In:* ROTHSTEIN, H. R., SUTTON, A. J. & BORENSTEIN, M. (eds.) *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments.* Hoboken, New Jersey, United States: Wiley.
DONALDSON, L. 2001. *The Contingency Theory of Organizations,* Thousand Oaks, California, United States, SAGE Publishing.
DONALDSON, T. & PRESTON, L. E. 1995. The stakeholder theory of the corporation: Concepts, evidence, and implications. *Academy of Management Review,* 20, 65-91.
EBRAHIM, A. & RANGAN, V. K. 2014. What impact? A framework for measuring the scale and scope of social performance. *California Management Review,* 56, 118-141.
EDMONDSON, A. C. & MCMANUS, S. E. 2007. Methodological fit in management field research. *Academy of Management Review,* 32, 1246-1264.
EISENHARDT, K. M. 1989. Building theories from case study research. *Academy of Management Review,* 14, 532-550.
ELKINGTON, J. 1994. Towards the sustainable corporation: Win-win-win business strategies for sustainable development. *California Management Review,* 36, 90.
ELMAN, J. L. 1990. Finding structure in time. *Cognitive Science,* 14, 179-211.
ENDRIKAT, J. 2016. Market reactions to corporate environmental performance related events: A meta-analytic consolidation of the empirical evidence. *Journal of Business Ethics,* 138, 535-548.
ETZION, D. 2007. Research on organizations and the natural environment, 1992-present: A review. *Journal of Management,* 33, 637-664.
FERDOWS, K. & DE MEYER, A. 1990. Lasting improvements in manufacturing performance: in search of a new theory. *Journal of Operations Management,* 9, 168-184.
FISHER, M. L. 1997. What is the right supply chain for your product? *Harvard Business Review,* 75, 105-117.
FLYNN, B. B., SAKAKIBARA, S., SCHROEDER, R. G., BATES, K. A. & FLYNN, E. J. 1990. Empirical research methods in operations management. *Journal of Operations Management,* 9, 250-284.
FRANZOSI, R. 2008. *Content Analysis (SAGE Benchmarks in Social Research Methods Series),* Thousand Oaks, California, United States, SAGE Publishing.
FREEMAN, R. E. 1984. *Strategic Management: A Stakeholder Approach,* Cambridge, United Kingdom, Cambridge University Press.
FRIEDMAN, M. 1970. The social responsibility of business is to increase its profits. *The New York Times Magazine.*
GENG, R., MANSOURI, S. A. & AKTAS, E. 2017. The relationship between green supply chain management and performance: A meta-analysis of empirical evidences in Asian emerging economies. *International Journal of Production Economics,* 183, 245-258.
GLASER, B. G. & STRAUSS, A. L. 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research,* Piscataway, New Jersey, United States, Transaction Publishers.
GORE, A. 1992. *Earth in the Balance: Ecology and the Human Spirit,* Boston, Massachussetts, United States, Houghton Mifflin.
GRALLA, E., GOENTZEL, J. & FINE, C. 2014. Assessing trade-offs among multiple objectives for humanitarian aid delivery using expert preferences. *Production and Operations Management,* 23, 978-989.
HAAVISTO, I. & KOVÁCS, G. 2014. Perspectives on sustainability in humanitarian supply chains. *Disaster Prevention and Management,* 23, 610-631.
HAAVISTO, I. & KOVÁCS, G. 2019. Sustainability in humanitarian supply chains. *Handbook on the Sustainable Supply Chain.* Cheltenham, United Kingdom: Edward Elgar Publishing.
HART, S. L. 1995. A natural-resource-based view of the firm. *Academy of Management Review,* 20, 986-1014.
HARTUNG, J., KNAPP, G. & SINHA, B. K. 2011. *Statistical Meta-analysis with Applications,* Hoboken, New Jersey, United States, Wiley.

HAYES, R. H. & PISANO, G. P. 1996. Manufacturing strategy: at the intersection of two paradigm shifts. *Production and Operations Management,* 5, 25-41.
HELP LOGISTICS, A., KUEHNE LOGISTICS UNIVERSITY, IFRC & ICRC. 2018. *Supply Chain Expenditure and Preparedness Investment Opportunities* [Online]. Available: https://www.help-logistics.org/fileadmin/user_upload/Dateien_HELP/documents/report/HELP_IFRC_ROI-Report_printFinal.pdf [Accessed].
HERZBERG, F. I. 1968. *Work and the Nature of Man,* Cleveland, Ohio, United States, World Publishing.
HOFFMAN, A. J. 2018. The next phase of business sustainability. *Stanford Social Innovation Review,* 16, 34-39.
HOFFMAN, A. J. & BANSAL, P. 2012. Retrospective, perspective, and prospective: Introduction to the oxford handbook on business and the natural environment. *In:* BANSAL, P. & HOFFMAN, A. J. (eds.) *The Oxford Handbook of Business and the Natural Environment.* Oxford: Oxford University Press.
HOFFMAN, A. J. & GEORG, S. 2012. A history of research on business and the natural Environment: Conversations from the field. *In:* GEORG, S. & HOFFMAN, A. J. (eds.) *Business and the Environment: Critical Perspectives in Business and Management.* Oxford: Routledge.
HOLGUÍN-VERAS, J., JALLER, M., VAN WASSENHOVE, L. N., PÉREZ, N. & WACHTENDORF, T. 2012a. Material convergence: Important and understudied disaster phenomenon. *Natural Hazards Review,* 15, 1-12.
HOLGUÍN-VERAS, J., JALLER, M., VAN WASSENHOVE, L. N., PÉREZ, N. & WACHTENDORF, T. 2012b. On the unique features of post-disaster humanitarian logistics. *Journal of Operations Management,* 30, 494-506.
HOLGUÍN-VERAS, J., PÉREZ, N., JALLER, M., VAN WASSENHOVE, L. N. & AROS-VERA, F. 2013. On the appropriate objective function for post-disaster humanitarian logistics models. *Journal of Operations Management,* 31, 262-280.
HUMANITARIAN RESEARCH GROUP 2011. Greening the Humanitarian Response: 2010 Haiti Earthquake
HUNTER, J. E. & SCHMIDT, F. L. 1990. *Methods of Meta-analysis: Correcting Error and Bias in Research Findings,* Thousand Oaks, California, United States, SAGE Publishing.
JAFFE, A. B., PETERSON, S. R., PORTNEY, P. R. & STAVINS, R. N. 1995. Environmental regulation and the competitiveness of US manufacturing: what does the evidence tell us? *Journal of Economic literature,* 33, 132-163.
JONSEN, A. R. 1986. Bentham in a box: technology assessment and health care allocation. *Law, Medicine and Health Care,* 14, 172-174.
JOSEPHSON, J. R. & JOSEPHSON, S. G. 1996. *Abductive Inference: Computation, Philosophy, Technology,* Cambridge, United Kingdom, Cambridge University Press.
KLEINDORFER, P. R., SINGHAL, K. & VAN WASSENHOVE, L. N. 2005. Sustainable operations management. *Production and Operations Management,* 14, 482-492.
KRETSCHMER, A., SPINLER, S. & VAN WASSENHOVE, L. N. 2014. A School Feeding Supply Chain Framework: Critical Factors for Sustainable Program Design. *Production and Operations Management,* 23, 990-1001.
LEE, H. L. 2002. Aligning supply chain strategies with product uncertainties. *California Management Review,* 44, 105-119.
LIPSEY, M. W. & WILSON, D. B. 2001. *Practical Meta-Analysis,* Thousand Oaks, California, United States, SAGE Publishing.
LITTELL, J. H., CORCORAN, J. & PILLAI, V. 2008. *Systematic Reviews and Meta-Analysis,* Oxford, Oxford University Press.
LUCAS, M. T. & NOORDEWIER, T. G. 2016. Environmental management practices and firm financial performance: The moderating effect of industry pollution-related factors. *International Journal of Production Economics,* 175, 24-34.
MASLOW, A. H. 1943. A theory of human motivation. *Psychological Review,* 50, 370.
MEREDITH, J. 1993. Theory building through conceptual methods. *International Journal of Operations & Production Management,* 13, 3-11.

MILES, R. E. 1978. *Organizational Strategy, Structure, and Process,* New York City, New York, United States, McGraw-Hill.
MINTZBERG, H. 1978. Patterns in strategy formation. *Management Science,* 24, 934-948.
MINTZBERG, H. 2000. *The Fall and Rise of Strategic Planning,* Upper Saddle River, New Jersey, United States, Pearson Education.
NOBLE, M. A. 1995. Manufacturing strategy: testing the cumulative model in a multiple country context. *Decision Sciences,* 26, 693-721.
OCHA. 2018. *Appeals and Response Plans 2018* [Online]. New York: United Nations Office for the Coordination of Humanitarian Affairs. Available: https://fts.unocha.org/appeals/overview/2018/plans [Accessed Jan 07 2019].
ORLITZKY, M., SCHMIDT, F. L. & RYNES, S. L. 2003. Corporate social and financial performance: A meta-analysis. *Organization Studies,* 24, 403-441.
ORSATO, R. J. 2006. Competitive environmental strategies: when does it pay to be green? *California Management Review,* 48, 127-143.
PEDRAZA-MARTINEZ, A. J., STAPLETON, O. & VAN WASSENHOVE, L. N. 2011. Field vehicle fleet management in humanitarian operations: A case-based approach. *Journal of Operations Management,* 29, 404-421.
PEDRAZA-MARTINEZ, A. J. & VAN WASSENHOVE, L. N. 2016. Empirically grounded research in humanitarian operations management: The way forward. *Journal of Operations Management,* 45, 1-10.
PELOZA, J. 2009. The challenge of measuring financial impacts from investments in corporate social performance. *Journal of Management,* 35, 1518-1541.
PEROTTI, S., ZORZINI, M., CAGNO, E. & MICHELI, G. J. L. 2012. Green supply chain practices and company performance: the case of 3PLs in Italy. *International Journal of Physical Distribution & Logistics Management,* 42, 640-672.
PORTER, M. E. 1979. The five competitive forces that shape strategy. *Harvard Business Review,* 86, 25-40.
PORTER, M. E. 1985. *Competitive Advantage: Creating and Sustaining Superior Performance,* New York, United States, Free Press.
PORTER, M. E. 1991a. America's green strategy. *Scientific American,* 264, 168-168.
PORTER, M. E. 1991b. Towards a dynamic theory of strategy. *Strategic Management Journal,* 12, 95-117.
PORTER, M. E. & VAN DER LINDE, C. 1995. Toward a new conception of the environment-competitiveness relationship. *Journal of Economic Perspectives,* 9, 97-118-118.
PRIMC, K. & CATER, T. 2015. Environmental proactivity and firm performance: a fuzzy-set analysis. *Management Decision,* 53, 648-667.
PRINCIPLES FOR RESPONSIBLE INVESTMENT. 2010. *Universal ownership: Why environmental externalities matter to institutional investors* [Online]. Available: https://www.unpri.org/download?ac=5875 [Accessed 24 July 2019].
QSR INTERNATIONAL, Q. 2020. *NVivo 12 Plus* [Online]. Available: https://www.qsrinternational.com/nvivo/nvivo-products/nvivo-12-plus [Accessed 25 July 2019].
RELIEFWEB. 2019. *The global climate in 2015–2019* [Online]. Available: https://reliefweb.int/report/world/global-climate-2015-2019 [Accessed 12 September 2019].
REUTERS, T. 2017. *Web of Science -The Complete Citation Connection* [Online]. Available: https://web.archive.org/web/20170224013916/http://wokinfo.com/citationconnection/realfacts [Accessed 17 May 2018].
RUSSO, M. V. & MINTO, A. 2012. Competitive strategy and the environment: A field of inquiry emerges. *In:* BANSAL, P. & HOFFMAN, A. J. (eds.) *The Oxford Handbook of Business and the Natural Environment.* Oxford: Oxford University Press.
SARKIS, J. 2012. Models for compassionate operations. *International Journal of Production Economics,* 139, 359-365.
SCHMENNER, R. W. & SWINK, M. L. 1998. On theory in operations management. *Journal of Operations Management,* 17, 97-113.

SHRIVASTAVA, P. 1995. Environmental technologies and competitive advantage. *Strategic Management Journal,* 16.
SIMCHI-LEVI, D., KAMINSKY, P. & SIMCHI-LEVI, E. 2004. *Managing the Supply Chain: The Definitive Guide for the Business Professional,* New York City, New York, United States, McGraw-Hill.
SIMCHI-LEVI, D., SCHMIDT, W. & WEI, Y. 2014. From superstorms to factory fires: Managing unpredictable supply chain disruptions. *Harvard Business Review,* 92, 96-101.
SKINNER, W. 1974. The focused factory. *Harvard Business Review*, 114-121.
SMITH, N. C. 2003. Corporate social responsibility: whether or how? *California Management Review,* 45, 52-76.
SODHI, M. S. 2016. Natural disasters, the economy and population vulnerability as a vicious cycle with exogenous hazards. *Journal of Operations Management,* 45, 101-113.
SOUSA, R. & VOSS, C. A. 2008. Contingency research in operations management practices. *Journal of Operations Management,* 26, 697-713.
STARIK, M. & KANASHIRO, P. 2013. Toward a theory of sustainability management: Uncovering and integrating the nearly obvious. *Organization & Environment,* 26, 7-30.
STARR, M. K. & VAN WASSENHOVE, L. N. 2014. Introduction to the special issue on humanitarian operations and crisis management. *Production and Operations Management,* 23, 925-937.
STERMAN, J. D. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World,* New York, United States, McGraw-Hill.
TEECE, D. J., PISANO, G. & SHUEN, A. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal,* 18, 509-533.
TOMASINI, R. & VAN WASSENHOVE, L. 2009. *Humanitarian Logistics,* London, Palgrave Macmillan.
TRANFIELD, D., DENYER, D. & SMART, P. 2003. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management,* 14, 207-222.
TRUMPP, C., ENDRIKAT, J., ZOPF, C. & GUENTHER, E. 2015. Definition, conceptualization, and measurement of corporate environmental performance: A critical examination of a multidimensional construct. *Journal of Business Ethics,* 126, 185-204.
TRUMPP, C. & GUENTHER, T. 2017. Too Little or too much? Exploring U-shaped relationships between corporate environmental performance and corporate financial performance. *Business Strategy and the Environment,* 26, 49-68.
VAN AKEN, J., CHANDRASEKARAN, A. & HALMAN, J. 2016. Conducting and publishing design science research: Inaugural essay of the design science department of the Journal of Operations Management. *Journal of Operations Management,* 47, 1-8.
VAN RIJSBERGEN, C. J. 1986. A non-classical logic for information retrieval. *The Computer Journal,* 29, 481-485.
VAN WASSENHOVE, L. N. 2019. Sustainable innovation: Pushing the boundaries of traditional operations management. *Production and Operations Management,* 28, 2930-2945.
VAN WASSENHOVE, L. N. & PEDRAZA MARTINEZ, A. J. 2012. Using OR to adapt supply chain management best practices to humanitarian logistics. *International Transactions in Operational Research,* 19, 307-322.
WADDOCK, S. A. & GRAVES, S. B. 1997. The corporate social performance-financial performance link. *Strategic Management Journal,* 18, 303-319.
WALKER, E., HERNANDEZ, A. V. & KATTAN, M. W. 2008. Meta-analysis: Its strengths and limitations. *Cleveland Clinic Journal of Medicine,* 75, 431.
WALLEY, N. & WHITEHEAD, B. 1994. It's not easy being green. *Reader in Business and the Environment,* 36, 81.
WILDING, R., WAGNER, B., ASHBY, A., LEAT, M. & HUDSON-SMITH, M. 2012. Making connections: a review of supply chain management and sustainability literature. *Supply Chain Management: An International Journal,* 17, 497-516.
WORLD COMMISSION ON ENVIRONMENT AND DEVELOPMENT 1987. *Our Common Future ("The Brundtland Report"),* Oxford, Oxford University Press.
YIN, R. K. 1994. *Case Study Research: Design and Methods,* Thousand Oaks, California, United States, SAGE Publishing.

Appendix I – GC Research Articles Reviewed

RF	AUTHOR(S)	YR	JR	HY	FD	G	GDT	C	CDT	ST	RG	FS	TF	TL	DA
1	Fogler & Nutt	1975	AMJ	1A_GC-+	Y	G4	E	C6	F	HPS	NA	Pb	LG	1Y	1971
2	Spicer	1978	JBFA	4A_GC++	Y	G4	E	C6	F	HPS	NA	Pb	LG	CCR	1970
3	Chen & Metcalf	1980	AR	4A_GC++	N	G4	E	C6	F	HPS	NA	Pb	LG	CCR	1970
4	Cormier et. Al	1993	EE	2A_GC--	Y	G5	P2	C8	F	HPS	NA	Pb	LG	CCR	1987
5	Shrivastava	1995	SMJ	4A_GC++	Y	G10	A	C9	A	Mfg	NA	Pb	CS	CCR	1993
6	Hart & Ahuja	1996	BSE	4A_GC++	Y	G5	E	C5	F	Mfg	NA	Pb	LG	1-3Y	1989
7	Klassen & McLaughlin	1996	MS	2A_GC--	Y	G9	C	C6	F	HPS	NA	Pb	LG	3Ds	1998
8	Russo & Fouts	1997	AMJ	4A_GC++	Y	G9	E	C5	F	Mfg	NA	Pb	CS	CCR	1992
9	Cordeiro & Sarkis	1997	BSE	3A_GC+-	Y	G5	P2	C8	E	HPS	NA	Pb	CS	CCR	1992
10	Bhat	1998	ETL	2A_GC--	Y	G8	E	C5	F	Mfg	NA	Pb	CS	CCR	1989
11	Judge & Douglas	1998	JMS	4A_GC++	Y	G10	P1	C9	P1	Mfg	NA	Mx	CS	CCR	1996
12	Sharma & Vredenburg	1998	SMJ	4A_GC++	Y	G10	P1	C9	P1	HPS	NA	Lg	CS	3Ys+	1995
13	Klassen & Whybark	1999	AMJ	3A_GC+-	Y	G10	P1	C1	P1	Mfg	NA	Mx	CS	CCR	1994
14	Khanna & Damon	1999	JEEM	4A_GC++	Y	G10	P2	C8	F	HPS	NA	Pb	LG	1Y	1991
15	Christmann	2000	AMJ	4A_GC++	N	G10	P1	C1	P1	HPS	NA	Lg	CS	CCR	1996
16	Hughes	2000	AR	2A_GC--	Y	G5	P2	C8	F	HPS	NA	Pb	LG	CCR	1990
17	Gilley et al.	2000	JOM	4A_GC++	N	G10	C	C6	F	M&S	NA	Pb	LG	2Ds	1990
18	Dowell et al.	2000	MS	4A_GC++	Y	G10	E	C6	F	Mfg	NA	Pb	LG	CCR	1996
19	Carter et al.	2000	TRPE	4A_GC++	Y	G7	P1	C9	F	Mfg	NA	Mx	CS	CCR	1995
20	Sarkis & Cordeiro	2001	EJOR	3A_GC+-	Y	G5	P2	C5	F	HPS	NA	Pb	LG	1Y	1992
21	Thomas	2001	BSE	1A_GC-+	Y	G10	P1	C6	F	M&S	EU	Pb	LG	CCR	1992
22	King & Lenox	2001	JIE	Excluded	Y	G5	P2	C8	F	Mfg	NA	Pb	LG	CCR	1992
23	Alvarez Gil et al.	2001	OMEGA	4A_GC++	Y	G10	P1	C5	P1	Svc	EU	Mx	CS	CCR	1999
24	Wagner et al.	2002	CSR&EM	3A_GC+-	Y	G5	P1	C5	P1	HPS	EU	Mx	CS	CCR	1996
25	King & Lenox	2002	MS	4A_GC++	Y	G5	P2	C8	F	HPS	NA	Pb	LG	2Ys	1994
26	Melnyk et al.	2003	JOM	4A_GC++	Y	G6	P1	C9	P1	Mfg	NA	Mx	CS	CCR	1998
27	Kassinis & Soteriou	2003	POM	4A_GC++	Y	G10	P1	C2	P1	Svc	EU	Sm	CS	CCR	1999
28	Bansal & Clelland	2004	AMJ	4A_GC++	Y	G10	C	C6	F	HPS	NA	Pb	LG	CCR	1992
29	Clarkson & Li	2004	AR	2A_GC--	Y	G1	F	C6	F	HPS	NA	Pb	LG	CCR	1995
30	Zhu & Sarkis	2004	JOM	4A_GC++	Y	G7	P1	C3	P1	Mfg	CN	Mx	CS	CCR	2003
31	Céspedes-Lorente & Galdeano-Gómez	2004	BSE	4A_GC++	Y	G1	F	C1	F	M&S	EU	Mx	LG	CCR	1999
32	Rao & Holt	2005	IJOPM	4A_GC++	Y	G7	P1	C9	P1	Mfg	APO	Mx	CS	CCR	2004
33	Zhu et al.	2005	IJOPM	4A_GC++	N	G7	P1	C9	P1	Mfg	CN	Mx	CS	CCR	2004
34	Baker & Sinkula	2005	JAMS	4A_GC++	Y	G10	P1	C2	P1	M&S	NA	Lg	CS	CCR	2004

35	Menguc & Ozanne	2005	JBR	4A_GC++	Y	G10	P1	C7	P1	Mfg	ANZ	Mx	CS	CCR	2003
36	Wagner	2005	JEM	3A_GC+-	Y	G10	P1	C5	P1	HPS	EU	Mx	CS	CCR	2004
37	Chan	2005	JMS	4A_GC++	Y	G10	P1	C7	P1	Mfg	CN	Mx	CS	CCR	2003
38	Gonzalez-Benito & Gonzalez-Benito	2005	OMEGA	4A_GC++	Y	G10	P1	C1	P1	Mfg	EU	Mx	CS	CCR	2003
39	Orsato	2006	CMR	4A_GC++	Y	G10	A	C9	A	M&S	CR	Mx	CS	CCR	2004
40	Craig & Dibrell	2006	FBR	4A_GC++	Y	G10	P1	C9	P1	M&S	NA	Sm	CS	CCR	2006
41	Link & Naveh	2006	IEEE	4A_GC++	N	G10	P1	C9	P1	M&S	MENA	Mx	CS	CCR	2005
42	Chen et al.	2006	JBE	4A_GC++	Y	G2	P1	C9	P1	Mfg	APO	Mx	CS	CCR	2003
43	Clemens	2006	JBR	4A_GC++	Y	G10	P1	C9	P1	Mfg	NA	Sm	CS	CCR	2003
44	Rusinko	2007	IEEE	2A_GC--	Y	G10	P1	C9	P1	Mfg	NA	Mx	CS	CCR	2005
45	Montabon et al.	2007	JOM	4A_GC++	Y	G10	C	C7	F	M&S	CR	Mx	CS	CCR	2004
46	Nakao et al.	2007	BSE	Excluded	Y	G9	E	C8	F	Mfg	JP	Pb	LG	1Y	2001
47	Wahba	2008	CSR&EM	4A_GC++	Y	G6	E	C8	F	M&S	MENA	Pb	LG	CCR	2004
48	Vachon & Klassen	2008	IJPE	4A_GC++	Y	G7	P1	C1	P1	Mfg	NA	Mx	CS	CCR	2006
49	Chen	2008	JBE	4A_GC++	Y	G10	P1	C9	P1	Mfg	APO	Mx	CS	CCR	2006
50	Darnall et al.	2008	JIM	4A_GC++	Y	G6	P1	C5	P1	Mfg	CR	Mx	CS	CCR	2003
51	Eiadat et al.	2008	JWB	4A_GC++	Y	G10	P1	C9	P1	HPS	MENA	Mx	CS	CCR	2005
52	Lucas & Wilson	2008	SB	4A_GC++	Y	G10	E	C5	F	Svc	NA	Pb	CS	CCR	2004
53	Sharfman & Fernando	2008	SMJ	4A_GC++	Y	G5	P2	C8	F	M&S	NA	Pb	LG	1Y	2001
54	Elsayed & Paton	2009	BSE	4B_CG++	Y	G10	P1	C8	P1	M&S	EU	Pb	LG	1Y	1997
55	Wagner	2009	BSE	4A_GC++	Y	G10	P1	C9	F	Mfg	EU	Mx	CS	CCR	2001
56	Sueyoshi & Goto	2009	EP	3A_GC+-	Y	G1	F	C5	F	HPS	NA	Lg	LG	CCR	1995
57	Fraj-Andres et al.	2009	JBE	4A_GC++	N	G10	P1	C2	P1	Mfg	EU	Mx	CS	CCR	2007
58	First & Khetriwal	2010	BSE	4A_GC++	N	G10	C	C9	E	Mfg	CR	Lg	CS	CCR	2006
59	Sueyoshi & Goto	2010	EJOR	4A_GC++	Y	G10	E	C8	F	Mfg	JP	Pb	LG	CCR	2006
60	Menguc et al.	2010	JBE	4A_GC++	Y	G10	P1	C4	P1	Mfg	ANZ	Mx	CS	CCR	2008
61	Zeng et al.	2010	JCP	4A_GC++	Y	G10	P1	C5	P1	Mfg	CN	Mx	CS	CCR	2009
62	Jacobs et al.	2010	JOM	4A_GC++	Y	G10	C	C6	F	M&S	NA	Mx	LG	2Ds	2005
63	Busch & Hoffmann	2011	B&S	4A_GC++	Y	G5	E	C8	F	M&S	CR	Pb	CS	1Y	2006
64	Delmas et al.	2011	B&S	4A_GC++	Y	G10	P1	C9	P1	HPS	EU	Mx	CS	CCR	2010
65	Iwata & Okada	2011	EE	4A_GC++	Y	G5	P2	C8	F	Mfg	JP	Pb	LG	CCR	2006
66	Paulraj & de Jong	2011	IJOPM	3A_GC+-	Y	G6	C	C6	F	M&S	NA	Pb	LG	3Ds	2002
67	Rennings & Rammer	2011	INI	4A_GC++	Y	G2	P1	C5	P1	M&S	EU	Mx	LG	1-3Y	2001
68	Chang	2011	JBE	4A_GC++	Y	G10	P1	C9	P1	Mfg	APO	Mx	CS	CCR	2010
69	Heras-Saizarbitoria et al	2011	JCP	4B_CG++	Y	G6	E	C7	F	M&S	EU	Mx	LG	1-3Y	2003
70	Nishitani et al.	2011	JCP	4A_GC++	Y	G5	E	C9	F	Mfg	JP	Pb	LG	CCR	2005
71	Fisher-Vanden & Thorburn	2011	JEEM	3A_GC+-	Y	G10	C	C6	F	M&S	NA	Pb	LG	3Ds	2001

72	Chiou et al.	2011	TRPE	4A_GC++	Y	G10	P1	C9	P1	Mfg	APO	Mx	CS	CCR	2009
73	Alvarez	2012	BSE	3A_GC+-	Y	G5	P2	C5	F	Mfg	CR	Pb	LG	1-3Y	2009
74	Nishitani & Kokubu	2012	BSE	4A_GC++	Y	G5	P2	C6	F	Mfg	JP	Pb	LG	CCR	2007
75	Segarra-Oña et al.	2012	CHQ	4A_GC++	Y	G6	E	C7	F	Svc	EU	Mx	CS	CCR	2008
76	Horvathova	2012	EE	2A_GC--	Y	G5	P2	C5	F	M&S	EU	Mx	LG	1-3Y	2006
77	Martinez-Del-Rio et al.	2012	HRM	4A_GC++	Y	G10	P1	C5	F	Mfg	EU	Mx	CS	1-3Y	2006
78	Perotti et al.	2012	IJPDLM	4A_GC++	Y	G7	P1	C9	P1	Svc	EU	Mx	CS	CCR	2010
79	Chan et al.	2012	IMM	4A_GC++	Y	G7	P1	C9	P1	Mfg	CN	Lg	CS	CCR	2011
80	Chien & Peng	2012	JBR	4A_GC++	Y	G10	F	C8	F	HPS	APO	Pb	LG	3Ys+	1998
81	Kroes et al.	2012	MSOM	3A_GC+-	Y	G10	P2	C8	F	HPS	NA	Pb	LG	1-3Y	2005
82	Lai & Wong	2012	OMEGA	4A_GC++	Y	G10	P1	C1	P1	Mfg	CN	Mx	CS	CCR	2008
83	Green et al.	2012	SCMAIJ	4A_GC++	Y	G7	P1	C1	P1	Mfg	NA	Lg	CS	CCR	2011
84	Humphrey et al.	2013	AJM	Excluded	Y	G9	E	C6	F	M&S	EU	Pb	LG	CCR	2005
85	Flammer	2013	AMJ	2A_GC--	Y	G10	C	C6	F	M&S	NA	Pb	LG	2Ds	1995
86	Sariannidis et al.	2013	BSE	2A_GC--	Y	G10	E	C6	F	M&S	CR	Pb	LG	1-3Y	2005
87	Forsman	2013	BSE	4A_GC++	Y	G2	E	C9	F	Mfg	EU	Mx	LG	1-3Y	2006
88	Fujiil et al.	2013	BSE	4A_GC++	Y	G5	P2	C7	F	Mfg	JP	Pb	LG	1Y	2007
89	de Burgos-Jiménez et al.	2013	IJOPM	4A_GC++	N	G10	P1	C7	F	M&S	EU	Mx	CS	1-3Y	2011
90	Leonidou & Katsikeas	2013	JAMS	4A_GC++	Y	G10	P1	C9	P1	M&S	EU	Mx	LG	1Y	2012
91	Fraj et al.	2013	JBIM	4A_GC++	Y	G10	P1	C9	P1	Mfg	EU	Lg	CS	CCR	2011
92	Zhu et al.	2013	JPSM	4A_GC++	Y	G7	P1	C3	P1	Mfg	CN	Mx	CS	CCR	2012
93	Aguilera-Caracuel & Ortiz-de-Mandojana	2013	O&E	4A_GC++	N	G2	P2	C5	F	M&S	CR	Pb	LG	1-3Y	2009
94	Ba et al.	2013	POM	4A_GC++	Y	G2	C	C6	F	Mfg	CR	Pb	LG	3Ds	2003
95	Wong et al.	2013	SS	4A_GC++	Y	G10	P1	C3	P1	Svc	CN	Mx	CS	CCR	2012
96	Yang	2013	IJSTL	4A_GC++	Y	G10	P1	C9	P1	Svc	APO	Mx	CS	CCR	2010
97	Matsumura et al.	2014	AR	2A_GC--	Y	G5	P2	C6	F	M&S	NA	Pb	LG	CCR	2007
98	Linder et al.	2014	BSE	3A_GC+-	Y	G10	P1	C5	P1	M&S	EU	Sm	CS	CCR	2012
99	Moon et al.	2014	BSE	4A_GC++	Y	G6	E	C9	F	M&S	NA	Pb	CS	CCR	1995
100	Wang et al.	2014	BSE	1A_GC-+	Y	G5	P2	C8	F	M&S	ANZ	Pb	CS	CCR	2010
101	Woo et al.	2014	BSE	4A_GC++	Y	G2	P1	C1	P1	M&S	APO	Mx	CS	2Ys	2010
102	Zhang et al.	2014	CHQ	4A_GC++	Y	G6	E	C3	F	Svc	NA	Mx	CS	CCR	2011
103	Dam & Petkova	2014	IJOPM	3A_GC+-	Y	G7	C	C6	F	M&S	NA	Pb	LG	1D	2009
104	Mitra & Ditta	2014	IJPR	4A_GC++	Y	G7	P1	C9	P1	Mfg	APO	Mx	CS	CCR	2013
105	de Jong et al.	2014	JBE	4A_GC++	Y	G6	E	C7	F	M&S	NA	Pb	LG	3Ys+	2001
106	Ghisetti & Rennings	2014	JCP	3A_GC+-	Y	G2	P1	C5	P1	M&S	EU	Mx	LG	2Ys	2010
107	Yu et al.	2014	SCMAIJ	4A_GC++	Y	G10	P1	C9	P1	Mfg	CN	Mx	CS	CCR	2013
108	Bottcher & Müller	2015	BSE	4A_GC++	N	G10	P1	C9	P1	Mfg	EU	Mx	CS	CCR	2013

109	Dangelico & Pontrandolfo	2015	BSE	4A_GC++	Y	G10	P1	C9	P1	M&S	EU	Mx	CS	1-3Y	2012
110	Pereira-Moliner et al.	2015	IJCHM	4A_GC++	Y	G10	P1	C9	P1	Svc	EU	Mx	CS	CCR	2012
111	Lee et al.	2015	IJPE	4A_GC++	Y	G5	P2	C6	F	Mfg	JP	Pb	LG	CCR	2007
112	Djupdal & Westhead	2015	ISBJ	4A_GC++	Y	G6	E	C9	F	M&S	EU	Sm	CS	1Y	2009
113	Cordeiro & Tewari	2015	JBE	4A_GC++	Y	G9	E	C6	F	M&S	NA	Pb	LG	1-3Y	2009
114	Jo et al.	2015	JBE	4A_GC++	Y	G10	E	C5	F	Svc	CR	Pb	LG	1-3Y	2007
115	Muhammad et al.	2015	JBE	4A_GC++	Y	G5	P2	C6	F	M&S	ANZ	Pb	LG	1Y	2006
116	Sanchez-Medina et al.	2015	JBE	4A_GC++	Y	G5	P1	C9	P1	Mfg	SA	Sm	CS	CCR	2012
117	Wagner	2015	JBR	4A_GC++	Y	G10	P1	C9	P1	Mfg	EU	Mx	CS	CCR	2013
118	Muhammad et al.	2015	JCP	4A_GC++	Y	G5	P2	C8	F	M&S	ANZ	Pb	LG	1Y	2004
119	Primc & Cater	2015	MD	Excluded	Y	G10	P1	C9	P1	HPS	ANZ	Mx	CS	CCR	2012
120	Delmas et al.	2015	O&E	2A_GC--	Y	G10	E	C8	F	M&S	NA	Pb	LG	1Y	2006
121	Martinez-Del-Rio et al.	2015	O&E	4A_GC++	Y	G10	P1	C9	P1	M&S	EU	Mx	CS	CCR	2007
122	Rivera-Torres et al.	2015	O&E	4A_GC++	Y	G10	P1	C9	P1	M&S	EU	Mx	LG	CCR	2013
123	Lee et al.	2015	PPC	4A_GC++	Y	G5	P1	C1	P1	Mfg	APO	Mx	CS	CCR	2013
124	Amores-Salvadó et al.	2015	TF&SC	4A_GC++	Y	G3	P1	C2	P1	HPS	EU	Mx	CS	2Ys	2014
125	Fraj et al.	2015	TM	4A_GC++	Y	G10	P1	C9	P1	Svc	EU	Mx	CS	CCR	2014
126	Molina-Azorín et al.	2015	TM	4A_GC++	Y	G10	P1	C9	P1	Svc	EU	Mx	CS	CCR	2014
127	Lee et al.	2016	BSE	4A_GC++	Y	G10	E	C5	F	M&S	APO	Pb	LG	CCR	2012
128	Yadav et al.	2016	BSE	4A_GC++	Y	G9	E	C6	F	M&S	NA	Pb	CS	3Ds	2012
129	Martín-de Castro et al.	2016	CSR&EM	4A_GC++	Y	G6	P1	C9	P1	HPS	EU	Mx	CS	CCR	2014
130	Chopra & Wu	2016	EJOR	4A_GC++	Y	G10	C	C7	F	Mfg	CR	Pb	LG	2Ys	2006
131	O'Donohue & Torugsa	2016	IJHRM	4A_GC++	Y	G10	P1	C9	P1	Mfg	ANZ	Sm	CS	CCR	2015
132	Graham & McAdam	2016	IJOPM	4A_GC++	Y	G10	P1	C7	P1	Mfg	EU	Mx	CS	CCR	2011
133	Esfahbodi et al.	2016	IJPE	4A_GC++	N	G7	P1	C7	P1	Mfg	CN	Lg	CS	CCR	2014
134	Lam et al.	2016	IJPE	3A_GC+-	Y	G10	C	C6	F	M&S	CN	Pb	LG	2Ds	2010
135	Lucas & Noordewier	2016	IJPE	4A_GC++	Y	G9	E	C5	F	Mfg	NA	Pb	CS	CCR	2013
136	Feng & Wang	2016	JBE	4A_GC++	Y	G6	P1	C2	P1	Mfg	CN	Mx	CS	CCR	2014
137	Singh et al.	2016	MD	4A_GC++	Y	G1	F	C5	F	M&S	CR	Lg	CS	1Y	2012
138	Primc & Čater	2016	O&E	4A_GC++	Y	G10	P1	C9	P1	M&S	ANZ	Mx	CS	CCR	2014
139	Capece et al.	2017	BSE	4A_GC++	Y	G5	P2	C5	F	Mfg	EU	Lg	LG	CCR	2011
140	Chiu et al	2017	BSE	4A_GC++	Y	G1	C	C8	P1	Mfg	APO	Pb	LG	CCR	2012
141	Hassan & Romilly	2017	BSE	4A_GC++	Y	G5	P2	C6	F	M&S	CR	Pb	LG	CCR	2010
142	Laari et al.	2017	BSE	4A_GC++	N	G7	P1	C5	F	Svc	EU	Mx	CS	CCR	2014
143	Lewandowski	2017	BSE	4A_GC++	Y	G5	P2	C8	F	M&S	CR	Pb	LG	1Y	2009
144	Li et al.	2017	BSE	4A_GC++	Y	G10	C	C8	F	M&S	NA	Pb	LG	1Y	2013
145	Trumpp & Guenther	2017	BSE	4A_GC++	Y	G5	P2	C8	F	M&S	CR	Lg	LG	1Y	2010

146	Walsh & Dodds	2017	BSE	4A_GC++	Y	G10	P1	C9	P1	Svc	NA	Mx	CS	CCR	2016
147	Yadav et al.	2017	BSE	4A_GC++	Y	G9	E	C5	F	M&S	NA	Pb	CS	CCR	2012
148	Marin & Lotti	2017	ICC	3A_GC+-	Y	G2	E	C9	P1	Mfg	EU	Mx	LG	CCR	2001
149	Dai et al.	2017	IJPE	4A_GC++	Y	G10	P1	C1	P1	Mfg	NA	Pb	CS	CCR	2016
150	Sadovnikova & Pujari	2017	JAMS	4A_GC++	Y	G10	C	C6	F	Mfg	NA	Pb	LG	1Y	2006
151	Du et al.	2017	JBE	4A_GC++	Y	G10	C	C6	F	M&S	CN	Pb	LG	1Y	2010
152	Huang & Li	2017	JBE	4A_GC++	Y	G2	P1	C9	P1	Mfg	APO	Mx	CS	CCR	2015
153	Leonidou et al.	2017	JBE	4A_GC++	Y	G10	P1	C9	P1	M&S	EU	Sm	CS	CCR	2011
154	Schmidt et al.	2017	JSCM	4A_GC++	Y	G7	P1	C9	P1	M&S	EU	Mx	CS	CCR	2016
155	Brulhart et al.	2017	MD	4A_GC++	Y	G10	P1	C5	F	Mfg	EU	Mx	CS	CCR	2010
156	Laguir et al.	2017	MD	4B_CG++	Y	G10	E	C5	F	Svc	EU	Mx	LG	CCR	2010
157	Broadstock et al.	2018	BAR	1B_CG+-	Y	G5	P2	C8	F	M&S	EU	Pb	LG	CCR	2008
158	Brouwers et al.	2018	BSE	4A_GC++	Y	G5	P2	C8	F	HPS	EU	Pb	LG	CCR	2009
159	Feng et al.	2018	BSE	4A_GC++	Y	G5	P1	C9	P1	Mfg	CN	Mx	CS	CCR	2017
160	Gatimbu et al.	2018	BSE	3A_GC+-	Y	G5	P1	C5	F	M&S	AF	Sm	LG	CCR	2014
161	Liao	2018	BSE	4A_GC++	Y	G2	P1	C9	P1	M&S	CN	Mx	CS	CCR	2017
162	Tang et al.	2018	BSE	4A_GC++	Y	G2	P1	C9	P1	Mfg	CN	Sm	CS	CCR	2018
163	Hu et al.	2018	CSR&EM	4A_GC++	Y	G9	E	C8	F	M&S	CN	Pb	LG	CCR	2013
164	Maaloul	2018	CSR&EM	2A_GC--	Y	G5	P2	C8	F	M&S	NA	Pb	LG	CCR	2014
165	El Ghoul et al.	2018	JBE	4A_GC++	Y	G9	E	C8	F	Mfg	CR	Mx	LG	1Y	2007
166	Ramanathan	2018	JBE	4A_GC++	Y	G10	P1	C2	P1	Mfg	EU	Mx	CS	CCR	2010
167	Hirunyawipada & Xiong	2018	JBR	4B_CG++	Y	G10	P2	C8	F	M&S	NA	Pb	LG	1Y	2010
168	Chen et al.	2018	MD	4A_GC++	N	G10	C	C8	F	M&S	CR	Pb	CS	CCR	2014
169	Lo et al.	2018	MSOM	2A_GC--	Y	G8	P2	C6	F	Mfg	CN	Pb	LG	3Ds	2010
170	Zhang et al.	2019	IEEE	4A_GC++	Y	G7	P1	C9	P1	Mfg	CN	Mx	CS	CCR	2014
171	Kularatne et al.	2019	TM	4A_GC++	Y	G10	P1	C4	P1	Svc	APO	Mx	LG	CCR	2015

Abbreviations

ABBR	JOURNAL TITLE
AJM	Australian Journal of Management
AMJ	Academy of Management Journal
AR	The Accounting Review
B&S	Business & Society
BAR	The British Accounting Review
BSE	Business Strategy and the Environment
CHQ	Cornell Hospitality Quarterly
CMR	California Management Review
CSR&EM	Corporate Social Responsibility and Environmental Management
EE	Ecological Economics
EJOR	European Journal of Operational Research
EP	Energy Policy
ETL	Ecotoxicology
FBR	Family Business Review
HRM	Human Resource Management
I&I	Industry and Innovation
ICC	Industrial and Corporate Change
IEEE	IEEE Transactions on Engineering Management
IJCHM	International Journal of Contemporary Hospitality Management
IJHRM	The International Journal of Human Resource Management
IJOPM	International Journal of Operations & Production Management
IJPDLM	International Journal of Physical Distribution & Logistics Management
IJPE	International Journal Production Economics
IJPR	International Journal of Production Research
IJSTL	International Journal of Shipping and Transport Logistics
ISBJ	International Small Business Journal
JAMS	Journal of the Academy of Marketing Science
JBE	Journal of Business Ethics
JBFA	Journal of Business Finance & Accounting
JBIM	Journal of Business & Industrial Marketing
JBR	Journal of Business Research
JCP	Journal of Cleaner Production
JEEM	Journal of Environmental Economics and Management
JEM	Journal of Environmental Management
JEP	Journal of Economic Perspectives
JIE	Journal of Industrial Ecology
JIM	Journal of International Management
JMS	Journal of Management Studies
JOM	Journal of Management
JOM	Journal of Operations Management
JPSM	Journal of Purchasing & Supply Management
JSCM	Journal of Supply Chain Management
JWB	Journal of World Business
MD	Management Decision
MS	Management Science
O&E	Organization & Environment
PPC	Production Planning & Control
SB	Service Business

SMJ	Strategic Management Journal
SS	Service Science
TF&SC	Technological Forecasting & Social Change
TM	Tourism Management
TRPE	Transportation Research Part E

ABBR	CODE	CODE SCALE
RF	Reference no.	Running no of article
AUTHOR(S)	Author(s)	Author(s)
YR	Year	Year of publication
JR	Journal	Journal of publication
HY	Hypothesis	1A_GC-+ (Grey Exploitation) 2A_GC-- (Grey Penalty) 3A_GC+- (Green Liability) 4A_GC++ (Green Reward) 1B_CG+- (Irresponsible Growth) 2B_CG-- (Resource Constraints) 3B_CG-+ (Resource Diversion) 4B_CG++ (Slack Resources)
FD	Finding	Y: Hypothesis supported N: Hypothesis rejected
G	EM/EP measures	G1_EM Investments G2_Green Innovation G3_Product Design, G4_Pollution Reduction Practices G5_Pollution Output G6_Env Certification G7_Green SCM G8_Incidents & Penalties G9_Rating, Ranking & Awards G10_Overall & Others
GDT	EM/EP data type	A: Anecdotal C: Communicative E: Evaluative F: Financial P1: Perceptual P2: Physical
C	BP measures	C1_Produtivity & Efficiency C2_Customer & Market C3_Costs C4_Revenue C5_Profitability C6_Market Valuation C7_Composite Accounting Measures C8_Composite Accounting & Market C9_Overall & Others

CDT	BP data type	A: Anecdotal C: Communicative E: Evaluative F: Financial P1: Perceptual
ST	Sector	HPS: High-polluting sectors Mfg: Manufacturing general M&S: Manufacturing & services Svc: Services sectors
RG	Region	AF: Africa ANZ: Australia & New Zealand APO: Asia Pacific Others CN: China CR: Cross-Regional EU: Europe JP: Japan MENA: Middle East & North Africa NA: North America SA: South America
FS	Firm size	Lg: Large, Mx: Mixed, Pb: Public, Sm: SMEs
TF	Data timeframe	CS: Cross-sectional LG: Longitudinal
TL	Data time-lag	CCR: Concurrent 1D: 1 Day 2Ds: 2 Days 3Ds: 3 Days 1Y: 1 Year 2Ys: 2 Years 1-3Ys: 1-3 Years 3Ys+: 3 Years +
DA	Data age	Cross-sectional data: as reported, or the latest citation year Longitudinal data: mean year round-up

Appendix II – GC Construct Operationalization

MEASURE	EXAMPLES AND DESCRIPTION
Green	
G1_Green Investments	Investment & expenditure on EM
G2_Green Innovation	R&D and innovation to reduce process/product EIs
G3_Product Design	Modification on product design to reduce EIs
G4_Pollution Reduction Practices	Preventive/corrective practices for pollution reduction
G5_Pollution Output	Pollution output after pollution prevention/control
G6_Environmental Certification	EMS 14001 & other environmental certifications
G7_Green SCM	Green procurement, logistics, customer collaboration
G8_Incidents & Penalties	Environmental incidents/penalties/lawsuits
G9_Rating, Ranking & Awards	Third-party environmental rating/ranking/awards
G10_Overall & Others (Not Elsewhere Classified)	Environmental strategies/policies/capabilities, staff training/education/pay-links, participation in environmental programs, and other measures not elsewhere classified
Competitiveness	
C1_Produtivity & Efficiency	Productivity, quality, delivery, and flexibility; may be referred to as *manufacturing performance*
C2_Customer & Market	Customer satisfaction & loyalty, success of new product launch, market share
C3_Costs	Cost-related measures, including accounting-based financial measures & non-financial measures
C4_Revenue	Revenue-related measures, including accounting-based measures & non-financial measures
C5_Profitability	Financial measures on profitability, such as ROS/ROA/ROE/ROIC
C6_Market Valuation	Financial measures on stock market performance
C7_Composite Accounting	Combinations of different accounting-based financial measures in C3/C4/C5, such as revenue growth & profitability
C8_Composite Accounting & Market	Combinations of different accounting-based & market measures in C6/C7, such as Tobin's q, P/E ratio, and market-to-book ratio
C9_Overall & Others (Not Elsewhere Classified)	Other measures not elsewhere classified