Green Procurement: Collecting Evidence on the Practice-Performance Link by Employing a Meta-Analysis

Alessa Kozuch, Maurice Langen, Christian von Deimling and Michael Eßig
Green procurement: Collecting evidence on the practice-performance link by employing a meta-analysis

Alessa Kozuch*, Maurice Langen, Christian von Deimling and Michael Essig
Purchasing Research Group
Bundeswehr University Munich, Neubiberg, Germany

Abstract
The need for a transformation to an environmentally friendly economy is undisputed. Procurement has a steering function, as considerable influence can be exerted on the upstream supply chain. The incorporation of environmental ambitions in purchasing strategies, policies and processes has gained widespread interest in academic research. In both private and public procurement, recent research examines whether the implementation and uptake of green purchasing practices not only takes place in an efficient and effective manner, but whether the implementation and uptake can also improve the performance of an organisation. While studies on measuring the influence of green procurement practices on organisations' performance are available, they come to different conclusions. Hence, this paper, examines 22 empirical studies by employing a meta-analytical methodology. The results demonstrate, that adopting green procurement practices has a positive effect on organizational performance.

Keywords: green procurement, green procurement practices, performance, meta-analysis, implementation

CRediT authorship contribution statement:
Alessa Kozuch: Conceptualization, Methodology, Investigation, Validation, Writing - Original Draft
Maurice Langen: Conceptualization, Methodology, Data Curation, Formal analysis, Investigation, Writing - Original Draft
Christian von Deimling: Validation, Writing - Review & Editing
Michael Essig: Writing - Review & Editing, Supervision

Declaration of competing interest:
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. Maurice Langen has created an earlier version of this inquiry as part of his Master Thesis.

Submission category: Academic working paper

* Corresponding author: alessa.kozuch@unibw.de
1 Introduction

The incorporation of environmental ambitions in the purchasing strategies, policies and processes has gained widespread interest in academic research (e.g. Khan et al. 2022a; Khan et al. 2022b; Hazaea et al. 2022). In both private and public procurement, recent research examines whether the implementation and uptake of green purchasing practices not only takes place in an efficient and effective manner, but whether the implementation and uptake can also improve the performance of an organisation (e.g. Cheng et al. 2018; Malacina et al. 2022; Johnson & Klassen 2022; Khan et al. 2022a; Khan et al. 2022b; Hazaea et al. 2022).

There are already studies that empirically investigate the performance of green procurement as part of green supply chain management (GSCM) to provide organisations with measuring constructive guidance for adopting certain practices (Fang & Zhang, 2018; Golicic & Smith, 2013; Govindan, Rajendran, Sarkis, & Murugesan, 2015; Zhu & Sarkis, 2004). The results of previous empirical studies are not always consistent with each other, which may confuse practitioners and policy-makers alike in the adoption of GSCM and green purchasing and hinder further development. Rao and Holt e.g. stated that companies with GSCM significantly increase both their competitiveness and economic performance (Rao & Holt, 2005). Zhu, Sarkis, & Lai (2007) on the other hand, concluded that there is hardly a significant improvement in economic performance among companies with GSCM. However, a meta-analytic summary of those results and an indication of their variability in terms of results seems to be missing to date (Stamm & Schwarb, 1995). Thus, the following question should be answered by employing a meta-analytic approach:

RQ: What effect do green procurement practices have on the performance of the organisation in total?

Meta-analysis is a statistical technique that aims to quantitatively summarise research findings from a large number of studies and has been used as an effective analytical tool in medical and clinical settings for over two decades (e.g., (Bowater, Hartley, & Lilford, 2015; Lau, Schmid, & Chalmers, 1995), and has more recently been used in management settings (Gerrish, 2016; Geyskens, Krishnan, Steenkamp, & Cunha, 2009; Melo, Graham, & Noland, 2009). In accordance with medical research evidence levels, this type research methodology is typically based on large-scale studies and previously published research results (Reay, Berta, & Kazman Kohn, 2009). Hence, a more comprehensive quantitative analysis of the extensive GSCM and green procurement literature will be undertaken that can shed light on these inconsistencies in the empirical findings. The present study is dedicated to a meta-analysis that measures the effect of environmentally sustainable procurement practices on company performance. In this study, the method of correlation-based bare-bones meta-analysis is used.

2 Understanding the practice-performance link when procuring green

2.1 Understanding the concept of ‘green procurement practice’

Overall, procurement ‘practices’ are defined as a set of "activities that relate to the interface between the supply base and purchasing" (Narasimhan & Das, 2001) and aim at altering
organisational performance (Das & Narasimhan, 2000). Such practices may entail designing a supply chain according to social, ecological, and economic aspects and the integration of sustainability criteria into the supplier selection process (Mahammadzadeh, 2012). It is thus important to implement concrete criteria in the individual processes from the organisational culture and strategy (Geroski, 1990). Hence ‘green procurement practice’ may entail the integration and adoption of sustainability and related criteria into purchasing activities, from a sound adoption in purchasing strategies to an integration of relevant criteria into executional purchasing procedures (Koplin, 2006).

If one looks at the procurement process, the green procurement practice can be implemented or adopted as follows: First, the needs assessment takes place by already considering the sustainability goals described by the company. This is followed by the definition of an order-related reference strategy, which serves as a guideline for sustainable procurement for the procurers. The next step is the search for "green" suppliers and "green" supplies and services (D’Agostini et al., 2017) before they are evaluated. One of the most important collaboration skills for sustainability and resource-based strategies in networks is supplier selection as a component of sustainable supply chain management (Almeida, Gohr, Morioka, & Da Medeiros Nóbrega, 2021). The selection of environmentally friendly suppliers is thus an important factor for cooperation and also the performance of the procuring company (Münch et al., 2022; Song, Yu, & Zhang, 2017; Yu, Chavez, & Feng, 2017). The evaluation of tenders and suppliers is carried out according to the predefined sustainability criteria, for which e.g. activities of recycling and reuse, eco-labels, and quality labels (e.g. the "Blue Angel", the "EU Ecolabel" or the "Nordic Swan") or the application of a life cycle cost analysis are declared as sustainability practices (Igarashi, Boer, & Michelsen, 2015). After a calculation has been made, a distinction can be made between a conventional and an environmentally friendly product and the extent to which the environmentally friendly variant performs better overall than the conventional one from an economic point of view (Jenssen & Boer, 2019). Furthermore, for example, the production methods of the company under consideration of the external costs in the event of negative external effects can be included in the decision (D’Agostini et al., 2017). After the tender and supplier selection decision has been made and a contract has been concluded, performance measurement and management is carried out based on "key performance indicators", also including green criteria (such as degree of water consumption/pollution, energy consumption or CO2-Emissions). How these are defined and configured exactly depends on the individual needs of the respective organization. The last step serves to review achievements and learn from the results (Zimmer, Fröhling, & Schultmann, 2016).

2.2 Understanding the concept of green procurement and organizational performance

Seuring and Müller particularly emphasise the role of the actors involved in a supply chain from the service provider, customer, and supplier perspectives (Seuring & Müller, 2008). Only in this way is appropriate performance measurement possible (Saeed & Kersten, 2020). Uncertainty still seems to prevail in the implementation of performance measurement (Bocken, Morgan, & Evans, 2013). To meet the interests of different stakeholders, (Green, Zelbst, Meacham, & Bhadauria, 2012) have formed four categories of performance measurement for sustainability: environmental, economic, organisational, and operational performance.
measurement. In the area of ecology, for example, emissions, water, or material savings can be measured. Economic performance ensures the survival of the organisation and is designed to reduce costs. In the operational area, scrap rates, quality requirements, and process efficiency could be measured. Organisational performance is influenced by the other categories in the context of reputation and customer satisfaction and reflects financial and market performance. For the following analysis, the operational, market and time-related performance of procurement play a particularly important role (see section 2.3). To be able to measure the far-reaching effects of sustainability measures, a multidisciplinary approach to measuring the performance of supply chains is needed (Bai & Sarkis, 2014). This could both optimise actions within the supply chain and improve sustainability performance (Ahi & Searcy, 2015).

2.3 Collecting evidence on the practice-performance link when procuring green

In recent years, some meta-analyses have been conducted on the practice-performance link in procurement and supply chain management (D’Agostini et al., 2017; Fang & Zhang, 2018; Geng, Mansouri, & Aktas, 2017; Golicic & Smith, 2013; Zimmermann & Foerstl, 2014). The link to environmental sustainability has also been included (D’Agostini et al., 2017; Fang & Zhang, 2018). The overall links have been reported as consistently positive. In addition, moderating effects such as relationships with suppliers (Zimmermann & Foerstl, 2014), organisational size (Geng et al., 2017), industry, culture, certification (Fang & Zhang, 2018), etc. were found.

Table 1: Evidence from existing Meta-Analyses about Practice-Performance-Link

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Practice (Practices)</th>
<th>Performance</th>
<th>n</th>
<th>Effect</th>
<th>Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golicic/Smith</td>
<td>2013</td>
<td>Environmentally sustainable SCM Practices (Upstream supplier-facing, Design, Production, downstream customer-facing)</td>
<td>Operational, Market &amp; accounting Performance</td>
<td>89</td>
<td>positive</td>
<td>Companies should consider environmental aspects when developing new products or processes and when contracting suppliers.</td>
</tr>
<tr>
<td>D’Agostini et al.</td>
<td>2017</td>
<td>Sustainable operations practices (Ecodesign, GSC, Cleaner Production, Reverse Logistics)</td>
<td>Environmental, economic, operational, organisational Performance</td>
<td>37</td>
<td>positive</td>
<td>Contingency influences the relationship</td>
</tr>
<tr>
<td>Geng et al.</td>
<td>2017</td>
<td>Green SCM (Intra-organisational management, supplier integration, eco-design, customer cooperation, reverse logistics)</td>
<td>Environmental, economic, operational, and Social Performance</td>
<td>50</td>
<td>positive</td>
<td>Industry affiliation, company size, ISO certification, and export orientation moderate some of the relationships between GSCM practices and performance.</td>
</tr>
<tr>
<td>Fang/Zhang</td>
<td>2018</td>
<td>Green SCM (Green Purchasing, Customer Cooperation, Investment Recovery, Eco-Design)</td>
<td>Environmental, economic &amp; operational Performance</td>
<td>54</td>
<td>positive</td>
<td>Industry type, ISO certification, export orientation, and the cultural dimension of uncertainty avoidance have a moderating effect.</td>
</tr>
</tbody>
</table>

However, there is not yet a study that is, on the one hand, very up-to-date and, on the other hand, quite explicitly and exclusively focused on green procurement practices. Thus, the
practice-performance link between green purchasing/green procurement (GP) and corporate performance (CP) within GSCM is quantified through meta-analysis using a population effect size. The concretisation of the GP measures and the "corporate performance" dimensions was based on the theoretical foundations of the studies by (Younis, Sundarakani, & O'Mahony, 2020) and (Golicic & Smith, 2013). The four main measures of the GSCM are "Eco Design", "Green Purchasing/Procurement", "Environmental Cooperation" and "Reverse Logistics". In contrast, the three most frequently observed dimensions of corporate performance within studies are "Operational Performance" (OP), "Market Performance" (MP), and "Financial Performance" (FP) (Zimmermann & Foerstl, 2014). Within the GSCM, only the measures of "Green Purchasing/Procurement" are considered for the literature review, which has already been explained and is defined by (Younis et al., 2020) as follows: "An environmental purchasing initiative that aims to ensure purchased products and material meet with environmental objectives set by the purchasing firm [...]". The composition of company performance consists of the three dimensions mentioned above. According to (Golicic & Smith, 2013) "Operational Performance" includes quality, capacity utilisation, flexibility, and speed of delivery. "Market performance" includes market share and customer satisfaction. Concerning "financial performance", financial indicators such as ROI, EBITDA, or profit, in general, are used. Figure 1 shows the conceptual framework on, which the following meta-analysis is based.

Figure 1: Conceptual Framework

The operationalization of the GSCM measures and the "corporate performance" dimensions was carried out deductively based on the studies by (Younis et al., 2020) and (Golicic & Smith, 2013). The result can be seen in Appendix A.

2.4 Theoretical insights from the natural resource-based view (NRBV) as underpinning of the practice-performance link when procuring green

As an accepted theory to explain the relationship between sustainability-related resources and capabilities and a company's competitive advantage, the natural resource-based view (NRBV) forms the basis for research in various supply chain and procurement-focused studies (AlNuaimi, Singh, & Harney, 2021; Andersén, 2021; Münch, Benz, & Hartmann, 2022; Ngo, 2021). (Hart, 1995) introduced the NRBV view first as a further development of Wernerfelt (1984) resource-based view (RBV).

The RBV has served for almost three decades as a basis for strategic management to explain the achievement of competitive advantage (Wernerfelt, 1984). Its origins can be traced back to the economic theories in which companies were seen as collections of resources. These resources can be human skills, physical assets, organisational processes as well as knowledge
held by the firm (Barney, 1991; Wernerfelt, 1984). According to the theory, the combination of different resources, as well as the way they are managed, influences the capabilities of the company. They get a competitive advantage when the resources are valuable and inimitable, so there is no equivalent substitute. Procurement practices can help companies collaborate and share key attributes (Barnes & Liao, 2012). According to RBV, the time-consuming development of extravagant attributes, that would be costly to replicate, creates performance differences due to specific sourcing practices (Knudsen, 2003; Zimmermann & Foerstl, 2014). The impact of environmental practices in the supply chain on performance is also supported by the RBV (Peng & Lin, 2008; Wu, Melnyk, & Calantone, 2008), with the performance measure being specifically tied to cost benefits achieved (Christmann, 2000) or competitive advantages (Chen, Lai, & Wen, 2006; Chiou, Chan, Lettice, & Chung, 2011). However, in Hart’s view (Hart, 1995), the RBV ignores the challenges and constraints imposed by the natural environment that should be taken into account when developing new resources and capabilities. The acceleration in the scale and scope of human activities require more than previous economic and organisational practices to continue to deliver equivalent outcomes. Building on these considerations, Hart proposed the NRBV, which assumes that future competitive advantage will be based on "capabilities that enable ecologically sustainable economic activity" (p. 991).

Research has repeatedly confirmed this link, meaning that supply chain processes have direct impacts on the natural environment and that practices to manage and reduce these impacts can develop capabilities to improve performance (Vachon & Klassen, 2008). The NRBV extends the RBV theory by incorporating a company's interaction with its natural environment and includes three interrelated strategic capabilities to achieve economic and environmental sustainability to create competitive advantage: Pollution prevention, product stewardship, and sustainable development (Hart, 1995). Pollution prevention is about reducing pollution or inefficient use of material and human resources in the production process, e.g. by avoiding waste and emissions. This leads to improved operational performance through better use of inputs, shorter cycle times, and lower costs. As green procurement, and in particular supplier selection, has a significant impact on pollution (Kondo, Kinoshita, & Yamada, 2019), it also has a positive impact on achieving this capability. Product stewardship aims to integrate stakeholder perspectives into the product. It considers activities at each stage of the value chain throughout the life cycle of a product, from design to disposal. The aim is to ensure access to resources for environmentally friendly materials from a procurement perspective to minimise environmental costs throughout the life cycle, thereby creating a potential competitive advantage (Hart, 1995). The scientific theoretical background of the positive influence of green procurement practices on organisational performance is thus already underpinned by the NRBV. Before implementation, however, it is still necessary to determine both green procurement practices and performance indicators (as indicated above).
3 Methodology

3.1 Overall research design

The term meta-analysis was first introduced in 1976 by Gene V. Glass who describes it as follows: “Meta-analysis refers to the analysis of analyses.” (GLASS, 1976). The core intention of meta-analysis is to create an overview of research by integrating empirical primary results based on an orientation towards empirical criteria. To produce a meta-analysis, a basic precondition applies to the empirical primary investigations. The observed primary investigations must have conceptual comparability, i.e., they must have a common research question (Drinkmann, 1990). Result bias can be explained by certain problems researchers face when preparing and assessing a meta-analysis (Stegenga, 2011). There are different ways of conducting a meta-analysis. This paper uses the methods of correlation-based bare-bones meta-analysis according to Hunter and Schmidt, which primarily uses the correlation coefficient “r” as an effective measure for further investigation. The bare-bones meta-analysis aims to calculate a population effect size, to free it from interfering influences to be able to make an overall statement about an effect under consideration at the end. The procedure of the bare-bones meta-analysis can be roughly divided into five steps (Stamm & Schwarb, 1995). Apart from formulating the research question the following steps are crucial for a meta-analysis:

1. Clarify prerequisites for conducting a meta-analysis
2. Literature analysis and preparation of a meta-analytical overview.
4. Correction of possible errors and biases.
5. Calculation of the population effect size and discussion.

3.2 Prerequisites for conducting a meta-analysis

In the literature on methodology, the following four issues are listed most frequently: The “Apples and Oranges” problem, the “Garbage in – Garbage out” problem, the “Publication Bias” and the “Non-independent Effects” or “Multiple Effect Sizes” problem. The “Apples and Oranges” problem refers to the literature search and the preparation of the meta-analytical balance. When searching for literature on the desired effect sizes, it is usually not possible to find identical replication studies. Thus, studies are mixed up that differ in terms of operationalisations, samples, or evaluation methods. This makes comparability between the studies more difficult (Sharpe, 1997). The results of meta-analyses that focus on narrowly defined effect sizes (only “apples”) are biased by the addition of other effects (“oranges”) (Card, 2012). However, combining multiple effect sizes (“apples and oranges”) can be useful for meta-analyses that use different effect sizes for multiple hypotheses. In other words, “Indeed the approach does mix apples and oranges, as one would necessarily do in studying fruit” (Smith et al., 1980). Two approaches to meta-analysis can be derived from this. The first approach is used primarily in the natural sciences and requires the exclusive use of perfect replications. The second approach allows the use of imperfect replications by explaining the heterogeneity based
on moderator variables and still allowing an interpretable result to be determined at the end. In addition to the different effect sizes of primary studies, the quality of the empirical methods used in the studies also differs. The variability of meta-analytical results is influenced by the quality of the individual studies (Kock, 2009). The combination of studies with different methodological qualities can lead to misleading results and is called “garbage in – garbage out” (Sharpe, 1997). To avoid this problem, the number of studies can be limited based on defined quality characteristics (Borenstein, 2009). Publication bias is considered one of the greatest threats to the validity of meta-analyses. It occurs when the summarised published research results are not representative of all studies and a different overall result emerges when all research is considered. Significant results are often promoted, while non-significant results are often not published and thus assigned to the grey literature. This creates an underreporting of primary studies, which are difficult or impossible to include in meta-analyses (Eisend, 2009). Publication bias can be visualised with the help of a funnel plot. The next problem that is often found in the literature is “non-independent effects” or “multiple effect sizes”. This problem refers to the fact that several effect sizes within a study with the same sample size are not statistically independent of each other. Possible solutions could be that only the most important effect sizes per study are selected, or the effect sizes are summarised through median or mean value formation (Döring & Bortz, 2016). The “practice-performance link” between “Green Purchasing/Green Procurement (GP) and “Corporate Performance (CP) within “Green Supply Chain Management” (GSCM) is quantified by meta-analysis using a population effect size.

3.3 Data collection and preparation

First, there has to be conducted a literature review for collecting relevant data for the meta-analysis. The exclusion criteria are divided into two categories. First, formal exclusion criteria (C1) are examined without specifically addressing the content of the studies. This is followed by the content-related criteria (C2), which must be fulfilled to include studies in the meta-analytical review. To create a representative overview, the criterion of methodological quality must not be taken into account. Methodological quality plays a role in the interpretation of the results (Fricke & Treinies, 1985). It is possible to establish further exclusion criteria. However, since the data collection has to be as comprehensive as possible, no further restrictions are imposed on regional affiliation or company size.

<table>
<thead>
<tr>
<th>Table 2: Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category</strong></td>
</tr>
</tbody>
</table>
| Formal Exclusion Criteria (C1) | • The primary studies must have an effect size.  
• The effect size must be a Bravais-Pearson correlation coefficient.  
• Studies must have been published after 2000 (minimising publication bias).  
• The studies must be written in English or German.  
• The sample size of the study under consideration must be > 10. |
| Content-Related Criteria (C2) | • The primary studies must include the investigated effect of the “practice-performance link” between green procurement measures and company performance.  
• The green procurement measures investigated only include measures that are also referred to as such in the literature.  
• For business performance, at least one of the three dimensions must be used as a variable. |
Search terms, search combinations, and keywords were collected for the data collection.

### Table 3: Search Strings

<table>
<thead>
<tr>
<th>Category</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sustainability</td>
<td>Green, Sustainable, Sustainability, Environmentally</td>
</tr>
<tr>
<td>Corporate Performance</td>
<td>Performance, Corporate Performance, Firm Performance, Market Performance, Financial Performance, Operational Performance, Company Performance, Organizational Performance, Competitive Advantages</td>
</tr>
<tr>
<td>Study Type:</td>
<td>Study, Primary Study, Quantitative Study, Journal Study, Research, Analysis</td>
</tr>
<tr>
<td>Search Combination:</td>
<td>Practice-Performance Link; Green Procurement/Purchasing, Study; Research; Analysis; Green Procurement/Purchasing</td>
</tr>
</tbody>
</table>

As databases, there were used EconBiz, GoogleScholar, Emerald, DeepDyve, Science Direct, and Wiley Online Library. For improved transparency and visualisation of the research process, the creation of a flow chart is recommended. The “PRISMA Checklist” was used to create the flow chart (DeSimone et al. 2020). The aim of the “PRISMA statement” is to improve the search for systematic reviews and meta-analyses (Moher, Liberati, Tetzlaff, & Altman, 2010). The flow chart for this meta-analysis was modified because the literature search is based exclusively on the databases mentioned above. Thus, the extended search from other sources is omitted in the identification phase. A total of 1686 studies were identified by entering the search strings.

![Figure 2: Data Collection Protocol](image)

### 3.4 Data analysis and validation

Lore Ipsum The methods of a weighted bare-bones meta-analysis according to Hunter and Schmidt (2004) are used to estimate the population correlation. If the population correlation is
assumed to be constant across studies, the estimation of this correlation is not recommended by a simple calculation of the mean, but by weighting with the number of samples. The best estimate of the population correlation is thus:

The associated variance is the frequency-weighted average squared deviation:

$$ \bar{r} = \frac{\sum [N_i r_i]}{\sum N_i} \quad (1) $$

$$ s^2_r = \frac{\sum [N_i (r_i - \bar{r})^2]}{\sum N} \quad (2) $$

An analysis by Hunter and Schmidt shows that only in very rare cases unweighted analyses are better than weighted ones. Moreover, the question arises whether a transformation of the correlation coefficient into Fisher’s z would be preferable. Such a transformation leads to distortions since larger correlations are weighted more strongly than smaller correlations and thus the population correlation turns out more positive overall (Hunter, Schmidt 2004). From the observed variance $s^2_r$, the variation of the population correlation on the one hand, and the variation of the sample correlation, due to sampling errors, on the other hand, can be taken. Thus, it is possible to estimate the variance of the population correlation by correcting the sampling errors of the observed variance $s^2_r$. For each study, a true population correlation $p$ can thus be estimated, which can be compared with the study correlation $r$. The difference between the two correlations is the population correlation.

The difference between the two correlations is the sampling error $e$.

$$ e = r - p $$

$$ r = p + e $$

If one ignores minor distortions of the correlation coefficient, the average sampling error is 0 and the standard deviation depends on the sample size. Consequently, if one calculates the variance of the study correlation, then this is the sum of the variance of the population correlation and the variance of the sampling errors.

$$ \Sigma^2_r = \sigma^2_p + \sigma^2_e $$

The formula consists of three variances. If only two are known, the third variance can be calculated. From the available data, the variance of the study correlations and the variance of the sampling errors can be calculated to calculate the variance of the population correlation at the end. For the calculation of the variance of the sampling errors, Hunter and Schmidt derive the following formula as an approximation:
where \( \bar{r}^2 \) denotes the squared empirical mean and \( \bar{N} \) the average sample size. If formula 4 is converted to \( \sigma_p^2 \) and formula 5 is substituted into formula 4, the equation is:

\[
\sigma_p^2 = \sigma_r^2 - (1 - \bar{r}^2)^2 / (\bar{N} - 1)
\] (6)

It is also possible to correct the individual effect sizes for measurement errors. Reliabilities in the form of the measure “Cronbach’s Alpha” are required for the correction (Grief, Altmann, & Bogaschewsky, 2017). The reliabilities are estimated with the help of “Cronbach’s Alpha” coefficient as a measure of the internal consistency of a scale. In this meta-analysis, this applies to both the independent variable “green procurement/purchasing” and the dependent variable “corporate performance”. From most studies, the reliabilities could be taken in the form of “Cronbach’s Alpha” or the “Composite Reliability”. Following the examples of Younis et al. (2020) and Song et al. (2017), the lowest, still acceptable reliability can be set at 0.7. The formula by Hunter and Schmidt can be used to correct measurement errors:

\[
r_c = \frac{r_{xy}}{\sqrt{r_{xx} r_{yy}}}
\] (7)

Here \( r_c \) denotes the corrected effect size, \( r_{xy} \) the uncorrected effect between the two variables under consideration and \( r_{xx} \), respectively \( r_{yy} \) the reliabilities. “Cronbach’s Alpha” coefficient is used for the reliabilities:

\[
r_c = \frac{r_{xy}}{\sqrt{\alpha_{GPP} \alpha_{CP}}}
\] (8)

By including the reliabilities, the weighting of the studies from Formula 1 can also be adjusted. A simple weighting as in formula 1 does not take into account the information quality of the different studies. Studies with high reliability should be weighted higher than studies with lower reliability. The formula of the new weighting is therefore:

\[
w_i = N_i A_i^2 \quad with \quad A_i^2 = \sigma_{GPP} \sigma_{CP}
\] (9)

\( w_i \) describes the new weighting, \( A_i^2 \) the squared composite “Artefact Attenuation Factor”, which is the product of the “Cronbach’s Alpha” coefficients of the “Green Purchasing/Procurement” variable and the “Corporate Performance” variable. \( N_i \) is replaced by \( N_i \) in all the formulas mentioned.
3.5 Methodological limitations

As described above, meta-analyses are frequently used in empirical sciences. Although empirical methods are an integral part of economic research, the frequency of perfect replication studies is low. Accordingly, meta-analysis as a research approach to integrating multiple types of research is rather rare in economics. Concerning the “apples and oranges” problem, it can be stated that the selected studies are not perfect replications. Thus, the framework conditions under which the studies were conducted are different. However, given the two approaches that have arisen, perfect replications are not necessary to conduct a meta-analysis, provided that the individual effect sizes of the respective studies are weighted differently. To avoid the “garbage in – garbage out” problem, formal and substantive exclusion criteria were established. Although it would be possible to exclude poor-quality studies, such a procedure is associated with a high loss of information. By using the reliability coefficient “Cronbach’s Alpha”, qualitative differences could be taken into account by correcting measurement errors (Grief et al., 2017). Some studies showed “Cronbach’s Alpha” reliabilities below 0.7. In the literature, a threshold value of $\alpha = 0.8$ is described as good, but for shorter scales values from $\alpha = 0.6$ are still used. The range of the “Cronbach’s Alpha” coefficient for the work is: $[0.604;0.97]$. Some studies showed multiple relevant effect sizes within the research results. The problem of “multiple effect sizes” was solved by averaging, as recommended in the literature. However, inferior bias cannot be prevented by averaging. After considering the meta-analytic balance, we can now describe the limitations and possible biases of the bare-bones meta-analysis approach. For these meta-analyses, the methods of bare-bones meta-analysis according to Hunter and Schmidt were used to correct for sampling errors. Furthermore, by specifying the reliabilities within the studies, attenuations due to measurement errors could be corrected. However, biases in the results could arise because the bare-bones meta-analysis does not correct all possible artefacts described by Hunter and Schmidt. Thus, a bare-bones meta-analysis forms the basis for an extended investigation if the data of the selected studies are sufficient and meet the increased requirements.

4 Preliminary findings

For the meta-analytical summary, the studies were classified with the respective author and year, the journal, the dimension of “corporate performance” considered within the studies, the region, the sample size N and the effect size. The studies were sorted in ascending order from top to bottom in terms of their effect size. The Bravais-Pearson correlation coefficient was used as the effect size. In the case of “Multiple Effect Sizes”, the mean of the relevant effects was formed. The sample can be seen in Table 5.
The individual reliability data as well as the corrected effects can be found in Table 6.

<table>
<thead>
<tr>
<th>Autor (Jahr)</th>
<th>Journal (Jahr)</th>
<th>Variablen</th>
<th>Region</th>
<th>Effektgröße</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhu u.a. (2011)</td>
<td>Journal of Industrial Ecology</td>
<td>FP</td>
<td>China</td>
<td>0,189</td>
<td>396</td>
</tr>
<tr>
<td>Kalyar u.a. (2019)</td>
<td>Sustainability Accounting, Management and Policy Journal</td>
<td>FP</td>
<td>Pakistan</td>
<td>0,223</td>
<td>387</td>
</tr>
<tr>
<td>Najmi u.a. (2020)</td>
<td>Int. J. Business Performance and Supply Chain Modelling</td>
<td>FP, MP, OP</td>
<td>Pakistan</td>
<td>0,245</td>
<td>207</td>
</tr>
<tr>
<td>Tan u.a. (2019)</td>
<td>Int. J. Sustainable Strategic Management</td>
<td>FP, OP</td>
<td>Malaysia</td>
<td>0,262</td>
<td>122</td>
</tr>
<tr>
<td>Çankaya u.a. (2019)</td>
<td>Journal of Manufacturing Technology Management</td>
<td>FP, OP</td>
<td>Turkey</td>
<td>0,274</td>
<td>281</td>
</tr>
<tr>
<td>Gosh (2018)</td>
<td>J. Manufacturing Technology Management</td>
<td>FP, MP, OP</td>
<td>Indien</td>
<td>0,315</td>
<td>80</td>
</tr>
<tr>
<td>Lee u.a. (2012)</td>
<td>Industrial Management &amp; Data Systems Journal</td>
<td>FP, OP</td>
<td>Südkorea</td>
<td>0,315</td>
<td>273</td>
</tr>
<tr>
<td>Namagambe u.a. (2019)</td>
<td>Int. J. Management of Environmental Quality</td>
<td>OP</td>
<td>Uganda</td>
<td>0,325</td>
<td>200</td>
</tr>
<tr>
<td>Yu u.a. (2014)</td>
<td>Int. J. Supply Chain Management</td>
<td>FP, OP</td>
<td>China</td>
<td>0,365</td>
<td>126</td>
</tr>
<tr>
<td>Jawaad, Zafar (2020)</td>
<td>Sustainable Development</td>
<td>FP, OP</td>
<td>Pakistan</td>
<td>0,390</td>
<td>272</td>
</tr>
<tr>
<td>Choi u.a. (2018)</td>
<td>Int. J. Logistics Management</td>
<td>MP, OP</td>
<td>Südkorea</td>
<td>0,373</td>
<td>300</td>
</tr>
<tr>
<td>Song u.a. (2015)</td>
<td>Int. J. Logistics Management</td>
<td>FP, OP</td>
<td>China</td>
<td>0,374</td>
<td>206</td>
</tr>
<tr>
<td>Ismail u.a. (2019)</td>
<td>Int. J. Logistics Systems and Management</td>
<td>FP, MP</td>
<td>Jordanien</td>
<td>0,388</td>
<td>79</td>
</tr>
<tr>
<td>Mishra u.a. (2019)</td>
<td>Theoretical Economic Letters</td>
<td>FP, MP, OP</td>
<td>Indien</td>
<td>0,389</td>
<td>374</td>
</tr>
<tr>
<td>Green Jr. u.a. (2012)</td>
<td>Int. J. Supply Chain Management</td>
<td>FP, MP, OP</td>
<td>USA</td>
<td>0,438</td>
<td>159</td>
</tr>
</tbody>
</table>

The individual reliability data as well as the corrected effects can be found in Table 6.

Table 5: Table of Results (Cronbach’s Alpha and Composite Reliability)

<table>
<thead>
<tr>
<th>Autor (Jahr)</th>
<th>α GPP</th>
<th>α CP</th>
<th>r</th>
<th>τe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhu u.a. (2011)</td>
<td>0,91</td>
<td>0,89</td>
<td>0,189</td>
<td>0,21</td>
</tr>
<tr>
<td>Kalyar u.a. (2019)</td>
<td>0,92</td>
<td>0,79</td>
<td>0,223</td>
<td>0,282</td>
</tr>
<tr>
<td>Najmi u.a. (2020)</td>
<td>0,852*</td>
<td>0,813*</td>
<td>0,245</td>
<td>0,294</td>
</tr>
<tr>
<td>Tan u.a. (2019)</td>
<td>0,895*</td>
<td>0,886*</td>
<td>0,262</td>
<td>0,294</td>
</tr>
<tr>
<td>Çankaya u.a. (2019)</td>
<td>0,792</td>
<td>0,91</td>
<td>0,274</td>
<td>0,323</td>
</tr>
<tr>
<td>Abdallah u.a. (2019)</td>
<td>0,809</td>
<td>0,878</td>
<td>0,287</td>
<td>0,34</td>
</tr>
<tr>
<td>Gosh (2018)</td>
<td>0,836*</td>
<td>0,856*</td>
<td>0,315</td>
<td>0,372</td>
</tr>
<tr>
<td>Lee u.a. (2012)</td>
<td>0,891</td>
<td>0,914</td>
<td>0,32</td>
<td>0,355</td>
</tr>
<tr>
<td>Namagambe u.a. (2019)</td>
<td>0,85</td>
<td>0,86</td>
<td>0,325</td>
<td>0,38</td>
</tr>
<tr>
<td>Yu u.a. (2014)</td>
<td>0,791</td>
<td>0,842</td>
<td>0,365</td>
<td>0,447</td>
</tr>
<tr>
<td>Jawaad, Zafar (2020)</td>
<td>0,916</td>
<td>0,816</td>
<td>0,368</td>
<td>0,423</td>
</tr>
<tr>
<td>Choi u.a. (2018)</td>
<td>0,775</td>
<td>0,679</td>
<td>0,373</td>
<td>0,514</td>
</tr>
<tr>
<td>Song u.a. (2015)</td>
<td>0,7</td>
<td>0,7</td>
<td>0,374</td>
<td>0,534</td>
</tr>
<tr>
<td>Ismail u.a. (2019)</td>
<td>0,768</td>
<td>0,604</td>
<td>0,398</td>
<td>0,587</td>
</tr>
<tr>
<td>Mishra u.a. (2019)</td>
<td>0,809</td>
<td>0,938</td>
<td>0,389</td>
<td>0,447</td>
</tr>
<tr>
<td>Green Jr. u.a. (2012)</td>
<td>0,853</td>
<td>0,938</td>
<td>0,436</td>
<td>0,461</td>
</tr>
<tr>
<td>Younis u.a. (2018)</td>
<td>0,7</td>
<td>0,7</td>
<td>0,549</td>
<td>0,784</td>
</tr>
<tr>
<td>Li u.a. (2018)</td>
<td>0,952</td>
<td>0,959</td>
<td>0,552</td>
<td>0,578</td>
</tr>
<tr>
<td>Chan u.a. (2012)</td>
<td>0,81*</td>
<td>0,81*</td>
<td>0,62</td>
<td>0,722</td>
</tr>
<tr>
<td>Green Jr. u.a. (2014)</td>
<td>0,97</td>
<td>0,915</td>
<td>0,62</td>
<td>0,658</td>
</tr>
</tbody>
</table>

For better comparability, both uncorrected and corrected effects were calculated. The uncorrected population correlation is approximately 0.374, while the corrected population correlation is higher, approximately 0.435. Using the data collected within the meta-analytical balance sheet, the variances of the study correlation and the variances of the sampling errors
could be calculated in order to calculate the variances of the population correlation from the difference. The respective standard deviations could then be calculated from the calculated variances.

Table 6: uncorrected and corrected effects

<table>
<thead>
<tr>
<th>Uncorrected Effects</th>
<th>Corrected Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{r} = 0.374005 )</td>
<td>( \bar{r}_c = 0.4354796 )</td>
</tr>
<tr>
<td>( \sigma_{r}^2 = 0.02047822 )</td>
<td>( \sigma_{r_c}^2 = 0.02682593 )</td>
</tr>
<tr>
<td>( \sigma_{e}^2 = 0.004899141 )</td>
<td>( \sigma_{e_c}^2 = 0.004348654 )</td>
</tr>
<tr>
<td>( \sigma_p^2 = 0.015579079 )</td>
<td>( \sigma_{p_c}^2 = 0.022477276 )</td>
</tr>
<tr>
<td>( \sigma_p = 0.1248162529 )</td>
<td>( \sigma_{p_c} = 0.149924342 )</td>
</tr>
</tbody>
</table>

The corrected standard deviations \( \sigma_p \) and \( \sigma_{p_c} \) can now be compared with the empirical mean values \( \bar{r} \) and \( \bar{r}_c \):

For \( \bar{r} \) applies: \( 0.374005 / 0.1248162529 = 2.996445 \).

For \( \bar{r}_c \) applies: \( 0.4354796 / 0.1499242342 = 2.904664 \).

This means that the empirical mean of the uncorrected effects, rounded up, is 3.00 standard deviations above 0. For the corrected effects, the value is 2.90. If all study correlations are normally distributed, the probability that another correlation is zero or below zero is also practically zero. It can therefore be concluded that the population correlation is positive in all studies, both for the uncorrected and the corrected effects. The individual effect sizes can also be multiplied by the number of corporate performance dimensions used. The effect sizes of studies that include all dimensions are thus weighted more heavily than studies that include only two or one dimensions. Thus, the following applies to \( r \):

\[
 r_{ig} = r_i \times \frac{\sum [CPD_i]}{3} \tag{11}
\]

\( r_{ig} \) stands for the dimension-weighted effect size and \( CPD_i \) for the number of “corporate performance” dimensions included in the respective study. There is no weighting between the three dimensions, so all dimensions are weighted equally. The results of the dimension-weighted population effects \( \bar{r}_g \) and \( \bar{r}_{cg} \) are as follows:

\( \bar{r}_g = 0.255057 \)

\( \bar{r}_{cg} = 0.2979706 \)

The two results of the dimension-weighted population effects are both positive and thus show a correlative relationship. The correlation between green procurement practices and corporate performance sought within the conceptual framework can thus be quantified. It can thus be concluded that green procurement practices have a positive effect on corporate performance. For comparison: The meta-analysis conducted by Golicic and Smith on the correlation between environmental supply chain practices and market-based (\( \bar{r} = 0.317 \)), operational-based (\( \bar{r} = 0.301 \)) and accounting-based (\( \bar{r} = 0.283 \)) performance shows comparably positive effects regarding the introduction of green measures.
5 Discussion and implications

From a total of 1686 studies, 22 studies were finally included in the meta-analytical review. From all selected studies, the relevant effect sizes of the "practice-performance link" under consideration could be taken. In addition, almost all studies provided the reliability coefficient "Cronbach's Alpha". All effect sizes used were positive and two positive overall effects can be presented as a meta-analytical result. Using the rule of thumb according to Kosfeld et al. for interpreting the Bravais-Pearson correlation coefficient, the dimension-weighted, corrected population effect describes a weak positive correlation, and the corrected population affects a medium positive correlation. Thus, concerning the research question, green purchasing and procurement measures have a positive influence on the performance of companies. By comparing the standard deviations with the empirical mean values, the positive influence was clarified by establishing that the probability of a correlation of zero or below zero must be practically zero. Looking at the funnel plot of the effect sizes taken from the studies, it can be assumed on the one hand that a publication bias could be present, as not a single effect size is negative. On the other hand, the funnel plot confirms the statement that the probability of negative effects must be practically zero. A potential publication bias can be determined using a "file drawer analysis" by calculating the characteristic value "Fail-Safe N". This indicates the number of studies with an effect size of zero that would have to be found before the overall probability of a 1st type error applies to a desired significance level.

It is noticeable when looking at the meta-analytical balance that most of the studies were regionally based in Asia. The large economic areas in North America and Central Europe were hardly or not at all considered in this analysis. Given the increasing climate awareness and the possible pioneering role of the European Union in climate protection, it can be stated that further quantitative studies of the considered "practice-performance link" from Europe are needed to be able to make an improved overall statement. Within the results, it could also be established that the variation of the individual effect sizes is not only due to sampling errors, but that moderator variables must be present. The exact moderator variables can be determined in further research, otherwise, they can only be conjectured. Since it can be assumed that the region in which the studies were conducted is a possible moderator variable, additional studies from Europe could have an increased positive or negative influence on the result.

However, considering the methodology, it must be mentioned that the present result is a preliminary one. As mentioned above, the bare-bones meta-analysis does not correct all possible artefacts. Accordingly, a research outlook can be mentioned that considers possible further research on the "practice-performance link" between green purchasing and sourcing practices and business performance. Thus, for future research, before meta-analytical considerations, it can be stated that as many regions as possible should be covered by the selected studies. It seems obvious that research should initially focus on individual empirical studies in regions that have not yet been covered. Here, the regions in Central Europe and North America would be of particular importance. Furthermore, some aspects within the meta-analysis can be mentioned that could be taken up by further research. Beyond the bare-bones meta-analysis, the already mentioned "file drawer analysis" can be mentioned first to check for publication bias. In addition, if the relevant information is given in the studies, further artefacts can be corrected according to the methods of Hunter and Schmidt. Particularly important for an
extended analysis could be the identification of moderator variables that explain the variation between the individual studies. For this purpose, "credibility intervals" according to Golicic and Smith (2013) can be used. Based on the meta-analytical results and the finding of the positive correlation of the "practice-performance link" under consideration, further research needs to examine which specific green procurement and green purchasing measures improve company performance and how strong the individual influences of the individual measures and instruments are. In summary, further and comprehensive empirical studies and analyses are needed to further consolidate the overall data situation. Consequently, a consolidation of the data situation can improve the significance of further meta-analyses on the "practice-performance link" under consideration to show that green measures in the area of procurement and purchasing are in line with company performance and even improve it.

6 Conclusion

To answer the research question, a bare-bones meta-analysis was conducted. Despite the small number of perfect replication studies within business economics, studies were identified that fulfilled the formal and content-related criteria. As a result, a positive correlation between the introduction of green procurement measures and company performance was identified. Wernerfelt's RBV and Hart's NRVB, which builds on it, state that future competitive advantages are based on environmentally sustainable corporate activities. Existing studies and meta-analyses already show positive results for green supply chain practices. With the results of this study, it is now possible to explicitly identify an increased company performance for the procurement organisation, should it introduce green procurement measures.

Considering the triple bottom line, companies need to look at the economic, environmental and social dimensions to increase sustainability and take them into account in their actions. The economic dimension refers to the goal of qualitative growth of the company (Koplin, 2006). The results show that the introduction of green procurement measures improves both the ecological and the economic dimension and that companies can thus position themselves more sustainably overall. It illustrates the importance of the procurement organisation as a whole and the significance of ecological and green considerations and measures in the company. By having a positive effect on the increase of company performance, the result of this work sets further incentives for companies to make measures greener and more sustainable.
Appendix – Detailing the concepts of the practice-performance link

<table>
<thead>
<tr>
<th>Deductive Codes</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eco Design</td>
<td>The actions taken during product development stage are targeted towards minimizing a product's environmental impact during its whole life cycle starting from acquiring raw material to manufacturing, use, and finally to its final disposal without compromising other essential product criteria such as performance and cost.</td>
</tr>
<tr>
<td>Green Purchasing/Procurement</td>
<td>An environmental purchasing initiative that aims to ensure purchased products and materials meet with environmental objectives set by the purchasing firm such as reducing sources of waste, encouraging recycling, reuse, and substitution of materials.</td>
</tr>
<tr>
<td>Environmental Cooperation</td>
<td>The activities that take place between the supply chain members for eco-design, cleaner production, green packaging, use of less energy during transportation of materials and goods, and working together on mutual environmental responsibilities and objectives.</td>
</tr>
<tr>
<td>Reverse Logistics</td>
<td>The return or take back of products and materials from the point of consumption to the forward supply chain for recycling, reuse, remanufacturing, repairing, refurbishing, or safe disposal of products and materials.</td>
</tr>
<tr>
<td>Operational Performance</td>
<td>The organization's capabilities to more efficiently produce and deliver products to customers with improved quality and reduced lead times ultimately lead to improving its position in the marketplace and increasing its chances of selling its products into international markets.</td>
</tr>
<tr>
<td>Market Performance</td>
<td>Market-based performance centers on financial indicators reflecting market goals concerning meeting customer needs and includes market share, competitive advantage, customer loyalty, brand equity, etc.</td>
</tr>
<tr>
<td>Financial Performance</td>
<td>Financial Performance refers to overall profitability as indicated by return ratios, earnings, and profit.</td>
</tr>
</tbody>
</table>
References


