Systematic review of 16 years of scientific literature on public procurement

Sandra Lange, Jan Telgen, Fredo Schotanus

As public procurement research is fragmented in many sub-topics, a clear synthesis is lacking, which in turn inhibits the establishment of a clear body of knowledge. To start filling this gap, this systematic review provides an overview of the most influential literature in the field of public procurement. The findings are aimed at providing other researchers with relevant information to synthesize existing findings.

We found that public procurement research is maturing, with increasing attention from diverse scientific disciplines. USA and UK are most productive publishers, while European countries have become increasingly active. Although a wide spectrum of research designs have been utilized, the reviewed articles focused on few. For example, although articles addressed twenty different topics, eleven were only studied once or twice, while 61.4% of papers researched the topic of procurement strategies. Considerable variations were observed across countries, indicating different research foci, as well as varying levels of maturity per research characteristic. Research seems to have underused existing scientific knowledge in that literature and meta-studies were only utilized in 13.2% and 5% of papers respectively. Practical applicability of research findings is inhibited by a detected imprecision of research, such as not specifying the procuring government level in 56.1% of reviewed papers. The overall conclusion with respect to the maturity level of public procurement research is that while various different paths have been laid, most researchers continued to walk the main roads. To develop the field further, we recommend to research the field from more diverse angles.

Introduction

Public procurement is a powerful tool to make governments more efficient. It accounts for 13% to 20% of worldwide GDP (OECD, 2013), meaning that a significant proportion of all produced products and services are bought by governments. Good public procurement policies and practices lower public expenditures and free them up to be allocated to other areas (Choi, 2010).

Many researchers emphasized the powerful effects that public procurement can have on fostering innovation (a.o. Aschhoff and Sofka, 2009; Edler and Georgiou, 2007) and green production (Day, 2005; Günther and Scheibe, 2006). To achieve such socially desirable outcomes, governments operate as both regulators, passing laws and regulations, and market participants (Choi, 2010; McCrudden, 2004). And in times of global supply chains public procurement is not only a means to improve upon social outcomes on domestic markets, but also internationally (McCrudden, 2004).

1 This paper is a reworked version of Sandra Lange's master thesis work
2 Corresponding author j.telgen@utwente.nl
3 All authors are associated with School of Management and Governance, University of Twente, PO box 217, 7500 AE, Enschede, The Netherlands
While public procurement is a highly fragmented field, a systematic overview of research is still lacking. Crossan and Apaydin (2010) state that “fragmentation of the field prevents us from seeing the relations between these facets and ultimately impedes consolidation of the field.” (p. 1154). By showing which topics have been addressed by research, to what extent, detecting possibly understudied, as well as mature sub-fields, and by enabling researchers with the tools to conduct syntheses on findings for sub-fields, the field is developed to a new state of more clarity and unification. The practical impact of this work is therefore indirect, by stimulating and enabling a research agenda to derive at generalizable findings.

This literature review provides an overview of the most influential scientific literature published on the topic of public procurement (PP). Moreover, the overall state of PP research is assessed, providing insights into the maturity of the field. According to Cheon, Grover and Sabherwal (1993) mature research fields are characterized by studying a variety of different topics and applying various research methods instead of narrowly focusing on few. Therefore, this review focuses on addressed topics, as well as employed methodologies and their development over time. The literature review is focused on the past sixteen years (1997 to 2012).

**Methodology**

The methodology was informed by Wynstra (2010). For reviewing the *Journal of Purchasing and Supply Management*’s publications of the years 1994 to 2009 he developed an extensive list of categories each article was classified into. The main categories employed by Wynstra are: topic, research strategy, data collection, data analysis, type of product, type of purchase, as well industry and sector. With respect to the topics, the author classified each article into a maximum of three subject categories, while another, similar review conducted by Carter and Ellram (2003) on the *Journal of Supply Chain Management* categorized each article into only one subject category which summarized the article’s focus the best. We believe that Wynstra’s approach yields a more accurate representation of research topics, since procurement subjects are often researched against a clear background and therefore categorization into only one subject field will under-represent the other(s).

Each of the main categories was further divided by Wynstra into sub-categories. He also gathered general article data including publication year, contributing authors, institutions and citations.

All main categories employed in his research were adopted with the exception of “type of purchase”, as a scoping study of the PP field had revealed that this category was irrelevant. While the main categories were mostly adopted, the sub-categories were modified. His topic sub-categories were to a great extent not applicable to this review, since his research mainly focused on private sector procurements. The research strategies were all adopted except for “expert interviews/Focus group” and “laboratory experiment”, which were instead grouped as data collection methods. Wynstra’s subcategories for data collection are very specific, such as distinguishing between four kinds of questionnaires. As this level of specificity may falsely create the impression of variety, the subcategories were simplified. The same applies to his thirty-nine items list of data analysis techniques, which was simplified to only distinguish between quantitative and qualitative research. The product types were amended to account for works while the product type combination “good/service” was omitted. Nine industry and sector categories were adopted from the Wynstra classification scheme yet further extended during the data extraction stage of the review. Instead of collecting institute data, the country of the institute was noted per article. While institutional data would have provided interesting insights into institutional collaborations, as well as most active institutions on the field, this review adopted a global perspective limited to cross-country as opposed
to cross-institutional differences. Finally, his categorizations were extended with data on studied country/ies, publishing journal and the procuring government level. This latter study characteristic is specific to public procurement and inclusion in the review provided valuable information on the context of conducted research. Specifying the research context is according to Denyer, Tranfield and van Aken (2008) supportive to practical relevance. The complete classification scheme may be obtained from the authors upon request. The classification scheme was directly transcribed into an SPSS data extraction form, which was later used to conduct the analyses.

The review's objective was to assess the overall status development of public procurement research, its predominant study characteristics and research designs, as well as addressed topics. The status of public procurement research was operationalized as annual publications. Publication counts are an effective outcome measure to assess the scientific importance of a research field (Crossan and Apaydin, 2010) and their development is a reliable indication whether the relevance of the field changed. To further characterize the time developments, publishing countries and journals were assessed over time as well as authorships. It is believed that these variables provide a meaningful overview of the main stakeholders of the field.

Predominant study characteristics were assessed against the industries and sectors from which the government procures, the procuring government level, studied country/ies and types of products procured. There was no limitation for those categories with regard to maximum sub-categories per article.

Most frequently employed research strategies were assessed against the research methodology, time dimension, research strategy and data collection methods. While there was no limit as to how many data collection methods each article was grouped into, the other categories held mutually exclusive sub-categories. This exclusiveness was only breached when articles clearly articulated to have utilized more than one strategy.

To detect developments over time, the sixteen years of research have been subdivided into equal time intervals of four years each, inspired by Wynstra (2010) and Carter and Ellram (2003).

Searches were conducted by use of the databases Scopus and Web of Science. Both databases are well-established, multi-disciplinary research platforms, holding a wide variety of peer-reviewed journals, and they are being kept up to date. We chose for two databases to ensure all relevant papers are included, since it is possible that one database omits relevant research (Crossan and Apaydin, 2010).

To assess whether high impact papers differ from low and medium impact publications with regard to study characteristics and research designs we conducted citation analyses based on mean scores of the Scopus and Web of Science citation counts. We included both databases' citation counts as citations differ per database and therefore reliance on only one source may over- or undervalue individual papers. A mean citation count is believed to provide a more realistic assessment of each paper's scientific impact. Also, not every paper is enlisted in both databases, therefore considering only one of them could mean that some papers could not be assigned an impact assessor although they may be of value to research. A shortcoming of citation analysis is that recent papers have had less time to accumulate citations. However, one paper from 2010 is still the twenty-first most often cited paper overall, and the third most cited article in its research sub-field. Therefore, it is believed that, while acknowledging the discrimination against recent papers, research with high relevance to science will have gathered notable citations in the past least one year (status March 2014).

Searches were limited to English articles, published in peer-reviewed journals. While some authors have critiqued peer-review to be an obscure process (Altman, 1996), potentially biased by knowing
the identity of the author (Relman, 1990), the limitations of this research did not allow for a quality appraisal of each included article. Therefore, it is believed that peer-review was the best available measure to have some quality appraisal in place. It is assumed that high impact research on the subject of public procurement will have been translated into English and that therefore no high impact papers have been disregarded from the review based on the language restriction.

The time frame chosen for this systematic review are the years between January 1, 1997 and December 31, 2012. Around the millennium a number of initiatives have been launched by the National Institute of Governmental Purchasing, Inc. (NIGP) to foster academia to pay more attention to the largely neglected field of public procurement (Thai, 2001; Carter and Grimm, 2001). Those included, beyond others, under a partnership agreement with the Florida Atlantic University (FAU) the establishment of the Public Procurement Research Center, as well as the launch of the first scholarly journal on the field, the Journal of Public Procurement. We decided on setting the cut-off year for this review a few years before the launch of those initiatives to, among other things, be able to assess their impact on the field. The final year of consideration, 2012, was the most current research year when this systematic review was initiated in 2013 and was thus chosen to represent the most recent developments.

Search terms were developed by testing individual term's effectiveness against batches of twenty search results. After the test batches had revealed many terms to be ineffective in that they did not add any new or relevant results, the following are the final search terms utilized in combination.

- Variations of public: public; government
- Variations of procurement: procurement; purchasing; contracting; buying; commissioning

For the database Scopus, searches for the search terms was restricted to title, abstract and keywords of the article. The proximity operator of W/5 was included between two consecutive search terms to include results where the two search terms appear within five words. Scopus advises researchers to use a proximity operator of either 3, 4 or 5, if they wished to find the search terms within phrases (Scopus, n.d). To lower the threat of falsely omitting relevant literature, we utilized the widest of the advised proximity operators. For the Web of Science database searches were restricted to the topic subject and title. In line with the Scopus searches, the proximity operator NEAR/5 was used.

For the Scopus database searches were restricted to the subject area of Social Sciences & Humanities. For the Web of Science database searches were restricted to the subject areas Science Citation Index Expanded and Social Sciences Citation Index.

Search results were assessed for relevance in a three-step process based on Bettany-Saltikov (2010) by comparing title, abstract and the full text against stipulated criteria for in- and exclusion. The main rationale was that we only wanted to include articles that were strictly on the topic of public procurement and which provided exemplars of current practices, best or worst, as well as guidelines for practice and research. The criteria may be requested from the authors.

Relevance assessments were partly conducted in a team of four. The other three researchers could not finish all batches due to personal agendas that no longer allowed for the time intensive effort. As a result, the title assessment was fully conducted in a team, while the abstract assessment was only partly conducted in a team, and the final full text assessment was conducted by the main author alone.
The remaining 378 articles were then coded against eleven main categories and subsequently analyzed by means of descriptive statistics using the software SPSS.

Findings and discussion

This review analyzed the status of public procurement research and how it developed globally between the years 1997 and 2012. We found that the relevance of the field, assessed by annual publications, increased significantly over the years, a raise which began in 2003 (Figure 1). High individual authors, who only published one or two articles, as well as the fact that 199 different journals were the publishers of the 378 articles included in this review, highlight that public procurement is no isolated research field, but instead highly cross-disciplinary.

The majority of journals, 68.5%, published only one article over the sixteen years, and a further 14.5% published two. Even the top ten publishing journals each account for a maximum of 3.7% of all papers included in this review (Table 2) and they often published merely one or two relevant articles within four year intervals. Due to the high requirements of the chosen databases, the Journal of Public Procurement is not listed in the Web of Science database, while Scopus only lists publications from 2012 onwards. Yet, in this final year the journal published eight relevant articles, making it the overall fourth most productive publisher and the most knowledgeable on the field of public procurement research.

Forty-eight countries published relevant articles, while eighty-five were studied. While this poses to variation, our analyses revealed that most influential countries, both publishing as - due to a strong home bias of researchers - studied, are mainly the USA and UK, while all other countries have very low relative shares. Yet, the importance of the top publishers decreased greatly as more countries entered the field over the course of time. Especially European countries have rapidly increased their research activities in the later years. If this trend is continued, they could become more knowledgeable on the field and PP research could mature to a state where phenomena are studied against more versatile backgrounds.

PP research is highly practice oriented, which manifests itself both through employed research strategies, which were mostly case studies and survey researches (Figure 2), as through utilized data sources, of which reviews of non-academic literature and survey methods were most prominent (Figure 3). Shortcomings of these trends refer, paradoxically, to practical relevance as all those measures have limited reliability and their findings are difficult to generalize. The mostly neglected measures to pool findings (meta-analyses) or apply them to further analyses (literature studies) disable the field from deriving at definite findings, which can be applied by practitioners.

Figure 1 Publication trend
What further inhibits practical application is that research was very unspecific with respect to context variables: 56.1% of papers did not specify a government level, 28.6% no product, and 60.6% of articles were grouped into the pooled categories of unspecified industries and sectors from which the public procures. Moreover, 20.6% of papers did not specify their data sources, which poses to dubious scientificity as findings cannot be verified by others.

The field addressed a wide range of twenty different topics, including nine that were only studied once to twice and which were grouped into a pooled category entitled “Other topic”. However, topics have been addressed by uneven proportions of papers: while the top studied topic of procurement strategies was studied in 61.4% of papers, the second most prominent topic, selection, was merely studied by 17.2% of the articles. Even more so, other topics’ relevance decreased over time, while procurement strategies’ continued to grow (Figure 4).

Of the twenty different procurement strategies addressed, research mainly focused on contracting-out and PPP, which combined were studied in 64.2% of strategy papers.

This apparent, yet fallacious, versatility was also observed for studied industrial and sectoral contexts: while thirty-two different industries and sectors were studied, research mainly focused on the construction industry, studied in 20.6% of papers, while the second most often studied sector, the health sector, was only studied by 6.9% (Figure 5).

The review concentrated on discovering prominent research designs and study characteristics. Overall, services were the most often studied product type, although over time their proportionate relevance decreased as particularly works became increasingly popular.
The most frequently studied government level is the local one. Over the sixteen years, the municipal level was researched increasingly more often, while the federal level slightly decreased in scientific relevance. Whereas during the years 1997 and 2008 the majority of articles conducted quantitative research, qualitative methods gained more attention in the final four years, resulting in both methodologies being applied almost equally often. This achieved status is a sign of maturity in that no method is under- nor overused. With respect to time dimension it was shown that PP research was focused on cross-sectional research, applied by 79.6% of reviewed articles, which inhibits generalizability of findings (Babbie, 2006).

While on the global level overall favorites per research variable could be detected, cross-country, as well as cross-topic comparisons showed great variations.

Research on the strategy of e-procurement showed noticeable limitations in that it was only assessed by means of survey research and case studies, while fifteen of the sixteen papers on the topic did not specify the private industry or sector from which the public procures. Utilizing the full spectrum of research strategies increases generalizability of findings. Furthermore, contextualization is paramount for evidence-based management, as it also enables other researchers to assess reasons for potentially contradicting findings.

High impact papers did not show differences to low impact papers with respect to contextual imprecision. A noticeable distinctiveness observed is that while low and medium impact papers focused on the more practice oriented research strategies of case studies and survey
research, while mostly neglecting literature studies and meta-studies, the high impact papers exclusively relied on the latter strategies as well as on survey research. This shows that high impact research utilized existing, scientific knowledge, which is important in developing the PP field to maturity.

The overall conclusion with respect to the maturity level of PP research is that while various different paths have been laid, researchers continued to walk the main roads. To develop the field further, researchers are strongly advised to research the field from more diverse angles. The data provided in the full review can provide them with the needed information as to which designs have been underused until today. Moreover, scientists should synthesize past findings for the sake of deriving at definite conclusions. Only when research findings are tested against various backgrounds and when past findings are validated, can definite findings be developed, which can eventually be consulted by practitioners.

**Review limitations**

A limitation of this review is that the categorizations of all included articles were done by the main author alone. To make the findings objective, systematic literature reviews should be conducted
within a team of researchers (Tranfield, Denyer and Smart, 2003; Rousseau, Manning and Denyer, 2008). As this limitation was known at the beginning of the research project, this paper aimed at making the review process highly transparent to enable other researchers to replicate the work and test the findings.

A further limitation regards the exclusion of articles published in other languages than English, which may have yielded a language bias (Egger et al., 1997).

Restricting the search for relevant papers to only two databases may have omitted relevant papers, since even the renowned databases Scopus and Web of Science do not hold all relevant articles. This apprehension was confirmed by the fact that ten years of publications from the *Journal of Public Procurement* were missed because the journal was rejected by Thomson Reuters, and only accepted by Elsevier in 2012.

It is generally considered important to include grey literature in a systematic literature review to develop a more complete overview (Tranfield et al., 2003; Rousseau et al., 2008; Hopewell, McDonald, Clarke and Egger, 2007). Grey literature refers to "multiple document types produced on all levels of government, academics, business, and organization in electronic and print formats not controlled by commercial publishing i.e. where publishing is not the primary activity of the producing body." (GreyNet International, n.d.). Due to its nature, grey literature is difficult to locate and can be abundant, which would have exceeded the time range of this review. Moreover, this type of literature does not satisfy the research aim of presenting an overview of only the most influential scientific literature. According to Davies (2000) publication bias may adversely affect the validity of findings as journals tend to favour publishing positive results.

**Directions for future research**

To address this review’s limitations, other researchers are encouraged to replicate the work to test the findings. Full replications should take account of the limitations of this review.

While the review provides a detailed overview of researched topics and the designs applied to study them, researchers should take these findings into account when designing upcoming studies. To increase versatility and increase generalizability of findings, subjects should be assessed against different backgrounds and by different means. As shown in the topic analyses, increase in versatility is especially needed for research on e-procurement, which was approached highly limited in the past.

While this systematic literature review concentrated upon cross-country differences, it would be interesting to detect most active institutions on the field, as well as inter-institutional collaborations.

The high amount of 199 individual journals publishing relevant research may be analyzed with regard to most prominent backgrounds to detect which scientific fields are the main stakeholders of PP research.

This review does not include an analysis of most relevant topics per publishing country. Since public procurement practices are embedded in local system contexts, such analyses may provide interesting insights as to how PP research foci differ globally.

The citation analysis conducted in this review was limited to citations of Scopus and Web of Science. As PP is very practice oriented, it is likely that paper impacts differ when non-scientific citations are included. Such analysis could provide valuable insights into the potential differences in topics' relevance to practitioners. Potentially observed differences could then inform future research topics, which would both make the field more diverse, as it would also increase practical relevance.
Finally, replication of this review in a few years’ time is encouraged to assess how the field developed after 2012. Since recent trends have been observed, such as the study of more diverse industries and sectors in the later time intervals and an increasing relevance of European publishers, the research field may appear quite different in the years after 2012.
References


