Is contracting out of public services still the great panacea?
A systematic review of studies on economic and quality effects from 2000-2014

Petersen, Ole Helby; Hjelmar, Ulf; Vrangbæk, Karsten

Published in:
Social Policy and Administration

DOI:
10.1111/spol.12297

Publication date:
2018

Document version
Også kaldet Forlagets PDF

Citation for published version (APA):
Is Contracting out of Public Services still the Great Panacea?
A Systematic Review of Studies on Economic and Quality Effects from 2000 to 2014

Ole Helby Petersen\textsuperscript{a}, Ulf Hjelmar\textsuperscript{b} and Karsten Vrangbæk\textsuperscript{c}

\textsuperscript{a}Department of Social Sciences and Business, Roskilde University, Roskilde, Denmark
\textsuperscript{b}Danish Institute for Local and Regional Government Research, Copenhagen, Denmark
\textsuperscript{c}Department of Political Science, Copenhagen University, Copenhagen, Denmark

Abstract

This article presents the results of a systematic review of international studies on economic and quality effects of contracting out published in the period from 2000 to 2014. We conducted a comprehensive search of the literature and identified 49 relevant studies. There are three main findings of the systematic review: (1) cost savings documented in international contracting out literature have been decreasing over time; (2) cost savings have been much greater in technical services than in social services; and (3) economic effects have been twice as large in Anglo-Saxon countries compared with other countries. With regard to measuring the effect of contracting out on service quality, which is a vital component of any service delivery arrangement, very few studies assess this issue in a comprehensive manner. There is also a significant lack of studies that include measures of transaction costs, thereby making it difficult to evaluate the impact of contracting out on overall cost-effectiveness of public service delivery. We conclude that generalization of effects from contracting out should be made with caution and are likely to depend, among other things, on the transaction costs characteristics of the service, the market situation and the institutional/regulatory setting.

Keywords
Contracting out; Systematic review; Evidence; Costs; Service quality

Introduction

Private delivery of government-funded services by means of contracting out constitutes an increasingly utilized alternative to public provision of technical

Author Emails: olehp@ruc.dk; ulhj@kora.dk; kv@ifs.ku.dk

© 2017 John Wiley & Sons Ltd
and social services in developed countries (Ohlsson 2003; Stolt et al. 2011; Rostgaard and Szebehely 2012). A number of studies have highlighted private production as more cost-effective than public production (Blom-Hansen 2003; Dijkgraaf and Gradus 2007; Gilmer 2010), though recent analyses have contested the magnitude of previously documented cost savings from contracting out (cf. Bel et al. 2010; Hartman 2011; Petersen and Hjelmar 2013). The widespread belief in the private market as superior to public service provision was reinforced by a number of much-cited reviews and meta-analyses of empirical studies published throughout the 1960s–90s. The findings emerging from these meta-reviews were that the cost saving potential of contracting out was often in the range of 15–20 per cent (Borcherding et al. 1982), and in the most optimistic accounts even up to as much as 30 per cent (Savas 1987; Domberger and Jensen 1997).

However, other and more recent assessments of the international empirical evidence have found smaller and more mixed effects of contracting out (Bel et al. 2010; Hartman 2011). In a comprehensive meta-analysis of international studies, Hodge (2000) found average cost savings in the range of 6–12 per cent after correcting for transaction costs. Moreover, cost savings were both larger and better documented in technical services (e.g. garbage collection and ground maintenance) than in social services, such as health and policing. Based on these results, Hodge noted that ‘expecting contracting to be a panacea for all public services, though, would be foolish’ (Hodge 2000: 246–7). In another seminal study, Boyne (1998) examined almost 40 empirical analyses from the UK and the USA, and found cost savings documented in the majority of studies. At the same time, however, Boyne (1998) criticized the existing literature for using inadequate methodology, lacking sufficient data and for paying too little attention to service quality, thus making it difficult to evaluate the impact of contracting out on overall cost-effectiveness of service delivery. Whereas contracting out has remained high on government agendas and the volume and scope of privately delivered services has been gradually extended, there has not been a corresponding development in systematic assessments of international experience since the turn of the millennium (see, however, Bel et al. 2010; Vrangbæk et al. 2015).

This article presents the results of a systematic review of international studies on economic and quality effects of contracting out published in the period 2000–14. The review is based on 49 studies, and systematically assesses the collected empirical evidence on economic and quality effects in international peer-reviewed journals since the last systematic review that was published at the turn of the millennium (Hodge 2000). The examination in this article is based on the following research question: ‘Does the evidence presented in internationally published articles in the 15-year period 2000–14 show that contracting out has resulted in cost savings?’ In addition to this general research question, we examine documented effects for service quality, differences between technical and social services, differences in effects across different national institutional contexts (in particular Anglo-Saxon countries versus other countries), and longitudinal developments in documented effects over the examined time period.
Carrying out an updated review of the international literature on contracting out is timely and warranted for several reasons.

First, whereas the assessed literature in previous reviews was published in the 1960s, 1970s, 1980s and early 1990s, a number of recent studies have pointed to more mixed effects of contracting out (Bel et al. 2010; Petersen and Hjelmar 2013), and even to a tendency towards declining effects in some service areas (Bel and Costas 2006; Hutchinson and Pratt 2007; Houlberg and Petersen 2015). A reason could be that contractual relationships have moved from first generation to second, third or fourth generation, and public organizations have now been exposed to competition for several decades (Bekken et al. 2006; Bae 2010). An updated review of recent experiences could thus provide other and more nuanced results than assessments of the early experience.

Second, while previous reviews mainly covered experiences from the USA and the UK, there is today a much broader literature covering more countries and geographical areas with different institutional and regulatory settings (Brennan et al. 2012). We can thus examine whether the early lessons from Anglo-Saxon countries also hold true for other countries with different regulatory, ideological and institutional contexts.

Third, whereas the majority of previous research has examined experiences with a focus on technical service areas, the scope of contracting out has gradually expanded and nowadays also includes welfare services, which are more labour intensive (Hartman 2011; Szebehely and Trydegård 2012) and tend to have higher transaction costs (Brown and Potoski 2003; Hefetz and Warner 2012). Welfare services were only marginally represented in previous reviews, and we can therefore extend previous insights by systematically scrutinizing the evidence on contracting out across technical services and welfare services.

To summarize, the aim of the article is to establish an updated assessment of documented effects of contracting out that covers developments since the turn of the millennium and covers a broader range of countries and service areas than previous systematic reviews of the literature. The article is organized as follows. The next section discusses different approaches to assessing and evaluating effects of contracting out based on previous studies and recent theoretical discussions. We then outline how the systematic collection and assessment of studies was carried out. This is followed by a presentation of the results of the systematic review. We then discuss the lessons learned and provide some reflections pertaining to comparison of public and private service delivery. Lastly, we conclude on the findings and consider future avenues of research.

**Theoretical Background: Effects of Contracting Out**

Theoretical arguments regarding effects of contracting out in the public sector have been contested for decades (Vining and Boardman 1992; Bel and Warner 2008). Theoretical claims of positive effects typically stem from *public choice* and *property rights* theory, which revolves around a competition and an ownership argument, respectively (Domberger and Jensen 1997; Blom-Hansen 2003; Alonso et al. 2015).
The *competition* argument states that private sector producers are forced by competitive pressure to optimize efficiency, while public organizations operate in a monopoly situation with the same risk of inefficiencies as private monopolies. The lack of competitive pressure means that public managers are unable to measure the efficiency of their organizations against other similar organizations. Decisions on resource allocation and survival of the organization are therefore left to public decision makers who are unable to rely on information about demand and efficient supply that markets can provide.

The *ownership* argument states that public sector agencies lack incentives to perform efficiently; they often have broad and ill-defined public service objectives, and they have no bankruptcy constraint. Therefore, public organizations can continue to perform at sub-optimal levels without the risk of going out of business, which is a threat that private companies are faced with (Alonso *et al.* 2015). Furthermore, public organizations are not accountable to shareholders and owners, which in turn is a characteristic of private companies that makes them focus on innovation and technological development in order to stay competitive (Benmarker *et al.* 2013).

Several theoretical attempts to nuance and broaden the expectations from public choice and property rights theory have been provided in the literature (Bel and Warner 2008; Vrangbæk *et al.* 2015). The first stems from transaction costs economics, which underlines the importance of asset specificity and measurability of the service that is to be contracted out (Williamson 1979; Brown and Potoski 2003). Rather than approaching public services as something that would by definition be more effectively produced in a private market, transaction cost economics hypothesizes that different service characteristics create more or less favourable conditions for in-house production and contracting out (Hefetz and Warner 2012). Economic benefits from private contracting are thus more likely to be realized if the quantity and quality of the service can be unambiguously described and measured; otherwise, the administrative costs of preparing tenders, evaluating bids, signing the contract and monitoring (and possibly sanctioning) service delivery are likely to be high. The largest economic effects are thus expected in technical services characterized by low asset specificity and high measurability, whereas smaller or even negative economic effects could be expected in complex services with high asset specificity and low (or highly expensive) measurability. By this token, transaction cost economics adds a more task specific theoretical perspective on contracting out which nuances the somewhat generalized view stemming from the competition and ownership argument.

Another major theoretical approach that extends public choice and property rights theory is *industrial organization literature*, which stresses a number of factors that make public markets distinct from traditional private markets and thus create less optimal conditions for contracting out than expected by public choice theory (Bel and Warner 2008). According to the industrial organizational perspective, many public services are characterized by natural monopolies and high entrance costs, which limit competition and make public markets function in a less perfect way than private markets (Gilmour and Jensen 1998). Quasi-market theory also stresses that public services are often subject to significant information deficiencies and asymmetries (Bartlett and
Le Grand 1993). This is particularly a problem in welfare services, such as nursing homes or childcare, where those buying the service (i.e. the public sector) have limited oversight of those delivering the services (i.e. the private company). The presence of information asymmetries can lead to goal avoidance and unwanted practices, such as cream-skimming (selection of the easiest tasks) and parking of the least profitable citizens/customers. Principal-agent problems can thus make quality measurement difficult and possibly erode cost savings or quality improvements that were foreseen by theories based on assumptions of more or less perfect market conditions.

Lastly, an argument about decreasing marginal effects from contracting out, which suggests that economic effects tend to decrease over time, has been proposed in the literature (Bekken et al. 2006; Hutchinson and Pratt 2007; Houlberg and Petersen 2015). There are two theoretical claims behind this argument. First, it is likely that rational public organizations will begin the contracting out of the services and tasks for which the largest gains are expected. Hence, once they have ‘harvested the low hanging fruits’ public organizations could face decreasing marginal benefits from additional contracting out (Bel and Costas 2006; Houlberg and Petersen 2015). Second, involvement of private providers creates a competitive pressure on public in-house production units, which may lead to more effective public production (Bae 2010). In contrast to property rights theory, the assertion is that public organizations can, in fact, cease to exist if their services are contracted out. The market mechanism and exposure to competition, according to this argument, not only increases the efficiency of the external (contracted out) services, but also of the internally produced services (Bel and Costas 2006). Moreover, quasi-markets characterized by high entrance costs and/or natural monopolies may be dominated by monopolies or oligopolies, which impede competition and create less ideal conditions for contracting out than envisioned by public choice and property rights theory.

This brief introduction to key theoretical arguments in the literature leads us to formulate several themes for the systematic review. First, we analyze whether studies from the period 2000–14 support the findings from earlier research regarding cost savings from contracting out. Second, we examine whether the studies include assessments of service quality, as this is a key parameter in evaluating the effects of contracting out on overall cost-effectiveness. Third, we investigate whether documented effects are different in technical services with low transaction costs compared to social services characterized by high transaction costs. Fourth, we examine the issue of regulatory and institutional context by comparing the evidence from studies in Anglo-Saxon countries with evidence from more coordinated (and regulated) market economies (Hall and Soskice 2001), which according to the ‘varieties of unionism’ literature also display different traditions for labour union power and involvement (Gumbrell-McCormick and Hyman 2013). Union power can be assessed according to membership density and bargaining coverage rates. Anglo-saxon countries, such as England and Ireland, have much lower bargaining coverage than social democratic/egalitarian countries (Denmark and Sweden), social partnership countries (Germany, the Netherlands,
Austria) and countries with strong socialist/communist influence on labour unions (France, Italy) (Gumbrell-McCormick and Hyman 2013).

Lastly, inspired by Bel and Costas (2006) and Dijkgraaf and Gradus (2013), among others, we adopt a longitudinal perspective and examine how the documented effects of contracting out have developed over the 15 years since the turn of the millennium.

**Systematic Review Method**

The aim of any systematic review is to collect, assess and synthesize the available knowledge produced within a specific field of literature. Here, the literature field here is confined to studies of contracting out, which investigate effects relating to costs and service quality and were published in the period 2000–14.

Prior to carrying out the systematic review, a number of definitions and decisions need to be clarified. Contracting out is here defined as the private provision of publicly funded services, whereby the public sector takes on the role of buyer, but retains the overall responsibility for financing the service (Petersen and Hjelmar 2013). In line with Hodge (2000) and Jensen and Stonecash (2005), we distinguish between privatization as sale of public assets (not our focus) and contracting out defined as a temporary transfer of production responsibility to a private company (for-profit or non-profit) in return for economic compensation. The latter process commonly takes place by means of competitive tendering, which enables the public sector to choose among various private bidders. Experiences with other types of marketization, such as public-private partnerships, free-choice markets, voucher-based systems and privatization, were not our focus here, and such studies were omitted from the assessment (see more details on inclusion and exclusion criteria below).

The systematic review is based on a methodological approach inspired by the British EPPI-Centre’s methodology (Gough 2004). This approach is somewhat broader in scope than review methodologies used by Cochrane and Campbell, which emphasize randomized controlled trials (RCT) as the golden standard for evaluation (Konnerup and Kongsted 2012). In contrast to medicine and psychology, which make extensive use of RCT in empirical studies, experimental research designs are very rare in the field of contracting out. Using RCT as the norm would thus exclude most analyses with relevance to our research question. The EPPI-Centre’s approach, in contrast, emphasizes a broader range of methodological designs, and includes both quantitative and qualitative studies (Gough 2004). Using this approach enabled us to collect and assess the collected empirical literature on the effects of contracting regardless of the type of data and qualitative/quantitative method applied. The search for studies for the systematic review consisted of three stepwise phases, as illustrated in figure 1.

In the first phase, we conducted a systematic search of literature in a broad range of relevant databases, which were not confined to a specific research field. Nor was it confined to studies in Social Sciences Citation Index (SSCI) journals, but included all articles in peer-reviewed journals, thus leaving out
books and book chapters as well as papers published in non-English journals. The following databases were used in the screening of relevant studies: Web of Science (SSCI), Econlit, Sociological Abstracts, Social Services Abstracts, ERIC (Education Resources Information Center), PILOT Database, Campbell Library and Cochrane Library. Our searches were based on the most central terms in relation to our research question: contracting, contracting out, outsourcing, tender, competitive tender, competitive bidding, marketisation/marketization and public. Moreover, in order to include studies that use the term ‘ownership’ (which is for example the case in some health care analyses), but actually examine experiences with contracting out as defined above, we carried out additional searches in Web of Science (SSCI) and Econlit on ownership, contracting, outsourcing or bid in combination with public and private and cost, saving, quality or efficiency. In total, the searches resulted in 7,217 studies published in the period 2000–14.

In the second phase, we examined the individual publications on the basis of their title and abstracts, and made an assessment of whether the publications addressed the research question and met the predefined list of inclusion criteria and were not subject to exclusion criteria. The specific inclusion criteria were: contracting out (sometimes also formulated as outsourcing) in the public sector in advanced industrialized countries (Europe, North America, Australia, New Zealand, Japan, South Korea, etc.), focus on effects for cost and/or service quality, and publication year from 2000 until 2014. The costs measured in the studies are those incurred by governments and which are directly related to production of the service, whereas possible costs to other parties, such as users or workers, are typically not included in the studies (for an overview of employee costs, see Vrangbæk et al. 2015).
Exclusion criteria were: studies from other countries than those stipulated by the inclusion criteria and studies focusing on other forms of public/private interaction than contracting out. Further exclusion criteria were studies with a strictly theoretical focus, studies of legal and regulatory frameworks, analyses of determinants/causes of contracting out and studies focusing on the process of contracting out rather than the effects. Lastly, we have included only primary empirical studies in the review, and not literature reviews and meta-analyses because they do not contain primary empirical data.

In the third phase, we assessed in detail the selected publications in order to determine whether the publications were relevant and whether the publications were of a satisfactory quality. A quality assessment scheme with 13 overall assessment criteria was developed and applied to all studies that were relevant according to our research questions and inclusion/exclusion criteria. The overall quality criteria used to assess each study were: (1) whether the study is based on an appropriate analytical design compared to the object of the analysis and the availability of relevant data in the field; (2) whether the data collection meets relevant standards. In quantitative studies the data collection should support statistical generalization and in qualitative research the collected data should be sufficient to support in-depth interpretations; and (3) whether the conclusions of the analysis are based on suitable methods for calculating the effects of contracting out and whether reporting of the findings provide relevant considerations regarding generalizability and interpretation of the findings.

On the basis of an assessment of the full-length articles we could determine which articles were relevant for our study and which articles could be excluded due to lack of relevance. The quality assessment led us to discard a number of studies due to insufficient design, methodology, data and/or problems with the technical execution of the analysis. The final set of studies consists of 49 publications that deal with the effects of contracting out in terms of costs and/or service quality in the period 2000–14. In comparison, Hodge’s widely cited meta-analysis was based on 28 publications (Hodge 2000) and Boyne’s (1998) on around 40 studies.

Although our final list of studies is based on systematic searches in a comprehensive list of databases and include all relevant and methodologically satisfactory publications that came up during our searches, it should be acknowledged that additional relevant studies could potentially have been missed due to the use of alternative wording and/or publication channels than those covered by this review. Moreover, publication bias could also be an issue, as found in the meta-regression analysis by Bel et al. (2010). The qualitative review method we use does not make it possible to carry out a formal test of publication bias, although the fact that a rather high proportion of the 49 studies found no significant cost difference makes us less concerned about this potential problem.

**Results of the Systematic Review**

Table 1 provides an overview of key information for the 49 studies in the systematic review.

© 2017 John Wiley & Sons Ltd
## Table 1

Overview of 49 studies in the systematic review

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Service area</th>
<th>Methods and data</th>
<th>Cost effects</th>
<th>Service quality</th>
<th>Transaction costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Kavanagh and Parker (2000)</td>
<td>Great Britain</td>
<td>Technical services (all technical services)</td>
<td>Qualitative (single case, n = 1)</td>
<td>Reduced costs (12.5–25%)</td>
<td>Higher user satisfaction. Other quality dimensions not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>2. Reeves and Barrow (2000)</td>
<td>Ireland</td>
<td>Technical services (waste collection)</td>
<td>Quantitative (survey data, n = 51), qualitative data (7 case studies)</td>
<td>Reduced costs (45%)</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>3. Keane et al. (2001)</td>
<td>USA</td>
<td>Social services (public health)</td>
<td>Quantitative (survey data, n = 347)</td>
<td>Reduced costs (effect size not measured)</td>
<td>Mixed quality effects (effect size not measured)</td>
<td>Increased transaction costs (size not measured)</td>
</tr>
<tr>
<td>4. McDavid (2001)</td>
<td>Canada</td>
<td>Technical services (waste collection)</td>
<td>Quantitative (survey data, n = 279)</td>
<td>Reduced costs (16.9%)</td>
<td>Not measured</td>
<td>Included for private producers (size not measured)</td>
</tr>
<tr>
<td>5. Camp and Gaes (2002)</td>
<td>USA</td>
<td>Social services (prisons)</td>
<td>Quantitative (survey data, n = 91)</td>
<td>Not measured</td>
<td>Quality decreased (effect size not measured)</td>
<td>Not measured</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Service area</th>
<th>Methods and data</th>
<th>Cost effects</th>
<th>Service quality</th>
<th>Transaction costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.</td>
<td>Denmark</td>
<td>Technical services</td>
<td>Quantitative (register data, panel organized, n = 270)</td>
<td>Reduced costs (15%)*</td>
<td>Not measured</td>
<td>Cost savings reported separately)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(road maintenance)</td>
<td></td>
<td></td>
<td></td>
<td>withstand control for service quality</td>
</tr>
<tr>
<td>8.</td>
<td>Netherlands</td>
<td>Technical services</td>
<td>Quantitative (survey data, n = 85)</td>
<td>Reduced costs (15%)</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(waste collection)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Sweden</td>
<td>Technical services</td>
<td>Quantitative (register data, n = 170)</td>
<td>Increasing costs (6%)</td>
<td>Not measured</td>
<td>Discussed but not measured</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(waste collection)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>USA</td>
<td>Social services</td>
<td>Quantitative (register data, n = 8.15 million)</td>
<td>Increasing costs (17%)</td>
<td>No difference (but very limited quality measures)</td>
<td>Not measured</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(health)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>USA</td>
<td>Social services</td>
<td>Quantitative (register data, panel organized, n = 3,122)</td>
<td>No direct cost comparison (5% lower spending on core instruction)</td>
<td>Quality decreased (only significant for large school districts)</td>
<td>Not measured</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(schools)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>Korea</td>
<td>Technical services</td>
<td>Qualitative (single case, n = 1)</td>
<td>Reduced costs (17.8%)</td>
<td>No difference</td>
<td>Not measured</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(wastewater services)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td>Spain</td>
<td>Social services</td>
<td>Quantitative (register data, survey data, n = 180)</td>
<td>Increasing costs (11.7%)</td>
<td>Not measured (some control variables for</td>
<td>Not measured</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(health)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Service area</td>
<td>Methods and data</td>
<td>Cost effects</td>
<td>Service quality</td>
<td>Transaction costs</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>--------------</td>
<td>------------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>14. Brudney et al. (2005)</td>
<td>USA</td>
<td>Technical services and social services</td>
<td>Quantitative (survey data, n = 1,175)</td>
<td>No significant difference</td>
<td>Higher quality</td>
<td>Not measured</td>
</tr>
<tr>
<td>15. Lazarus and McCullough (2005)</td>
<td>USA</td>
<td>Technical services (school buses)</td>
<td>Quantitative (register data for school year 1999–2000 and survey data [2001], n = 343)</td>
<td>Increasing costs (10%)</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>16. Bekken et al. (2006)</td>
<td>Norway</td>
<td>Technical services (public buses)</td>
<td>Quantitative (register data, panel organized, n = 285)</td>
<td>Reduced costs (10%)</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>17. Bel and Costas (2006)</td>
<td>Spain</td>
<td>Technical services (waste collection)</td>
<td>Quantitative (register data, survey data, n = 186)</td>
<td>No significant difference</td>
<td>Not measured (some control variables for quality included in cost estimate)</td>
<td>Not measured</td>
</tr>
<tr>
<td>18. Pina and Torres (2006)</td>
<td>Europe, Africa, America, Asia, Australia</td>
<td>Technical services (public buses)</td>
<td>Quantitative (register data, n = 73)</td>
<td>No significant difference</td>
<td>One quality measure included as control variable in cost estimation, but no estimate of quality effect</td>
<td>Not measured</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Service area</td>
<td>Methods and data</td>
<td>Cost effects</td>
<td>Service quality</td>
<td>Transaction costs</td>
</tr>
<tr>
<td>-------</td>
<td>--------------</td>
<td>-------------------------------------</td>
<td>------------------------------------------</td>
<td>----------------------------</td>
<td>----------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>19.</td>
<td>Christoffersen et al. (2007)</td>
<td>Denmark</td>
<td>Technical services (cleaning of public buildings)</td>
<td>Quantitative (survey data, n = 1,081)</td>
<td>Reduced costs (25.7–29.6%)</td>
<td>Not measured (quality is unobserved but assumed to be unchanged)</td>
</tr>
<tr>
<td>20.</td>
<td>Dijkgraaf and Gradus (2007)</td>
<td>Netherlands</td>
<td>Technical services (waste collection)</td>
<td>Quantitative (register data, n = 453)</td>
<td>Reduced costs (14–17%)</td>
<td>Not measured</td>
</tr>
<tr>
<td>21.</td>
<td>Gilmer (2007)</td>
<td>USA</td>
<td>Social services (mental health)</td>
<td>Quantitative (register data, n = 20,562)</td>
<td>Reduced costs (16.4%)</td>
<td>Higher quality but few quality dimensions measured</td>
</tr>
<tr>
<td>22.</td>
<td>Kuhlmann (2008)</td>
<td>Germany, France</td>
<td>Technical services, social services (not specified)</td>
<td>Quantitative (survey data, n = 817), qualitative data (cases, n = 4)</td>
<td>No cost difference</td>
<td>Not measured</td>
</tr>
<tr>
<td>23.</td>
<td>Lalive and Schmutzler (2008)</td>
<td>Germany</td>
<td>Technical services (railways)</td>
<td>Register (n = 80)</td>
<td>Not measured</td>
<td>Quality increased (effect size not quantified, measured only as frequency of service)</td>
</tr>
<tr>
<td>24.</td>
<td>Marques (2008)</td>
<td>Portugal</td>
<td>Technical services (water services)</td>
<td>Quantitative (official documents, n = 56)</td>
<td>Conflicting results with different methods (private productivity 20% higher but higher quality on several process indicators: testing quality: 9.1 percentage)</td>
<td>Higher quality</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Service area</td>
<td>Methods and data</td>
<td>Cost effects</td>
<td>Service quality</td>
<td>Transaction costs</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>--------------</td>
<td>------------------</td>
<td>--------------</td>
<td>----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>25. Bel and Mur (2009)</td>
<td>Spain</td>
<td>Technical services (waste collection)</td>
<td>Survey data and register data (n = 56)</td>
<td>No significant difference</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>26. Chong et al. (2009)</td>
<td>Australia</td>
<td>Technical services (public sector audits)</td>
<td>Quantitative (register data, survey data, n = 178)</td>
<td>Increasing costs for specific types of agencies (effect size not specified)</td>
<td>Not measured (assumes that quality is unchanged after contracting)</td>
<td>Not measured</td>
</tr>
<tr>
<td>27. Purse (2009)</td>
<td>Australia</td>
<td>Social services (employment services)</td>
<td>Qualitative (single case, n = 1)</td>
<td>Increasing costs (effect size not specified)</td>
<td>Lower quality (but few quality measures and effect size not quantified)</td>
<td>Not measured</td>
</tr>
<tr>
<td>28. Leland and Smimova (2009)</td>
<td>USA</td>
<td>Technical services (public buses)</td>
<td>Quantitative (register data, n = 545)</td>
<td>No significant difference</td>
<td>No difference (measured only as safety and frequency of service)</td>
<td>Not measured</td>
</tr>
<tr>
<td>29. Tiemann and Schreyögg (2009)</td>
<td>Germany</td>
<td>Social services (hospitals)</td>
<td>Quantitative (register data, 1.8–2.3%)</td>
<td>Increasing costs (effect size not specified)</td>
<td>Not measured (some control)</td>
<td>Not measured</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Service area</td>
<td>Methods and data</td>
<td>Cost effects</td>
<td>Service quality</td>
<td>Transaction costs</td>
</tr>
<tr>
<td>------------------</td>
<td>---------</td>
<td>-----------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-----------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>30. Bae (2010)</td>
<td>USA</td>
<td>Technical services</td>
<td>Panel organized, n = 1,046</td>
<td>No significant difference</td>
<td>Variables for quality included in cost estimate)</td>
<td>Not measured</td>
</tr>
<tr>
<td>31. Gilmer (2010)</td>
<td>USA</td>
<td>Social services</td>
<td>Quantitative data (register data, n = 252)</td>
<td>Reduced costs (31%)</td>
<td>Mixed quality effects (but few quality measures)</td>
<td>Not measured</td>
</tr>
<tr>
<td>32. Iseki (2010)</td>
<td>USA</td>
<td>Technical services</td>
<td>Quantitative data (register data, n = 24,595)</td>
<td>Reduced costs (5.5–7.8%)</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>33. Correia and</td>
<td>Portugal</td>
<td>Technical services</td>
<td>Quantitative (official documents, etc. panel organized, n = 131)</td>
<td>Reduced costs (2.1%)</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>Marquez (2011)</td>
<td></td>
<td>(water services)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34. Stolt et al.</td>
<td>Sweden</td>
<td>Social services</td>
<td>Quantitative (survey data, n = 2,629)</td>
<td>Not measured</td>
<td>Mixed effects (varying results for structural and processual quality)</td>
<td>Not measured</td>
</tr>
<tr>
<td>(2011)</td>
<td></td>
<td>(nursing homes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35. Thompson</td>
<td>USA</td>
<td>Technical services</td>
<td>Quantitative (register data, panel organized, n = 2,011)</td>
<td>Increasing costs (15.8–20.7%)</td>
<td>Quality decreased (but only measured as number of)</td>
<td>Not measured</td>
</tr>
<tr>
<td>(2011)</td>
<td></td>
<td>(school buses)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Service area</td>
<td>Methods and data</td>
<td>Cost effects</td>
<td>Service quality</td>
<td>Transaction costs</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>--------------</td>
<td>------------------</td>
<td>--------------</td>
<td>----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>36. Walter (2011)</td>
<td>Germany</td>
<td>Technical services (public transport: buses, tram, light railway and metro)</td>
<td>Quantitative (register, panel organized, n = 254)</td>
<td>Reduced costs (0.7–1.9%)</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>37. Tuominen et al. (2012)</td>
<td>Finland</td>
<td>Social services (dental care)</td>
<td>Quantitative (register data, randomized experiment, n = 85,100)</td>
<td>Increasing costs (14.4%)</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>38. Bennmarker et al. (2013)</td>
<td>Sweden</td>
<td>Social services (employment services)</td>
<td>Quantitative (register data and two surveys, randomized experiment, n = 2,410)</td>
<td>Not measured</td>
<td>No significant overall quality difference (but variations within sub-groups of the sample)</td>
<td>Not measured</td>
</tr>
<tr>
<td>39. Boitani et al. (2013)</td>
<td>9 European countries</td>
<td>Technical services (public transport)</td>
<td>Quantitative (register data, panel organized, n = 437)</td>
<td>Reduced costs (effect size not specified)</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>40. Dijkgraaf and Gradus (2013)</td>
<td>Netherlands</td>
<td>Technical services (waste collection)</td>
<td>Quantitative (register data, panel organized, n = 5,886)</td>
<td>No significant difference</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>41. García-Valiñas et al. (2013)</td>
<td>Spain</td>
<td>Technical services (water services)</td>
<td>Quantitative (register data, n = 386)</td>
<td></td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Service area</td>
<td>Methods and data</td>
<td>Cost effects</td>
<td>Service quality</td>
<td>Transaction costs</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>--------------</td>
<td>------------------</td>
<td>--------------</td>
<td>----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>42. González-Gómez <em>et al.</em> (2013)</td>
<td>Spain</td>
<td>Technical services (water services)</td>
<td>Quantitative (register data, n = 80)</td>
<td>No significant difference</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>43. Mouwen and Rietveld (2013)</td>
<td>Netherlands</td>
<td>Technical services (public transport)</td>
<td>Quantitative (survey data, n = 883,009)</td>
<td>No significant difference</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>44. Pérotin <em>et al.</em> (2013)</td>
<td>England</td>
<td>Social services (hospitals)</td>
<td>Quantitative (survey data, n = 38,447)</td>
<td>Not measured</td>
<td>No significant difference</td>
<td>Not measured</td>
</tr>
<tr>
<td>45. Rho (2013)</td>
<td>USA</td>
<td>Social services (schools)</td>
<td>Quantitative (register data, panel organized, n = 12,368)</td>
<td>No direct cost comparison (5% lower spending on core instruction)</td>
<td>Quality increased (2% higher test score)</td>
<td>Higher transaction costs, but size not measured</td>
</tr>
<tr>
<td>46. Scheffler <em>et al.</em> (2013)</td>
<td>Germany</td>
<td>Technical services (public buses)</td>
<td>Quantitative (register data, n = 54,6)</td>
<td>Reduced costs (3%)</td>
<td>Not measured</td>
<td>Not measured</td>
</tr>
<tr>
<td>47. Stanley <em>et al.</em> (2013)</td>
<td>UK</td>
<td>Social services (children’s out-of-home care)</td>
<td>Qualitative and quantitative (two surveys, n = 17,82)</td>
<td>No significant difference</td>
<td>Mixed quality effects (varying effects for</td>
<td>Not measured</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Service area</td>
<td>Methods and data</td>
<td>Cost effects</td>
<td>Service quality</td>
<td>Transaction costs</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------</td>
<td>-------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>--------------------------</td>
<td>-----------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>48.</td>
<td>Sweden</td>
<td>Social services</td>
<td>Quantitative (register data, randomized experiment, n = 4,090)</td>
<td>No significant difference</td>
<td>No difference</td>
<td>Not measured</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(vocational rehabilitation)</td>
<td>and n = 1676; interviews, n = 225)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49.</td>
<td>France</td>
<td>Technical services</td>
<td>Quantitative (register data, n = 177)</td>
<td>Increasing costs (6%)</td>
<td>No difference</td>
<td>Not measured</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(water services)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *The effect estimate in this study is calculated as the cost effect of a specific percentage point increase in the level of private production. To make the estimate comparable to other studies that use a public/private dummy, the cost effect of a full transition from public to private production (i.e. from 0 to 100 per cent private production) was calculated.*
Focusing first on the documented effects of contracting out on costs, table 1 shows a relatively large diversity across studies that measure cost effects. There are all in all 37 studies documenting either cost reductions, cost increases or no significant differences between public (in-house) production and contracting out. Moreover, four additional studies consider economic effects, but do not provide an estimate of the effect size and therefore cannot be included in our analysis of average effects below. The magnitude of the reported cost effects varies from a cost reduction of 45 per cent to a cost increase of 18.3 per cent, which documents a very broad span in the economic effects of contracting out in studies published from 2000 to 2014. Of the 37 studies that provide estimates of cost effects, 17 studies document cost savings, while nine studies document cost increases and 11 studies find no significant cost changes. The average cost effect from contracting reported in studies published in the period 2000–14 is a cost reduction of 4.2 per cent.

Another finding relating to the cost of producing public services is the absence of studies that measure transaction costs and take these costs into account. Of the 49 studies there is only one study that includes transaction costs for private production, but not for public production, and the actual size of transaction costs is not measured (McDavid 2001). There are also three studies asserting that transaction costs are likely to be higher when services are contracted out than with public production, but none of them measures the size of these costs (Keane et al. 2001; Ohlsson 2003; Rho 2013). The general lack of measurement of transaction costs in the vast majority of studies clearly makes a comparison of the full costs of public and contracted out service provision difficult. The average cost saving of 4.2 per cent is essentially a comparison of direct production costs, whereas in the majority of research other relevant (but indirect) costs relating to tendering, bids evaluation, contract signature, monitoring and sanctioning of contracts are not included in the cost estimate.

Moreover, any assessment of the economic effects of contracting out should, ideally, also take into consideration outcomes relating to service quality. Altogether, 19 studies out of 49 consider service quality as a dependent variable, though most of them do not provide quantitative measures of the effect, but merely evaluate that quality has either improved, decreased or has not been affected. Of the 19 studies, six report positive effects on quality as a result of contracting out, though only two studies quantify the quality improvement (Marques 2008; Rho 2013). There are four studies which document negative effects on service quality as a result of contracting out, whereas nine studies find mixed quality effects or reach no clear conclusion. Lastly, six studies include service quality as a control variable (some with more data than others), while 24 studies do not consider service quality at all. Hence, approximately half of the 49 studies do not examine effects relating to service quality at all, and the details given in table 1 show that merely a handful of studies measure service quality in a comprehensive and encompassing manner (Marques 2008; Stolt et al. 2011; Pérotin et al. 2013; Stanley et al. 2013; Laun and Thoursie 2014).
The secondary research themes of the systematic review were: (1) developments in cost effects over time; (2) possible differences in effects across technical and social services; and (3) geographical variations and institutional context. These results are reported in table 2. The table provides noteworthy insights that extend previous knowledge on the internationally documented effects of contracting out. Looking first at longitudinal developments, table 2 shows a declining trend in the documented cost difference between public and private production over time. In studies from the period 2000–04 cost savings are 8.5 per cent on average, while cost savings dropped to 4.8 per cent from the period 2005–09 and further decreased to 0.4 per cent from the period 2010–14. Cost savings from contracting out have thus been significantly declining over the 15-year period and are only marginally different from zero in the last five-year period. The decrease in effects over time might be due to either increased efficiency of public provision or decreased efficiency of private provision, or a combination of the two. Which explanation actually applies is not addressed in most of the studies included in the review (see, however, Bel and Costas 2006), which means that we cannot draw any final conclusions regarding the specific reasons behind the declining effects from contracting out since the turn of the millennium.

Furthermore, the results for various types of services show that cost savings from the contracting out of technical services have on average been 5.7 per cent over the period. This compares to average cost savings of 0.2 per cent in social services, which for all practical purposes (and given the relatively small n) is not significantly different from zero. The varying cost effects in technical and social services are well in line with our theoretical expectations based on transaction cost theory and the difference in asset specificity and measurability, which made us expect smaller and more diverse cost effects in social services than in technical services. The differences in cost effects between social and technical services are noteworthy because they indicate that savings from contracting out previously found in systematic reviews of

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of studies (N = 37)</th>
<th>Cost difference public versus private production (per cent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000–2004</td>
<td>11</td>
<td>-8.5</td>
</tr>
<tr>
<td>2005–2009</td>
<td>12</td>
<td>-4.8</td>
</tr>
<tr>
<td>2010–2014</td>
<td>14</td>
<td>-0.4</td>
</tr>
<tr>
<td>Social services</td>
<td>11</td>
<td>-0.2</td>
</tr>
<tr>
<td>Technical services</td>
<td>27</td>
<td>-5.7</td>
</tr>
<tr>
<td>Anglo-Saxon countries</td>
<td>14</td>
<td>-6.4</td>
</tr>
<tr>
<td>Other countries</td>
<td>23</td>
<td>-2.9</td>
</tr>
</tbody>
</table>

Note: Studies of social and technical services sum to 38 and not 37 because one study covers both services.
studies in technical services cannot simply be generalized to social services. These insights add to previous literature by showing that the documented effects of contracting out depend highly on the transaction cost characteristics of the service.

Lastly, in relation to the geographical variable, the last two rows of table 2 show that cost savings from contracting out have been on average 6.4 per cent in Anglo-Saxon countries and 2.9 per cent in other countries. Interestingly, documented cost effects from contracting out between 2000 and 2014 are more than twice as large in Anglo-Saxon countries as in other countries. These findings support our theoretical expectations about larger effects from contracting out in liberal market economies with less regulated labour markets and weaker unions compared with more regulated market economies (Gumbrell-McCormick and Hyman 2013). The findings from our systematic review of studies from the period 2000–14 thus indicate that previous meta-analyses from the 1980s and 1990s, which mainly included technical services and Anglo-Saxon countries, might have found larger cost savings than would have been the case if the experiences from a broader range of countries and social services were to a larger extent included in the assessment of the evidence. These insights underpin that generalization of economic effects across geographic regions should be made with caution in the same way as generalization across service areas with different transaction cost characteristics.

Discussion

The review of international studies on the effects of contracting out illustrates that the studies are of variable quality, as also noted almost 20 years ago in the review by Boyne (1998). There is (with some notable exceptions) a distinct lack of studies with a solid methodological foundation documenting effects of contracting out in relation to costs, quality and transaction costs. Most studies focus on cost effects, while many entirely leave out or only scarcely examine service quality outcomes. As a consequence, there are very few studies that enable a comprehensive assessment of the effect of contracting out on overall effectiveness of service delivery. Some studies simply assume that quality remains unchanged (e.g. Christoffersen et al. 2007), while many others fail to consider service quality altogether. These weaknesses and limitations are problematic because they can lead to simplified or even erroneous interpretations of the overall effects of contracting out.

In the following, we discuss a selection of the themes that have become apparent in our review, and how these issues may have influenced the results that were reported in the studies.

The challenge of comparability

There is a fundamental challenge when comparing effects of public and private organizations. Ideally, assessments of the effects should compare private and public suppliers that provide exactly the same services under similar
organizational and regulatory conditions. This situation is rarely present in practice, as industrial organization theory has underlined. There may be differences in the scope of services, the regulatory settings and other task-specific elements that frame the trade-off between in-house provision and contracting out (Levin and Tadelis 2010). Basically, we demand that our public organizations deliver a broader range of services than their private counterparts. Also, public organizations are often required to serve all citizens in a given area, while private companies can often choose their markets and services.

Moreover, if it is not ownership (property rights theory), but rather the competitive environment (public choice theory) that determines the outcome of contracting out, it also becomes important to measure competition in public service markets. Most studies measuring competition find that competition for public services is limited (Dijkgraaf and Gradus 2007; Levin and Tadelis 2010; Hefetz and Warner 2012), although the more precise link between market characteristics, institutional settings and outcomes of contracting out still needs further scrutiny (Brown and Potoski 2003). These and related differences in the production context of public and private organizations potentially create problems of comparability, which scholars of public service delivery need to take into account when evaluating the costs and quality of public and private service provision.

The challenge of short time spans

The importance of viewing effects of contracting out as a dynamic phenomenon developing over time has been stressed in recent research (Bel and Costas 2006; Hutchinson and Pratt 2007; Bae 2010) and has been highlighted by our findings as well. However, most empirical studies identified in our review limit themselves to look at a short time period and consider first generation contracting only. The short time perspective is problematic because the first contract period can be considered a learning period for both the public and the private party. In this learning period transaction costs are likely to be higher than in subsequent tenders. It is also problematic because during the first tender the private party may be tempted or pressured to set a low price to enter the market and typically feel compelled to raise the price in the next contract round (Bel and Costas 2006).

A basic problem that is seldom accounted for in current studies is the need to carry out a fair comparison between public and private organizations at the same point in time. In many qualitative studies of contracting out it is common to compare the performance of public organizations at an earlier point in time than the performance of private organizations. However, in a pre-post study it is methodologically problematic to make an uncontrolled comparison between public and private organizations without taking into account the development over time which may have been taking place after the contracting out. It is thus crucial to avoid using the wrong public sector benchmark as the basis for a comparison with private provision because such effects can also be dynamic and change over time (Bel and Costas 2006; Houlberg and Petersen 2015).
Inclusion of all direct and indirect costs

Only few studies in the review provide an examination of transaction costs related to tendering, contracting and subsequent monitoring of service delivery. Yet, when measuring the economic effects of contracting out, it is vital to include transaction costs. Such costs should, ideally, be compared to the administrative costs related to producing services in-house by the public sector. However, among the 49 studies we have not identified any study making such a comparison.

There are also other possible indirect costs that should be considered. This applies to the costs related to maintaining a preparedness to take over in case of private sector bankruptcy and/or non-compliance with a contract. Furthermore, as highlighted by Bel and Mur (2009), it is important to take into account different contextual and market conditions in urban and rural settings. A well-functioning private market requires several providers that compete for contracts. That is not always the case in sparsely populated rural areas. Effects of contracting out obtained in urban areas do not necessarily apply in rural areas where infrastructure, population density and potential for recruitment are very different, and the cost structure for both public and private actors vary significantly compared to an urban context.

Measuring service quality: a challenging yet important task

Most scholars agree that service quality constitutes an integral part of service delivery, yet few studies accomplish to measure it. Service quality, moreover, is an ambiguous and methodologically difficult concept to measure in public and private services alike. Quality can refer to different things and include user perceptions as well as more objective and standardized measures. A commonly used terminology distinguishes between input, process and outcome dimensions of quality (Stolt et al. 2011). Input is typically measured in terms of resources (e.g. human and technical resources). Process refers to the procedures and to the direct output (e.g. frequency of snow-ploughing, cutting the grass, collection of garbage and so on). Lastly, outcome is the most important measure of quality as it relates to the actual impact (e.g. long-term health after hospitalization or student performance in schools). In technical areas it is relatively easy to specify process and output measures and argue that they are valid indicators of quality. Repairing pot-holes and providing frequent snow-ploughing probably have a fairly straightforward causal link to number of road accidents, which could be a solid outcome indicator.

When it comes to more complex welfare services, however, it is not that simple. One can count the number of hours provided in day care or the frequency of visits to elderly persons, but this does not necessarily provide information about the effects on the social and intellectual skills conveyed to children in day care or the health and quality of life for the elderly. In addition, there are often conflicting views on what the primary outcome should be. Preschool day care can, for instance, be judged in terms of intellectual progress, quality of life, social skills or parent satisfaction. Such conceptual and methodological challenges of measuring quality can perhaps
explain why this dimension is rarely included, but it can hardly justify this practice. Obviously, more thorough studies about changes in quality resulting from contracting out are needed, before general conclusions about effects of contracting out for service quality can be drawn (see also Boyne 1998).

**Concluding Remarks**

The systematic review of international studies in the period 2000–14 shows that there is empirical evidence of cost savings from contracting out in technical areas, but not in social services. It is important to stress, however, that the evidence in social services is much scarcer and more difficult to uncover than in technical services. There is a remarkable lack of studies thoroughly examining effects of contracting out for service quality, and the findings in the few studies that do examine service quality are mixed. Furthermore, hardly any studies include empirical estimates of transaction costs. Transaction cost theory teaches us that the full cost of producing any good or service consists of the sum of production costs and transaction costs (Williamson 1979; Brown and Potoski 2003; Hefetz and Warner 2012). Excluding the latter from the cost-estimation is therefore an important limitation of the current literature on effects of contracting out.

Most of the 49 studies are carried out as static investigations with limited consideration of the dynamic and longer-term effects, for instance by repeated contracting out (for exceptions, see Bel and Costas 2006; Dijkgraaf and Gradus 2013). In these instances, private companies typically hold the contract as a result of an earlier tender and other private companies are the main competitors. This is, obviously, a very different scenario compared to first generation tenders. The few studies that do include a longer-term perspective indicate that marginal benefits of contracting out tend to be decreasing over time (Bel and Costas 2006; Hutchinson and Pratt 2007; Houlberg and Petersen 2015), which was also a main result of our assessment of economic effects in studies from the periods 2000–04, 2005–09 and 2010–14, respectively.

Overall, the assessment of the contracting out literature from the period 2000–14 illustrates that cost savings (average 4.2 per cent) have been documented, though mainly in technical services, in Anglo-Saxon countries, and in the beginning of the examined time period. Results from existing studies are fairly heterogeneous (from cost savings of 45 per cent to cost increases of 18 per cent) and generalization of findings across service areas, between countries with differing regulatory and institutional settings, and across time should thus be made with caution. Cost savings are rather limited, they seem to become smaller over time, and other important aspects, such as service quality and transaction costs, are scarcely measured. The lack of solid documentation of service quality and transactions costs highlights the need for a balanced and pragmatic view on the respective pros and cons of contracting out.

The results from the systematic review are generally in line with previous review studies; hence, like Boyne (1998), Hodge (2000) and Bel et al. (2010), we conclude the article by emphasizing the methodological limitation
of most existing studies, particularly with regard to examining outcomes for service quality. Admittedly, there are some studies that scrutinize contracting out and service quality in sufficient detail (cf. Stolt et al. 2011; Pérotin et al. 2013; Laun and Thoursie 2014), but most analyses examine this important element superficially or not at all. It is unlikely that anybody visiting the local grocery store would purchase a bottle of milk or a packet of meat without being concerned with the relation between price and quality. However, when contracting out services to the private market the public sector often lacks this fundamental information. Unfortunately, contemporary academic literature provides limited guidance to public policymakers on this important topic, since hardly any studies evaluate contracting out with regard to the impact on overall cost-effectiveness.

For future research, a more context-oriented approach, which measures both direct production costs, transaction costs and service quality, is timely and warranted. Other dimensions including effects on users and employees also constitute important avenues for further research. Based on current knowledge from international research, we conclude that contracting out has not been the great panacea, though, in specific cases and settings, contracting out has proven to be an effective tool for the public sector.

Acknowledgements
The authors would like to thank the anonymous reviewers for their constructive comments on an earlier version of the manuscript. Karsten Vrangbæk’s contribution was made possible by a grant from The Danish Research Council (DFF – 1327-00122).

Parts of the collection and assessment of empirical studies for the systematic review were carried out during a research project that was partially funded by a number of public service unions in Denmark.

References

© 2017 John Wiley & Sons Ltd


Konnerup, M. and Kongsted, H. C. (2012), Do Cochrane reviews provide a good model for social science? The role of observational studies in systematic reviews, Evidence & Policy, 8, 1: 79–96.


