An Optimization Framework for Efficient Information Supply in the Academic Sector

Daniel Beimborn
Johann Wolfgang Goethe University Frankfurt,

Rainer Fladung
Johann Wolfgang Goethe University Frankfurt,

Wolfgang Koenig
Johann Wolfgang Goethe University Frankfurt,

Follow this and additional works at: http://aisel.aisnet.org/pacis2004

Recommended Citation
http://aisel.aisnet.org/pacis2004/61

This material is brought to you by the Pacific Asia Conference on Information Systems (PACIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in PACIS 2004 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
An Optimization Framework for Efficient Information Supply in the Academic Sector

Daniel Beimborn, Rainer B. Fladung, Wolfgang Koenig
Institute of Information Systems
Johann Wolfgang Goethe University Frankfurt, Germany
{beimborn|rfladung|koenig}@wiwi.uni-frankfurt.de

Abstract

In recent years, increasing monopolization tendencies in the market for electronic information resources in the academic sector have led to the formation of consortia by university libraries and other academic institutions, which are intended to act as an antipole to increase their bargaining power against the monopolistic structures of the publisher market. The aim of this research paper is the development of a model that determines minimal-cost structures for buying consortia in the context of academic libraries. Two different points of view are discussed: on the one hand the optimal choice of the pricing model for purchasing electronic journals and on the other hand the optimization of consortia structures for bundling bargaining power. For the first part of the framework a break-even based approach is developed and applied to empirical data. It can be shown, that German academic libraries can cut their procurement costs by canceling subscriptions and switching to pay per view by about 30%. The second topic – the consortia structure problem – is solved by a binary nonlinear optimization model, which also is applied to a real situation. In future it is able to determine, for example, the optimality of a Germany-wide library consortium and to show efficient alternatives.

Keywords: Academic Libraries, Publishers, Procurement Consortia, Optimization, Electronic Journals

1. Introduction

In recent years, increasing monopolization tendencies in the market for electronic information resources in the academic sector have led to the formation of consortia by university libraries and other academic institutions, which are intended to act as an antipole to increase their bargaining power against the also monopolistic structures of the publisher market (McCabe 2002). This trend can be seen in several countries like Germany, the United Kingdom, Switzerland and the Netherlands (Okerson 1996; Wiesner and Dugall 2002). In the United States library consortia can only be found rarely, the most important one being OHIO-Link (http://www.lib.ohio-state.edu/).

In Germany, due to the federal financing structure, most of the consortia are organized regionally; procurement cooperation with a thematic focus is not common (Degkwitz and Andermann 2003). Recently, such trends have not had any counteractive effect on the increasing prices in the academic information market, e.g. for electronic journals and databases. The NESLI-Project in the United Kingdom, which had the aim of negotiating the general terms for libraries and publishers, failed in 2002 (Walport 2003).

The aim of this research paper is the development of a model that determines minimal-cost structures for buying consortia in the context of academic libraries. In the process two different points of view will be discussed: on the one hand the optimal choice of the pricing model
for purchasing electronic journals and on the other hand the optimization of the consortia structures for bundling bargaining power.

Both parts of the framework are applied to a real scenario: the procurement relationship of the Hessian library consortium HeBIS, which consists of 12 academic libraries, and a large international publisher of scientific journals. Section 2 describes the actual situation of scientific journal procurement in Germany and provides some basic empirical data. In section 3 the model is developed in three steps – optimization of the pricing portfolio, optimization of consortia structures, and simultaneous optimization of both – and applied to empirical data, which was extracted from earlier work (Dugall and Fladung 2002a). Section 4 gives a short summary and an overview over further research.

2. Electronic Procurement of Scientific Information

2.1 Participating Actors and Market Structure

The market for academic journals basically consists of four major actors: scientists who produce the academic content, publishers who collect this content, libraries who buy the publishers’ products, and finally again researchers and students who use the academic literature in libraries (Odlyzko 1997). The actual situation is characterized by the paradox that the publicly funded academic libraries have to repurchase the digital content which is mostly “produced” by also publicly financed scientists (Degkwitz and Andermann 2003).

Libraries which claim to have a virtually complete information stock or the fact that research and teaching are confronted with a very problematic development, is also known as the crisis of academic information provision (Bergstrom and Bergstrom 2002; Wiesner and Dugall 2002). While the prices of electronic journals and databases have continuously escalated during the last several years, the libraries’ budgets have moved in a downward direction. Major commercial publishers like Reed Elsevier or Kluwer/Springer, who publish lots of journals with a high reputation, can raise their content prices without fearing that any library will cancel the subscription. The main reason for this effect is the lack of competition along the highly specialised journals and a lack of elasticity of demand for the “need-to-know”-product academic information (Degkwitz and Andermann 2003). Normally, scientific journals do not compete by pricing; the libraries’ procurement strategy is determined by the demand of the end users who do not usually have to pay for the journals (Odlyzko 1997). A further important trend in the market for scientific communication is the enormously growing relevance of electronic information since the late 90s. Electronically-accessible information has a lot of advantages in comparison to the printed equivalents; e.g. electronic content can be more easily found and searched, can be theoretically accessed anytime and from anywhere, and be handled more flexibly. Online access to academic information can lead to an acceleration and increase of efficiency in intellectual production by the scientific community (NN 2001). On the one hand in the present time of decreasing budgets academic libraries have to cancel existing subscriptions in order to be able to finance other relevant literature (Wiesner and Dugall 2002), and on the other hand publishers try to protect their revenue while bundling electronic access with hardcopies of the same journals. Often, libraries can only purchase a pre-selected set of journals bundled by the publishers and more often the basic contract fixes the price of cancelling subscriptions (Degkwitz and Andermann 2003).

Recently, academic libraries have reacted by reorganization and by forming buying consortia to bundle the demand and to get a better bargaining position (Wiesner and Dugall 2002).

---

1 Hesse is the fifth largest German state with about 6 M inhabitants. The largest city is Frankfurt/Main.
2 The number of participating libraries can differ from year to year and publisher-related contract.
2.2 Academic Library Consortia in Germany

At present, German academic libraries are primarily organized in 16 consortia with around 90 participating libraries (Reinhardt 2003). They differ in size, in their importance as negotiators, in their legal form, in their financial base and in their human resources. The larger German consortia are organized in regional structures because of the federal financial public system in Germany (Figure 1).

![Figure 1: Map of states and active consortia in Germany (in 2003) [Reinhardt 2003]](image)

Publishers are afraid of losing sales volume in the print-media area when providing the contents on the internet, whereas electronic resources will not unwind these losses (Phillips and Phillips 2002). Because of its medium the electronic version has two main characteristics that influence the sales situation of the publishers significantly:

- Due to the possibility of electronic access there is no necessity to subscribe to multiple copies of the same printed journal. This fact is amplified by establishing consortia (cross access, cf. 2.3.1).
- The distribution channel causes additional costs for software development, care and maintenance as well as for hardware acquisition.

The first aspect in particular causes the publishers to connect prices for electronic access to the journals subscribed to in the print-version (Meier 2002). Normally, these offers are combined with a clause that makes the printed journals non-unsubscribable to prevent sales decreases (Anglada and Comellas 2002). The frequently found high complexity of contract conditions for consortial subscription does not allow the advantages to be estimated. Often the pricing makes a fair assignment of costs related to single access impossible (Dugall and Fladung 2002b).

Former research approaches were limited to the description of present pricing models and their economic implications for the publishers’ earnings (Getz 1999; Odlyzko 1997; Okerson 1996; Varian 1996). Odlyzko addresses the question of competition on the scientific literature market and the implications of the quasi-monopolistic price ranges of single publishers as well as alternative methods of publishing. He comes to the result that the present publishing system cannot be financed in future, and that scientists will have to support alternative options opened up by electronic publication (Odlyzko 1997), e.g. academic self-organized repositories. Further on Odlyzko sees the possibility of cost reduction by using variably priced document delivery services (Odlyzko 1996).

Okerson mentions new challenges for libraries and publishers in terms of the increased amount of work in information editing and processing as well as in pricing (Okerson 1996). McCabe examines the mergers of publishers and the consequences for prices (McCabe 2002). His examination provides evidence for the fact that the high concentration on the content supplier side makes monopoly profits possible (McCabe 2002).
All previous works in this field have in common that there is neither a decision making model for the libraries or consortia nor an empirical investigation that shows at least rudimentary consequences of alternative pricing models.

One of our earlier research works dealt with the usage of electronic journals (Dugall and Fladung 2002a). One major result over all investigated libraries and publisher is that 50% of the full text accesses within a period are limited to around 7% of the titles. 80% of the accesses are taken up by around 25% of the titles. Further, the analysis showed that around 63% of the relevant journals on the average were recalled under 12 (!) times per year and library. On average 97% are requested fewer than 120 times per year (in full text, i.e. one article). (Dugall and Fladung 2002a).

2.3 Existing Pricing Models
The currently available pricing models for procuring electronic journals are the fixed-price calculation based on existing subscriptions to paper-based journals and pay per view (ppv).

2.3.1 Subscription-based Fixed Price Calculation
The pricing model based on subscribed print-media (surcharge model) is commonly used for electronic journals and will be explained by an example in the following (Dugall and Fladung 2002a):

The monetary amount of subscriptions in the HeBIS-consortium for printed journals of a large sized publisher was around € 1,605,586 (net) in 2001 including multiple subscriptions of the same journal. The license (€ 189,250 (net)) included cross access, so consorted libraries were allowed to access issues they did not have physically in stock at their own location electronically. The license surcharge for electronic access was 15%. This shows that the basic cost for subscribed print-media does not include the printed version of the journal only, but an option for electronic usage, too. Taking the option causes additional costs amounting to the surcharge.

The license allows all members of the consortium electronic access to the journals subscribed to by at least one member of the consortium in printed form (cross access) and named in the contract.

In general, the basic costs for consortia are calculated on multiple print subscriptions, while libraries that act alone have only to take into account one copy of the relevant journal to determine the basic costs.

2.3.2 Pay per View
In contrast to the fixed-price surcharge model presented in section 2.3.1, the pay per view-model (ppv) assumes that every single access to a complete journal article (“full text”) is charged with a specific amount. The costs for electronic access to a journal result from the number of calls and the price per call. The cost rate can differ from journal to journal, but presently there is no price differentiation for ppv-journals.

3. A Framework for Optimizing the Procurement of Scientific Journals
To optimize the procurement of electronic journals from a global view, i.e. the minimization of the aggregate expenses of all considered libraries, multiple strategies are imaginable.

Firstly, by evaluating the demand statistics costs-cutting decisions about a change in the pricing model to ppv could be made. In this model subscriptions to printed journal are one of the optimization parameters and therefore endogenous. consider

Secondly, the size and structure of consortia for bundling the bargaining power against the publishers could be optimized. In contrast to the first examination subscriptions to the hard-copy version of the relevant journals are only input parameters and therefore exogenous.
Finally, we will give a possible approach to minimizing costs by a simultaneous solution method. In this model the print subscriptions are endogenous and object of the optimization again.

3.1 Optimization of Individual Journal Portfolios

3.1.1 Model Description
A change of the currently existing procurement practice from a pure surcharge model to a mixed pricing portfolio, including ppv for the less demanded journals, can have an impact on different elements of the subscription license. If we assume that all journals of a single publisher would be bought by a fixed-price subscription, all printed copies form the base for calculating the price for the electronic version. If a journal is not demanded often a library might change the annual subscription to ppv purchasing. In that case the library also abandons the printed journal. This effect can be neglected; empirical studies showed that the demand for printed journals will be almost non-existent, if the electronic version is available (Keller 2002).

In essence, the decision model consists in essence of a break-even analysis, which determines the cost-minimal pricing alternative for every journal. To apply a break-even analysis the expected usage degree \(d_k\) for every journal has to be estimated, which will be done by the linear extrapolation of historical data, and be varied within a sensitivity analysis to decrease the probability of wrong decisions.

3.1.2 Model Application
The following application of the model is based on an empirical survey of the electronic journals usage within the Hessian library consortium HeBIS in 2001 (Dugall and Fladung 2002a).

Although only the journal portfolio of one exemplary publisher is presented here, the journals of different publishers were included in the survey; the results are transferable to the other ones.

In 2001, 965 periodic journal subscriptions existed for different printed journals published by the particular publisher. When removing multiple subscriptions to the same journal, 550 titles will remain.

The basic subscription costs for the printed journals amounted to € 1,605,586 and the licensing costs were at a level of € 189,250, so in total the procurement costs added up to € 1,794,836. Divided by the total number of accesses (229,898) the average costs per access to an electronic article amounted to € 7.81.

In the following the break-even analysis is applied to this data base. The first calculation shows the resulting costs when assuming a (quite realistic) ppv fee of \(c' = € 25.00\), a license cost factor of \(\alpha = 0.15\) (15%) and an estimated demand growth factor of \(\gamma = 0.0\).
Figure 3: Procurement costs subject to pricing portfolio

Figure 3 shows the costs which result, when switching the fixed price fee for the x% least used journals to ppv. For the explored journal portfolio under the given conditions it would be cost-minimal to switch the pricing model to ppv for about the 24% (132 journals) least demanded journals. By realizing this, portfolio costs could be reduced by about more than 15%.

To analyze the impact of $c_v$ in the next step it will be varied between € 20 and 50 per access. $\gamma$ and $\alpha$ remain constant at their previous values.

In Figure 4 it can be seen, that even with high access fees total savings of over 5.6% could be achieved, assuming the demand level of 2001. Depending on the level of $c_v$ savings of between € 93.700 and € 274.678 could be achieved. Again, it should be mentioned that this is only for the journal portfolio of one publisher. Even if € 50 had to be paid per view, 15% of the journals should be bought in that way.

Figure 4: Savings and minimal costs at different ppv fees

The most stochastic parameter is the access frequency $d_k$. To test the sensitivity of this parameter it will be varied between $\gamma = 0.0$ and 2.0, i.e. a tripling of the demand. Again, the fee is set to $c_v = € 25$.

Figure 5 shows the results. Even when the usage of electronic journals is tripled switching of some journals (11.6%) to ppv is efficient and leads to savings up to € 50,000 (i.e. 3%). However, longitudinal studies showed that the number of requests didn’t vary strongly between several years (Dugall and Fladung 2002a).
Figure 5: Procurement costs and savings at different demand growth rates

3.2 Optimization of Consortia Structures

3.2.1 Model Description

Based on the general conditions described in chapter 2 in the following we present a normative model, which optimizes – given a set of libraries – the number and structure of consortia, the libraries should be organized in, and minimizes the total (i.e. over all libraries) costs for procuring electronic journals. The optimization is done only from a centralized perspective; individual incentives and goals of the different libraries will not be considered. The total costs, which are minimized by the model, consist of all libraries’ procurement costs as well as their administrative costs, which result from managing the procurement process and coordinating consortia.

The model is based on the following assumptions, which have partly simplifying character:

- The number of potential consortium members is known.
- A particular library can only be member of one consortium (for the procurement of journals from one publisher).
- The existing subscriptions of printed journals are known.
- The demand for every journal at every library is known or can be estimated adequately.
- The costs for print-subscriptions of each title are known.
- Every library offers each journal electronically, independent of an existing print-subscription.
- The surcharge for electronic access is known and identical for each journal published by a particular publisher.
- The licensing costs are lower for libraries which are not members of a consortium.
- For consortia multiple subscriptions of the same title will be taken into account when calculating the costs base; if the library does not act as a consortium member only one subscription will be considered.
- Journals, which are not subscribed to in print, will be accounted for by ppv, when demanded electronically. The ppv fee is known and identical for every journal.
- The administrative costs increase super-proportionally with the number of consortium participants.

The procurement costs as first part of the cost function are aggregated over all libraries and depend on the coordination form – stand alone procurement or concerted procurement by a consortium. Further, the costs depend on the subscriptions which exist for printed journals. If
one library holds a print subscription for a particular journal, none of the other consortium members has to pay ppv fees when accessing the electronic version (cross access, cf. 2.3.1). Libraries, which are not members of a consortium, have to pay the subscription fees for the printed journals (just for one copy) as well as pay-per-view costs for demanding journals which are not in the library’s stock.

\[ C_{\text{Procurement}}^\text{Network} = \sum_{i=1}^{I} \left( (\alpha - \theta \cdot (1 - z_i)) \right) \sum_{k=1}^{K} \left( (1 - z_i) g_{ik} + z_i h_{ik} \right) c_k^a + \sum_{i=1}^{I} x_{ij} \sum_{k=1}^{K} c_k^v \right) \]

**Equation 1: Procurement costs**

Equation 1 describes the calculus of the procurement costs by using different decision variables and input parameters. These are explained in Table 1.

Variable \( z_i \) is equal to one, if library \( i \) is member of a consortium; in this case the complete surcharge has to be paid. If it does not cooperate, \( z_i \) is 0 and the surcharge \( \alpha \) will be diminished by \( \theta \) (see 8th assumption).

The level of the basic costs also depends on the coordination form. If the library acts on its own (\( z_i = 0 \)) possible multiple copies of printed journals are not taken into account (the binary variable \( g_{ik} \) indicates, if a (printed) journal \( k \) is in stock at library \( i \)). Otherwise (the library is member of a consortium) all print subscriptions (including multiple copies of the same journal within the whole consortium) will be the basis for calculating the basic costs (number of copies (or only one) multiplied by the journal’s subscription fee \( c_k^a \)).

To describe consortia in the model we introduced an index \( j \), which stands for procurement entities. These entities could either be libraries which act alone or consortia. The binary decision variable \( x_{ij} \) describes which library is assigned to which procurement entity in the optimal solution. If for a particular \( j \) only one \( x_{ij} \) is equal to one, library \( i \) manages its procurement independently; in the other case it is member of a consortium.

<table>
<thead>
<tr>
<th>Input Parameters</th>
<th>Decision Variables</th>
<th>Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = ) relative surcharge for consortia</td>
<td>( z_i = \left{ \begin{array}{ll} 1 &amp; \text{if library } i \text{ is member of a consortium} \ 0 &amp; \text{else} \end{array} \right. )</td>
<td>( i = \text{library index} \quad \forall \ i = {1, \ldots, I} )</td>
</tr>
<tr>
<td>( \alpha - \theta = ) relative surcharge for stand-alone acting</td>
<td>( x_{ij} = \left{ \begin{array}{ll} 1 &amp; \text{if library } i \text{ is member of procurement entity } j \ 0 &amp; \text{else} \end{array} \right. )</td>
<td>( j = \text{procurement entity index} \quad \forall \ j = {1, \ldots, J} )</td>
</tr>
<tr>
<td>( c_k^a = ) costs for a print subscription of journal ( k ) per period</td>
<td>( y_{jk} = \left{ \begin{array}{ll} 1 &amp; \text{if journal } k \text{ is subscribed by at least one member of procurement entity } j \ 0 &amp; \text{else} \end{array} \right. )</td>
<td>( k = \text{journal index} \quad \forall \ k = {1, \ldots, K} )</td>
</tr>
<tr>
<td>( c^v = ) fee per access (pay per view)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( d_{ik} = ) demand for journal ( k ) in library ( i )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( g_{ik} = \left{ \begin{array}{ll} 1 &amp; \text{if library } i \text{ has subscribed the printed version of journal } k \ 0 &amp; \text{else} \end{array} \right. )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( h_{ik} = ) number of printed copies of journal ( k ) which are subscribed by library ( i )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In the last term the ppv costs are formulated. If \( y_{jk} = 0 \), then the journal is subscribed to by none of the libraries in the considered consortium. In this case the demand has to be provided by paying for every access.

The second cost factor is circumscribed as the administrative costs. These depend on the size of the libraries and of the consortia. We here assume that libraries can be classified into three different groups of size, associated with different levels of administrative costs. The “large”-group consists of the university libraries, while the state libraries as well as the technical universities’ libraries are assigned to the “medium” class. The libraries of the universities of applied sciences (UAS) form the group of small libraries.

Depending on the coordination form (stand-alone activity vs. consortium) the administrative costs vary. For each group of libraries therefore two different administrative cost factors have to be determined: \( c_{ads}^i \) for acting alone and \( c_{adk}^i \) for being a consortium member. The values for \( c_{ads}^i \) were acquired by interviews with experts (Table 2).

| Administrative costs per library, year, and contract \( c_{ads}^i \) |
|--------------------------|-----------------|
| Small library            | 1250            |
| Medium-size library      | 2500            |
| Large library            | 3750            |

Table 2: Stand-alone administrative costs

Next, the administrative costs for coordinating in a consortium have to be determined. Empirically it can be shown, that the average administrative costs per member in a consortium initially decrease with growing consortium size. But the more participants the consortium contains, the larger is the necessary coordination effort, so that the average costs outbalance the stand-alone administrative costs above a certain consortium size.

To formalize this effect we use the consortia administrative cost coefficient \( c_{adk}^i \) \( (s \in \{ \text{small, medium, large} \}) \) and exponentiate the cumulated cost coefficients by a constant parameter \( \beta \). The resulting administrative cost function is given by Equation 2, the variables and parameters used are described in Table 3.

\[
C_{\text{Network}}^{Ad} = \sum_{j=1}^{J} \left( (1 - u_j) \sum_{i=1}^{I} c_{ads}^i + u_j \left( \sum_{i=1}^{I} x_{ij} \cdot c_{adk}^i \right)^{\beta} \right)
\]

Equation 2: Administrative costs

Variable \( u_j \) is equal to 1, if procurement entity \( j \) consists of more than one member (i.e. a consortium). The parameter \( \beta \) has to be larger than 1.0 to ensure the described cost relationship. It can not be ascertained empirically; therefore, it is varied within a sensitivity analysis later.
Table 3: Input parameters and decision variables for administrative costs

In a last step the consortium administrative costs parameters $c^{adk}_s$ for the different libraries’ sizes have to be estimated. For that, experts were asked, at which consortium size $size^*_s$ ($s \in \{\text{small, medium, large}\}$) – assumed that a consortium only consists of libraries of the same size – the consortia administrative costs per member have the same level as the stand-alone administrative costs.

This break-even point was estimated as a critical consortium size of 10 large, 15 medium-size, or 25 small libraries. Given the $c^{adk}_s$ values and a hypothetic determination of $\beta$ the values of $c^{adk}_s$ can be computed by Equation 3.

\[
(3) \quad c^{adk}_s = \beta \frac{\text{size}^*_s \cdot c^{adk}_s}{\text{size}^*_s} \quad \forall \ s \in \{\text{small, medium, large}\}
\]

Equation 3: Derivation of cost coefficients

Figure 6 shows the average administrative costs per library for varying the size of a homogeneous consortium of large libraries. The minimal administrative costs will be reached with a consortium size of 4 libraries, the stand-alone administrative costs are again reached at the break-even $size^*_s = 10$.

Taking the presented cost functions an optimization model can be developed, which minimizes the aggregate costs and determines the optimal consortia structure of the libraries we took into account.

\[
\text{Figure 6: Exemplary average administrative costs / library for different consortium sizes (homogeneous consortium with only large libraries; } \beta=1.3)\]

The aggregate of the cost functions is given in the model’s objective in Equation 4. The given constraints are added to ensure, that only proper variable values occur. The resulting binary non-linear optimization program (INLP) can be solved by appropriate software.
\begin{equation}
\text{(4 - 1) Minimize}
\text{ } \sum_{i=1}^{n} \left( (\alpha - \theta \cdot (1 - z_i)) \sum_{j=1}^{m} k_i j \cdot (x_{ij} + z_j) \cdot c_{ij} \right) + \sum_{j=1}^{m} \left( \sum_{i=1}^{n} k_i j \cdot (x_{ij} + z_j) \cdot c_{ij} \right)^{\beta} + \sum_{i=1}^{n} \left( \sum_{j=1}^{m} k_i j \cdot z_j \cdot \left( \sum_{k=1}^{n} z_k \cdot c_{ik} \right)^{\gamma} \right)
\end{equation}

\text{subject to:}
\begin{align*}
(4 - 2) \quad & \sum_{j=1}^{m} z_{ij} = 1 \quad \forall i \\
(4 - 3) \quad & \sum_{j=1}^{m} x_{ij} - 1 \geq u_j \quad \forall j \\
(4 - 4) \quad & \frac{1}{2} \left( \sum_{j=1}^{m} x_{ij} - 1 \right) \leq u_j \quad \forall j \\
(4 - 5) \quad & \sum_{j=1}^{m} g_{ik} \cdot x_{ij} \geq y_{ik} \quad \forall j, k \\
(4 - 6) \quad & 2z_i \geq x_{ij} + u_j \quad \forall i, j \\
(4 - 7) \quad & z_i \leq x_{ij} \quad \forall i, j \\
(4 - 8) \quad & z_i \leq u_j \quad \forall i, j \\
(4 - 9) \quad & z_i \geq 0 \quad \forall i \\
(4 - 10) \quad & y_{ik} \geq 0 \quad \forall j, k \\
(4 - 11) \quad & x_{ij} \in [0, 1] \quad \forall i, j \\
(4 - 12) \quad & u_j \in [0, 1] \quad \forall j
\end{align*}

\textbf{Equation 4: Optimization model (INLP)}

Constraint (4-2) ensures that every library becomes member of exactly one procurement entity (stand-alone or member of max. one consortium). (4-3) and (4-4) ensure the correct use of \( u_j \) while the same for \( z_j \) takes place in conditions (4-6) to (4-9). \( x_{ij} \) and \( u_j \) have to be declared as binary variables, while the other variables will implicitly accept only values of 0 or 1, caused by the given constraints and the impact of the minimization objective.

\textbf{3.2.2 Model Application}

The presented model in the following is applied to a small real problem to answer the question of how the twelve libraries, which are at present organized in the Hessian journal procurement consortium HeBIS, should be linked to minimize the procurement costs for the 514 scientific journals that are offered by a large international publisher in the market for academic content.

To acquire the necessary input parameters (\( \alpha, \theta, \beta, c_{ads}^{\text{adk}}, c_{ads}^{\text{vk}}, c_{ads}^{\text{gk}}, h_{ik}, \text{ and } d_{ik} \)) several empirical studies were conducted and acquired most of the necessary data for the year 2002. The participating libraries and their stand-alone administrative costs are given in Table 4. The empirical work discovered the relevant stocks of subscriptions \( h_{ik} \) as well as the number of electronic calls \( d_{ik} \) of every journal in every library.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
Size & Library & Stand-alone administrative costs per library, year, and contract, \( c_{ads}^{\text{adk}} \) \\
\hline
large & StUB Frankfurt/Main & 3750 \\
& UB Mainz & \\
& UB Gießen & \\
& UB Kassel & \\
& UB Marburg & \\
medium-size & LB Wiesbaden & 2500 \\
& HLB Fulda & \\
& LHB Darmstadt & \\
small & FHB Frankfurt/Main & 1250 \\
& FHB Wiesbaden & \\
& FHB Darmstadt & \\
& FHB Gießen & \\
\hline
\end{tabular}
\caption{Libraries which were included in the investigation and their stand-alone administrative costs}
\end{table}

The following pricing models are offered by the publisher: based on existing print subscriptions the electronic access will be granted for a surcharge of 10\% (\( \alpha = 0.1 \)) on top of the print

\footnote{The German abbreviations mean: StUB = City and University Library; UB = University Library; LB = State Library, LHB = State University Library; FHB = Academy (i.e. UAS) Library}
subscription fees to consortia and for 5% ($\theta=0.05$) to stand-alone libraries. Second, journals can be accessed by ppv; for every full article access a fee of $c^v=10 \, \text{€}$ has to be paid.

The only parameter which could not be gained empirically is $\beta$; therefore, we vary its value between 1.05 and 1.5, which again results in varying consortia administrative costs $c_i^{adk}$ (also depending on the library size).

The problem we focused on results in a non-linear binary program (Equation 4) with 6,336 variables (156 binaries) and 12,984 constraints. The 12 libraries can be organized in consortia in about 4,213,597 different constellations. The efficient solution method for that problem size was a restricted enumeration algorithm, which computed the optimal solution within eight hours on an Intel Dual Xeon 2.8 GHz machine.

The calculations for all $\beta$ between 1.05 and 1.5 resulted in the same solution: all 12 libraries should be organized in one large consortium to minimize the procurement costs for electronic journals published by the publisher in question, given the demand counts and the subscriptions of 2002.

![Figure 7: Average administrative cost coefficients for different library sizes and different $\beta$](image)

Figure 7: Average administrative cost coefficients for different library sizes and different $\beta$

Figure 8 shows the resulting administrative costs and total costs of the optimal solution for different $\beta$. Increasing $\beta$-values correlate with decreasing administrative costs. The reason therefore is that for only 12 libraries the administrative costs are always lower within a consortium because of the rather high $size^v_s$ values. Furthermore the calculations showed that it would be always efficient to realize the cross access by increased subscriptions costs instead of pay per view fees.

![Figure 8: Administrative and total costs in optimum (for different $\beta$)](image)

Figure 8: Administrative and total costs in optimum (for different $\beta$)

When varying the ppv fee ($c^v$) instead of $\beta$ it can be seen that even for halved $c^v$ of $5 \, \text{€}$ the solution of only one large consortium stays stable and leads to minimal total costs.
For the viewed real situation we can deduce that the consortia structure – as it is today – achieves minimal total costs (sum of procurement and administrative costs). Of course, this result can not simply be transferred to another situation. In each case the existing print subscriptions have to be taken into account. So even for the journal portfolio of another publisher the picture can change. If taking more libraries into account, e.g. when thinking about bundling bargaining power by a – currently discussed – Germany-wide consortium, the super-proportionally increasing administrative costs as well as the higher number of parallel printed journal subscriptions which increase the cost base of additional access costs will lead to a more fragmented consortia structure.

3.3 Simultaneous Optimization of Portfolios and Consortia Structures

To develop a simultaneous optimization model, which determines the optimal procurement strategy as well as the efficient consortia structure from a total-cost-minimizing view, just the corresponding parameters, which describe the libraries’ demand of journals \( (h_{ik}, g_{ik}) \), have to be endogenized. Because for using electronic journals only one physical copy of the journal has to be held, \( h_{ik} \) will not become larger than 1. Further on, it is dominant to hold every wanted journal at most once in the whole consortium. So, \( y_{jk} \) = 1 represents that consortium \( j \) subscribes to one printed copy of journal \( k \). Because the model minimizes the total costs it is irrelevant which particular library holds the journal. In reality the paper-based journal often is even not delivered by the publisher, because the library has only storage expenses although nobody uses it. For that reason the model does not assign the journal to a specific library within the consortium. Thus, in this model we can abandon \( g_{ik} \) and \( h_{ik} \). The delivery costs (first sum term in the objective) now are based on the index \( j \) (i.e. the procurement entities) and the following model formulation results:

\[
\begin{align*}
(5-1) \text{Min } & \sum_{j \in J} \sum_{i \in I} \left( (a - \theta \cdot (1 - z_i)) \cdot y_{jk} \cdot c_{ik} + \sum_{i \in I} \left( c_i \cdot d_{ik} (1 - y_{jk}) \right) \right) + \sum_{j \in J} \left( \sum_{i \in I} (1 - u_{ij}) \cdot c_{ik} + z_j \left( \sum_{i \in I} x_{ij} \cdot c_{ik} \right)^{\alpha} \right) \\
\mbox{u.d.N.:} & \sum_{j \in J} x_{ij} = 1 & \forall i & \quad (5-2) \quad \sum_{j \in J} x_{ij} - u_{ij} \geq 0 & \forall j & \quad (5-3) \quad \sum_{j \in J} x_{ij} \leq u_{ij} & \forall j & \quad (5-4) \quad \frac{1}{\theta} \left( \sum_{j \in J} x_{ij} - 1 \right) \leq u_{ij} & \forall j & \quad (5-5)
\end{align*}
\]

Equation 5: Simultaneous Optimization Model

In fact, the optimal solution can only be realized in long term because of the contractual relationships between libraries and publishers, which usually ensure, that libraries can not reduce their stock of printed journals even within a few years.

An application of this model to our empirical data will be the next step of our work. Due to the increased complexity, which is caused by the explicitly binary declaration of \( y_{jk} \), first of all appropriate solving mechanisms have to be found and to be evaluated.

4. Conclusion and Further Research

The presented paper gave an approach to a solution for optimizing the expenses which libraries spend for procuring electronic journals.

Based on empirical research it could be shown, that libraries can cut their procurement costs by canceling subscriptions and switching to ppv. For the second focus – optimizing the consortia structure – a binary non-linear optimization model was developed and applied to a small real example, taken from our empirical research.
In the next steps of our work, based on the presented optimization model we will develop a branch & bound algorithm, which is able to handle larger instances of the consortia structure problem, presented in 3.2. Besides that, the simultaneous model (Equation 5) will also be applied to the chosen example.

To provide the data, which is necessary to optimize the procurement type and the consortia structure, within our research project we began to establish a database, which prepares and pre-processes the monthly access counts \((d_{ik})\) for each library and journal. To import this data, converters have to be established to the publishers’ systems, because in most cases the electronic journals are hosted on their servers. Once this database is established, the necessary demand parameters for the optimization could be extracted as well as usage trend estimations could be made. Using an integrated database and optimization tool, an important contribution to improving the public supply of academic content can be provided.

5. Acknowledgment
This research is supported by grants from the German National Science Foundation (DFG) (Project “Perspectives for the Purchase of Electronic Information Resources in Germany”, Grant No. 554 922 (2)) and PhD Program 492 “Enabling Technologies for Electronic Commerce”). We gratefully acknowledge the financial support.

6. References


