UNIVERSITY OF HOHENHEIM

FACULTY OF BUSINESS, ECONOMICS AND SOCIAL SCIENCES

HOHENHEIM DISCUSSION PAPERS

IN BUSINESS, ECONOMICS AND SOCIAL SCIENCES

Institute of Interorganizational Management & Performance

DISCUSSION PAPER **28**-2017

CLUSTERING SURGICAL PROCEDURES FOR MASTER SURGICAL SCHEDULING

Alexander Kressner

University of Hohenheim

Katja Schimmelpfeng

University of Hohenheim

State: September 2017

www.wiso.uni-hohenheim.de

Discussion Paper 28-2017

Clustering Surgical Procedures for Master Surgical Scheduling

Alexander Kressner, Katja Schimmelpfeng

ISSN 2364-2084

Clustering Surgical Procedures for Master Surgical Scheduling

Alexander Kressner · **Katja Schimmelpfeng**

Abstract The sound management of operating rooms is a very important task in each hospital. To use this crucial resource efficiently, cyclic master surgery schedules are often developed. To derive sensible schedules, high-quality input data are necessary. In this paper, we focus on the (elective) surgical procedures' stochastic durations to determine reasonable, cyclically scheduled surgical clusters. Therefore, we adapt the approach of van Oostrum et al (2008), which was specifically designed for clustering surgical procedures for master surgical scheduling, and present a two-stage solution approach that consists of a new construction heuristic and an improvement heuristic. We conducted a numerical study based on real-world data from a German hospital. The results reveal clusters with considerably reduced variability compared to those of van Oostrum et al (2008).

Keywords master surgery scheduling (MSS) · stochastic surgery duration · surgery types · clustering

A. Kressner, K. Schimmelpfeng
Chair of Procurement and Production, University of Hohenheim, 70599 Stuttgart, Germany
Tel.: +49 711 459-23358
Fax: +49 711 459-23232
E-mail: {Alexander.Kressner|Katja.Schimmelpfeng}@uni-hohenheim.de

1 Introduction and Problem Description

Over recent decades, the demand for health care services in industrialized countries has been constantly rising (OECD, 2011). Simultaneously, most countries limit public health spending. Therefore, hospitals face the challenge of using scarce resources even more efficiently. One of these resources is the operating theater, which generates the largest part of the cost and revenues in a hospital (Cardoen et al, 2010). To manage its operations and processes successfully, adequate planning and scheduling approaches are crucial. Generally, planning and scheduling tasks in the context of the operating theater belong to a specific level of the decision hierarchy: the strategic, tactical or operational level (Guerriero and Guido, 2011; Hans et al, 2012). At the strategic level, a hospital determines the capacity dimensions, such as the number of operating rooms (OR) or the technical equipment that each OR contains. Allocating available OR capacities to specialties or surgery types belongs to the tactical level, whereas the operational level addresses short-term scheduling and the rescheduling of patients.

Among others, van Oostrum et al (2008) proposed a so-called cyclic master surgery scheduling approach for tactical planning tasks that can be used in hospitals with a stable volume of elective surgical procedures during consecutive weeks. The idea is to aggregate surgical procedures to some reasonable surgery types and to determine the number of slots allocated to each type for any OR and day within one cycle. After a fixed cycle length of typically one or two weeks, the schedule is repeated until a new (cyclic) schedule seems to be necessary. Figure 1 shows an example of such a master surgery schedule for the working days of Monday to Friday, using three operation rooms. Using a master surgery schedule (MSS)

- lowers the managerial burden of developing new schedules every week,
- makes is possible to coordinate technical and personnel resources early and
- guides patient scheduling such that hospitals use their ORs efficiently (van Oostrum et al, 2010).

MSS objectives cover maximizing utilization, minimizing cost, controlling overtime or leveling workloads in the ORs and subsequent departments, for example, the intensive care units (ICU) or wards (Beliën and Demeulemeester, 2007; Fügener et al, 2014; van Oostrum et al, 2008).

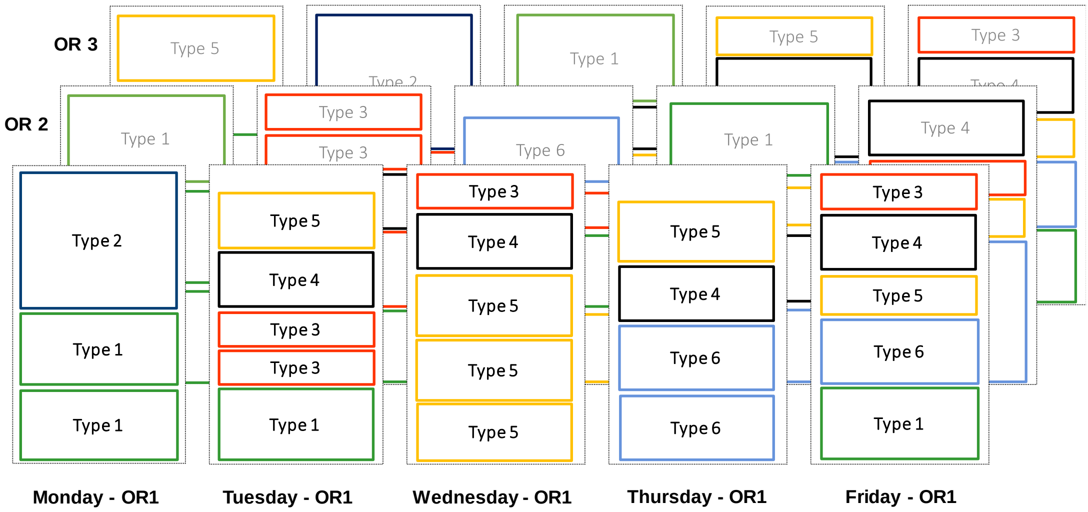

Fig. 1: Example of a master surgery schedule for three operation rooms operating from Monday to Friday

Cyclically scheduled surgery types represent the building block of a MSS. Taking the German situation as an example, a surgery type consists of the surgical procedures defined by the so-called German "Operationen- und Prozedurenschlüssel" (OPS) – a modification of the International Classification of Procedures in Medicine (ICPM). The OPS defines surgeries at the lowest level of aggregation. In its current version, it consists of 28,800 different codes (DMDI, 2014). Grouping surgical procedures to construct logistically homogenous surgery types is a very challenging planning task. Logistical homogeneity is measurable based on attributes such as surgery duration, length of stay, staff requirements and use of medical devices (van Oostrum et al, 2011). For master surgery scheduling, constructing surgery types with little variability in the surgery duration is very important: poorly grouped surgical procedures exhibit high variability, which will make building master schedules with high OR utilization and little overtime on this basis almost impossible. From this perspective, a very granular grouping with a large number of surgery types appears to be beneficial. However, we must also account for the cyclic nature of a MSS and ensure that integer numbers of surgery slots are scheduled to obtain a valid solution, which calls for a substantial degree of aggregation.

Generally, when constructing a new schedule, forecasting the number of cases per surgery type divided by the number of cycle repetitions yields the corresponding cyclic number of slots. In most cases, the resulting number is not an integer. One intuitive method of addressing the problem may be rounding up to the next integer value. Certainly, doing so would heavily overestimate the capacity demand for the operating theater, and it may even be impossible to find feasible plans. However, rounding down to the next integer value may underestimate the resource requirements. Somewhat infrequent surgery types that have a cyclic number of slots smaller than one may not even occur in the schedule. Instead of relying on (arbitrary) rounding, a so-called dummy surgery type can be introduced and used to pool all of the fractional parts of the surgery types' cyclic number of slots, as proposed by van Oostrum et al (2011). Finally, rounding up this (pooled) number to the next integer results in an almost perfectly matched demand. Using this specific concept, it is mandatory to consider the composition and the resulting variability in the dummy type compared to the regular surgery types.

The task of grouping surgical procedures is basically a clustering problem that is not exclusively relevant in master surgery scheduling. In the healthcare literature, several papers that address operating theater planning consider aggregated surgical procedures (Adan and Vissers, 2002; Santibáñez et al, 2007; Adan et al, 2009; Ma and Demeulemeester, 2013). Nevertheless, despite its importance and the diversity of well-elaborated clustering algorithms (for an overview see, Xu and Wunsch (2005)), tailored approaches to generating appropriate surgery types/clusters for strategic and tactical operating theater planning in the literature are rare (Dilts et al, 1995). Furthermore, except for the work by van Oostrum et al (2011), we are not aware of any approach to surgery type clustering in a cyclic planning environment.

In this paper, we show how operations research techniques can be applied to solve a specific clustering problem for a cyclic planning problem, namely, master surgery scheduling. Similar to the work by van Oostrum et al (2011) our approach is based on the concept of a dummy surgery type and aims to minimize the variability in the clusters regarding surgery durations. In addition, we make the following contributions: we present a two-stage clustering algorithm with a new constructive heuristic compared to van Oostrum et al (2011) and a heuristic that improves initial partitions. The specific feature of the latter is the use of a non-linear optimization model integrated in an algorithmic framework. To solve the model with a commercial MIP-solver, we present a linear reformulation and introduce some intuitive simple inequalities to accelerate computation times. In a numerical study with real-world data from a German hospital, we show that our clustering algorithm is able to find partitions with considerably reduced variability.

The remainder of this paper is structured as follows: we dedicate Section 2 to a brief overview of the relevant literature. In Section 3, we illustrate the clustering problem and a corresponding mathematical optimization model. In Section 4, we present a two-stage solution approach that aims at homogenous clusters. In Section 5, our algorithm is applied to real-world data from a German hospital. Furthermore, the results of the numerical study are presented. Finally, Section 6 recapitulates the paper's most important findings and outlines some ideas for future research.

2 Related Literature

Clustering as a main data mining task refers to descriptive modeling (Meisel and Mattfeld, 2010). Its objective is to partition a given set of objects into subgroups such that the objects within a subgroup are similar to each other and separable from objects in other groups according to some similarity measure. Before defining such an adequate similarity measure, relevant object features must be selected. Discovering these relevant features primarily depends on the underlying decision problem. To ease the computational burden of any clustering algorithm and allow an intuitive comprehension of the results, only the most relevant features should be used. After having determined the relevant features, carefully defining an adequate similarity measure is crucial. In most cases, it is possible to define (dis-)similarity based on well-known distance measures, for example, Euclidian, city block or Mahalanobis distance (for a general overview, see Xu and Wunsch (2005); Jain et al (1999)). Finally, constructing a function to evaluate the partition's quality is necessary. In this sense, it seems natural to represent a clustering problem as a mathematical optimization problem. Hansen and Jaumard (1997) illustrate various optimization criteria and the formulation of clustering problems as mathematical programs. Recent review papers by Olafsson et al (2008); Meisel and Mattfeld (2010); Corne et al (2012) highlight this relationship and emphasize the synergies of the well-elaborated domains of operations research and data mining. Sağlam et al (2006); Inniss (2006); Romanowski et al (2006); Kulkarni and Fathi (2007) present examples that apply operations research techniques to clustering problems.

To solve a clustering problem, two types of algorithms are available: hierarchical and partitional algorithms (Jain et al, 1999). The latter start with an initial partition of objects, choosing the number of clusters in advance. Subsequently, objects are assigned to clusters to optimize a given objective function. Most likely, the best-known partitional clustering method is *k-means* (MacQueen, 1967). It begins by randomly picking *k* cluster centers and then assigns each object to the closest center. Then, the cluster centers are recomputed. The algorithm iterates until no more changes in the cluster centers occur. Variants of the basic *k-means* algorithm attempt to find good initial partitions to accelerate convergence or to allow a dynamic number of clusters by splitting and merging procedures (Jain et al, 1999).
Hierarchical algorithms generate a series of partitions organized in a hierarchical manner. With agglomerative and divisive methods, two variants of hierarchical algorithms exist. The former starts with a partition where each object forms an individual cluster. Given some distance matrix, the two clusters closest to each other are merged. This process is repeated until all objects lie within one cluster. Divisive methods work in the opposite direction. Initially, all objects belong to a single cluster. In the next iterations, the algorithm successively divides partitions until each object forms its own cluster. Due to the computational complexity of divisive hierarchical algorithms, it is common to use agglomerative methods (Xu and Wunsch, 2005). One popular approach in this domain is *Ward's method*, which uses the sum of error squares to evaluate different partitions. In each step of the algorithm, cluster pairs that lead to the objective function's minimal increase are merged. Finally, the decision maker can appropriately choose out of the derived partitions (Ward, 1963).

As described in Section 1, papers that address strategic and tactical planning in the operating theater typically only assume that surgery types with a low variability of resource consumption exist. However, reviewing the literature related to healthcare management, we only identified the approach of van Oostrum et al (2011) that groups surgical procedures for a specific OR planning task. As in our case, the authors perform clustering with the goal of allowing master surgery scheduling. Typically, authors consider the features "surgery durations" and "lengths of stay" when constructing surgery types. Conceptually, the employed clustering algorithm is a variant of Ward's method that uses a modified distance matrix (van Oostrum et al, 2011). The distance between a cluster pair is computed in three steps:

1. First, van Oostrum et al (2011) compute the sum of squared errors regarding the surgery duration and the length of stay in each cluster.
2. Second, they determine the number of dummy surgeries associated with each cluster. By using a scalarization function, squared error sums and dummy surgeries are aggregated per cluster.
3. Finally, summing up over all clusters allows to evaluate the partition's quality. The entries of the distance matrix represent the change in the objective function for each possible merger of two clusters in some iteration of the algorithm.

In their case study, van Oostrum et al (2011) show the influence of different parametrizations of the scalarizing function on the partitioning of the data set.

3 Detailed Problem Description and Model Formulation for the Clustering Problem

3.1 Definition of an Appropriate Evaluation and Objective Function

The main goal in master surgery scheduling is to ensure a high utilization of ORs without having excessive overtime. Because surgery durations exhibit a distinct natural variability, planning approaches that anticipate this uncertainty are very well suited. However, defining surgery types with little variability in surgery durations is a prerequisite to obtain good-quality planning results: the higher the surgery durations' variability is, the more additional slack capacity in the ORs is necessary to buffer against overtime. Consequently, this slack has a negative effect on the OR utilization. Thus, given a historical record of surgical procedures with corresponding realizations of the random duration of individual surgeries, we strive to find a partition of procedures that minimizes the overall sum of squared errors. Such a partition defines the surgery types (clusters of procedures) used in MSS. Figure 2 shows the hierarchical relationship between individual surgeries i, surgical procedures p and the surgery types c we are aiming for.

In our approach, we account for the cyclic nature of MSS and build on the concept of a dummy surgery type adjacent to the regular surgery types. However, unlike van Oostrum et al (2011), we precisely evaluate not only the total number of dummy surgeries but also the sum of

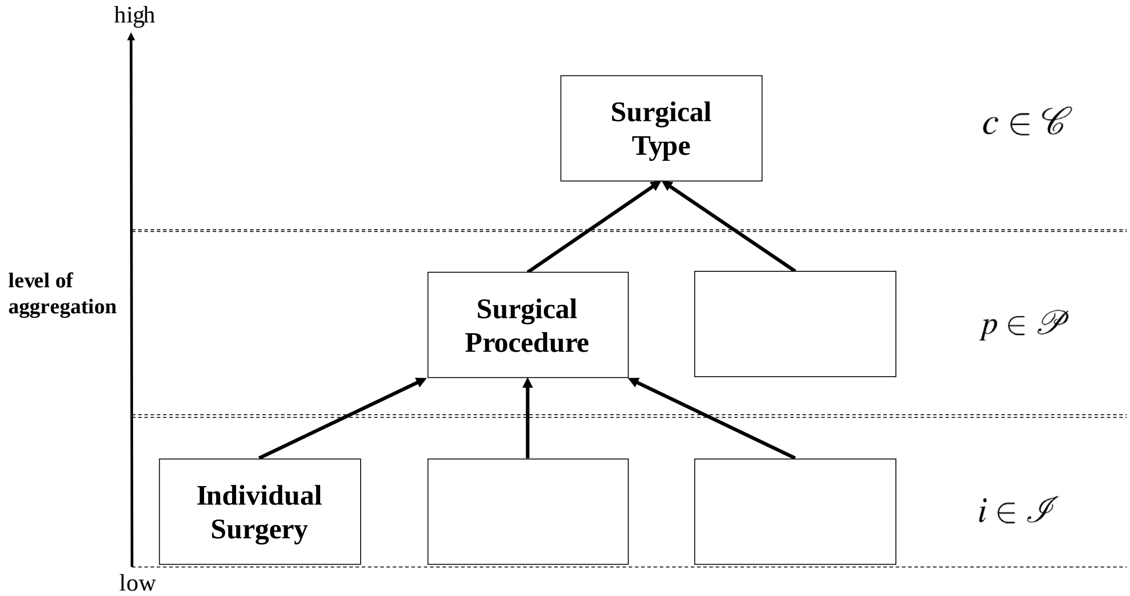

Fig. 2: Hierarchical relationship between i, p and c

squared errors. Hence, we avoid using an arbitrarily chosen scalarization function to summarize two distinct variables, i.e., the squared error sum of surgery durations and the number of dummy surgeries.

To quantify the loss of information resulting from clustering surgical procedures to surgical types, we need a function that evaluates the sum of squared errors over all clusters, including the dummy cluster for any grouping of surgical procedures. In the following, we derive such an evaluation function step by step, using the subsequent notation summarized in Table 1 in alphabetical order.

Therefore, we denote the number of MSS cycle repetitions by r. Let us assume that we have a sample $\mathscr{I}_p$ of individual (recorded) surgeries $i \in \mathscr{I}$ each associated with a surgical procedure $p \in \mathscr{P}$. In addition, let a_{pi} be the recorded duration of an individual surgery i associated with procedure p. The parameter n_p denotes the (forecasted) number of surgeries of procedure p over the planning horizon of typically one year (for example, see van Oostrum et al (2008)). $\mathscr{Z}_c$ defines a cluster of surgical procedures p associated with surgery type $c \in \mathscr{C}$.

Sets and indices	
$c \in \mathscr{C}$	set of surgery types c
$i \in \mathscr{I}$	set of individual (recorded) incidents/surgeries i
$p \in \mathscr{P}$	set of recorded surgical procedures p, defined according to the OPS
$\mathscr{I}_p$	sample of individual surgeries $i \in \mathscr{I}$ each associated with a surgical procedure $p \in \mathscr{P}$
$\mathscr{Z}_c$	set of surgical procedures p associated with surgery type $c \in \mathscr{C}$
Parameters	
$\hat{a}_D$	average surgery duration in the dummy cluster
$\bar{a}_c$	average surgery durations in the regular clusters
a_{pi}	recorded surgery duration of surgery i of procedure p
$ESS_{c,D}$	squared error for each cluster c regarding the average dummy surgery duration
ESS_D	squared error sum in dummy cluster D
ESS_c	expected squared error for each regular cluster c
ESS_T	total squared errors over all clusters
n_p	(forecasted) number of surgeries of procedure p over the planning horizon
r	number of MSS cycle repetitions
V_c	number of surgeries of type $c \in \mathscr{C}$ moved into the dummy cluster

Table 1: Notation in Section 3.1

- First, we calculate the squared error sum in the dummy cluster ESS_D. Therefore, we determine the number of surgeries from each regular cluster V_c moving into the dummy cluster:

$$V_c = \sum_{p \in \mathscr{Z}_c} n_p - \left\lfloor \frac{\sum\limits_{p \in \mathscr{Z}_c} n_p}{r} \right\rfloor r, \qquad c \in \mathscr{C} \quad (1)$$

  The first term represents the volume of cases in cluster c over the complete planning horizon, whereas the second term yields the corresponding number of cases if $\left\lfloor \frac{\sum_{p \in \mathscr{Z}_c} n_p}{r} \right\rfloor$ slots of surgery type c are scheduled in each cycle. The remaining difference reveals the number of dummy surgeries originating from surgery type c.

- To assess ESS_D, it is necessary to calculate the average surgery duration in the dummy cluster $\hat{a}_D$ (which is also denoted as the cluster centroid). We use a weighted sum of the average surgery durations from the regular clusters $\bar{a}_c$:

$$\hat{a}_D = \frac{\sum\limits_{c \in \mathscr{C}} V_c \bar{a}_c}{\sum\limits_{c \in \mathscr{C}} V_c} \quad (2)$$

– Next, we compute the sum of squared errors for each cluster regarding the average dummy surgery duration $ESS_{c,D}$:

$$ESS_{c,D} = \sum_{p \in \mathscr{Z}_c} \sum_{i \in \mathscr{I}_p} (a_{pi} - \hat{a}_D)^2, \qquad c \in \mathscr{C} \quad (3)$$

– Naturally, only a certain fraction of that variability can be attributed to the dummy cluster. Therefore, we scale $ESS_{c,D}$ according to the number of dummy surgeries and regular surgeries in each cluster and compute ESS_D:

$$ESS_D = \sum_{c \in \mathscr{C}} \left(ESS_{c,D} \cdot \frac{V_c}{\sum\limits_{p \in \mathscr{Z}_c} n_p} \right) \quad (4)$$

– The computation of the squared error sum ESS_c for each regular cluster c must consider that V_c of the $\sum_{p \in \mathscr{Z}_p} n_p$ surgeries move in the dummy cluster. Consequently, the original variability in each cluster can only be considered proportionally to the number of surgeries remaining in the regular cluster:

$$ESS_c = \sum_{p \in \mathscr{Z}_c} \sum_{i \in \mathscr{I}_p} (a_{pi} - \bar{a}_c)^2 \left(1 - \frac{V_c}{\sum\limits_{p \in \mathscr{Z}_c} n_p} \right), \qquad c \in \mathscr{C} \quad (5)$$

– Finally, we can aggregate the sum of squared errors over all clusters and obtain the total squared error sum ESS_T:

$$ESS_T = \sum_{c \in \mathscr{C}} ESS_c + ESS_D, \quad (6)$$

In the following sections, we present the assumptions of our model and a mathematical model that groups surgical procedures to clusters minimize the evaluation function value ESS_T.

3.2 Assumptions

First, we cluster only within a specialty, mainly for organizational reasons, because sharing slots for surgeries among specialties is very conflicting. Second, we assume the number of clusters and average surgery durations in each cluster to be known a priori. In doing so, we face a reduced

Indices and index sets:	
$p, p' \in \mathscr{P}$	surgical procedures according to OPS
$c \in \mathscr{C}$	regular surgery types/ clusters
$i \in \mathscr{I}$	individual surgeries
$\mathscr{I}_p$	subset of surgeries assigned to procedure p
Parameters:	
a_{pi}	recorded duration of indvidual surgery i belonging to procedure p
n_p	forecasted number of surgeries of procedure p
$\bar{a}_c$	average surgery duration of surgery type/ cluster c
$\hat{a}_D$	average surgery duration of the dummy-surgery type/ dummy-cluster
r	number of MSS cycle repetitions
Decision variables:	
X_{pc}	$= \begin{cases} 1, & \text{if procedure } p \text{ is assigned to surgery type/ cluster } c \\ 0, & \text{else} \end{cases}$
$V_c \geq 0$	number of dummy surgeries originating from surgery type/ cluster c
$X_c^{Int} \in \mathbb{N}_0$	integer number of slots of surgery type/ cluster c in one MSS-cycle

Table 2: Notation for the mathematical model

complexity of the optimization model, and its solution becomes tractable. However, this is a simplification because we cannot compute the optimal number of clusters in advance. Additionally, even if we could somehow identify the optimal cluster number, the average surgery durations in each cluster would still depend on the grouping of surgical procedures. Thus, starting from predefined cluster centroids, we cannot guarantee finding the optimal solution. We discuss how to address these problems in Section 4, where we present our solution approach.

3.3 Notation and Mathematical Model

When constructing surgery types, it must be ensured that each surgical procedure is assigned exclusively to one cluster. The objective is to minimize the sum of squared errors over the dummy and all regular clusters. Hence, we obtain the following mixed integer non-linear optimization program, using the notation given in Table 2.

Model NLCM

$$\text{Min } ESS_T = \sum_{c\in\mathscr{C}} \sum_{p\in\mathscr{P}} \sum_{i\in\mathscr{I}_p} \left[(a_{pi} - \bar{a}_c)^2 \left(1 - \frac{V_c}{\sum_{p'} n_{p'} X_{p'c}}\right) X_{pc} \right] + \left[(a_{pi} - \hat{a}_D)^2 \frac{V_c X_{pc}}{\sum_{p'} n_{p'} X_{p'c}} \right] \tag{7}$$

$$\sum_{c \in \mathscr{C}} X_{pc} = 1, \qquad p \in \mathscr{P} \quad (8)$$

$$V_c = \sum_{p \in \mathscr{P}} n_p X_{pc} - r X_c^{Int}, \qquad c \in \mathscr{C} \quad (9)$$

$$X_c^{Int} > \frac{\sum_{p \in \mathscr{P}} n_p X_{pc}}{r} - 1, \qquad c \in \mathscr{C} \quad (10)$$

In the objective function (7), the first term considers the sum of squared errors over all clusters except the dummy cluster. In case procedure p is assigned to cluster c, i.e., decision variable X_{pc} equals one, the corresponding squared error sum is taken into account according to the portion of surgeries $(1 - \frac{V_c}{\sum_{p'} n_{p'} X_{p'c}})$ in that cluster. We consider the dummy cluster's variability in the second term. Again, we compute the sum of squared errors with respect to the cluster centroid for each procedure. In case either X_{pc} or V_c is zero, i.e., procedure p is not assigned to cluster c or there are no dummy surgeries from cluster c, there is no contribution to the overall variability. In all remaining cases, the squared error sum associated with procedure p is scaled proportionally to the dummy surgeries originating from cluster c. Constraints (8) ensure that any procedure is assigned to exactly one of the pre-specified clusters. Constraints (9) and (10) serve to compute the number of dummy surgeries from cluster c. Constraints (10) reveal the maximum integer number of slots per surgery type scheduled in the MSS. Hence, they basically model the supposed rounding procedure and, in combination with (9), derive the number of dummy surgeries attributed to cluster c. We omit restrictions on the decision variables' domains, given that they are provided in Table 2. Finally, it is worth noting that the model allows a flexible number of active clusters, i.e., not all of the $|C|$ clusters must be used. However, generating additional clusters is not possible.

4 Two-Stage Solution Approach

The model presented in the previous section is non-linear and assumes a predefined number of clusters with corresponding centroids. To find good partitions of surgical procedures, we apply a two-stage solution approach. The goal of the first stage is to construct promising initial clusters and to initialize the optimization model. For this purpose, we employ an adjusted version of Ward's method (Ward, 1963). In the second stage, we use an improvement heuristic based on our optimization model to reassign surgical procedures to clusters to decrease the objective function value of the initial partitions.

4.1 Stage 1: Constructing an Initial Solution

Promising initial solutions are generated by a constructive heuristic that is a modified version of the agglomerative hierarchical clustering algorithm of Ward (1963) and closely related to van Oostrum et al (2011). The main steps of the procedure are highlighted in algorithm 1.

Algorithm 1: Constructive Heuristic

main input : $\mathscr{P}$, $\mathscr{I}_p$, a_{pi}, n_p, r

main output: partition of surgical procedures ($Z_c^{j^*}$) of iteration $j^* = min_j\{ESS_T^j\}$

1 **begin**
2 $\quad j = |\mathscr{P}|$, $\mathscr{C} = \mathscr{P}$, $Z_c^j = \{c\}\ \forall c$;
3 $\quad$ Calculate ESS_T^j;
4 $\quad j = j - 1$;
5 $\quad$ **while** $j \geq 1$ **do**
6 $\quad\quad$ Calculate $\Delta ESS_T^j(c,c')\ \forall c, c' > c$;
7 $\quad\quad (c^*, c'^*) = min_{(c,c')}\{\Delta ESS_T^j(c,c')\}$;
8 $\quad\quad Z_c^j = Z_c^{j+1}\ \forall c \neq c^*, c'^*$;
9 $\quad\quad Z_{c^*}^j = Z_{c^*}^{j+1} \cup Z_{c'^*}^{j+1}$;
10 $\quad\quad ESS_T^j = ESS_T^{j+1} + \Delta ESS_T^j(c^*, c'^*)$;
11 $\quad\quad \mathscr{C} = \mathscr{C} \setminus \{c'^*\}$;
12 $\quad\quad j = j - 1$;
13 $\quad$ **end**
14 **end**

Please note that we employ the heuristic for each specialty. Running the algorithm results in a series of $|\mathscr{P}|$ different partitions, each indicated with index j. At the beginning of the algorithm, the number of clusters equals the number of surgical procedures, and each procedure constitutes its own cluster (line 2). An evaluation of this first partition is performed in line 3. The following while-loop returns a new partition by merging exactly two clusters in each iteration (lines 5-13). To find two promising candidates for each two clusters c and c' that can possibly be merged, the overall change in the objective function denoted by $\Delta ESS_T^j(c,c')$ is determined (line 6). In computational terms, this part is the most expensive part of the algorithm: for each possible merge, the centroids and the sum of squared errors in the newly built cluster and the dummy cluster must be calculated. Finally, the pair with the minimal increase in the sum of squared errors forms the new cluster (lines 7-9). After having evaluated this new partition (line 10), the set of clusters is redefined (line 11). The main output of the algorithm is a partition of surgical procedures from which relevant parameters, for example, the number of clusters or cluster centroids used in the improvement heuristic, can be derived.

4.2 Stage 2: Improving the Initial Solution

The constructive heuristic presented in the previous section iteratively changes the number of clusters and assigns surgical procedures to clusters. A major drawback of such a procedure is the fact that assignments performed in earlier iterations are fixed and cannot be resolved later. Thus, starting from an initial solution of algorithm 1, it is advisable to rearrange objects to further decrease the overall variability. In this sense, a popular approach is the classical *k-means algorithm*, which allocates an object to the most appropriate cluster according to some similarity measure (see, e.g., Dilts et al (1995)). This allocation is particularly easy when the assignment decision for each object can be made independent of all others. In our case, due to the one dummy cluster concept, this decision is not possible. Reassignments of surgical procedures alter the clusters' size and thus the number of dummy surgeries originating from the clusters. Consequently, the composition of the dummy cluster and its associated squared error sum changes. Hence, to optimally rearrange surgical procedures, we apply a linear reformulation of the mathematical model of Section 3.3 that also considers the effect of assignments on the variability in the dummy cluster. The relevant model inputs are provided by the constructive heuristic. Furthermore, we embed the optimization model in an algorithmic procedure closely related to k-means, which successively improves the previous partitions.

Indices and index sets:

$\mathscr{H}$	$=\{0,1,\ldots,r-1\}$, equal to the domain of V_c

Parameters:

K,M	big number

Decision variables:

δ_c	non-negative variables, reciprocal of the number of surgeries in cluster c
λ_{ch}	auxilliary binary variable
θ_{pc}	reciprocal of the number of surgeries in the corresponding cluster $= \begin{cases} 1, & \text{if surgical procedure p is assigned to the corresponding cluster} \\ 0, & \text{else} \end{cases}$
$\hat{\theta}_{pch}$	non-negative variables, share of dummy surgeries with respect to the total number of surgeries in cluster c if procedure p is assigned to that cluster

Table 3: Additional notation for the mathematical model

4.2.1 Linearization of the Base Model

The objective function (7) of our original mathematical model is non-linear. To obtain a linear MIP, the expression $\frac{V_c}{\sum_{p'} n_{p'} X_{p'c}} X_{pc}$ must be linearized. Therefore, we perform the following four steps, using the additional notation given in Table 3.

- First, we address the term $\frac{X_{pc}}{\sum_p n_p X_{pc}}$. According to an idea of Li (1994), we introduce non-negative variables δ_c, which are defined as the reciprocal of the number of surgeries in cluster c:

$$\delta_c = \frac{1}{\sum\limits_{p \in \mathscr{P}} n_p X_{pc}} \tag{11}$$

  Adding the constraints (12), we guarantee that the new variables take the appropriate values:

$$\sum_{p \in \mathscr{P}} n_p X_{pc} \delta_c = 1, \qquad c \in \mathscr{C} \tag{12}$$

- Clearly, this procedure does not dissolve the non-linearity of the formulation, given that we end up with products of the form $X_{pc}\delta_c$. However, it is now possible to apply the approach by Wu (1997) that makes it possible to linearize the product of two variables. Again, we define new non-negative variables $\theta_{pc} = X_{pc}\delta_c$. To adequately model this equality, we introduce a set of linear constraints:

$$\sum_{p \in \mathscr{P}} n_p \theta_{pc} = 1, \qquad c \in \mathscr{C} \tag{13}$$

$$\delta_c - \theta_{pc} \leq K(1 - X_{pc}), \qquad p \in \mathscr{P}, c \in \mathscr{C} \tag{14}$$

$$\theta_{pc} \leq \delta_c, \qquad p \in \mathscr{P}, c \in \mathscr{C} \tag{15}$$

$$\theta_{pc} \leq K X_{pc}, \qquad p \in \mathscr{P}, c \in \mathscr{C} \tag{16}$$

$$\theta_{pc} \geq 0, \qquad p \in \mathscr{P}, c \in \mathscr{C} \tag{17}$$

$$\delta_c \geq 0, \qquad c \in \mathscr{C} \tag{18}$$

To ensure that θ_{pc} equals the reciprocal of the number of surgeries in the corresponding cluster (denoted by δ_c) only if surgical procedure p is assigned to the corresponding cluster and zero otherwise, we use a Big-M formulation. A valid upper bound for K is:

$$K = \frac{1}{\min_p\{n_p\}} \tag{19}$$

This becomes clear by the following consideration: assume $X_{pc} = 0$ and $\delta_c > 0$ for some p and c, i.e., procedure p is not assigned to the active cluster c. Then, constraint (16) forces $\theta_{pc} = 0$, and constraint (14) becomes $\delta_c \leq K$. Because δ_c must not necessarily be restricted, it must be ensured that the surgical procedure with the smallest record of surgeries can exclusively form a surgery type that gives the maximum reciprocal of the number of surgeries in a cluster.

– Applying this reformulation, we still end up with a non-linear term of the form $\theta_{pc}V_c$, i.e., products of a rational and an integer variable. Therefore, we model V_c in a third step with the help of binary variables λ_{ch} and the constraints $\sum_{h \in \mathscr{H}} h\lambda_{ch} = V_c$ and $\sum_{h \in \mathscr{H}} \lambda_{ch} = 1$ for each $c \in \mathscr{C}$. The variables λ_{ch} equal one if the number of dummy variables in cluster c equals h and zero otherwise. This relationship is ensured by the two constraints established for each cluster and setting $\mathscr{H} = \{0, 1, \ldots, r-1\}$, which represents the domain of V_c. Using this formulation, the variables λ_{ch} are defined as variables belonging to an ordered set of type 1 (see Beale and Tomlin (1970)). Additionally, please note that the number of additional binary variables remains relatively small because most instances consist of only a few clusters and the maximum number of dummy surgeries attributed to a cluster equals $r-1$, where r typically represents the number of weeks for which a MSS is valid.

– Finally, we linearize the latest reformulation of the form $\theta_{pc}h\lambda_c$ by defining non-negative variables $\hat{\theta}_{pch} = \theta_{pc}h\lambda_{ch}$, indicating the share of dummy surgeries with respect to the total number of surgeries in cluster c in case procedure p is assigned to that cluster and the following set of constraints:

$$\sum_{h\in\mathscr{H}} h\lambda_{ch} = \sum_{p\in\mathscr{P}} n_p X_{pc} - rX_c^{Int}, \qquad c\in\mathscr{C} \qquad (20)$$

$$\sum_{h\in\mathscr{H}} \lambda_{ch} = 1, \qquad c\in\mathscr{C} \qquad (21)$$

$$h\theta_{pc} - \hat{\theta}_{pch} \le M(1-\lambda_{ch}), \qquad p\in\mathscr{P}, c\in\mathscr{C}, h\in\mathscr{H} \qquad (22)$$

$$\hat{\theta}_{pch} \le h\theta_{pc}, \qquad p\in\mathscr{P}, c\in\mathscr{C}, h\in\mathscr{H} \qquad (23)$$

$$\hat{\theta}_{pch} \le \lambda_{ch}, \qquad p\in\mathscr{P}, c\in\mathscr{C}, h\in\mathscr{H} \qquad (24)$$

$$\hat{\theta}_{pch} \ge 0, \qquad p\in\mathscr{P}, c\in\mathscr{C}, h\in\mathscr{H} \qquad (25)$$

Constraints (20) and (21) store the number of dummy surgeries from each cluster in appropriate binary variables, as outlined above. (22), (23) and (24) enforce the equality $\hat{\theta}_{pch} = h\theta_{pc}$ if λ_{ch} and X_{pc} equal one. In case λ_{ch} is zero, the constraints (24) force $\hat{\theta}_{pch}$ to zero as well. Ultimately, the necessary variable definitions are given. With reasoning analogous to that in the first Big-M formulation, we set $M = (r-1)\frac{1}{\min_p\{n_p\}}$.

Putting it all together, the previous considerations result in the following MILP formulation of the non-linear base model:

Model LCM

$$\text{Min } ESS_T = \sum_{c\in\mathscr{C}} \sum_{p\in\mathscr{P}} \sum_{i\in\mathscr{I}_p} \left[(a_{pi} - \bar{a}_c)^2 \, (X_{pc} - \sum_{h\in\mathscr{H}} \hat{\theta}_{pch}) + \sum_{h\in\mathscr{H}} (a_{pi} - \hat{a}_D)^2 \, \hat{\theta}_{pch} \right] \tag{26}$$

s.t.

$$\sum_{c\in\mathscr{C}} X_{pc} = 1, \qquad p\in\mathscr{P} \tag{27}$$

$$X_c^{Int} > \frac{\sum_{p\in\mathscr{P}} n_p X_{pc}}{r} - 1, \qquad c\in\mathscr{C} \tag{28}$$

and the constraints (13)-(18) and (20)-(25). Table 4 illustrates the number of variables used in the non-linear and linearized models. Clearly, to linearize the model, we must accept a considerable number of additional continuous variables and only a few binary variables.

	non-linear model	linear model
# binary variables	$\lvert\mathscr{P}\rvert \times \lvert\mathscr{C}\rvert$	$\lvert\mathscr{P}\rvert \times \lvert\mathscr{C}\rvert + \lvert\mathscr{C}\rvert \times \lvert\mathscr{H}\rvert$
# integer variables	$\lvert\mathscr{C}\rvert$	$\lvert\mathscr{C}\rvert$
# continuous variables	$\lvert\mathscr{C}\rvert$	$\lvert\mathscr{C}\rvert + \lvert\mathscr{P}\rvert \times \lvert\mathscr{C}\rvert + \lvert\mathscr{P}\rvert \times \lvert\mathscr{C}\rvert \times \lvert\mathscr{H}\rvert$

Table 4: Number of variables in the non-linear and linearized models

4.2.2 Simple Inequalities

Implementing and testing the linearized model reveal poor LP-Relaxations. Specifically, we observe that the objective function value became negative when relaxing integrality, which naturally is not feasible in any solution of the MIP. Hence, we introduce simple valid inequalities of the following type:

$$X_{pc} \geq \sum_{h\in\mathscr{H}} \hat{\theta}_{pch}, \qquad p\in\mathscr{P}, c\in\mathscr{C} \tag{29}$$

Furthermore, investigating LP-Relaxations shows that the relations between the decision variables and between the domains of the decision variables were violated on a constant basis. To avoid this problem, the following intuitive constraints are formulated:

$$\theta_{pc} \geq \frac{1}{\sum\limits_{p' \in \mathscr{P}} n_{p'}} X_{pc}, \qquad p \in \mathscr{P}, c \in \mathscr{C} \quad (30)$$

$$\sum_{h \in \mathscr{H}} \hat{\theta}_{pch} \geq \frac{1}{\sum\limits_{p' \in \mathscr{P}} n_{p'}} (1 - \lambda_{c0}), \qquad p \in \mathscr{P}, c \in \mathscr{C} \quad (31)$$

$$\sum_{h \in \mathscr{H}} \hat{\theta}_{pch} \leq (r-1)\theta_{pc}, \qquad p \in \mathscr{P}, c \in \mathscr{C} \quad (32)$$

Inequalities (30) ensure that, in case procedure p is assigned to cluster c, the reciprocal of the number of surgeries in that cluster is greater than zero and equals at least the smallest possible reciprocal (all surgeries in one cluster) for procedure p. Constraints (31) enforce the fraction of dummy surgeries associated with a procedure and cluster to be greater than zero if λ_{c0} equals zero, i.e., the number of surgeries in a cluster is not a multiple of the number of MSS-cycle repetitions. The last inequalities (32) prohibit domain violations with respect to the fraction of dummy surgeries.

4.2.3 Improvement Heuristic

In the previous sections, we presented a mathematical model and corresponding exact and heuristic solution approaches to find adequate clusters of surgical procedures. As outlined, the optimization model relies on input parameters, namely, the number and centroids of clusters, as calculated by algorithm 1. Solving the optimization model with these input parameters can alter the assignment of surgical procedures to clusters and thus the cluster centroids. Given the new cluster centers, it may be beneficial to reassign some procedures to further decrease the overall sum of squared errors. For this purpose, the clustering problem is solved again. From the model solution, new cluster centroids can be extracted and used in the next optimization run. This procedure can be repeated until there is no further improvement or only a marginal improvement in the objective function. The basic idea of the solution approach is illustrated in algorithm 2. The repeat loop in lines 3-8 represents the iterative nature of the improvement heuristic. In line 5, the optimization model is solved. Additionally, please note that, for any calculation of the relative improvement in line 8, the true overall sum of squared errors in the current iteration j (ESS_T^j) is considered, i.e., the assignment is evaluated with updated cluster centroids (lines 6-7). The algorithm terminates when the relative improvement drops below a predefined level defined by the parameter *gap* and returns the best assignment found (X^*_{pc}).

Algorithm 2: Improvement Heuristic

main input : solution by algorithm 1 (objective function value denoted by ESS^0_T)
main output: partition of surgical procedures (X^*_{pc})

begin
  $j = 0$;
  repeat
    $j = j + 1$;
    Solve clustering problem $\Rightarrow X^j_{pc}$;
    Update $\bar{a}_c$ and $\hat{a}_D$;
    Calculate ESS^j_T;
  until $\frac{ESS^{j-1}_T - ESS^j_T}{ESS^{j-1}_T} \leq gap$;
  $X^*_{pc} = X^{j-1}_{pc}$;
end

5 Numerical Study

To assess the benefit of our approach, we tested our two-stage clustering algorithm on real-world data provided by a German hospital. The data set contains all elective surgeries from January to November 2013. Different surgeries with respect to the same patient are documented individually. Furthermore, for each surgery the data set includes the main surgical procedure according to the OPS but not the whole set of procedures. Thus, multiple surgeries can exhibit the same surgical procedure. The hospital has five different specialties: general surgery (GS), orthopedic surgery (OS), vascular surgery (VS), neurosurgery (NS) and plastic surgery (PS). For each surgery exists a record of the specialty in charge and the corresponding duration he/she occupied the OR, i.e., the anesthesia time plus surgery duration. Table 5 summarizes the data set. For the numerical study, we assume the planning horizon and cycle length of the MSS to be eleven months and one week, respectively. Thus, the parameter r, which represents the number of cycle repetitions, equals 46. Additionally, we set the number of forecasted surgeries equal to the number of surgeries observed in the sample, given that the hospital could not provide a forecast.

Specialty	# surgeries	# surgical procedures	surgery duration (min)		
			mean	std. dev.	coef. of var.
GS	2159	144	142.5	86.5	0.607
OS	1618	104	131.7	67.6	0.513
VS	1317	62	146.5	78.5	0.536
NS	1190	63	223.5	105.4	0.472
PS	827	70	100.5	35.1	0.349

Table 5: Summarized hospital data for elective surgical inpatients from January to November 2013 provided by a German hospital; std. dev. = standard deviation, coef. of var. = coefficient of variation

All numerical experiments were performed on an Intel(R) Xeon(R) CPU E5-1620v2 3.70 GHz with 64 GB RAM. The optimization model was coded in the General Algebraic Modeling System (GAMS) software version 24.2.3 and solved with ILOG CPLEX version 12.6. We implemented the constructive heuristics in Scilab version 5.5. The presented results of our numerical experiments base on

- the comparison between the constructive heuristic presented by van Oostrum et al (2011) and our algorithm 1 (Section 5.1),
- the application of our exact solution approach to our proposed mathematical model together with the evaluation of the solutions' quality and computational times (Section 5.2), and
- our improvement heuristic (algorithm 2) (Section 5.3).

5.1 Constructive Heuristic's Results

To compare our constructive heuristic to that presented in van Oostrum et al (2011), we must parametrize the scalarizing function, aggregating distinct variables (the sum of squared errors and dummy surgeries) to evaluate the quality of a partition. For this purpose, van Oostrum et al (2011) introduce two scaling parameters. k_1 scales the number of dummy surgeries and k_2 the variability in the regular clusters. In accordance with the original paper, we fix k_2 to one and vary the value of k_1 to find distinct partitions. After some test runs, we find the following domain of k_1 to be most appropriate:

$$D_{k_1} = \begin{cases} 0-20, & \text{stepsize } 2 \\ 30-200, & \text{stepsize } 10 \\ 250-800, & \text{stepsize } 50 \\ 900-2,000, & \text{stepsize } 100 \\ 2,250-3,500, & \text{stepsize } 250 \\ 4,000-8,000, & \text{stepsize } 500 \\ 9,000-18,000, & \text{stepsize } 1,000 \end{cases}$$

The tests' results for both heuristics shown in Tables 6 and 7 include:

- the number of clusters (#clusters),
- the sum of squared errors in total (ESS_T) and in the regular clusters (ESS_R) and the dummy cluster (ESS_D),
- the number of surgeries in the dummy cluster (#dummy surgeries), and
- the runtime in seconds (CPU (sec.)) for each specialty.

For the approach by van Oostrum et al (2011), we display the CPU time to run the heuristic for all parameter combinations of k_1 and k_2 and the value of k_1 for which the best solution is found (k^*).

In both tables, we find moderate numbers of clusters generated by both algorithms. The only exception is the OS specialty when applying the approach by van Oostrum et al (2011). The variability of the partitions found reveals the superiority of the algorithm by van Oostrum et al (2011). For each specialty, the overall sum of squared errors is smaller compared to our approach. In general, our algorithm produces partitions with a considerably higher number of dummy surgeries, which results in high variability in the dummy cluster. Again, the OS specialty is an exception. The computation times compared to van Oostrum et al (2011) are relatively moderate, given that only one solution is generated. Altogether, the results of our computational study concerning the alternative constructive heuristics suggest that, in case long computation times are not a matter of concern, it is sufficient to use an algorithm such as that presented by van Oostrum et al (2011). Such an algorithm creates multiple solutions by biasing the number of dummy surgeries over a wide set of bias factors instead of using a more elaborate procedure that considers the dummy cluster's variability but only provides one solution such as our approach.

In the following section, we will investigate the extent to which the initial partitions can be improved by reassigning surgical procedures to surgery types.

Specialty	#clusters	ESS_T	ESS_R	ESS_D	#dummy surgeries	k^*	CPU (sec.)
GS	5	6,482,958	6,255,642	227,316	43	17,000	17,155.8
OS	20	3,893,939	3,007,621	886,318	100	100	4,016.2
VS	6	2,934,448	2,627,135	307,313	29	3,500	523.2
NS	6	7,080,227	6,782,912	297,315	40	5,500	440.1
PS	6	613,929	550,418	63,511	45	450	531.2

Table 6: Results of the constructive heuristics by van Oostrum et al (2011)

Specialty	#clusters	ESS_T	ESS_R	ESS_D	#dummy surgeries	CPU (sec.)
GS	9	9,415,300	4,753,067	4,662,233	319	1,254.0
OS	4	5,092,175	4,382,600	709,575	100	221.2
VS	3	4,369,533	3,811,314	558,219	75	22.6
NS	5	8,963,437	6,535,974	2,427,463	132	19.3
PS	7	881,877	380,525	50,1352	229	91.8

Table 7: Results of our constructive heuristics

5.2 Results Obtained by Our Exact Solution Approach

To initialize our optimization model LCM presented in Section 4.2, we use the number of clusters with the corresponding centroids from the constructive heuristics. First, we demonstrate the usefulness of the simple inequalities. Table 8 highlights the results that we obtain when solving the linearized optimization model without these inequalities, whereas Table 9 contains the results when considering inequalities (29), (31) and (32), which yield the best results. For all computations, we define a time limit of 3,600 seconds. Running CPLEX on the linearized model without any simple inequalities reveals poor results. Except for the case where we solve the model for the VS specialty, given the initial solution as described in Section 5.1, the lower bounds are negative, resulting in huge gaps. For the OS specialty, we cannot even find a feasible integer solution when initializing our model with the solution from the constructive heuristic of van Oostrum et al (2011). A considerable improvement in solution quality and computation times is achieved by adding simple inequalities (29), (31) and (32), as shown in Table 9. For problem instances with a moderate number of clusters and surgical procedures, we find optimal or near-optimal solutions within the time limit (GAP(%) $= (\frac{UB-LB}{LB}) \cdot 100$). However, for the OS specialty, which has 20 different surgery types, the resulting solution gap is 24.4 %, and the best solution is even worse than that found by the constructive heuristic. In the following, we will evaluate the ability of the optimization model integrated in the algorithmic framework of the improvement heuristic to further reduce the variability of the initial partitions.

Specialty	Initial solution					
	Our approach			van Oostrum et al. (2011)		
	UB	LB	CPU (sec.)	UB	LB	CPU (sec.)
GS	6,230,259	< 0	3,600.0	6,425,906	< 0	3,600.0
OS	3,954,792	< 0	3,600.0	n.s.	< 0	3,600.0
VS	4,057,646	4,057,646	71.7	2,929,527	< 0	3,600.0
NS	7,461,922	3,367,180	3,600.0	7,132,961	< 0	3,600.0
PS	690,030	< 0	3,600.0	678,911	< 0	3,600.0

Table 8: Results of solving the linearized model **without** simple inequalities, given different initial solutions

5.3 Improvement Heuristic's Results

For the improvement heuristic presented in Section 4.2, we use different time limits per iteration (3,600; 1,800; 900; and 450 seconds) and terminate the algorithm when the objective function value's improvement with respect to the previous iteration is less than 0.5 %. Tables 10 and 11 show the objective function value (ESS_T) of the clusters found for different runtime settings and initial solutions. In each table, for any specialty, the underlined value represents the best solution found in the shortest time, and the additional bold numbers indicate the overall best partition.

Specialty	Initial solution: Our approach			Initial solution: van Oostrum et al. (2011)		
	ESS_T	GAP(%)	CPU (sec.)	ESS_T	GAP(%)	CPU (sec.)
GS	5,323,914	0.49	3,600.0	6,098,256	0.00	1,800.1
OS	3,948,699	0.00	480.5	4,533,797	24.4	3,600.0
VS	4,057,646	0.00	50.7	2,761,389	1.24	3,600.0
NS	7,444,881	0.00	307.6	7,001,679	0.00	448.9
PS	649,928	1.32	3,600.0	591,224	0.32	3,600.0

Table 9: Results of solving the linearized model **with** simple inequalities (29), (31) and (32), given different initial solutions

Both tables reveal that the solution quality only slightly deteriorates with shrinking computation times when the problem size is moderate. Interestingly, in the case of the GS, OS and PS specialties in Table 10 and the PS specialty in Table 11, the best clusters are identified when the maximum time allowed per iteration is less than 3,600 seconds. Thus, we observe that it is not mandatory to solve our optimization model to (near-)optimality at each iteration of the improvement heuristic but instead to discover different solution paths by changing the computation time limits (or solution gaps). However, in examining the largest problem instance, namely, the OS specialty in Table 10, we observe a considerable increase in the overall variability as the available time to solve the optimization model drops below 1,800 seconds and the solution gaps remain high at each iteration (also, see Table 8). Based on the numerical tests, we further conclude that a good starting point for our improvement heuristic is not necessarily a first partition with a small overall sum of squared errors. For the GS and PS specialties, we find the best partition initializing our model with the (poor) solutions created by our constructive algorithm. For the other specialties, starting with the algorithm by van Oostrum et al (2011) yields the best results. Hence, we observe that, to find high-quality solutions, starting the improvement heuristic with distinct initial partitions seems to be important.

In the following, we compare the best initial partition with the best improved partition. Table 12 illustrates the corresponding results. The second column displays for each specialty the relative change in the objective function value of the best initial partition compared to the best improved partition ($\Delta_{ESS_T}(\%)$). In the case of the NS specialty, we find that ESS_T decreases only by 1.3 %, i.e., the constructive algorithm by van Oostrum et al (2011) already yields good results. For the other surgical departments, the improvements are more considerable, whereas for the GS specialty, the application of the improvement heuristic is most beneficial. Going into more detail, we have a closer look at the effects of reassignments on the relative change in the sum of squared errors in the regular clusters ($\Delta_{ESS_R}(\%)$) and the dummy cluster ($\Delta_{ESS_D}(\%)$). For the latter, large improvements can be stated, at maximum 77.8 % for the OS specialty. Table 12 also indicates that, in some cases (OS and NS), it can be favorable to increase the variability in the regular clusters to allow assignments that reduce the variability in the dummy cluster to a large extent. Furthermore, we observe that for each specialty, the clusters used to initialize the optimization model are active, i.e., at least one surgical procedure is assigned to a cluster. Hence, we end up with 48 clusters (surgery types) in total. Investigating the number of surgeries in the dummy cluster highlights the fact that good partitions are characterized by a small-sized dummy cluster. To assess the variability in the clusters, for each specialty, we present the mean and standard deviation of the coefficient of variation. As with the sum of squared errors, both variables reveal evidence of the effectiveness of the presented clustering approach. Finally, we make an overall assessment of our algorithm with respect to its ability to reduce the variability of the initial partitions. Therefore, we refer to Table 13. In the second row, the sum of squared errors summed up over all clusters and specialties is shown – for the best solution found by the constructive heuristic and the improvement heuristic. We clearly observe that the application of our algorithm is beneficial, given that the variability, as measured by the sum of squared errors, decreases by 9.28 % in total.

Based on the results given in this section, we can conclude the following: regarding the two constructive heuristics considered in this paper, the heuristic by van Oostrum et al (2011) yields solutions with better objective function values compared to our algorithm. However, the computation times are considerably longer, especially for large problem instances, given that the algorithm must be run with different parameter combinations. In such cases, our approach may be preferable. Another advantage of our approach concerns the fact that there is no need to use a scalarization function and derive proper values for the weights. For the application of the improvement heuristic, we demonstrated that the proposed simple inequalities accelerate the solution times of the embedded optimization model. Only for the largest problem instances

Specialty	Time limits per iteration							
	3,600 sec.		1,800 sec.		900 sec.		450 sec.	
	ESS_T	CPU	ESS_T	CPU	ESS_T	CPU	ESS_T	CPU
GS	5,867,379	2,160.7	5,867,379	2,163.2	5,867,266	1,303.7	5,882,256	1,350.0
OS	3,613,236	10,800.0	**3,549,686**	5,400.0	6,094,830	1800.0	6,675,952	900.0
VS	2,728,992	6,157.7	2,728,992	4,274.3	2,728,992	2,474.1	**2,728,992**	1,350.0
NS	6,988,316	1,079.1	6,988,316	1,079.1	6,988,316	1,079.1	**6,988,316**	887.6
PS	589,658	4,593.9	589,658	2,796.9	589,658	1,800.0	583,325	883.1

Table 10: Results of the improvement heuristic with different runtime limits per iteration, given the initial solution of van Oostrum et al (2011)

the remaining gap between the best lower and upper bound stayed substantial. Running the optimization model in the developed algorithmic framework showed that the initial solutions can be drastically improved. We further observed that the best initial partition will not always result in the ultimate best partition for a specialty. In addition, our numerical experiments revealed (with one exception) the robustness of the solution quality with respect to the computation time limits. In some cases, the partition with the best objective function value was even found when less computation time was allowed. Thus, it is important to initialize our optimization model with different partitions and to allow different solution paths in the execution of the improvement algorithm by controlling the computation times and optimality gaps, respectively.

Specialty	Time limits per iteration: 3,600 sec.		1,800 sec.	
	ESS_T	CPU	ESS_T	CPU
GS	**5,211,220**	7,200.0	5,245,819	3,600.0
OS	3,811,406	580.5	3,811,406	580.5
VS	3,912,755	84.5	3,912,755	84.5
NS	7,196,539	394.5	7,196,539	394.5
PS	578,658	10,800.0	578,658	5,400.0

Specialty	Time limits per iteration: 900 sec.		450 sec.	
	ESS_T	CPU	ESS_T	CPU
GS	5,331,244	1,800.0	5,397,374	900.0
OS	3,811,406	580.5	3,811,406	554.0
VS	3,912,755	84.5	3,912,755	84.5
NS	7,196,539	394.5	7,196,539	394.5
PS	**578,658**	2,700.0	584,179	1,350.0

Table 11: Results of the improvement heuristic with different runtime limits per iteration, given our initial solution

Spe-cialty	$\Delta(\%)$ ESS_T	ESS_R	ESS_D	#clus-ters	#dummy surgeries	CV mean	std. dev.
GS	- 19.62	- 18.58	- 48.03	9	43	0.32	0.05
OS	- 8.84	+ 11.48	- 77.80	20	100	0.32	0.05
VS	- 7.00	- 0.92	- 59.00	6	29	0.29	0.03
NS	- 1.30	+ 0.01	- 31.21	6	40	0.32	0.03
PS	- 5.75	- 2.63	- 32.70	7	45	0.25	0.02

Table 12: Best solution found by the improvement heuristic in comparison to the best solution found by the constructive heuristics; std. dev. = standard deviation, CV = coefficient of variation

6 Conclusion

In this paper, we developed a non-linear model (NLCM) and a linearized model (LCM) to determine clusters with minimal ESS_T for the MSS. Furthermore, we presented a two-stage clustering

$\sum$ Specialty	Constructive heuristic	Improvement heuristic	$\Delta(\%)$
ESS_T	21,005,500	19,065,327	**- 9.28**
ESS_R	19,223,728	18,365,527	**- 4.45**
ESS_D	1,781,772	699,801	**- 61.38**

Table 13: Comparison of the best constructive and improvement heuristics with respect to the sum of squared errors over all specialties

algorithm specifically designed for an aggregate cyclic planning environment. We provided an example of its application in master surgery scheduling by aggregating surgical procedures to surgery types while minimizing the sum of squared errors regarding the surgery duration. In the algorithm's first stage, we employed two different constructive heuristics that create the initial partitions of surgical procedures. These results were used to initialize the algorithm's second stage, which finds the assignments of surgical procedures to clusters with a better objective function value. In the numerical study, we illustrated that the improvement heuristic is able to find partitions with considerably reduced variability, which is beneficial for master surgery scheduling.

We envision different directions for further research: the fast construction of promising initial partitions is very important for the results of the improvement heuristic. Especially for large problem instances, the time effort of the constructive heuristics presented in this paper is prohibitive, and new approaches are required. The same holds true for the optimization model. For instances with a large number of clusters, tailored solution approaches are necessary to reduce computation times. These approaches would also allow to evaluate more partitions in the improvement step and to identify better clusters. Further modeling extensions are possible. For example, clustering attributes such as setup times, special surgical equipment or surgeon requirements may be included. In case the influence on other units such as recovery rooms, intensive care units or wards is considered in master surgery scheduling, attributes that represent the flow of patients must be considered when clustering surgical procedures.

References

Adan I, Vissers J (2002) Patient mix optimisation in hospital admission planning: a case study. International Journal of Operations & Production Management 22(4):445–461

Adan I, Bekkers J, Dellaert N, Vissers J, Yu X (2009) Patient mix optimisation and stochastic resource requirements: A case study in cardiothoracic surgery planning. Health Care Management Science 12(2):129–141

Beale E, Tomlin J (1970) Special facilities in a general mathematical programming system for nonconvex problems using ordered sets of variables. In: Lawrence J (ed) Proceedings of the fifth international conference on operational research, Tavistock Publications, London, pp 447–454

Beliën J, Demeulemeester E (2007) Building cyclic master surgery schedules with leveled resulting bed occupancy. European Journal of Operational Research 176(2):1185–1204

Cardoen B, Demeulemeester E, Beliën J (2010) Operating room planning and scheduling: A literature review. European Journal of Operational Research 201(3):921–932

Corne D, Dhaenens C, Jourdan L (2012) Synergies between operations research and data mining: The emerging use of multi-objective approaches. European Journal of Operational Research 221(3):469–479

Dilts D, Khamalah J, Plotkin A (1995) Using cluster analysis for medical resource decision making. Medical Decision Making 15(4):333–346

DMDI (2014) Operationen- und prozedurenschlüssel: Deutsches institut für medizinische dokumentation und information

Fügener A, Hans EW, Kolisch R, Kortbeek N, Vanberkel PT (2014) Master surgery scheduling with consideration of multiple downstream units. European Journal of Operational Research

Guerriero F, Guido R (2011) Operational research in the management of the operating theatre: a survey. Health Care Management Science 14(1):89–114

Hans EW, van Houdenhoven M, Hulshof, Peter J H (2012) A framework for healthcare planning and control. In: Hall R (ed) Handbook of Healthcare System Scheduling, International Series in Operations Research & Management Science, vol 168, Springer US, pp 303–320, URL http://dx.doi.org/10.1007/978-1-4614-1734-7_12

Hansen P, Jaumard B (1997) Cluster analysis and mathematical programming. Mathematical programming 79(1-3):191–215

Inniss TR (2006) Seasonal clustering technique for time series data. European Journal of Operational Research 175(1):376–384

Jain AK, Murty MN, Flynn PJ (1999) Data clustering: a review. ACM computing surveys (CSUR) 31(3):264–323

Kulkarni G, Fathi Y (2007) Integer programming models for the q-mode problem. European Journal of Operational Research 182(2):612–625

Li HL (1994) A global approach for general 0–1 fractional programming. European Journal of Operational Research 73(3):590–596

Ma G, Demeulemeester E (2013) A multilevel integrative approach to hospital case mix and capacity planning. Computers & Operations Research (40):2198–2207

MacQueen J (ed) (1967) Some methods for classification and analysis of multivariate observations, Fifth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley and Calif, URL http://projecteuclid.org/euclid.bsmsp/1200512992

Meisel S, Mattfeld D (2010) Synergies of operations research and data mining. European Journal of Operational Research 206(1):1–10

OECD (2011) Health at a glance 2011: Oecd indicators

Olafsson S, Li X, Wu S (2008) Operations research and data mining. European Journal of Operational Research 187(3):1429–1448

van Oostrum JM, van Houdenhoven M, Hurink JL, Hans EW, Wullink G, Kazemier G (2008) A master surgical scheduling approach for cyclic scheduling in operating room departments. OR Spectrum 30(2):355–374

van Oostrum JM, Parlevliet T, Wagelmans, Albert P M, Kazemier G (2011) A method for clustering surgical cases to allow master surgical scheduling. INFOR: Information Systems and Operational Research 49(4):254–260

Romanowski CJ, Nagi R, Sudit M (2006) Data mining in an engineering design environment: Or applications from graph matching. Computers & Operations Research 33(11):3150–3160

Sağlam B, Salman FS, Sayın S, Türkay M (2006) A mixed-integer programming approach to the clustering problem with an application in customer segmentation. European Journal of Operational Research 173(3):866–879

Santibáñez P, Begen M, Atkins D (2007) Surgical block scheduling in a system of hospitals: an application to resource and wait list management in a british columbia health authority. Health Care Management Science 10(3):269–282

van Oostrum, Jeroen M, Bredenhoff E, Hans EW (2010) Suitability and managerial implications of a master surgical scheduling approach. Annals of Operations Research 178(1):91–104

Ward J (1963) Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58 (1963):236–244

Wu TH (1997) A note on a global approach for general 0–1 fractional programming. European Journal of Operational Research 101(1):220–223

Xu R, Wunsch D (2005) Survey of clustering algorithms. Neural Networks, IEEE Transactions on 16(3):645–678

Hohenheim Discussion Papers in Business, Economics and Social Sciences

The Faculty of Business, Economics and Social Sciences continues since 2015 the established “FZID Discussion Paper Series” of the “Centre for Research on Innovation and Services (FZID)” under the name “Hohenheim Discussion Papers in Business, Economics and Social Sciences”.

Institutes

510 Institute of Financial Management
520 Institute of Economics
530 Institute of Health Care & Public Management
540 Institute of Communication Science
550 Institute of Law and Social Sciences
560 Institute of Economic and Business Education
570 Institute of Marketing & Management
580 Institute of Interorganizational Management & Performance

Research Areas (since 2017)

INEPA “Inequality and Economic Policy Analysis”
TKID “Transformation der Kommunikation – Integration und Desintegration”
NegoTrans “Negotiation Research – Transformation, Technology, Media and Costs”
INEF “Innovation, Entrepreneurship and Finance”

Download Hohenheim Discussion Papers in Business, Economics and Social Sciences
from our homepage: https://wiso.uni-hohenheim.de/papers

No.	Author	Title	Inst
01-2015	Thomas Beissinger, Philipp Baudy	THE IMPACT OF TEMPORARY AGENCY WORK ON TRADE UNION WAGE SETTING: A Theoretical Analysis	520
02-2015	Fabian Wahl	PARTICIPATIVE POLITICAL INSTITUTIONS AND CITY DEVELOPMENT 800-1800	520
03-2015	Tommaso Proietti, Martyna Marczak, Gianluigi Mazzi	EUROMIND-D: A DENSITY ESTIMATE OF MONTHLY GROSS DOMESTIC PRODUCT FOR THE EURO AREA	520
04-2015	Thomas Beissinger, Nathalie Chusseau, Joël Hellier	OFFSHORING AND LABOUR MARKET REFORMS: MODELLING THE GERMAN EXPERIENCE	520
05-2015	Matthias Mueller, Kristina Bogner, Tobias Buchmann, Muhamed Kudic	SIMULATING KNOWLEDGE DIFFUSION IN FOUR STRUCTURALLY DISTINCT NETWORKS – AN AGENT-BASED SIMULATION MODEL	520
06-2015	Martyna Marczak, Thomas Beissinger	BIDIRECTIONAL RELATIONSHIP BETWEEN INVESTOR SENTIMENT AND EXCESS RETURNS: NEW EVIDENCE FROM THE WAVELET PERSPECTIVE	520
07-2015	Peng Nie, Galit Nimrod, Alfonso Sousa-Poza	INTERNET USE AND SUBJECTIVE WELL-BEING IN CHINA	530

No.	Author	Title	Inst
08-2015	Fabian Wahl	THE LONG SHADOW OF HISTORY ROMAN LEGACY AND ECONOMIC DEVELOPMENT – EVIDENCE FROM THE GERMAN LIMES	520
09-2015	Peng Nie, Alfonso Sousa-Poza	COMMUTE TIME AND SUBJECTIVE WELL-BEING IN URBAN CHINA	530
10-2015	Kristina Bogner	THE EFFECT OF PROJECT FUNDING ON INNOVATIVE PERFORMANCE AN AGENT-BASED SIMULATION MODEL	520
11-2015	Bogang Jun, Tai-Yoo Kim	A NEO-SCHUMPETERIAN PERSPECTIVE ON THE ANALYTICAL MACROECONOMIC FRAMEWORK: THE EXPANDED REPRODUCTION SYSTEM	520
12-2015	Volker Grossmann Aderonke Osikominu Marius Osterfeld	ARE SOCIOCULTURAL FACTORS IMPORTANT FOR STUDYING A SCIENCE UNIVERSITY MAJOR?	520
13-2015	Martyna Marczak Tommaso Proietti Stefano Grassi	A DATA–CLEANING AUGMENTED KALMAN FILTER FOR ROBUST ESTIMATION OF STATE SPACE MODELS	520
14-2015	Carolina Castagnetti Luisa Rosti Marina Töpfer	THE REVERSAL OF THE GENDER PAY GAP AMONG PUBLIC-CONTEST SELECTED YOUNG EMPLOYEES	520
15-2015	Alexander Opitz	DEMOCRATIC PROSPECTS IN IMPERIAL RUSSIA: THE REVOLUTION OF 1905 AND THE POLITICAL STOCK MARKET	520
01-2016	Michael Ahlheim, Jan Neidhardt	NON-TRADING BEHAVIOUR IN CHOICE EXPERIMENTS	520
02-2016	Bogang Jun, Alexander Gerybadze, Tai-Yoo Kim	THE LEGACY OF FRIEDRICH LIST: THE EXPANSIVE REPRODUCTION SYSTEM AND THE KOREAN HISTORY OF INDUSTRIALIZATION	520
03-2016	Peng Nie, Alfonso Sousa-Poza	FOOD INSECURITY AMONG OLDER EUROPEANS: EVIDENCE FROM THE SURVEY OF HEALTH, AGEING, AND RETIREMENT IN EUROPE	530
04-2016	Peter Spahn	POPULATION GROWTH, SAVING, INTEREST RATES AND STAGNATION. DISCUSSING THE EGGERTSSON-MEHROTRA-MODEL	520
05-2016	Vincent Dekker, Kristina Strohmaier, Nicole Bosch	A DATA-DRIVEN PROCEDURE TO DETERMINE THE BUNCHING WINDOW – AN APPLICATION TO THE NETHERLANDS	520
06-2016	Philipp Baudy, Dario Cords	DEREGULATION OF TEMPORARY AGENCY EMPLOYMENT IN A UNIONIZED ECONOMY: DOES THIS REALLY LEAD TO A SUBSTITUTION OF REGULAR EMPLOYMENT?	520

No.	Author	Title	Inst
07-2016	Robin Jessen, Davud Rostam-Afschar, Sebastian Schmitz	HOW IMPORTANT IS PRECAUTIONARY LABOR SUPPLY?	520
08-2016	Peng Nie, Alfonso Sousa-Poza, Jianhong Xue	FUEL FOR LIFE: DOMESTIC COOKING FUELS AND WOMEN'S HEALTH IN RURAL CHINA	530
09-2016	Bogang Jun, Seung Kyu-Yi, Tobias Buchmann, Matthias Müller	THE CO-EVOLUTION OF INNOVATION NETWORKS: COLLABORATION BETWEEN WEST AND EAST GERMANY FROM 1972 TO 2014	520
10-2016	Vladan Ivanovic, Vadim Kufenko, Boris Begovic Nenad Stanisic, Vincent Geloso	CONTINUITY UNDER A DIFFERENT NAME. THE OUTCOME OF PRIVATISATION IN SERBIA	520
11-2016	David E. Bloom Michael Kuhn Klaus Prettner	THE CONTRIBUTION OF FEMALE HEALTH TO ECONOMIC DEVELOPMENT	520
12-2016	Franz X. Hof Klaus Prettner	THE QUEST FOR STATUS AND R&D-BASED GROWTH	520
13-2016	Jung-In Yeon Andreas Pyka Tai-Yoo Kim	STRUCTURAL SHIFT AND INCREASING VARIETY IN KOREA, 1960–2010: EMPIRICAL EVIDENCE OF THE ECONOMIC DEVELOPMENT MODEL BY THE CREATION OF NEW SECTORS	520
14-2016	Benjamin Fuchs	THE EFFECT OF TEENAGE EMPLOYMENT ON CHARACTER SKILLS, EXPECTATIONS AND OCCUPATIONAL CHOICE STRATEGIES	520
15-2016	Seung-Kyu Yi Bogang Jun	HAS THE GERMAN REUNIFICATION STRENGTHENED GERMANY'S NATIONAL INNOVATION SYSTEM? TRIPLE HELIX DYNAMICS OF GERMANY'S INNOVATION SYSTEM	520
16-2016	Gregor Pfeifer Fabian Wahl Martyna Marczak	ILLUMINATING THE WORLD CUP EFFECT: NIGHT LIGHTS EVIDENCE FROM SOUTH AFRICA	520
17-2016	Malte Klein Andreas Sauer	CELEBRATING 30 YEARS OF INNOVATION SYSTEM RESEARCH: WHAT YOU NEED TO KNOW ABOUT INNOVATION SYSTEMS	570
18-2016	Klaus Prettner	THE IMPLICATIONS OF AUTOMATION FOR ECONOMIC GROWTH AND THE LABOR SHARE	520
19-2016	Klaus Prettner Andreas Schaefer	HIGHER EDUCATION AND THE FALL AND RISE OF INEQUALITY	520
20-2016	Vadim Kufenko Klaus Prettner	YOU CAN'T ALWAYS GET WHAT YOU WANT? ESTIMATOR CHOICE AND THE SPEED OF CONVERGENCE	520

No.	Author	Title	Inst
01-2017	Annarita Baldanzi Alberto Bucci Klaus Prettner	CHILDRENS HEALTH, HUMAN CAPITAL ACCUMULATION, AND R&D-BASED ECONOMIC GROWTH	INEPA
02-2017	Julius Tennert Marie Lambert Hans-Peter Burghof	MORAL HAZARD IN VC-FINANCE: MORE EXPENSIVE THAN YOU THOUGHT	INEF
03-2017	Michael Ahlheim Oliver Frör Nguyen Minh Duc Antonia Rehl Ute Siepmann Pham Van Dinh	LABOUR AS A UTILITY MEASURE RECONSIDERED	520
04-2017	Bohdan Kukharskyy Sebastian Seiffert	GUN VIOLENCE IN THE U.S.: CORRELATES AND CAUSES	520
05-2017	Ana Abeliansky Klaus Prettner	AUTOMATION AND DEMOGRAPHIC CHANGE	520
06-2017	Vincent Geloso Vadim Kufenko	INEQUALITY AND GUARD LABOR, OR PROHIBITION AND GUARD LABOR?	INEPA
07-2017	Emanuel Gasteiger Klaus Prettner	ON THE POSSIBILITY OF AUTOMATION-INDUCED STAGNATION	520
08-2017	Klaus Prettner Holger Strulik	THE LOST RACE AGAINST THE MACHINE: AUTOMATION, EDUCATION, AND INEQUALITY IN AN R&D-BASED GROWTH MODEL	INEPA
09-2017	David E. Bloom Simiao Chen Michael Kuhn Mark E. McGovern Les Oxley Klaus Prettner	THE ECONOMIC BURDEN OF CHRONIC DISEASES: ESTIMATES AND PROJECTIONS FOR CHINA, JAPAN, AND SOUTH KOREA	520
10-2017	Sebastian Till Braun Nadja Dwenger	THE LOCAL ENVIRONMENT SHAPES REFUGEE INTEGRATION: EVIDENCE FROM POST-WAR GERMANY	INEPA
11-2017	Vadim Kufenko Klaus Prettner Vincent Geloso	DIVERGENCE, CONVERGENCE, AND THE HISTORY-AUGMENTED SOLOW MODEL	INEPA
12-2017	Frank M. Fossen Ray Rees Davud Rostam-Afschar Viktor Steiner	HOW DO ENTREPRENEURIAL PORTFOLIOS RESPOND TO INCOME TAXATION?	520
13-2017	Steffen Otterbach Michael Rogan	SPATIAL DIFFERENCES IN STUNTING AND HOUSEHOLD AGRICULTURAL PRODUCTION IN SOUTH AFRICA: (RE-) EXAMINING THE LINKS USING NATIONAL PANEL SURVEY DATA	INEPA
14-2017	Carolina Castagnetti Luisa Rosti Marina Töpfer	THE CONVERGENCE OF THE GENDER PAY GAP – AN ALTERNATIVE ESTIMATION APPROACH	INEPA

No.	Author	Title	Inst
15-2017	Andreas Hecht	ON THE DETERMINANTS OF SPECULATION – A CASE FOR EXTENDED DISCLOSURES IN CORPORATE RISK MANAGEMENT	510
16-2017	Mareike Schoop D. Marc Kilgour (Editors)	PROCEEDINGS OF THE 17TH INTERNATIONAL CONFERENCE ON GROUP DECISION AND NEGOTIATION	NegoTrans
17-2017	Mareike Schoop D. Marc Kilgour (Editors)	DOCTORAL CONSORTIUM OF THE 17TH INTERNATIONAL CONFERENCE ON GROUP DECISION AND NEGOTIATION	NegoTrans
18-2017	Sibylle Lehmann-Hasemeyer Fabian Wahl	SAVING BANKS AND THE INDUSTRIAL REVOLUTION IN PRUSSIA SUPPORTING REGIONAL DEVELOPMENT WITH PUBLIC FINANCIAL INSTITUTIONS	520
19-2017	Stephanie Glaser	A REVIEW OF SPATIAL ECONOMETRIC MODELS FOR COUNT DATA	520
20-2017	Dario Cords	ENDOGENOUS TECHNOLOGY, MATCHING, AND LABOUR UNIONS: DOES LOW-SKILLED IMMIGRATION AFFECT THE TECHNOLOGICAL ALIGNMENT OF THE HOST COUNTRY?	INEPA
21-2017	Micha Kaiser Jan M. Bauer	PRESCHOOL CHILD CARE AND CHILD WELL-BEING IN GERMANY: DOES THE MIGRANT EXPERIENCE DIFFER?	INEPA
22-2017	Thilo R. Huning Fabian Wahl	LORD OF THE LEMONS: ORIGIN AND DYNAMICS OF STATE CAPACITY	520
23-2017	Matthias Busse Ceren Erdogan Henning Mühlen	STRUCTURAL TRANSFORMATION AND ITS RELEVANCE FOR ECONOMIC GROWTH IN SUB-SHARAN AFRICA	INEPA
24-2017	Sibylle Lehmann-Hasemeyer Alexander Opitz	THE VALUE OF POLITICAL CONNECTIONS IN THE FIRST GERMAN DEMOCRACY – EVIDENCE FROM THE BERLIN STOCK EXCHANGE	520
25-2017	Samuel Mburu Micha Kaiser Alfonso Sousa-Poza	LIFESTOCK ASSET DYNAMICS AMONG PASTORALISTS IN NORTHERN KENYA	INEPA
26-2017	Marina Töpfer	DETAILED RIF DECOMPOSITION WITH SELECTION – THE GENDER PAY GAP IN ITALY	INEPA
27-2017	Robin Jessen Maria Metzing Davud Rostam-Afschar	OPTIMAL TAXATION UNDER DIFFERENT CONCEPTS OF JUSTNESS	INEPA
28-2017	Alexander Kressner Katja Schimmelpfeng	CLUSTERING SURGICAL PROCEDURES FOR MASTER SURGICAL SCHEDULING	580

FZID Discussion Papers

(published 2009-2014)

Competence Centers

IK	Innovation and Knowledge
ICT	Information Systems and Communication Systems
CRFM	Corporate Finance and Risk Management
HCM	Health Care Management
CM	Communication Management
MM	Marketing Management
ECO	Economics

Download FZID Discussion Papers from our homepage: https://wiso.uni-hohenheim.de/archiv_fzid_papers

Nr.	Autor	Titel	CC
01-2009	Julian P. Christ	NEW ECONOMIC GEOGRAPHY RELOADED: Localized Knowledge Spillovers and the Geography of Innovation	IK
02-2009	André P. Slowak	MARKET FIELD STRUCTURE & DYNAMICS IN INDUSTRIAL AUTOMATION	IK
03-2009	Pier Paolo Saviotti, Andreas Pyka	GENERALIZED BARRIERS TO ENTRY AND ECONOMIC DEVELOPMENT	IK
04-2009	Uwe Focht, Andreas Richter and Jörg Schiller	INTERMEDIATION AND MATCHING IN INSURANCE MARKETS	HCM
05-2009	Julian P. Christ, André P. Slowak	WHY BLU-RAY VS. HD-DVD IS NOT VHS VS. BETAMAX: THE CO-EVOLUTION OF STANDARD-SETTING CONSORTIA	IK
06-2009	Gabriel Felbermayr, Mario Larch and Wolfgang Lechthaler	UNEMPLOYMENT IN AN INTERDEPENDENT WORLD	ECO
07-2009	Steffen Otterbach	MISMATCHES BETWEEN ACTUAL AND PREFERRED WORK TIME: Empirical Evidence of Hours Constraints in 21 Countries	HCM
08-2009	Sven Wydra	PRODUCTION AND EMPLOYMENT IMPACTS OF NEW TECHNOLOGIES – ANALYSIS FOR BIOTECHNOLOGY	IK
09-2009	Ralf Richter, Jochen Streb	CATCHING-UP AND FALLING BEHIND KNOWLEDGE SPILLOVER FROM AMERICAN TO GERMAN MACHINE TOOL MAKERS	IK

Nr.	Autor	Titel	CC
10-2010	Rahel Aichele, Gabriel Felbermayr	KYOTO AND THE CARBON CONTENT OF TRADE	ECO
11-2010	David E. Bloom, Alfonso Sousa-Poza	ECONOMIC CONSEQUENCES OF LOW FERTILITY IN EUROPE	HCM
12-2010	Michael Ahlheim, Oliver Frör	DRINKING AND PROTECTING – A MARKET APPROACH TO THE PRESERVATION OF CORK OAK LANDSCAPES	ECO
13-2010	Michael Ahlheim, Oliver Frör, Antonia Heinke, Nguyen Minh Duc, and Pham Van Dinh	LABOUR AS A UTILITY MEASURE IN CONTINGENT VALUATION STUDIES – HOW GOOD IS IT REALLY?	ECO
14-2010	Julian P. Christ	THE GEOGRAPHY AND CO-LOCATION OF EUROPEAN TECHNOLOGY-SPECIFIC CO-INVENTORSHIP NETWORKS	IK
15-2010	Harald Degner	WINDOWS OF TECHNOLOGICAL OPPORTUNITY DO TECHNOLOGICAL BOOMS INFLUENCE THE RELATIONSHIP BETWEEN FIRM SIZE AND INNOVATIVENESS?	IK
16-2010	Tobias A. Jopp	THE WELFARE STATE EVOLVES: GERMAN KNAPPSCHAFTEN, 1854-1923	HCM
17-2010	Stefan Kirn (Ed.)	PROCESS OF CHANGE IN ORGANISATIONS THROUGH eHEALTH	ICT
18-2010	Jörg Schiller	ÖKONOMISCHE ASPEKTE DER ENTLOHNUNG UND REGULIERUNG UNABHÄNGIGER VERSICHERUNGSVERMITTLER	HCM
19-2010	Frauke Lammers, Jörg Schiller	CONTRACT DESIGN AND INSURANCE FRAUD: AN EXPERIMENTAL INVESTIGATION	HCM
20-2010	Martyna Marczak, Thomas Beissinger	REAL WAGES AND THE BUSINESS CYCLE IN GERMANY	ECO
21-2010	Harald Degner, Jochen Streb	FOREIGN PATENTING IN GERMANY, 1877-1932	IK
22-2010	Heiko Stüber, Thomas Beissinger	DOES DOWNWARD NOMINAL WAGE RIGIDITY DAMPEN WAGE INCREASES?	ECO
23-2010	Mark Spoerer, Jochen Streb	GUNS AND BUTTER – BUT NO MARGARINE: THE IMPACT OF NAZI ECONOMIC POLICIES ON GERMAN FOOD CONSUMPTION, 1933-38	ECO

Nr.	Autor	Titel	CC
24-2011	Dhammika Dharmapala, Nadine Riedel	EARNINGS SHOCKS AND TAX-MOTIVATED INCOME-SHIFTING: EVIDENCE FROM EUROPEAN MULTINATIONALS	ECO
25-2011	Michael Schuele, Stefan Kirn	QUALITATIVES, RÄUMLICHES SCHLIEßEN ZUR KOLLISIONSERKENNUNG UND KOLLISIONSVERMEIDUNG AUTONOMER BDI-AGENTEN	ICT
26-2011	Marcus Müller, Guillaume Stern, Ansger Jacob and Stefan Kirn	VERHALTENSMODELLE FÜR SOFTWAREAGENTEN IM PUBLIC GOODS GAME	ICT
27-2011	Monnet Benoit, Patrick Gbakoua and Alfonso Sousa-Poza	ENGEL CURVES, SPATIAL VARIATION IN PRICES AND DEMAND FOR COMMODITIES IN CÔTE D'IVOIRE	ECO
28-2011	Nadine Riedel, Hannah Schildberg-Hörisch	ASYMMETRIC OBLIGATIONS	ECO
29-2011	Nicole Waidlein	CAUSES OF PERSISTENT PRODUCTIVITY DIFFERENCES IN THE WEST GERMAN STATES IN THE PERIOD FROM 1950 TO 1990	IK
30-2011	Dominik Hartmann, Atilio Arata	MEASURING SOCIAL CAPITAL AND INNOVATION IN POOR AGRICULTURAL COMMUNITIES. THE CASE OF CHÁPARRA - PERU	IK
31-2011	Peter Spahn	DIE WÄHRUNGSKRISENUNION DIE EURO-VERSCHULDUNG DER NATIONALSTAATEN ALS SCHWACHSTELLE DER EWU	ECO
32-2011	Fabian Wahl	DIE ENTWICKLUNG DES LEBENSSTANDARDS IM DRITTEN REICH – EINE GLÜCKSÖKONOMISCHE PERSPEKTIVE	ECO
33-2011	Giorgio Triulzi, Ramon Scholz and Andreas Pyka	R&D AND KNOWLEDGE DYNAMICS IN UNIVERSITY-INDUSTRY RELATIONSHIPS IN BIOTECH AND PHARMACEUTICALS: AN AGENT-BASED MODEL	IK
34-2011	Claus D. Müller-Hengstenberg, Stefan Kirn	ANWENDUNG DES ÖFFENTLICHEN VERGABERECHTS AUF MODERNE IT SOFTWAREENTWICKLUNGSVERFAHREN	ICT
35-2011	Andreas Pyka	AVOIDING EVOLUTIONARY INEFFICIENCIES IN INNOVATION NETWORKS	IK
36-2011	David Bell, Steffen Otterbach and Alfonso Sousa-Poza	WORK HOURS CONSTRAINTS AND HEALTH	HCM
37-2011	Lukas Scheffknecht, Felix Geiger	A BEHAVIORAL MACROECONOMIC MODEL WITH ENDOGENOUS BOOM-BUST CYCLES AND LEVERAGE DYNAMICS	ECO
38-2011	Yin Krogmann, Ulrich Schwalbe	INTER-FIRM R&D NETWORKS IN THE GLOBAL PHARMACEUTICAL BIOTECHNOLOGY INDUSTRY DURING 1985–1998: A CONCEPTUAL AND EMPIRICAL ANALYSIS	IK

Nr.	Autor	Titel	CC
39-2011	Michael Ahlheim, Tobias Börger and Oliver Frör	RESPONDENT INCENTIVES IN CONTINGENT VALUATION: THE ROLE OF RECIPROCITY	ECO
40-2011	Tobias Börger	A DIRECT TEST OF SOCIALLY DESIRABLE RESPONDING IN CONTINGENT VALUATION INTERVIEWS	ECO
41-2011	Ralf Rukwid, Julian P. Christ	QUANTITATIVE CLUSTERIDENTIFIKATION AUF EBENE DER DEUTSCHEN STADT- UND LANDKREISE (1999-2008)	IK

Nr.	Autor	Titel	CC
42-2012	Benjamin Schön, Andreas Pyka	A TAXONOMY OF INNOVATION NETWORKS	IK
43-2012	Dirk Foremny, Nadine Riedel	BUSINESS TAXES AND THE ELECTORAL CYCLE	ECO
44-2012	Gisela Di Meglio, Andreas Pyka and Luis Rubalcaba	VARIETIES OF SERVICE ECONOMIES IN EUROPE	IK
45-2012	Ralf Rukwid, Julian P. Christ	INNOVATIONSPOTENTIALE IN BADEN-WÜRTTEMBERG: PRODUKTIONSCLUSTER IM BEREICH „METALL, ELEKTRO, IKT“ UND REGIONALE VERFÜGBARKEIT AKADEMISCHER FACHKRÄFTE IN DEN MINT-FÄCHERN	IK
46-2012	Julian P. Christ, Ralf Rukwid	INNOVATIONSPOTENTIALE IN BADEN-WÜRTTEMBERG: BRANCHENSPEZIFISCHE FORSCHUNGS- UND ENTWICKLUNGSAKTIVITÄT, REGIONALES PATENTAUFKOMMEN UND BESCHÄFTIGUNGSSTRUKTUR	IK
47-2012	Oliver Sauter	ASSESSING UNCERTAINTY IN EUROPE AND THE US - IS THERE A COMMON FACTOR?	ECO
48-2012	Dominik Hartmann	SEN MEETS SCHUMPETER. INTRODUCING STRUCTURAL AND DYNAMIC ELEMENTS INTO THE HUMAN CAPABILITY APPROACH	IK
49-2012	Harold Paredes-Frigolett, Andreas Pyka	DISTAL EMBEDDING AS A TECHNOLOGY INNOVATION NETWORK FORMATION STRATEGY	IK
50-2012	Martyna Marczak, Víctor Gómez	CYCLICALITY OF REAL WAGES IN THE USA AND GERMANY: NEW INSIGHTS FROM WAVELET ANALYSIS	ECO
51-2012	André P. Slowak	DIE DURCHSETZUNG VON SCHNITTSTELLEN IN DER STANDARDSETZUNG: FALLBEISPIEL LADESYSTEM ELEKTROMOBILITÄT	IK
52-2012	Fabian Wahl	WHY IT MATTERS WHAT PEOPLE THINK - BELIEFS, LEGAL ORIGINS AND THE DEEP ROOTS OF TRUST	ECO
53-2012	Dominik Hartmann, Micha Kaiser	STATISTISCHER ÜBERBLICK DER TÜRKISCHEN MIGRATION IN BADEN-WÜRTTEMBERG UND DEUTSCHLAND	IK
54-2012	Dominik Hartmann, Andreas Pyka, Seda Aydin, Lena Klauß, Fabian Stahl, Ali Santircioglu, Silvia Oberegelsbacher, Sheida Rashidi, Gaye Onan and Suna Erginkoç	IDENTIFIZIERUNG UND ANALYSE DEUTSCH-TÜRKISCHER INNOVATIONSNETZWERKE. ERSTE ERGEBNISSE DES TGIN-PROJEKTES	IK
55-2012	Michael Ahlheim, Tobias Börger and Oliver Frör	THE ECOLOGICAL PRICE OF GETTING RICH IN A GREEN DESERT: A CONTINGENT VALUATION STUDY IN RURAL SOUTHWEST CHINA	ECO

Nr.	Autor	Titel	CC
56-2012	Matthias Strifler Thomas Beissinger	FAIRNESS CONSIDERATIONS IN LABOR UNION WAGE SETTING – A THEORETICAL ANALYSIS	ECO
57-2012	Peter Spahn	INTEGRATION DURCH WÄHRUNGSUNION? DER FALL DER EURO-ZONE	ECO
58-2012	Sibylle H. Lehmann	TAKING FIRMS TO THE STOCK MARKET: IPOS AND THE IMPORTANCE OF LARGE BANKS IN IMPERIAL GERMANY 1896-1913	ECO
59-2012	Sibylle H. Lehmann, Philipp Hauber and Alexander Opitz	POLITICAL RIGHTS, TAXATION, AND FIRM VALUATION – EVIDENCE FROM SAXONY AROUND 1900	ECO
60-2012	Martyna Marczak, Víctor Gómez	SPECTRAN, A SET OF MATLAB PROGRAMS FOR SPECTRAL ANALYSIS	ECO
61-2012	Theresa Lohse, Nadine Riedel	THE IMPACT OF TRANSFER PRICING REGULATIONS ON PROFIT SHIFTING WITHIN EUROPEAN MULTINATIONALS	ECO

Nr.	Autor	Titel	CC
62-2013	Heiko Stüber	REAL WAGE CYCLICALITY OF NEWLY HIRED WORKERS	ECO
63-2013	David E. Bloom, Alfonso Sousa-Poza	AGEING AND PRODUCTIVITY	HCM
64-2013	Martyna Marczak, Víctor Gómez	MONTHLY US BUSINESS CYCLE INDICATORS: A NEW MULTIVARIATE APPROACH BASED ON A BAND-PASS FILTER	ECO
65-2013	Dominik Hartmann, Andreas Pyka	INNOVATION, ECONOMIC DIVERSIFICATION AND HUMAN DEVELOPMENT	IK
66-2013	Christof Ernst, Katharina Richter and Nadine Riedel	CORPORATE TAXATION AND THE QUALITY OF RESEARCH AND DEVELOPMENT	ECO
67-2013	Michael Ahlheim, Oliver Frör, Jiang Tong, Luo Jing and Sonna Pelz	NONUSE VALUES OF CLIMATE POLICY - AN EMPIRICAL STUDY IN XINJIANG AND BEIJING	ECO
68-2013	Michael Ahlheim, Friedrich Schneider	CONSIDERING HOUSEHOLD SIZE IN CONTINGENT VALUATION STUDIES	ECO
69-2013	Fabio Bertoni, Tereza Tykvová	WHICH FORM OF VENTURE CAPITAL IS MOST SUPPORTIVE OF INNOVATION? EVIDENCE FROM EUROPEAN BIOTECHNOLOGY COMPANIES	CFRM
70-2013	Tobias Buchmann, Andreas Pyka	THE EVOLUTION OF INNOVATION NETWORKS: THE CASE OF A GERMAN AUTOMOTIVE NETWORK	IK
71-2013	B. Vermeulen, A. Pyka, J. A. La Poutré and A. G. de Kok	CAPABILITY-BASED GOVERNANCE PATTERNS OVER THE PRODUCT LIFE-CYCLE	IK
72-2013	Beatriz Fabiola López Ulloa, Valerie Møller and Alfonso Sousa-Poza	HOW DOES SUBJECTIVE WELL-BEING EVOLVE WITH AGE? A LITERATURE REVIEW	HCM
73-2013	Wencke Gwozdz, Alfonso Sousa-Poza, Lucia A. Reisch, Wolfgang Ahrens, Stefaan De Henauw, Gabriele Eiben, Juan M. Fernández-Alvira, Charalampos Hadjigeorgiou, Eva Kovács, Fabio Lauria, Toomas Veidebaum, Garrath Williams, Karin Bammann	MATERNAL EMPLOYMENT AND CHILDHOOD OBESITY – A EUROPEAN PERSPECTIVE	HCM

Nr.	Autor	Titel	CC
74-2013	Andreas Haas, Annette Hofmann	RISIKEN AUS CLOUD-COMPUTING-SERVICES: FRAGEN DES RISIKOMANAGEMENTS UND ASPEKTE DER VERSICHERBARKEIT	HCM
75-2013	Yin Krogmann, Nadine Riedel and Ulrich Schwalbe	INTER-FIRM R&D NETWORKS IN PHARMACEUTICAL BIOTECHNOLOGY: WHAT DETERMINES FIRM'S CENTRALITY-BASED PARTNERING CAPABILITY?	ECO, IK
76-2013	Peter Spahn	MACROECONOMIC STABILISATION AND BANK LENDING: A SIMPLE WORKHORSE MODEL	ECO
77-2013	Sheida Rashidi, Andreas Pyka	MIGRATION AND INNOVATION – A SURVEY	IK
78-2013	Benjamin Schön, Andreas Pyka	THE SUCCESS FACTORS OF TECHNOLOGY-SOURCING THROUGH MERGERS & ACQUISITIONS – AN INTUITIVE META-ANALYSIS	IK
79-2013	Irene Prostolupow, Andreas Pyka and Barbara Heller-Schuh	TURKISH-GERMAN INNOVATION NETWORKS IN THE EUROPEAN RESEARCH LANDSCAPE	IK
80-2013	Eva Schlenker, Kai D. Schmid	CAPITAL INCOME SHARES AND INCOME INEQUALITY IN THE EUROPEAN UNION	ECO
81-2013	Michael Ahlheim, Tobias Börger and Oliver Frör	THE INFLUENCE OF ETHNICITY AND CULTURE ON THE VALUATION OF ENVIRONMENTAL IMPROVEMENTS – RESULTS FROM A CVM STUDY IN SOUTHWEST CHINA –	ECO
82-2013	Fabian Wahl	DOES MEDIEVAL TRADE STILL MATTER? HISTORICAL TRADE CENTERS, AGGLOMERATION AND CONTEMPORARY ECONOMIC DEVELOPMENT	ECO
83-2013	Peter Spahn	SUBPRIME AND EURO CRISIS: SHOULD WE BLAME THE ECONOMISTS?	ECO
84-2013	Daniel Guffarth, Michael J. Barber	THE EUROPEAN AEROSPACE R&D COLLABORATION NETWORK	IK
85-2013	Athanasios Saitis	KARTELLBEKÄMPFUNG UND INTERNE KARTELLSTRUKTUREN: EIN NETZWERKTHEORETISCHER ANSATZ	IK

Nr.	Autor	Titel	CC
86-2014	Stefan Kirn, Claus D. Müller-Hengstenberg	INTELLIGENTE (SOFTWARE-)AGENTEN: EINE NEUE HERAUSFORDERUNG FÜR DIE GESELLSCHAFT UND UNSER RECHTSSYSTEM?	ICT
87-2014	Peng Nie, Alfonso Sousa-Poza	MATERNAL EMPLOYMENT AND CHILDHOOD OBESITY IN CHINA: EVIDENCE FROM THE CHINA HEALTH AND NUTRITION SURVEY	HCM
88-2014	Steffen Otterbach, Alfonso Sousa-Poza	JOB INSECURITY, EMPLOYABILITY, AND HEALTH: AN ANALYSIS FOR GERMANY ACROSS GENERATIONS	HCM
89-2014	Carsten Burhop, Sibylle H. Lehmann-Hasemeyer	THE GEOGRAPHY OF STOCK EXCHANGES IN IMPERIAL GERMANY	ECO
90-2014	Martyna Marczak, Tommaso Proietti	OUTLIER DETECTION IN STRUCTURAL TIME SERIES MODELS: THE INDICATOR SATURATION APPROACH	ECO
91-2014	Sophie Urmetzer, Andreas Pyka	VARIETIES OF KNOWLEDGE-BASED BIOECONOMIES	IK
92-2014	Bogang Jun, Joongho Lee	THE TRADEOFF BETWEEN FERTILITY AND EDUCATION: EVIDENCE FROM THE KOREAN DEVELOPMENT PATH	IK
93-2014	Bogang Jun, Tai-Yoo Kim	NON-FINANCIAL HURDLES FOR HUMAN CAPITAL ACCUMULATION: LANDOWNERSHIP IN KOREA UNDER JAPANESE RULE	IK
94-2014	Michael Ahlheim, Oliver Frör, Gerhard Langenberger and Sonna Pelz	CHINESE URBANITES AND THE PRESERVATION OF RARE SPECIES IN REMOTE PARTS OF THE COUNTRY – THE EXAMPLE OF EAGLEWOOD	ECO
95-2014	Harold Paredes-Frigolett, Andreas Pyka, Javier Pereira and Luiz Flávio Autran Monteiro Gomes	RANKING THE PERFORMANCE OF NATIONAL INNOVATION SYSTEMS IN THE IBERIAN PENINSULA AND LATIN AMERICA FROM A NEO-SCHUMPETERIAN ECONOMICS PERSPECTIVE	IK
96-2014	Daniel Guffarth, Michael J. Barber	NETWORK EVOLUTION, SUCCESS, AND REGIONAL DEVELOPMENT IN THE EUROPEAN AEROSPACE INDUSTRY	IK

IMPRINT

University of Hohenheim
Dean's Office of the Faculty of Business, Economics and Social Sciences
Palace Hohenheim 1 B
70593 Stuttgart | Germany
Fon +49 (0)711 459 22488
Fax +49 (0)711 459 22785
E-mail wiso@uni-hohenheim.de
Web www.wiso.uni-hohenheim.de