Matching with Indifferences:
A Comparison of Algorithms in the Context of Course Allocation

Franz Diebold and Martin Bichler[a,2,]

[a]*Technical University of Munich, Department of Informatics, Boltzmannstraße 3, 85748 Garching, Germany*

Abstract

We evaluate six one- and eight two-sided matching mechanisms with preferences based on a collection of 28 field data sets. Although important properties of matching mechanisms such as strategy-proofness or Pareto-efficiency can be shown by formal proofs, the size, the average rank, and the popularity of matchings ask for an empirical evaluation. We introduce different metrics to compare the results. The study shows trade-offs between various design desiderata, which are relevant in the field. The results provide guidelines for the selection of matching mechanisms.

Keywords: matching under preferences, course assignment, strategy-proofness, Pareto efficiency

Email address: `diebold@in.tum.de, bichler@in.tum.de` (Franz Diebold and Martin Bichler)

[1]Corresponding Author, Phone: +49 89 289 17500, Fax: +49 89 289 17504

[2]We thank David Manlove and Thayer Morrill for many helpful comments. This project is supported by the Deutsche Forschungsgemeinschaft (DFG) (BI 1057/7-1).

1. Introduction

In 2012 the Royal Swedish Academy of Sciences awarded Alvin E. Roth and Lloyd S. Shapley with the Nobel Memorial Prize in Economic Science "for the theory of stable allocations and the practice of market design" (Riksbank 2012). Many algorithms have been developed in this field starting with the seminal work by Gale and Shapley (1962) who introduced the *Deferred Acceptance* (DA) algorithm for the *Stable Marriage problem* (SM). The problem is two-sided as men and women have preferences over each other. The one-to-many generalization of SM is known as the *Hospitals / Residents problem* (HR), because the assignment of residents or students to hospitals has become an influential application. The one-to-many version of the DA algorithm, where students are proposing to the hospitals, is also referred to as *Student Optimal Stable Mechanism* (SOSM) algorithm. Gale and Shapley (1962) showed that when preferences of both sides are strict in a student-proposing version, a DA algorithm yields the unique stable matching that is Pareto superior for students to any other stable matching. A matching is *stable* if no student-school pair exists, where both prefer one another to their current assignment. The algorithm has a number of additional properties if preferences are strict: It is strategy-proof for students, and it is not Pareto-dominated by any other Pareto efficient mechanism that is strategy-proof. The DA algorithm is also favorable from a computational point of view, because the worst-case complexity is only quadratic.

Note that every stable matching is Pareto efficient with strict preferences of both sides. In course assignment, both course organizers and students might have preferences over the opposite side. In many cases, welfare analysis takes only students preferences into account, and treat course rankings over students as exogenous constraints, often referred to as priorities. In this paper, we focus on Pareto efficiency from the students' perspective.

In contrast to two-sided matching problems, in *one-sided matching problems* only one set of participants has preferences. In the *House Allocation problem* (HA) only applicants have preferences over houses. The one-to-many version where multiple applicants can be assigned to one house is called the *Capacitated House Allocation problem* (CHA) and houses are characterized by a fixed capacity. Some of the mechanisms for CHA can be adapted such that they allow for the houses or objects to have priorities over the agents, which is then similar to the school choice or course allocation problem. In addition to these wide-spread problems, many-to-many versions of the bipartite matching problem under preferences have been studied as well as problems which are not based on a bipartite graph, such as the *Stable Roommates* problem. Manlove (2013) provides an excellent overview of the various matching problems under preferences and respective algorithms to solve these problems. The literature is close to established fields in operational research, but adds stability and incentive compatibility as design desiderata.

In particular, one-to-many matching under preferences has found wide-spread application for *school choice*, *college admission*, or the *assignment of residents to hospitals*. In the school choice problem, for example, public school districts give parents the opportunity to choose the public school their child will attend. The problem is solved by considering preferences of both families and schools (Abdulkadiroğlu and Sönmez 2003). The assignment of students to courses at universities is similar and even more wide-spread. Such problems can be found in almost every university, although they are often solved via first-come first-served mechanisms in spite of the advantages that stable matching mechanisms provide (Diebold et al. 2014). All of these prominent applications are versions of HR or CHA problems, which we will focus on in this paper. Properties such as efficiency, incentives to report preferences truthfully, and stability of the outcome in two-sided markets are success factors in practical applications. But there are several other design desiderata, which asks for an empirical evaluation, in particular when preferences include indifferences.

1.1. Matching with Indifferences

The DA algorithm and its one-to-many extension require strict preferences. In most applications, however, preferences include ties. For example, course organizers might be indifferent among all students in a particular semester. Also, in school choice, school priorities are typically determined according to criteria that do not provide a strict ordering of all students, often leading to large indifference classes. We will refer to the two-sided version of the HR problem with ties as HRT. In school choice such ties in preferences are typically broken randomly to achieve a fixed strict ordering of the schools. Although random tie-breaking procedures preserve stability and strategy-proofness of the matching, they adversely affect the efficiency of the outcome as is illustrated in Example 1.

Example 1 Consider three course organizers c_1, c_2, and c_3 with a capacity of one each and three students s_1, s_2, and s_3 with the following preferences in Table 1. For example, course organizer c_2 is indifferent between the students s_1 and s_3, but prefers student s_2.

Table 1: Exemplary matching instance I with three students and three course organizers. Entries in round brackets are tied.

Students s_i	Preference list $\geq_{s_i}$			Course organizers c_j	Capacity q_j	Preference list $\geq_{c_j}$	
s_1:	c_2	c_1	c_3	c_1:	1	(s_1, s_2, s_3)	
s_2:	c_3	c_2	c_1	c_2:	1	s_2	(s_1, s_3)
s_3:	c_2	c_3	c_1	c_3:	1	(s_1, s_3)	s_2

If ties are broken such that s_1 is preferred to s_2, and s_2 is preferred to s_3, then the student optimal stable matching M would assign s_1 to c_1, s_2 to c_2, and s_3 to c_3. However, considering the original course organizer preferences with indifferences there are Pareto improvements by assigning student s_3 to course c_2 and student s_2 to course c_3. The new matching Pareto dominates M from the students' point of view and it is stable with respect to the original preferences of the course organizers.

Note that there are two notions of Pareto efficiency in two-sided matching, one where both sides are considered and one where only the students are considered in a student optimal stable matching. Every stable matching with strict preferences is Pareto efficient when both sides are considered, because no participant can be made better off without making some other worse off. However, when only the students are considered in the student optimal matching, then there can be Pareto improvements for students in general (Roth 2008) and also when we have preferences with ties as Example 1 illustrates. As indicated earlier, we will focus on the Pareto improvements for students in this paper.

1.2. Related Literature

Abdulkadiroğlu et al. (2009) show that every student optimal stable matching can be obtained with single tie-breaking, but the outcome might not be Pareto efficient. Single tie-breaking means that there is a single tie-breaking rule that is used for all courses in a course assignment problem for example. They show that there exists no strategy-proof mechanism (stable or not) that Pareto improves on the DA algorithm with single tie-breaking. This means that no other strategy-proof mechanisms such as serial dictatorship or top trading cycles Pareto dominate the DA algorithm. Ashlagi and Nikzad (2015) compare single and multiple tie-breaking rules for the DA algorithm in school choice depending on the balance between supply and demand in the matching market.

However, there may be stable matching mechanisms, which are not strategy-proof. Erdil and Ergin (2008) and Kesten (2010) introduced mechanisms (ESMA, WOSMA, and EADAM), which correct for efficiency losses in SOSM. When preferences involve ties, there is a non-empty set of stable matchings that are Pareto optimal for the students. These stable matchings may have different sizes and one might want to find the *maximum stable matching*, even if respective mechanisms are not strategy-proof. Manlove et al. (2002) already showed that ties can lead to computationally hard problems. Kwanashie and Manlove (2013) proposed an integer program to find the *maximum-cardinality stable matching* for the HR problem with ties (MAX HRT). This maximum-cardinality stable matching is not necessarily efficient, however. Empirical results on MAX HRT can be found in (Irving and Manlove 2010).

There is a trade-off between strategy-proofness, the maximization of the cardinality of a matching, and the efficiency of different stable matching mechanisms, but little is known about the order of magnitude of these trade-offs. Obviously, the cost of strategy-proofness depends on the distribution of preferences and can be analyzed using field data. Abdulkadiroğlu et al. (2009) analyzed a data set on student preferences from 2003-04 in New York City and compared algorithms developed by Erdil and Ergin (2008) to find student optimal matchings based on the preferences with ties and the outcome of the SOSM algorithm with single tie-breaking. They found a significant efficiency loss due to tie-breaking, which is in contrast to data from Boston's assignment system in 2005-06, where they report that the costs of strategy-proofness were negligible. It is important to better understand the cost of strategy-proofness and the differences between matching mechanisms. This has also been posed as a central research question in the seminal paper by Abdulkadiroğlu et al. (2009).

The same question can be asked for one-sided matching problems. For example, in course allocation students often need to be assigned to tutor groups at different times of the week. The material discussed in these tutor groups is identical within a week and tutors do not have preferences, but students do have preferences for different times in their weekly schedule. This is an example of a one-sided one-to-many matching, where students are often indifferent between several tutor groups. We refer to the problem as CHAT if ties are included. Stability is not an issue in these settings, but Pareto efficiency is. Apart from this the *popularity* and the *profile* of a matching matter. A matching is popular, if there is no other matching that is preferred by a majority of the agents. The profile describes the distribution of ranks resulting from a matching. We compare *Random Serial Dictatorship* (RSD), a well-known strategy-proof mechanism, to alternative algorithms which maximize popularity or the profile to estimate the cost of strategy-proofness.

1.3. Contributions

In this paper, we contribute to the central question raised by Abdulkadiroğlu et al. (2009) and provide an empirical analysis based on 28 data sets from course assignment applications, both one- and two-sided. Currently, there is limited literature comparing matching algorithms based on field data (Abdulkadiroğlu et al. 2009, Irving and Manlove 2010), and we are not aware of an analysis of the cost of strategy-proofness for matching with indifferences beyond the paper by Abdulkadiroğlu et al. (2009). We elicited the data and assigned courses using SOSM with random tie-breaking such that students had incentives to reveal their preferences truthfully. We calculate the cost of strategy-proofness by using the same preferences to compute matchings using different matching mechanisms. We analyze the existing one-sided and two-sided mechanisms, but also contribute two derivatives of existing mechanisms.

Apart from the theoretical properties, we use a number of different metrics to better understand the trade-offs between different mechanisms. We discuss the time complexity, the size, the popularity, and the average rank of resulting matchings. We also introduce the "*Area under the Profile Curve Ratio*" (AUPCR) as a convenient way to compare rank profiles of different matchings. The AUPCR up to a specific rank describes the probability that a matching mechanism will rank a randomly chosen student higher than his *n*-th preference. The study reveals trade-offs between design desiderata such as strategy-proofness, efficiency, size, and popularity of different mechanisms.

Apart from SOSM, we use the Top Trading Cycles algorithm (TTC) as a strategy-proof mechanism (for students) for two-sided matching problems. If one is willing to give up on strategy-proofness, but still wants weaker forms of truthfulness and Pareto efficiency for students in two-sided matching problems, then EADAM and WOSMA are close in terms of average size, average rank, and popularity. They provide Pareto improvements regarding students compared to SOSM with tie-breaking. IP MAX HRT computes a maximum cardinality stable matching via integer programming, but it is not truthful. We also introduce MESMA and MWOSMA, which are both based on IP MAX HRT, but compute Pareto improvements like in ESMA and WOSMA. Such mechanisms might well be an alternative in some applications, where participants have little prior knowledge.

In one-sided matching applications RSD is a simple and strategy-proof mechanism. The Top Trading Cycle (TTC) and successors provide strategy-proof alternatives. Probabilistic Serial (PS) is a non-strategy-proof but envy-free randomized alternative. Algorithms such as Pop-CHAT and ProB CHAT are more popular and have a better average rank. MPO CHA and MPO CHAT increase the size and the AUPCR of the matchings. RSD is the only strategy-proof one-sided mechanism in this study.

For both, the one-sided and two-sided mechanisms, we compute the cost of strategy-proofness with respect to size and AUPCR.

2. One-to-Many Matching Mechanisms

In what follows, we introduce definitions, metrics, and matching mechanisms analyzed in this paper.

2.1. Matchings, Matching Mechanisms, and their Properties

A *one-to-many matching problem* consists of two finite sets of *agents* $S = \{s_1, s_2, \ldots, s_n\}$ and $C = \{c_1, c_2, \ldots, c_m\}$ with $S \cap C = \emptyset$. The agents $s_i \in S$ have a unit capacity while the agents $c_j \in C$ have a positive integer *maximum capacity* $q_j \in \mathbb{Z}^+$. Let $Q = \sum_{j=1}^{m} q_j$ denote the total capacity of all agents $c_j \in C$. We will alternatively refer to students s_i and courses or course organizers c_j in this paper. To ensure that a feasible matching exists we assume $n \leq Q$.

Let $E \subseteq S \times C$ denote the set of *acceptable s-c-pairs*. For each agent in $s_i \in S$ (resp. $c_j \in C$) the set of *acceptable agents* in C (resp. S) is defined by:

$$\begin{aligned} A(s_i) &= \left\{c_j \in C : \left(s_i, c_j\right) \in E\right\}, \\ A\left(c_j\right) &= \left\{s_i \in S : \left(s_i, c_j\right) \in E\right\}. \end{aligned}$$

Each agent $a_k \in S \cup C$ has a *preference list* $\succeq_{a_k}$ ranking the acceptable agents in $A(a_k)$. For $a_u, a_v \in A(a_k)$ agent a_k is said to *strictly prefer* a_u to a_v if $a_u \succ_{a_k} a_v$ and *indifferent* between if $a_u \sim_{a_k} a_v$. Indifferences are also called *ties*. The *rank* of an agent a_u in the preference list $\succeq_{a_k}$ of an agent a_k is defined as one plus the number of agents that a_k prefers to a_u:

$$rank(a_k, a_u) = 1 + \left|\{a_v \in A(a_k) : a_v \succ_{a_k} a_u\}\right|.$$

Definition 2.1 (Matching) *A* matching *M is a subset of E: $M \subseteq E$. The agents $s_i \in S$ and $c_j \in C$ are said to be* assigned *if $\left(s_i, c_j\right) \in M$. Let $M(s_i)$ and $M\left(c_j\right)$ denote the set of* assignees *of s_i and c_j in M, resp.:*

$$\begin{aligned} M(s_i) &= \left\{c_j \in C : \left(s_i, c_j\right) \in M\right\}, \\ M\left(c_j\right) &= \left\{s_i \in S : \left(s_i, c_j\right) \in M\right\}. \end{aligned}$$

A matching M is feasible *if:*

(i) $|M(s_i)| \leq 1 \quad \forall\, s_i \in S$ *and*

(ii) $\left|M\left(c_j\right)\right| \leq q_j \quad \forall\, c_j \in C.$

Let $\mathcal{M}$ denote the set of all feasible matchings. One desirable property of matchings is *Pareto efficiency*, meaning that no student can be made better off without making any other student worse off:

Definition 2.2 (Pareto efficiency of matchings) *A matching M is* Pareto efficient *with respect to the students if there is no other feasible matching M' such that $M'(s) \succeq_{s_i} M(s_i)$ for all students $s_i \in S$ and $M'(s_i) \succ_{s_i} M(s_i)$ for some $s_i \in S$.*

Stability means that there should be no unmatched pair of a student and a course $\left(s_i, c_j\right)$ where student s_i prefers course c_j to her current assignment and she has higher priority than some other student who is assigned to course c_j. Stability can be seen as a property of a solution that has *no justified envy*:

Definition 2.3 (Stability) *A matching M is* stable *if $M(s') \succ_s M(s)$ implies $s' \succ_{M(s')} s$ for all $s, s' \in S$.*

The set of stable matchings is a subset of the Pareto optimal matchings considering both sides. Note that the notion of Pareto efficiency introduced above concerns only the students. In the absence of ties, a stable matching is Pareto optimal for both sides, but not necessarily for the students only, as we showed in Example 1; see Roth (1982) and Ergin (2002). An example in Abdulkadiroğlu and Sönmez (2003) also shows that there are strict preferences where a matching can either be stable or Pareto efficient, but not both.

Next we will discuss mechanisms to compute matchings and their properties. A mechanism computes a matching for given preferences of students and courses. More formally, a *mechanism* χ is a function $\chi : \mathcal{P}^{|A|} \rightarrow \mathcal{M}$ that returns a feasible matching $M \in \mathcal{M}$ of students to courses for every preference profile $\succeq_A \in \mathcal{P}^{|A|}$ of students and courses. For a submitted preference profile $\succeq_A \in \mathcal{P}^{|A|}$ of the students, $\chi(\succeq_A)$ is the associated matching. For a student $s \in S$ the assigned course is $\chi_s(\succeq_A) \in C$.

A mechanism is *Pareto efficient* if it always selects a Pareto efficient matching. Also, a mechanism is *stable* if it always selects a stable matching. Another important property of a mechanism is *strategy-proofness*. This means that there is no incentive for any student not to submit her true preferences, no matter what the other students submit:

Definition 2.4 (Strategy-proofness) *A mechanism χ is* strategy-proof *if for any preference profile $\succeq_A \in \mathcal{P}^{|A|}$, any student $s \in S$ and $\succeq_s'$ we have $\chi_s(\succeq_A) \succeq_s \chi_s(\succeq_s', \succeq_{A\setminus\{s\}})$.*

$(\succeq_s', \succeq_{A\setminus\{s\}})$ denotes the preference profile, where the preferences $\succeq_s'$ of student s differ from her true preferences $\succeq_s$.

Note that Roth (1982) has shown that there cannot be a stable and strategy-proof matching mechanism for both sides. For one-sided matching the result for strategy-proof and non-dictatorial mechanisms follows from Gibbard (1973, 1977). In the literature on school choice, schools are typically assumed to have publicly known priorities, but they are assumed not to be strategic, while the students are possibly strategic. These are reasonable assumptions also for many course assignment problems. In the following subsections, we describe metrics for matchings and matching mechanisms to solve one-to-many matching problems with preferences and summarize their properties. We start with two-sided matching mechanisms, then we continue with mechanisms for one-sided problems and metrics relevant to one-sided matching.

2.2. *Metrics for Matchings*

Efficiency and strategy-proofness are relevant to one-sided and two-sided matching problems. However, stability is not meaningful for one-sided matching problems. We will now introduce four additional metrics that are particularly useful for one-sided matchings.

The *size* of a matching simply describes the number of matched agents. Another possibility to compare different matchings is the *average rank* and the *median rank*. However, these metrics are only meaningful in combination with the size of the matching, because a smaller matching could easily have a smaller average rank if only certain students would be matched. We report the average average rank, because it has been used as a metric to gauge the difference in welfare of matching algorithms in Budish and Cantillon (2012) and Abdulkadiroğlu et al. (2009), two of the few experimental papers on matching mechanisms. Alternatively, one could compute the maximum k such that some student has their kth-choice course.

The *profile* contains more information as it compares how many students were assigned to their first choice, how many to their second choice, and so on. Similarly, for courses it compares how many students were assigned to the first choice of a course (see Figure 4 for an example of profiles for different matching mechanisms).

The profile of two matchings is not straightforward to compare. Second order stochastic dominance is one possibility (Levy 1992).[3] Risk-averse expected-utility maximizers prefer a second-order stochastically dominant gamble to a dominated one. We want to compare multiple profiles based on a single metric, and decided to use a metric similar to the Area under the Curve of a Receiver Operating Characteristic in signal processing (Hanley and McNeil 1982). The *Area Under the Profile Curve Ratio* (AUPCR) describes the area which is covered by the profile of a particular matching mechanism and is scaled between 0 and 100. The AUPCR up to a specific rank is equal to the probability that a matching mechanism will match a randomly chosen student higher than his n-th preference.

Definition 2.5 (Area Under Profile Curve Ratio) *The* Area Under the Profile Curve Ratio *(AUPCR) is the ratio of the Area Under the Profile Curve (AUPC) and the total area (TA), where the AUPC describes the integral below the profile curve.*
Regarding the students it is defined as follows:

$$
\begin{aligned}
TA_S(M) &= |C| \cdot \min(|S|, Q) \\
AUPC_S(M) &= \sum_{r=1}^{|C|} \left| \left(s_i, c_j\right) \in M : rank\left(s_i, c_j\right) \leq r \right| \\
AUPCR_S(M) &= \frac{AUPC_S(M)}{TA_S(M)}
\end{aligned}
$$

Finally, the *popularity* helps to compare two matchings with each other. A matching M is *popular* if it is preferred by the majority of the agents. The formal definition is as follows:[4]

Definition 2.6 (Popularity (Manlove 2013, Abraham et al. 2007)) *Let $M, M' \in \mathcal{M}$ and $P(M, M')$ be the number of agents preferring M to M'. M is* more popular than *M' ($M \blacktriangleright M'$), M is* equally popular than *M' ($M \approx M'$) and M is* popular *if the following equivalences hold:*

$$
\begin{aligned}
M \blacktriangleright M' \quad &\Leftrightarrow \quad P(M, M') > P(M', M), \\
M \approx M' \quad &\Leftrightarrow \quad P(M, M') = P(M', M), \\
M \in \mathcal{M} \text{ is popular} \quad &\Leftrightarrow \quad \forall\, M' \in \mathcal{M} : M \blacktriangleright M'.
\end{aligned}
$$

[3] In contrast to cumulative distribution functions, a more preferable profile is to the left and not to the right of a profile.

[4] An alternative to popularity are the unpopularity factor or unpopularity margin described in (Manlove 2013, Sec. 7.2.7).

Please note that the *more popular than* relation $\blacktriangleright$ is not transitive (Abraham et al. 2007).

2.3. Two-Sided Matching Mechanisms

Let us now briefly introduce the two-sided matching mechanisms analyzed in our study.

2.3.1. Gale-Shapley Student-Optimal Stable Mechanism (SOSM)

The *Gale-Shapley Student-Optimal Stable Mechanism* (SOSM) is an extended version of the *Deferred Acceptance algorithm* from Gale and Shapley (1962), which allows for one-to-many assignments. SOSM is *stable* and *strategy-proof* (Abdulkadiroğlu and Sönmez 2003). The SOSM is stable when all preferences are strict. Moreover it is constrained efficient for students in the sense that the outcome is not Pareto dominated by any other stable matching. If there are ties, independent of how ties are broken the outcome is still stable, however is not guaranteed to be constrained efficient any more (Erdil and Ergin 2008). Also, the SOSM matching with ties does not need to be maximal, because with ties multiple stable matchings with different sizes may exist. The time complexity of SOSM is $O(|E|)$ as every acceptable student-course pair ($E \subseteq S \times C$) is at most considered once (Gusfield and Irving 1989).

2.3.2. Efficiency Adjusted Deferred Acceptance Mechanism (EADAM)

Kesten (2010) developed the *Efficiency Adjusted Deferred Acceptance Mechanism* (EADAM), in order to recover the efficiency losses in SOSM. EADAM takes a SOSM matching and identifies student-course pairs (called *interrupting pairs*) that may harm the welfare when constructing the SOSM matching. Therefore, the preferences of the student are modified if the student consents to waive her preferences, and SOSM is executed on the updated preferences. This procedure is repeated until no interrupting pair exists any more. The algorithm only discovers Pareto improvements such that no student would be worse off compared to the SOSM matching, even if she waives preferences. There is no other fair matching that Pareto dominates the EADAM outcome. A similar algorithm can also be used to handle inefficiency due to ties. One can combine the EADAM algorithm proposed for the case when priority orders are strict with the one to handle ties as a way to achieve full Pareto efficiency.

EADAM achieves Pareto efficiency in presence of ties regarding students, but it relaxes strategy-proofness and stability. The matching may become unstable concerning the original preferences. EADAM is not strategy-proof, but Kesten (2010) proved that truth-telling is an ordinal Bayesian Nash equilibrium in EADAM. A strategy is an ordinal Bayesian Nash equilibrium if it is a Bayesian Nash equilibrium for every possible von Neumann-Morgenstern utility representation of students true preferences. The time complexity of EADAM is $O\left(|E|^2\right)$ (Kesten 2010).

2.3.3. Top Trading Cycle

The Top Trading Cycle (TTC) mechanism was described by Shapley and Scarf (1974) and designed for housing markets, where agents initially hold an object. Subsequently, the agents trade among themselves to return the final allocation. TTC satisfies individual rationality, Pareto-efficiency, strategy-proofness, and the outcome is in the *core* with strict preferences. The TTC algorithm can also be adapted for one-to-many and *two-sided matching problems* such as the course allocation problem, where students and course organizers have preferences, and more than only a single course seat is available. While this TTC algorithm for two-sided matching problems is Pareto efficient and strategy-proof, it is not stable (Abdulkadiroğlu and Sönmez 2003). It is also not necessarily Pareto efficient in the presence of ties.

2.3.4. Efficient and Stable Matching Algorithm (ESMA)

Erdil and Ergin (2006, 2008) developed the *Efficient and Stable Matching Algorithm* (ESMA) to address ties in the preferences. ESMA also takes a SOSM matching and performs Pareto improvements, which preserve stability by exchanging course seats in the presence of ties. As a result, there is no stable matching which Pareto dominates the outcome. However, ESMA does not necessarily lead to a maximum matching, and it is not strategy-proof. The authors showed, however, that truth-telling is also an ordinal Bayesian Nash equilibrium in ESMA in an environment with symmetric information. The time complexity of ESMA is $O\left(|S|^3 \cdot Q\right)$.

2.3.5. Worker Optimal Stable Matching Algorithm (WOSMA)

Erdil and Ergin (2006, 2008) also suggested the *Worker Optimal Stable Matching Algorithm* (WOSMA), which is similar to ESMA. It computes a student-optimal stable matching by carrying out *stable student improvement* chains and cycles based on a SOSM matching. WOSMA is again stable, but not Pareto efficient for students (Erdil and Ergin 2006, 2008). Similar to ESMA, WOSMA does not necessarily lead to a maximum matching and it is not strategy-proof. In contrast to EADAM, WOSMA does not change the students preference lists. The time complexity of WOSMA is $O\left(|S|^3 \cdot |C|\right)$.

2.3.6. Integer Program for MAX HRT (IP MAX HRT*)*

Stable matchings in HRT instances may have different sizes. For the problem of finding a maximum cardinality stable matching (MAX HRT) (Rothblum 1992) already provided integer programming formulations. In addition, there were constraint programming approaches to HRT and related problems (see Manlove (2013) for an overview). In our analysis, we draw on a recent integer programming formulation by Kwanashie and Manlove (2013) for MAX HRT (IP MAX HRT). The solution of this integer program is a stable matching of maximum cardinality. Even though the problem is NP-complete, typically instances with hundreds of students can be solved to optimality in practice. The IP MAX HRT is an NP-complete problem (Manlove et al. 2002), and there is some literature on approximation algorithms for this problem (Király 2013). Note that the outcome does not need to be Pareto efficient for students.

2.3.7. Maximum Efficient and Stable Matching Algorithm (MESMA)

Similar to ESMA, Pareto improvement chains and cycles in an IP MAX HRT matching can be found using the ESMA approach. The difference is that the algorithm starts with the result of IP MAX HRT rather than SOSM. We refer to this procedure as *Maximum Efficient and Stable Matching Algorithm* (MESMA). This way we can gain stability and Pareto efficiency with maximum cardinality matchings.

Let us briefly describe how ESMA and MESMA find Pareto improvements in a stable matching. Regarding the matching instance I in Example 1 and a possible stable matching $M_0 = \{(s_1, c_3), (s_2, c_2), (s_3, c_1)\}$ computed by IP MAX HRT, the stable Pareto improvement would work as follows. The idea is to construct a 2-labeled graph Γ^M for a stable matching M.

Erdil and Ergin (2006) proved that a stable matching is Pareto efficient if and only if it does not admit a Pareto improvement cycle or chain. The 2-labeled directed graph Γ^{M_0} corresponding to the matching M_0 is illustrated in Figure 1.

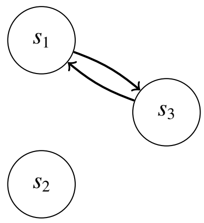


Figure 1: Example of MESMA applied to the matching instance I in Table 1

For example, there is an edge from s_1 to s_3 because $M_0(s_3) \succeq_{s_1} M_0(s_1)$ and $s_1 \succeq_{M_0(s_3)} s_3$. The graph contains a cycle $C = \langle s_1, s_2 \rangle$ representing a stable Pareto improvement cycle. It can be carried out by exchanging the course seats accordingly. Therefore, student s_1 is assigned to course c_1 in M_1, where s_3 was assigned to in M_0. Both, M_0 and

Table 2: Matching instance I from Example 1 with $\underline{M_0}$ and $\boxed{M_1}$

Students s_i	Preference list $\succeq_{s_i}$			Course organizers c_j	Capacity q_j	Preference list $\succeq_{c_j}$	
s_1:	c_2	$\boxed{c_1}$	$\underline{c_3}$	c_1:	1	$\left(\boxed{s_1}, s_2, \underline{s_3}\right)$	
s_2:	c_3	$\underline{\boxed{c_2}}$	c_1	c_2:	1	$\underline{\boxed{s_2}}$	(s_1, s_3)
s_3:	c_2	$\boxed{c_3}$	$\underline{c_1}$	c_3:	1	$\left(\underline{s_1}, \boxed{s_3}\right)$	s_2

M_1 are shown in Table 2. Hence, the matching $M_1 = \{(s_1, c_1), (s_2, c_2), (s_3, c_3)\}$ Pareto dominates M_0 regarding both students and course organizers while still being stable.

2.3.8. Maximum Worker Optimal Stable Matching Algorithm (MWOSMA)

Just as MESMA improves on the IP MAX HRT matching using the ESMA approach, MWOSMA does the same using stable student improvement chains and cycles from WOSMA. Therefore, the result is also stable and of maximum size. However, MWOSMA is not strategy-proof since IP MAX HRT is not strategy-proof either.

A summary of the theoretical properties and time complexity as well as empirical results is given in Table 20 in the results section.

2.4. *One-Sided Matching Mechanisms*

One-sided matching mechanisms are often described in the context of house allocation problems. We will continue to use course assignment problems as an example, but assume that course organizers do not have preferences in this context.

2.4.1. *Serial Dictatorship*

The *Serial Dictatorship* mechanism (SD) can be considered as a classical mechanism to construct a Pareto optimal matching. Based on an ordering of the students S the mechanism takes each student in turn and assigns her to the most-preferred under-subscribed course. The ordering can be based on priority or the outcome of a lottery. In the latter case the mechanism is called *Random Serial Dictatorship* (RSD).

The RSD mechanism is Pareto efficient and strategy-proof. The size of the matching is not necessarily maximum, due to the greedy approach of assigning students to courses and can have different sizes depending on the ordering. The time complexity of RSD is $O(|E|)$, because every acceptable student-course pair is considered at most once. The RSD mechanism can also be applied to two-sided matching problems. However, it may be unstable for two-sided problems.

In the presence of ties, RSD is not necessarily efficient. Suppose there is a student s_1, who is indifferent between courses c_1 and c_2, and a student s_2 only interested in c_1. If the tie in s_1s list is broken in favor of c_1, and the policy for RSD is $< s_1, s_2 >$, then s_1 is matched to c_1 which is not Pareto optimal.

2.4.2. *Popular-CHAT*

The *Popular-CHAT* mechanism (Pop-CHAT) (Manlove and Sng 2006) aims to find the largest popular matching in a CHAT instance, if a popular matching exists. This is not always the case (see for example the results on data set OS2 in our empirical analysis). The idea of Pop-CHAT is derived from a Theorem proposed by Manlove and Sng (2006), which characterizes popular matchings. As every popular matching is Pareto optimal, Popular-CHAT is also Pareto optimal as long as a popular matching exists. However, Popular-CHAT is not strategy-proof, meaning that a student could benefit from concealing her true preferences. The time complexity of Popular-CHAT is $O\left(\left(\sqrt{Q} + |S|\right)|E|\right)$ (Manlove and Sng 2006).

2.4.3. *Profile-based optimal CHAT*

A rank-maximal matching can be constructed using the *Profile-based optimal CHAT* mechanism (ProB CHAT) by Sng (2008). A matching is rank-maximal if its profile is lexicographically maximum over all possible matchings. ProB CHAT proceeds incrementally by regarding ranks in an ascending manner. ProB CHAT is Pareto efficient but not strategy-proof. The time complexity is $O\left(\min\left(z^* \sqrt{Q}, Q + z^*\right)|E|\right)$ where z^* is the maximal rank of an edge in an optimal solution (Sng 2008). A recent extension of the ProB CHAT algorithm to compute a maximum cardinality matching considering the profile can be found in Kwanashie et al. (2015). Our analysis is based on the original paper by Sng (2008).

2.4.4. Maximum Pareto optimal CHA

The *Maximum Pareto optimal CHA* mechanism (MPO CHA) by Sng (2008) aims to find a maximum matching which is Pareto optimal. It proceeds in three phases satisfying three conditions for a Pareto optimal matching (maximal, trade-in-free and cyclic coalition-free) stated by Sng (2008) and Manlove (2013). MPO CHA is not strategy-proof. The time complexity is $O\left(\sqrt{n}\,|E|\right)$ (Sng 2008, Manlove 2013). However, MPO CHA cannot handle ties in preference lists.

2.4.5. Maximum Pareto optimal CHAT

The *Maximum Pareto optimal CHAT* mechanism (MPO CHAT) is an extension of the MPO CHA mechanism where the phase for searching cyclic coalitions is adapted to handle ties. Therefore, a 2-labeled envy graph, comparable to Erdil and Ergin (2006) (WOSMA), is built and searched for cycles. These cycles represent cyclic coalitions. The time complexity of MPO CHAT is $O\left((|E| + |S| + |C|)\,|E|\right)$. A very recent strategy-proof and Pareto efficient mechanism for CHAT is given in Cechlárová et al. (2015).

2.5. Mechanisms Not Considered in the Study

There is a rapidly growing literature on one-sided and two-sided matching mechanisms. While we tried to cover the most important mechanisms in the context of course allocation problems, we cannot possibly analyze all mechanisms. For example, there are several new developments in the field of randomized assignment problems. Probabilistic Serial (PS) is a randomized matching mechanism, which yields a lottery on a set of objects in the house allocation problem. The PS mechanism produces an envy-free assignment with respect to the reported preferences, but it is not strategy-proof (Bogomolnaia and Moulin 2001). There are various extensions of the initial house allocation problem (Kojima 2009), but we decided to focus largely on deterministic algorithms in our study.

A few articles address housing markets with indifferences. The Top Trading Absorbing Sets (TTAS) (Alcalde-Unzu and Molis 2011) and the Top Cycle Rules (TCR) (Jaramillo and Manjunath 2012) are strategy-proof, core selecting, and Pareto optimal. Aziz and De Keijzer (2012) highlights commonalities among these mechanisms. Housing markets require an initial endowment, which is also different to the assignment problems discussed in this paper.

3. Field Data

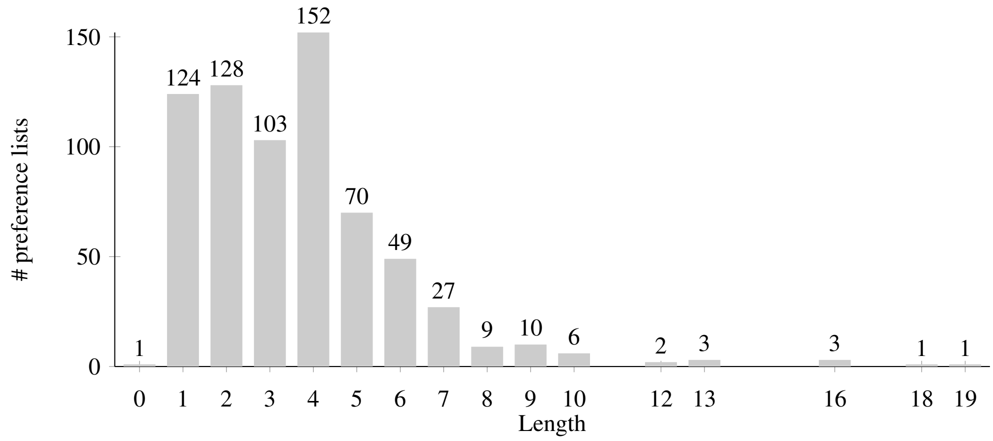
The data sets describe different matching instances in the context of course allocation. We collected the data sets from a course allocation system that is used at the Technical University of Munich at the Department of Informatics. The participants were told that the SOSM will be used to match students to courses. Students had a short introduction and they knew that the mechanism is strategy-proof for students. Strategic manipulation by course organizers is unlikely because they did not have information about student preferences and therefore, we assume that all participants revealed their true preferences.

3.1. Two-Sided Preferences

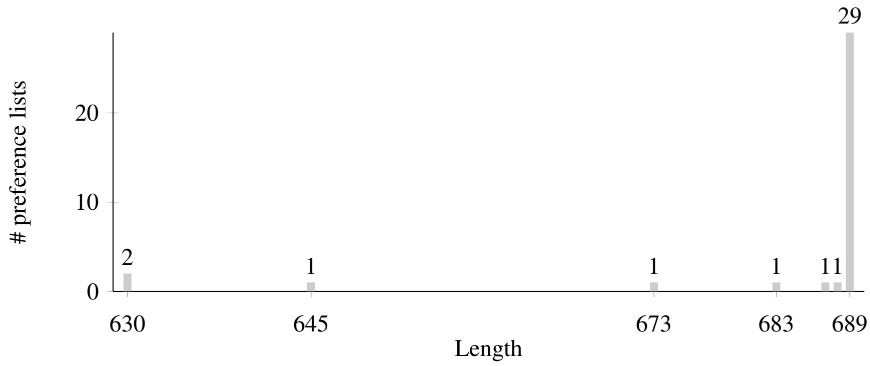
Table 3: Summary of Two-Sided data sets

Name	Matching instance	#Students	#Courses	Total capacity	Ties	Incomplete lists	
						Students	Courses
TS1	HRT	539	26	575	✓	✓	
TS2	HRT	113	6	72	✓	✓	
TS3	HRT	557	38	459	✓	✓	
TS4	HRT	27	6	59	✓	✓	
TS5	HRT	88	12	116	✓	✓	
TS6	HRT	689	36	726	✓	✓	✓
TS7	HRT	662	39	637	✓	✓	✓
TS8	HRT	314	13	292	✓	✓	✓
TS9	HRT	27	13	94	✓	✓	✓
TS10	HRT	57	9	86	✓	✓	✓
TS11	HRT	18	3	36	✓	✓	✓
TS12	HRT	636	40	758	✓	✓	✓
TS13	HRT	595	41	549	✓	✓	✓
TS14	HRT	105	11	110	✓	✓	✓
TS15	HRT	78	16	253	✓	✓	✓
TS16	HRT	731	40	775	✓	✓	✓
TS17	HRT	733	43	753	✓	✓	✓
TS18	HRT	426	14	264	✓	✓	✓
TS19	HRT	99	4	44	✓	✓	✓

For the group of two-sided matchings we collected 19 data sets from the registration for seminars and practical courses for both bachelor and master students. Both students and course organizers could provide their preferences by



(a) Students' preference lists



(b) Preference lists of course organizers

Figure 2: Histograms of the length of preference lists in data set TS6

submitting a sorted list containing indifferences. The data sets were collected in the time period between June 2014 and March 2016. The characteristics of the data sets are described in Table 3. Seven data sets (TS2, TS3, TS7, TS8, TS13, TS18, TS19) represent competitive matching instances, where the number of students exceeds the total course capacity. In the data sets TS1 to TS5, all students are acceptable for course organizers, meaning that no incomplete course organizer preference list exists.

Let us provide a few additional statistics on the distribution of preferences in data set TS6, which is representative for other data sets as well. Figure 2 shows histograms of the length of the preference lists of students (a) and course organizers (b) in data set TS6. The majority of students submitted preferences for one to four courses, marking all other courses as unacceptable. One student submitted an empty preference list, meaning that all courses are unacceptable, while one student gave preferences over 19 courses. 29 out of 36 course organizers in TS6 did not mark any student as unacceptable. Two course organizers submitted preference lists of length 630, meaning that 59 students were unacceptable.

3.2. *One-Sided Preferences*

For matchings with one-sided preferences we collected nine data sets from the registrations to tutorials for both bachelor and master students in the period between October 2012 and October 2015. Table 4 shows a summary of the nine one-sided data sets. All data sets contain incomplete lists, meaning that students were able to mark courses as unacceptable in their preference lists. All but two data sets (OS1 and OS2) contain ties.

Table 4: Summary of One-Sided data sets

Name	Matching instance	#Students	#Courses	Total capacity	Ties	Incomplete lists
OS1	CHA	136	8	136		✓
OS2	CHA	418	47	429		✓
OS3	CHAT	915	51	1080	✓	✓
OS4	CHAT	114	7	266	✓	✓
OS5	CHAT	156	6	180	✓	✓
OS6	CHAT	1035	39	1282	✓	✓
OS7	CHAT	522	21	626	✓	✓
OS8	CHAT	248	8	336	✓	✓
OS9	CHAT	106	5	130	✓	✓

Again, we describe one representative large data set. The histogram of the length of the preference lists of students in data set OS3 is illustrated in Figure 3. It shows that the majority of students gave preferences over 10 to 15 tutorials, while 6 students submitted preferences over all 51 tutorials.

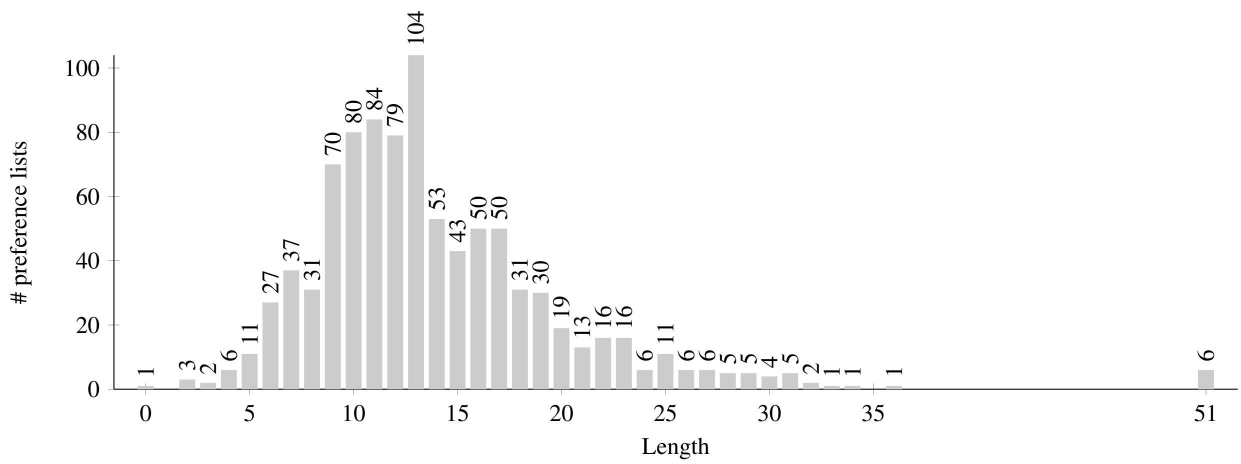


Figure 3: Histogram of the length of the preference lists of students in data set OS3

4. Results of Field Experiments

In what follows, we report on the main results of our analysis. All matching mechanisms were implemented in Python 3.4.2 and run on a machine with a 2.6 GHz dual-core processor with 16 GB of RAM running OS X 10.11.3. The IP Solver used in IP MAX HRT MESMA and MWOSMA was the open-source solver COIN CBC[5] using the PuLP[6] LP modeler. For breaking ties we used multiple tie breakers for both students and course organizers. This means that for every participant (student and course organizer) ties were broken independently. Since some of the matching mechanisms depend on chance due to random tie breaking and may result in different outcomes in different runs we repeatedly executed each mechanism. The results presented in the following sections show the average metrics of 100 runs of each mechanism.

4.1. Two-Sided Matching

We applied the mechanisms described in Section 2.3 to the 19 data sets introduced in Section 3.1, assuming that participants would also be truthful in mechanisms that are not strategy-proof. We will first analyze the results of a single but representative data set (TS6) before we provide summary results of metrics for all 19 data sets for two-sided matchings.

Table 5: Comparison of profiles for students in data set TS6. Absolute numbers and percentages in parenthesis.

Rank	SOSM		EADAM		TTC		ESMA		WOSMA		IP MAX HRT		MESMA		MWOSMA	
1	492.4	(71.5)	514.2	(74.6)	**525.0**	(**76.2**)	498.2	(72.3)	514.0	(74.6)	490.0	(71.1)	491.0	(71.3)	498.0	(72.3)
2	81.9	(11.9)	68.2	(9.9)	**51.5**	(**7.5**)	81.5	(11.8)	70.0	(10.2)	93.0	(13.5)	92.0	(13.4)	89.0	(12.9)
3	24.0	(3.5)	18.2	(2.7)	**15.6**	(**2.3**)	23.2	(3.4)	21.8	(3.2)	36.0	(5.2)	36.0	(5.2)	34.0	(4.9)
4	12.8	(1.9)	11.3	(1.6)	**10.5**	(**1.5**)	11.8	(1.7)	10.8	(1.6)	20.0	(2.9)	20.0	(2.9)	18.0	(2.6)
5	2.7	(0.4)	2.3	(0.3)	**2.3**	(**0.3**)	2.4	(0.4)	2.2	(0.3)	6.0	(0.9)	6.0	(0.9)	6.0	(0.9)
6	0.3	(0.0)	0.1	(0.0)	**0.1**	(**0.0**)	0.1	(0.0)	0.0	(0.0)	-		-		-	
8	0.1	(0.0)	0.0	(0.0)	-		0.1	(0.0)	-		-		-		-	
9	0.4	(0.1)	0.3	(0.0)	**0.3**	(**0.1**)	0.3	(0.0)	0.2	(0.0)	3.0	(0.4)	3.0	(0.4)	3.0	(0.4)
10	0.0	(0.0)	-		-		-		-		-		-		-	
∞	74.5	(10.8)	74.5	(10.8)	**83.6**	(**12.1**)	71.3	(10.4)	70.0	(10.2)	41.0	(6.0)	41.0	(6.0)	41.0	(6.0)

Table 5 describes the profiles for the eight matching mechanisms executed on data set TS6. For example, 525.0 students on average are assigned to their first choice using TTC. Furthermore, 83.6 students on average are not matched (∞) using TTC, which is maximum compared to the other mechanisms. The numbers in parentheses denote the percentage of all students. Table 5 shows that TTC results in the rank-maximal profile, because most students are

[5] `http://www.coin-or.org/projects/Cbc.xml`, last accessed on March 31st, 2016

[6] `https://github.com/coin-or/pulp`, last accessed on March 31st, 2016

matched to their first choice course. Regarding the AUPCR metric, however, MWOSMA shows the best results since it performed best regarding the entire profile including the number of unmatched students.

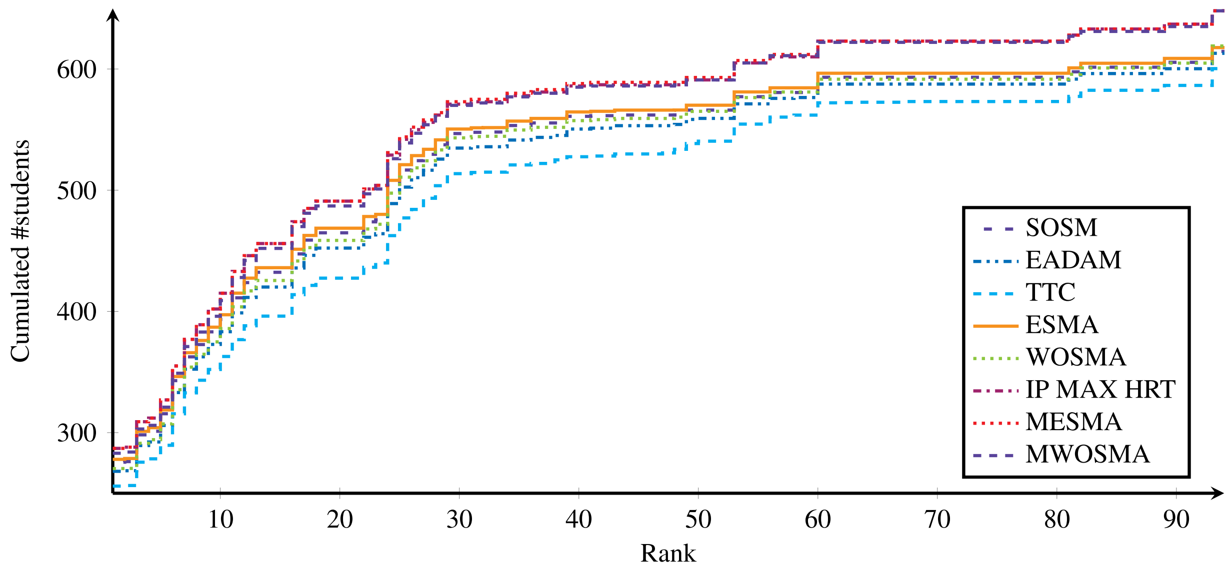


Figure 4: Comparison of profiles for courses in data set TS6

A similar set of profiles can be computed for the course organizers. The profiles for course organizers for data set TS6 can be found in detail in the Appendix A.1. Figure 4 illustrates the cumulated profiles for course organizers as profile curves from which the AUPCR metric is computed. For instance, about 470 students are assigned to rank 20 or lower using ESMA.

Table 6: Comparison of popularity for students in data set TS6

	EADAM	TTC	ESMA	WOSMA	IP MAX HRT	MESMA	MWOSMA
SOSM	◀ (0.0, 25.6)	◀ **(31.3, 62.4)**	◀ (0.0, 9.0)	◀ (0.0, 27.2)	▶ (48.8, 45.8)	▶ (48.5, 46.5)	◀ (45.2, 49.5)
EADAM		◀ **(35.1, 44.8)**	▶ (22.3, 5.4)	◀ (12.0, 13.3)	▶ (64.5, 39.8)	▶ (63.6, 39.9)	▶ (58.1, 40.5)
TTC			▶ **(58.2, 34.4)**	▶ **(47.6, 38.9)**	▶ **(84.4, 53.1)**	▶ **(83.4, 53.1)**	▶ **(78.9, 53.8)**
ESMA				◀ (3.1, 21.4)	▶ (50.8, 40.0)	▶ (50.4, 40.6)	▶ (46.8, 43.7)
WOSMA					▶ (61.7, 35.2)	▶ (60.7, 35.3)	▶ (54.7, 35.8)
IP MAX HRT						◀ (0.0, 1.0)	◀ (0.0, 10.0)
MESMA							◀ (0.0, 9.0)

Table 6 describes the popularity regarding students in data set TS6. For example, 48.8 students on average prefer the SOSM matching to the IP MAX HRT matching, while only 45.8 students on average prefer the IP MAX HRT matching to the SOSM matching. Thus SOSM is more popular than IP MAX HRT in this application. SOSM is more popular than MESMA as well, but the difference is smaller ($48.5 - 46.5 = 2.0$) than compared to IP MAX HRT and

SOSM (48.8 − 45.8 = 3.0). The reason for this is that MESMA improves on the IP MAX HRT matching. When comparing IP MAX HRT and MESMA that MESMA is more popular than IP MAX HRT with one student prefering the MESMA matching to the IP MAX HRT matching. The improvement of this student also becomes evident in the comparison of the profiles of these two mechanisms, where one more student is matched to her first choice using MESMA (see Table 5).

In terms of popularity, IP MAX HRT is dominated by all other matching mechanisms, while TTC dominates all other mechanisms in TS6.

It is now interesting to look at summary statistics across all data sets to better understand the trade-offs. We will first look at the AUPCR metric for students and for course organizers to estimate the cost of strategy-proofness. Thereafter, we will consider the average ranks of all mechanisms, the sizes, their popularity, and the runtimes to compute them.

Table 7 shows the AUPCR for students, which is highest for MWOSMA in all but three (TS2, TS18, TS19) of the 19 data sets. In TS2 and TS18 EADAM achieves the highest AUPCR, while TTC results in the best AUPCR in TS19. These three data sets (TS2, TS18, TS19) represent the most competitive matching instances of all data sets and thus lead to slightly different results. Note that all EADAM, ESMA and WOSMA show better results than SOSM regarding AUPCR, because they all improve efficiency of the SOSM matching. The same holds for the MESMA and MWOSMA compared to IP MAX HRT. Overall, MWOSMA results in the highest average AUPCR of 89.70% (see bottom row in Table 7).

We also report the results of a one-sided Wilcoxon signed-rank test to see if the differences in the ranks of different matching mechanisms are statistically significant (at a significance level of 5%). The results are shown in Table 8. The test supports the findings of the AUPCR. For example EADAM shows significantly better results than IP MAX HRT in one of the 19 data sets, whereas IP MAX HRT is significantly better in four of the 19 data sets. In the remaining 14 data sets no statistically significant result can be observed. Table 8 indicates that MWOSMA performs best compared to the other mechanisms because it computes significantly better results in many data sets and no other mechanism shows significantly better results than MWOSMA in any data set.

The sizes of the matchings are shown in Table 9, and the results are in line with those of the AUPCR. All IP MAX HRT, MESMA and MWOSMA result in the maximum matchings, because they optimize the cardinality of the matchings. Table 9 also illustrates, that ESMA and WOSMA can increase the size of the SOSM matching. The last row shows the average of the size compared to the maximum size. Note that TTC results in the smallest average matching with 94.39% of the maximum size.

The average rank (see Table 10) is an alternative metric to compare different matchings. It does not consider the entire profile curve and yields different results. Typically, either EADAM, TTC or WOSMA result in the best (lowest) average rank across all mechanisms for the different data sets. The median rank is almost always 1 for our data sets.

Table 11 describes the popularity of the different mechanisms, i.e., the students' point of view comparing the outcome of two different matchings. It shows in how many data sets one mechanism is more popular than another

Table 7: Comparison of AUPCR for students for two-sided data sets

	SOSM	EADAM	TTC	ESMA	WOSMA	IP MAX HRT	MESMA	MWOSMA
TS1	92.54%	92.82%	91.28%	93.21%	93.39%	97.72%	98.27%	**98.40**%
TS2	79.82%	**87.69**%	87.28%	83.18%	85.94%	82.18%	85.42%	87.50%
TS3	89.77%	90.43%	89.37%	90.29%	90.65%	94.26%	94.64%	**95.09**%
TS4	68.83%	69.16%	69.14%	69.73%	69.65%	74.69%	**75.93**%	**75.93**%
TS5	90.63%	90.96%	90.60%	91.25%	91.42%	93.94%	95.17%	**95.36**%
TS6	88.45%	88.58%	87.35%	88.94%	89.21%	92.95%	92.95%	**93.01**%
TS7	86.31%	86.67%	85.87%	86.78%	87.04%	94.89%	94.98%	**95.01**%
TS8	91.25%	92.43%	89.86%	92.26%	93.22%	94.20%	95.26%	**95.94**%
TS9	48.41%	48.41%	48.41%	48.41%	48.41%	**50.43**%	**50.43**%	**50.43**%
TS10	95.29%	95.31%	95.26%	95.34%	96.36%	**97.47**%	**97.47**%	**97.47**%
TS11	80.56%	80.56%	81.33%	80.56%	**83.33**%	**83.33**%	**83.33**%	**83.33**%
TS12	92.36%	92.43%	91.40%	92.85%	93.18%	97.04%	**97.11**%	**97.11**%
TS13	82.66%	83.19%	81.74%	83.58%	83.68%	87.90%	88.47%	**88.74**%
TS14	83.19%	83.67%	82.37%	83.89%	84.20%	87.27%	87.53%	**88.05**%
TS15	97.88%	97.88%	97.92%	97.88%	99.27%	99.60%	99.60%	**99.68**%
TS16	88.43%	88.53%	87.51%	88.92%	89.20%	93.45%	93.45%	**93.48**%
TS17	85.30%	85.49%	84.12%	85.92%	86.38%	**92.20**%	**92.20**%	**92.20**%
TS18	90.03%	**95.33**%	94.59%	93.82%	94.61%	90.58%	94.40%	95.16%
TS19	74.81%	86.47%	**87.50**%	82.55%	83.43%	73.30%	81.82%	82.39%
Avg.	84.55%	86.11%	85.42%	85.76%	86.45%	88.28%	89.39%	**89.70**%

one. For example EADAM is more popular than WOSMA for students in four of the 19 data sets, while WOSMA is more popular than EADAM in 14 of the 19 data sets. In the remaining data set they are equally popular. In total, WOSMA is more popular than all other mechanisms in most of the cases.

The runtimes reported in Table 12 are consistent with the theoretical complexities. The runtimes are all within seconds except for IP MAX HRT, MESMA and MWOSMA in TS6 and TS16, where the integer program took more than six hours to solve to optimality. All other data sets took significantly less time, often only seconds or minutes.

We'll provide an overall summary of AUPCR, popularity, average rank, and average size in Table 20.

Table 8: Comparison of one-sided Wilcoxon signed-rank tests for students for two-sided data sets

	EADAM	TTC	ESMA	WOSMA	IP MAX HRT	MESMA	MWOSMA
SOSM	(0, 12)	(0, 3)	(0, 13)	(0, 14)	(0, 5)	(0, 9)	(**0**, **10**)
EADAM		(0, 0)	(4, 0)	(0, 3)	(1, 4)	(0, 6)	(**0**, **6**)
TTC			(0, 0)	(0, 1)	(1, 5)	(0, 7)	(**0**, **7**)
ESMA				(0, 9)	(0, 5)	(0, 5)	(**0**, **8**)
WOSMA					(2, 0)	(0, 1)	(**0**, **1**)
IP MAX HRT						(0, 10)	(**0**, **11**)
MESMA							(**0**, **4**)

Table 9: Comparison of sizes for two-sided data sets

	#Students	Capacity	SOSM	EADAM	TTC	ESMA	WOSMA	IP MAX HRT	MESMA	MWOSMA
TS1	539	575	505.26	505.26	497.02	507.48	507.38	**538.00**	**538.00**	**538.00**
TS2	113	72	**72.00**	**72.00**	**72.00**	**72.00**	**72.00**	**72.00**	**72.00**	**72.00**
TS3	557	459	422.76	422.76	417.69	423.95	424.26	**449.00**	**449.00**	**449.00**
TS4	27	59	22.83	22.83	22.69	22.99	22.93	**26.00**	**26.00**	**26.00**
TS5	88	116	82.06	82.06	81.78	82.24	82.25	**87.00**	**87.00**	**87.00**
TS6	689	726	614.52	614.52	605.38	617.67	619.01	**648.00**	**648.00**	**648.00**
TS7	662	637	559.49	559.49	554.40	561.20	561.81	**622.00**	**622.00**	**622.00**
TS8	314	292	279.02	279.02	271.03	280.20	280.91	**292.00**	**292.00**	**292.00**
TS9	27	94	13.41	13.41	13.41	13.41	13.41	**14.00**	**14.00**	**14.00**
TS10	57	86	54.87	54.87	54.92	54.87	55.29	**56.00**	**56.00**	**56.00**
TS11	18	36	14.50	14.50	14.71	14.50	15.00	**15.00**	**15.00**	**15.00**
TS12	636	758	591.68	591.68	584.82	594.60	596.19	**623.00**	**623.00**	**623.00**
TS13	595	549	463.31	463.31	455.39	466.26	466.04	**499.00**	**499.00**	**499.00**
TS14	105	110	90.24	90.24	88.71	90.70	90.65	**96.00**	**96.00**	**96.00**
TS15	78	253	76.60	76.60	76.64	76.60	77.66	**78.00**	**78.00**	**78.00**
TS16	731	775	652.60	652.60	644.88	656.04	657.29	**692.00**	**692.00**	**692.00**
TS17	733	753	634.58	634.58	624.05	638.41	641.26	**690.00**	**690.00**	**690.00**
TS18	426	264	**264.00**	**264.00**	**264.00**	**264.00**	**264.00**	**264.00**	**264.00**	**264.00**
TS19	99	44	**44.00**	**44.00**	**44.00**	**44.00**	**44.00**	**44.00**	**44.00**	**44.00**
Avg.			95.12%	95.12%	94.39%	95.40%	95.74%	**100.00%**	**100.00%**	**100.00%**

Table 10: Comparison of average ranks for students for two-sided data sets

	#Courses	SOSM	EADAM	TTC	ESMA	WOSMA	IP MAX HRT	MESMA	MWOSMA
TS1	26	1.333	1.256	1.263	1.261	**1.206**	1.546	1.403	1.368
TS2	6	2.211	**1.738**	1.763	2.009	1.844	2.069	1.875	1.750
TS3	38	1.961	1.689	**1.679**	1.852	1.732	2.401	2.218	2.105
TS4	6	2.114	2.091	**2.062**	2.084	2.077	2.346	2.269	2.269
TS5	12	1.337	1.294	1.301	1.282	**1.262**	1.598	1.448	1.425
TS6	36	1.299	1.245	**1.209**	1.286	1.252	1.421	1.420	1.400
TS7	39	1.675	1.514	1.521	1.583	**1.513**	2.101	2.064	2.053
TS8	13	1.585	1.424	1.414	1.501	**1.403**	1.753	1.616	1.527
TS9	13	**1.328**	**1.328**	**1.328**	**1.328**	**1.328**	1.357	1.357	1.357
TS10	9	1.091	1.089	1.102	1.086	**1.059**	1.071	1.071	1.071
TS11	3	**1.000**	**1.000**	1.014	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
TS12	40	1.288	1.258	1.239	1.275	**1.241**	1.376	1.347	1.353
TS13	41	1.842	1.584	1.597	1.651	**1.583**	2.351	2.094	1.972
TS14	11	1.353	1.291	1.275	1.317	**1.272**	1.500	1.469	1.406
TS15	16	1.054	1.054	1.054	1.054	**1.048**	1.064	1.064	1.051
TS16	40	1.380	1.333	1.320	1.369	**1.319**	1.513	1.512	1.500
TS17	43	1.633	1.539	**1.515**	1.582	1.541	1.886	1.923	1.891
TS18	14	2.396	**1.654**	1.757	1.865	1.754	2.318	1.784	1.678
TS19	4	2.007	**1.541**	1.500	1.698	1.663	2.068	1.727	1.705
Avg.		1.573	1.417	**1.416**	1.478	1.426	1.723	1.614	1.573

Table 11: Comparison of popularity for students for two-sided data sets

	EADAM	TTC	ESMA	WOSMA	IP MAX HRT	MESMA	MWOSMA
SOSM	(0, 16)	(1, 15)	(0, 16)	(**0**, **18**)	(10, 8)	(6, 12)	(0, 18)
EADAM		(3, 13)	(13, 3)	(**4**, **14**)	(15, 3)	(14, 4)	(12, 6)
TTC			(13, 3)	(**8**, **10**)	(15, 3)	(14, 4)	(12, 6)
ESMA				(**0**, **18**)	(15, 3)	(11, 7)	(6, 12)
WOSMA					(**16**, **0**)	(**15**, **1**)	(**13**, **3**)
IP MAX HRT						(0, 15)	(0, 16)
MESMA							(0, 14)

Table 12: Comparison of runtimes (in seconds) for two-sided data sets

	#Students	#Courses	Capacity	SOSM	EADAM	TTC	ESMA	WOSMA	IP MAX HRT	MESMA	MWOSMA
TS1	539	26	575	**0.018**	1.268	0.241	7.020	0.750	79.545	88.580	80.589
TS2	113	6	72	**0.004**	0.337	0.011	0.117	0.044	5.339	5.564	5.395
TS3	557	38	459	**0.021**	8.127	0.227	5.109	0.844	68.384	73.862	68.775
TS4	27	6	59	**0.001**	0.009	0.002	0.008	0.004	0.312	0.349	0.335
TS5	88	12	116	**0.002**	0.026	0.014	0.128	0.013	1.937	2.169	1.947
TS6	689	36	726	**0.017**	2.401	0.364	7.444	0.505	24061	24345	24292
TS7	662	39	637	**0.021**	5.528	0.307	13.214	0.921	129.444	144.179	130.016
TS8	314	13	292	**0.009**	0.980	0.076	2.206	0.234	51.427	54.651	51.577
TS9	27	13	94	**0.000**	0.001	0.001	0.003	0.002	0.066	0.069	0.068
TS10	57	9	86	**0.001**	0.002	0.005	0.034	0.005	0.362	0.365	0.364
TS11	18	3	36	**0.000**	0.001	0.001	0.002	0.001	0.050	0.055	0.053
TS12	636	40	758	**0.015**	1.358	0.345	5.148	0.484	57.581	66.833	57.854
TS13	595	41	549	**0.020**	7.031	0.269	9.265	1.184	69.336	84.760	70.757
TS14	105	11	110	**0.002**	0.044	0.013	0.088	0.014	2.559	2.611	2.573
TS15	78	16	253	**0.001**	0.003	0.009	0.046	0.008	0.886	0.924	0.896
TS16	731	40	775	**0.018**	3.597	0.464	7.021	0.643	21547	21827	21748
TS17	733	43	753	**0.021**	6.454	0.399	13.598	1.091	550.366	623.715	587.004
TS18	426	14	264	**0.034**	8.125	0.112	4.921	1.019	130.306	134.276	131.940
TS19	99	4	44	**0.003**	0.201	0.008	0.126	0.053	2.662	2.758	2.714

4.2. One-Sided Matching

Next, we applied the mechanisms described in Section 2.4 to the nine data sets introduced in Section 3.1. Apart from RSD, Pop-CHAT, ProB CHAT, MPO CHA, and MPO CHAT, we also report on WOSMA for one-sided matching. Again, we will first analyze the results of a single but representative data set (OS3) to illustrate the metrics in detail before we provide summary results of metrics for all data sets with one-sided matchings.

Figure 5 illustrates the profile curves for the different matching mechanisms. With ProB CHAT 685 students are assigned to their first choice. Moreover, the right side of the figure shows the size of the matchings, since the profile curves reflect the cumulated profiles. It indicates that MPO CHA and MPO CHAT result in the largest matchings, because they optimize on the size of the matching. Detailed rank profiles for data set OS3 can be found in the

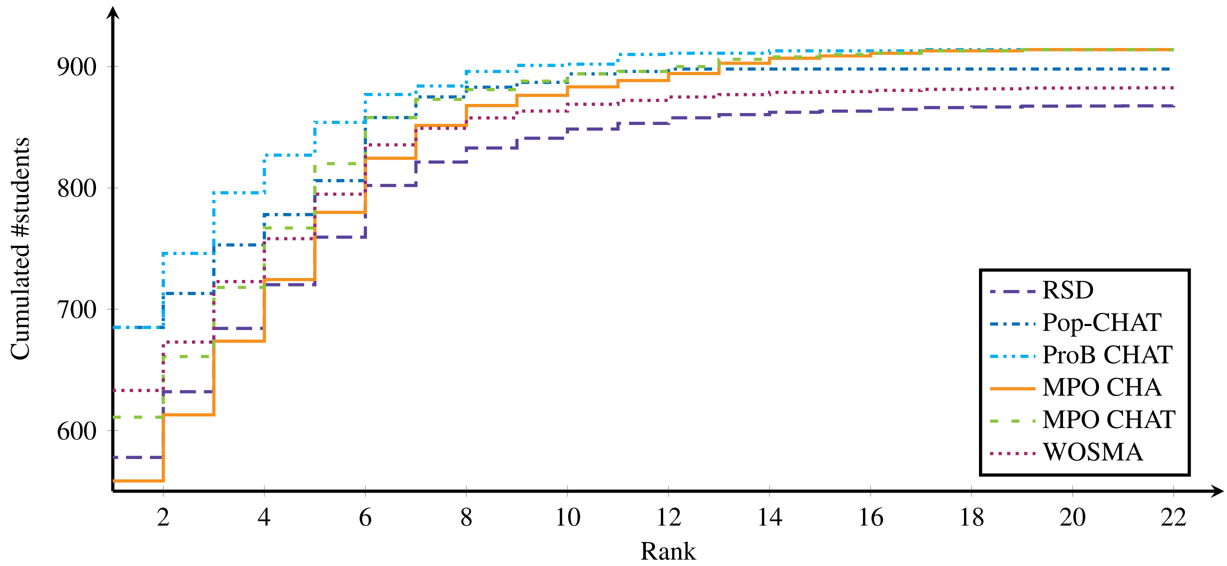


Figure 5: Comparison of profiles for students in data set OS3

Appendix A.2.

Table 13 describes the popularity. Not surprisingly, Pop-CHAT is most popular among all mechanisms followed by ProB CHAT.

Let us now discuss summary statistics across the different matching mechanisms and data sets. In terms of the average AUPCR (see Table 14) Pop-CHAT, ProB CHAT, MPO CHA, and MPO CHAT all perform well. In most cases ProB CHAT performs best.

Since no popular matching exists for data set OS2, the values of Pop-CHAT for OS2 are missing.[7]

The results of the pairwise one-sided Wilcoxon signed-rank tests are illustrated in Table 15. It demonstrates that ProB CHAT computes significantly better results than the other mechanisms in four of the nine data sets and is not

[7]For this reason, we removed the results of OS2 from all computations of averages reported in the paper.

[8]OS2 is omitted in the computation of the average in the last row.

Table 13: Comparison of popularity for students in data set OS3

	Pop-CHAT	ProB CHAT	MPO CHA	MPO CHAT	WOSMA
RSD	◀ (**120.9**, **216.9**)	◀ (120.25, 216.05)	▶ (215.15, 203.55)	◀ (169.0, 218.1)	◀ (11.15, 96.0)
Pop-CHAT		▶ (**105.0**, **95.0**)	▶ (**228.0**, **127.85**)	▶ (**184.0**, **134.0**)	▶ (**167.75**, **127.0**)
ProB CHAT			▶ (227.95, 110.5)	▶ (172.0, 110.0)	▶ (164.5, 124.55)
MPO CHA				◀ (48.6, 128.65)	◀ (160.25, 231.3)
MPO CHAT					◀ (167.8, 180.95)

Table 14: Comparison of AUPCR for one-sided data sets[8]

	RSD	Pop-CHAT	ProB CHAT	MPO CHA	MPO CHAT	WOSMA
OS1	91.77%	93.75%	**94.67**%	92.19%	92.37%	91.77%
OS2	85.06%	-	92.10%	95.84%	**96.08**%	85.06%
OS3	92.26%	96.32%	**98.21**%	96.55%	97.27%	94.30%
OS4	96.29%	**97.87**%	**97.87**%	97.70%	**97.87**%	96.62%
OS5	97.59%	**98.93**%	**98.93**%	98.41%	98.82%	98.30%
OS6	92.69%	97.60%	**98.14**%	96.74%	97.25%	94.09%
OS7	94.73%	98.06%	**98.70**%	97.46%	97.81%	95.63%
OS8	96.89%	97.83%	**98.03**%	97.71%	97.78%	97.20%
OS9	96.60%	**98.11**%	**98.11**%	97.36%	97.55%	96.81%
Avg.	94.85%	97.31%	**97.83**%	96.77%	97.09%	95.59%

dominated in any data set.

Table 16 describes the average size of the matchings. It is remarkable, that ProB CHAT results in the largest matching in eight out of the nine data sets. This is particularly interesting, because ProB CHAT only optimizes the profile of the matching, in contrast to MPO CHA and MPO CHAT. Furthermore, the largest differences in size can be observed in data set OS2. The last row shows the average of the size compared to the maximum size, with data set OS2 being omitted in this computation.

Regarding the average rank (see Table 17) the ProB CHAT mechanism performs best, except for the data sets OS4 and OS9, where WOSMA results in the lowest average rank. The largest differences in the average rank metric can observed with regard to data set OS2. This could be due to the large differences in the size of the matchings (recall Table 16). The median rank was 1 in all applications.

[9] OS2 is omitted in the computation of the average in the last row.
[10] OS2 is omitted in the computation of the average in the last row.

Table 15: Comparison of one-sided Wilcoxon signed-rank tests for one-sided data sets

	Pop-CHAT	ProB CHAT	MPO CHA	MPO CHAT	WOSMA
RSD	(1, 3)	(**0**, **4**)	(0, 1)	(0, 2)	(0, 3)
Pop-CHAT		(**0**, **4**)	(3, 1)	(2, 1)	(2, 1)
ProB CHAT			(**5**, **0**)	(**5**, **0**)	(**4**, **0**)
MPO CHA				(0, 3)	(0, 1)
MPO CHAT					(0, 0)

Table 16: Comparison of sizes for one-sided data sets[9]

	#Students	Capacity	RSD	Pop-CHAT	ProB CHAT	MPO CHA	MPO CHAT	WOSMA
OS1	136	136	129.10	**132**	**132**	**132**	132.00	129.10
OS2	418	429	363.32	-	392	**418**	**418**	363.32
OS3	915	1080	867.85	898	**914**	**914**	**914**	882.50
OS4	114	266	111.03	**113**	**113**	**113**	**113**	111.34
OS5	156	180	154.51	**156**	**156**	**156**	**156**	155.14
OS6	1035	1282	984.10	**1031**	**1031**	**1031**	**1031**	994.05
OS7	522	626	501.71	**519**	**519**	**519**	**519**	505.00
OS8	248	336	242.79	**245**	**245**	**245**	**245**	243.28
OS9	106	130	104.46	**106**	**106**	**106**	**106**	104.57
Avg.			97.48%	99.78%	**100.00**%	**100.00**%	**100.00**%	98.00%

Both, Pop-CHAT and ProB CHAT are more popular compared to other mechanisms (see Table 18) with Pop-CHAT being more popular than ProB CHAT in two of the nine data sets.

The average computation time is shown in Table 19. It is less than a minute for all mechanisms.

Table 17: Comparison of average ranks for one-sided data sets[10]

	#Courses	RSD	Pop-CHAT	ProB CHAT	MPO CHA	MPO CHAT	WOSMA
OS1	8	1.266	1.273	**1.197**	1.402	1.386	1.266
OS2	47	2.007	-	**1.842**	2.957	2.844	2.007
OS3	51	2.392	1.947	**1.859**	2.708	2.339	2.137
OS4	7	1.079	1.088	1.088	1.101	1.088	**1.075**
OS5	6	1.088	**1.064**	**1.064**	1.096	1.071	1.069
OS6	39	1.982	1.789	**1.577**	2.127	1.924	1.795
OS7	21	1.301	1.289	**1.154**	1.415	1.341	1.241
OS8	8	1.083	1.078	**1.061**	1.087	1.082	1.074
OS9	5	1.099	1.094	1.094	1.132	1.123	**1.093**
Avg.		1.41	1.33	**1.26**	1.51	1.42	1.34

Table 18: Comparison of popularity for one-sided data sets

	Pop-CHAT	ProB CHAT	MPO CHA	MPO CHAT	WOSMA
RSD	(**0**, **8**)	(0, 9)	(8, 1)	(4, 5)	(0, 7)
Pop-CHAT		(**2**, **0**)	(**8**, **0**)	(**7**, **0**)	(**7**, **0**)
ProB CHAT			(9, 0)	(8, 0)	(8, 0)
MPO CHA				(0, 9)	(0, 9)
MPO CHAT					(1, 7)

Table 19: Comparison of runtimes (in seconds) for one-sided data sets

	#Students	#Courses	Capacity	RSD	Pop-CHAT	ProB CHAT	MPO CHA	MPO CHAT	WOSMA
OS1	136	8	136	**0.002**	0.017	0.089	0.037	0.046	0.009
OS2	418	47	429	**0.006**	0.177	2.550	0.154	0.862	0.113
OS3	915	51	1080	**0.014**	2.458	33.852	0.522	36.009	39.512
OS4	114	7	266	**0.001**	0.022	0.077	0.037	0.039	0.011
OS5	156	6	180	**0.002**	0.026	0.080	0.049	0.044	0.015
OS6	1035	39	1282	**0.015**	1.359	34.468	0.568	16.079	18.329
OS7	522	21	626	**0.006**	0.220	1.867	0.168	2.198	1.824
OS8	248	8	336	**0.003**	0.076	0.188	0.063	0.076	0.027
OS9	106	5	130	**0.001**	0.016	0.075	0.018	0.035	0.018

5. Summary and Conclusions

In this section, we summarize the results from the previous section and provide some conclusions. Again, we start with two-sided matching mechanisms before we discuss one-sided matching mechanisms.

5.1. Two-sided Matching

Table 20 provides an overview of the theoretical and empirical properties of the two-sided matching mechanisms. The overview illustrates the trade-offs. All mechanisms, except for EADAM and TTC are stable. No mechanism is Pareto efficient regarding students. However, ESMA, WOSMA, MESMA and MWOSMA are Pareto stable, meaning that there is no stable matching which Pareto dominates the outcome. Only SOSM and TTC are strategy-proof. However, in EADAM, ESMA, and WOSMA truth-telling is an ordinal Bayesian Nash equilibrium (indicated by the square sign ■in the Strategy-proofness column). IP MAX HRT, MESMA and MWOSMA are not strategy-proof.

Table 20: Summary of two-sided matching mechanisms for HRT

Matching mechanism	Stability	Pareto efficiency[11]	Strategy-proofness	Maximum size	Time complexity	Average AUPCR(S)	More popular (S)	Average rank (S)	Average size
SOSM	✓	✗[12]	✓	✗	$O(\|E\|)$	84.55%	1	1.573	95.12%
EADAM	✗[13]	✓	■	✗	$O\left(\|E\|^2\right)$	86.11%	5	1.417	95.12%
TTC	✗	✗[12]	✓	✗	$O(\|E\|)$	85.42%	6	**1.416**	94.39%
ESMA	✓	■[14]	■	✗	$O\left(\|S\|^3 \cdot Q\right)$	85.76%	3	1.478	95.40%
WOSMA	✓	■[14]	■	✗	$O\left(\|S\|^3 \cdot \|C\|\right)$	86.45%	**7**	1.426	95.74%
IP MAX HRT	✓	✗	✗	✓	exponential-time	88.28%	0	1.723	**100%**
MESMA	✓	■[14]	✗	✓	exponential-time	89.39%	2	1.614	**100%**
MWOSMA	✓	■[14]	✗	✓	exponential-time	**89.70%**	4	1.573	**100%**

IP MAX HRT, MESMA and MWOSMA are designed to compute the maximum cardinality matching (see column ”Maximum size”) and thus allow computing the cost of strategy-proofness with respect to the size of a matching. The column ”Average size” in Table 20 shows the differences in average size as a percentage of the maximum size, while Table 9 shows details across all data sets. The average *cost of strategy-proofness in terms of size* is 4.88% across all data sets compared with SOSM.

The ”Average AUPCR(S)” in Table 20 describes the average Area Under the Profile Curve Ratio for students. With AUPCR as primary criterion MWOSMA, MESMA and IP MAX HRT are preferred to WOSMA, EADAM,

[11] Throughout, we mean Pareto efficiency for students.
[12] Due to random tie-breaking in presence of ties.
[13] The outcome is stable only if all residents consent to waive their preferences.
[14] Pareto stable: there is no stable matching which Pareto dominates the outcome

ESMA, and TTC which are again preferred to SOSM. The ranking is correlated with the average size ranking. On average the cost of strategy-proofness (while still maintaining stability) in terms of AUPCR for students is 4.28%, which is the difference of MWOSMA to TTC, the best strategy-proof mechanism.

Popularity takes the point of view of the students and it gets to a different ranking. The column ”More popular (S)” in Table 20 shows in how many comparisons one matching mechanism is more popular than another one. WOSMA is more popular than all other seven mechanisms, while IP MAX HRT is dominated by all other mechanisms in terms of popularity. We also report average rank, however, we argue that this metric is more difficult to interpret. If size or AUPCR matter, it is also important to remember that the computation of the maximum cardinality matching is NP-complete, which can become an issue for larger data sets, as we have seen in Table 12.

In summary, if one is willing to give up on strategy-proofness, but still wants weaker forms of truthfulness and Pareto efficiency for students, then EADAM and WOSMA are close in terms of average size, average rank, and popularity. All these metrics are improved compared to SOSM with random tie-breaking from the students point of view. Both mechanisms satisfy a weaker form of incentive compatibility. Note that in EADAM the outcome might become unstable with respect to the original preferences (■). IP MAX HRT, MESMA and MWOSMA are best in terms of average size and AUPCR, but worse than TTC, EADAM and WOSMA in terms of popularity and average rank. However, they are not incentive-compatible. Although, it might be hard to manipulate such mechanisms in applications where participants have little prior information, incentive-compatibility might be very important in others.

5.2. One-sided Matching

Table 21 provides an overview of theoretical and empirical properties of one-sided matching mechanisms. WOSMA and other mechanisms can be used for one-sided problems as well. All mechanisms, except for RSD and MPO CHA are Pareto efficient when applied to CHAT instances. Only RSD is strategy-proof and in WOSMA truth-telling is an ordinal Bayesian Nash equilibrium strategy (■). MPO CHA and MPO CHAT are designed for maximum size matchings.

Table 21: Summary of one-sided matching mechanisms for CHAT

Matching mechanism	Pareto efficiency	Strategy-proofness	Maximum size	Time complexity[15]	Average AUPCR	More popular	Average rank	Average size
RSD	✗[16]	✓	✗	$O(\|E\|)$	94.85%	1	1.41	97.48%
Pop-CHAT[17]	✓[18]	✗	✗	$O\left(\left(\sqrt{Q}+\|S\|\right)\|E\|\right)$	97.31%	**5**	1.33	99.78%
ProB CHAT	✓	✗	✗	$O\left(\min\left(z^*\sqrt{Q}, Q+z^*\right)\|E\|\right)$	**97.83%**	4	**1.26**	**100%**
MPO CHA	✗[16]	✗	✓	$O\left(\sqrt{\|S\|}\,\|E\|\right)$	96.77%	0	1.51	**100%**
MPO CHAT	✓	✗	✓	$O((\|E\|+\|S\|+\|C\|)\|E\|)$	97.09%	2	1.42	**100%**
WOSMA	✗	■	✗	$O\left(\|S\|^3\|C\|\right)$	95.59%	3	1.34	98.00%

The cost of strategy-proofness in terms of size comparing RSD with these mechanisms is 2.52%.

In terms of the AUPCR metric ProB CHAT has the highest value, but this is not significantly different from the other non-strategy-proof mechanisms. The cost of strategy-proofness in terms of the average AUPCR is 2.98% comparing ProB CHAT with the strategy-proof RSD mechanism. In terms of popularity Pop-CHAT stands out as it is more popular than all five other mechanisms. The average rank is lowest for ProB CHAT.

To sum up, if incentive compatibility and the AUPCR are an issue, WOSMA might be a candidate for one-sided matching as it improves over RSD and is better in terms of average rank. Pop-CHAT and ProB CHAT can be alternatives if manipulability is less of a concern.

[15] z^* is the maximal rank of an edge in a rank-maximal solution.

[16] Due to random tie-breaking in presence of ties efficiency is not guaranteed.

[17] There was no popular matching for OS2.

[18] As long as a popular matching exists.

References

Abdulkadiroğlu, A., Pathak, P., Roth, A., 2009. Strategy-proofness versus efficiency in matching with indifferences: Redesigning the NYC high school match. The American Economic Review 88 (5), 1954–1978.

Abdulkadiroğlu, A., Sönmez, T., 2003. School choice: A mechanism design approach. The American Economic Review 93 (3), 729–747.

Abraham, D. J., Irving, R. W., Kavitha, T., Mehlhorn, K., 2007. Popular Matchings. SIAM Journal on Computing 37 (4), 1030–1045.

Alcalde-Unzu, J., Molis, E., 2011. Exchange of indivisible goods and indifferences: The top trading absorbing sets mechanisms. Games and Economic Behavior 73 (1), 1–16.

Ashlagi, I., Nikzad, A., 2015. What matters in tie-breaking rules? how competition guides design. Unpublished working paper.

Aziz, H., De Keijzer, B., 2012. Housing markets with indifferences: A tale of two mechanisms. In: Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence. AAAI'12. AAAI Press, pp. 1249–1255.
URL `http://dl.acm.org/citation.cfm?id=2900728.2900905`

Bogomolnaia, A., Moulin, H., 2001. A new solution to the random assignment problem. Journal of Economic Theory 100 (2), 295–328.

Budish, E., Cantillon, E., 2012. The multi-unit assignment problem: Theory and evidence from course allocation at harvard. American Economic Review 102, 2237–71.

Cechlárová, K., Eirinakis, P., Fleiner, T., Magos, D., Manlove, D. F., Mourtos, I., Oceľáková, E., Rastegari, B., 2015. Algorithmic Game Theory: 8th International Symposium, SAGT 2015, Saarbrücken, Germany, September 28-30, 2015. Proceedings. Springer Berlin Heidelberg, Berlin, Heidelberg, Ch. Pareto Optimal Matchings in Many-to-Many Markets with Ties, pp. 27–39.
URL `http://dx.doi.org/10.1007/978-3-662-48433-3_3`

Diebold, F., Aziz, H., Bichler, M., Matthes, F., Schneider, A., 2014. Course Allocation via Stable Matching. Business & Information Systems Engineering 6 (2), 97–110.
URL `http://dx.doi.org/10.1007/s12599-014-0316-6`

Erdil, A., Ergin, H., 2006. Two-sided matching with indifferences. Unpublished mimeo, Harvard Business School.

Erdil, A., Ergin, H., 2008. What's the matter with tie-breaking? Improving efficiency in school choice. The American Economic Review 98 (3), 669–689.

Ergin, H. I., 2002. Efficient resource allocation on the basis of priorities. Econometrica 70 (6), 2489–2497.

Gale, D., Shapley, L. S., 1962. College Admissions and the Stability of Marriage. The American Mathematical Monthly 69 (1), 9–15.

Gibbard, A., 1973. Manipulation of voting schemes: a general result. Econometrica 41, 587–601.

Gibbard, A., 1977. Manipulation of schemes that mix voting with chance. Econometrica, 665–681.

Gusfield, D., Irving, R. W., 1989. The Stable Marriage Problem: Structure and Algorithms. Vol. 54. MIT press Cambridge.

Hanley, J. A., McNeil, B. J., 1982. The meaning and use of the area under a receiver operating characteristic (roc) curve. Radiology 143 (1), 29–36.

Irving, R. W., Manlove, D. F., Jan. 2010. Finding large stable matchings. J. Exp. Algorithmics 14, 2:1.2–2:1.30. URL http://doi.acm.org/10.1145/1498698.1537595

Jaramillo, P., Manjunath, V., 2012. The difference indifference makes in strategy-proof allocation of objects. Journal of Economic Theory 147 (5), 1913–1946.

Kesten, O., 2010. School choice with consent. The Quarterly Journal of Economics 125 (3), 1297–1348.

Király, Z., 2013. Linear time local approximation algorithm for maximum stable marriage. Algorithms 6 (3), 471–484.

Kojima, F., 2009. Random assignment of multiple indivisible objects. Mathematical Social Sciences 57 (1), 134–142.

Kwanashie, A., Irving, R. W., Manlove, D. F., Sng, C. T. S., 2015. Combinatorial Algorithms: 25th International Workshop, IWOCA 2014, Duluth, MN, USA, October 15-17, 2014, Revised Selected Papers. Springer International Publishing, Cham, Ch. Profile-Based Optimal Matchings in the Student/Project Allocation Problem, pp. 213–225. URL http://dx.doi.org/10.1007/978-3-319-19315-1_19

Kwanashie, A., Manlove, D. F., 2013. An Integer Programming Approach to the Hospital/Residents Problem with Ties. CoRR abs/1308.4064.

Levy, H., 1992. Stochastic dominance and expected utility: survey and analysis. Management Science 38 (4), 555–593.

Manlove, D. F., 2013. Algorithmics of Matching Under Preferences. Series on Theoretical Computer Science. World Scientific Publishing Company Incorporated. URL http://books.google.de/books?id=7wGJMAEACAAJ

Manlove, D. F., Irving, R. W., Iwama, K., Miyazaki, S., Morita, Y., 2002. Hard variants of stable marriage. Theoretical Computer Science 276 (1), 261–279.

Manlove, D. F., Sng, C., September 2006. Popular Matchings in the Capacitated House Allocation Problem. In: Azar, Y., Erlebach, T. (Eds.), Proceedings of ESA 2006: the 14th Annual European Symposium on Algorithms. Vol. 4168 of Lecture Notes in Computer Science. Springer, pp. 492–503.

Riksbank, S., 2012. The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2012. http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2012/, [Online; accessed 16-May-2015].

Roth, A. E., 1982. The economics of matching: Stability and incentives. Mathematics of operations research 7 (4), 617–628.

Roth, A. E., 2008. Deferred acceptance algorithms: History, theory, practice, and open questions. International Journal of Game Theory 36 (3-4), 537–569.

Rothblum, U. G., 1992. Characterization of stable matchings as extreme points of a polytope. Mathematical Programming 54 (1-3), 57–67.

Shapley, L. S., Scarf, H., 1974. On cores and indivisibility. Journal of Mathematical Economics 1 (1), 23–37.

Sng, C., 2008. Efficient Algorithms for Bipartite Matching Problems with Preferences. Ph.D. thesis, University of Glasgow.

Appendix A. Results

Appendix A.1. Two-Sided Matching: Data Set TS6

Appendix A.2. One-Sided Matching: Data Set OS3

Table A.22: Comparison of profiles for course organizers in data set TS6

Rank	SOSM		EADAM		TTC		ESMA		WOSMA		IP MAX HRT		MESMA		MWOSMA	
1	275.61	(37.96)	268.09	(36.93)	255.88	(35.25)	277.96	(38.29)	270.36	(37.24)	287.00	(39.53)	**287.00**	**(39.53)**	283.00	(38.98)
2	0.49	(0.07)	0.49	(0.07)	0.49	(0.07)	0.69	(0.10)	0.11	(0.02)	1.00	(0.14)	**1.00**	**(0.14)**	1.00	(0.14)
3	22.07	(3.04)	20.75	(2.86)	19.25	(2.65)	22.43	(3.09)	20.73	(2.86)	21.00	(2.89)	**21.00**	**(2.89)**	19.00	(2.62)
4	2.97	(0.41)	2.97	(0.41)	2.76	(0.38)	3.01	(0.41)	2.99	(0.41)	3.00	(0.41)	**3.00**	**(0.41)**	3.00	(0.41)
5	14.45	(1.99)	13.61	(1.87)	11.17	(1.54)	14.60	(2.01)	13.01	(1.79)	15.00	(2.07)	**15.00**	**(2.07)**	15.00	(2.07)
6	27.30	(3.76)	27.18	(3.74)	26.02	(3.58)	27.61	(3.80)	28.24	(3.89)	28.00	(3.86)	**28.00**	**(3.86)**	28.00	(3.86)
7	19.63	(2.70)	19.26	(2.65)	17.24	(2.37)	19.63	(2.70)	18.61	(2.56)	22.00	(3.03)	**22.00**	**(3.03)**	22.00	(3.03)
8	10.15	(1.40)	10.21	(1.41)	10.42	(1.44)	10.12	(1.39)	10.61	(1.46)	12.00	(1.65)	**12.00**	**(1.65)**	12.00	(1.65)
9	10.47	(1.44)	10.24	(1.41)	8.89	(1.22)	10.96	(1.51)	10.23	(1.41)	13.00	(1.79)	**13.00**	**(1.79)**	13.00	(1.79)
10	10.24	(1.41)	10.45	(1.44)	10.70	(1.47)	10.25	(1.41)	10.98	(1.51)	13.00	(1.79)	**13.00**	**(1.79)**	13.00	(1.79)
11	17.83	(2.46)	15.54	(2.14)	13.90	(1.91)	17.83	(2.46)	17.89	(2.46)	18.00	(2.48)	**18.00**	**(2.48)**	19.00	(2.62)
12	12.57	(1.73)	12.75	(1.76)	11.44	(1.58)	12.38	(1.71)	13.18	(1.82)	13.00	(1.79)	**13.00**	**(1.79)**	14.00	(1.93)
13	8.50	(1.17)	8.58	(1.18)	7.97	(1.10)	8.61	(1.19)	8.59	(1.18)	10.00	(1.38)	**10.00**	**(1.38)**	10.00	(1.38)
16	15.16	(2.09)	15.71	(2.16)	17.69	(2.44)	15.22	(2.10)	16.08	(2.21)	18.00	(2.48)	**18.00**	**(2.48)**	18.00	(2.48)
17	11.41	(1.57)	10.43	(1.44)	7.64	(1.05)	11.37	(1.57)	10.99	(1.51)	11.00	(1.52)	**11.00**	**(1.52)**	11.00	(1.52)
18	6.00	(0.83)	6.00	(0.83)	6.00	(0.83)	6.00	(0.83)	6.00	(0.83)	6.00	(0.83)	**6.00**	**(0.83)**	6.00	(0.83)
22	9.06	(1.25)	8.91	(1.23)	8.94	(1.23)	9.63	(1.33)	9.41	(1.30)	10.00	(1.38)	**10.00**	**(1.38)**	10.00	(1.38)
23	1.65	(0.23)	2.93	(0.40)	3.58	(0.49)	1.77	(0.24)	3.90	(0.54)	3.00	(0.41)	**3.00**	**(0.41)**	4.00	(0.55)
24	27.68	(3.81)	24.87	(3.43)	22.54	(3.10)	28.07	(3.87)	25.67	(3.54)	25.00	(3.44)	**27.00**	**(3.72)**	25.00	(3.44)
25	13.41	(1.85)	13.59	(1.87)	14.62	(2.01)	12.91	(1.78)	13.19	(1.82)	13.00	(1.79)	**13.00**	**(1.79)**	13.00	(1.79)
26	7.73	(1.06)	7.79	(1.07)	7.01	(0.97)	7.48	(1.03)	7.69	(1.06)	8.00	(1.10)	**8.00**	**(1.10)**	8.00	(1.10)
27	5.23	(0.72)	6.16	(0.85)	9.20	(1.27)	5.22	(0.72)	5.80	(0.80)	6.00	(0.83)	**6.00**	**(0.83)**	7.00	(0.96)
28	7.86	(1.08)	9.15	(1.26)	10.37	(1.43)	7.79	(1.07)	8.98	(1.24)	6.00	(0.83)	**6.00**	**(0.83)**	7.00	(0.96)
29	9.28	(1.28)	9.10	(1.25)	9.93	(1.37)	8.93	(1.23)	9.94	(1.37)	9.00	(1.24)	**9.00**	**(1.24)**	9.00	(1.24)
31	1.03	(0.14)	1.03	(0.14)	1.00	(0.14)	1.03	(0.14)	1.03	(0.14)	2.00	(0.28)	**2.00**	**(0.28)**	2.00	(0.28)
32	0.22	(0.03)	0.07	(0.01)	0.25	(0.03)	0.15	(0.02)	0.27	(0.04)	-		-		-	
34	5.37	(0.74)	5.54	(0.76)	5.96	(0.82)	5.37	(0.74)	5.19	(0.71)	5.00	(0.69)	**5.00**	**(0.69)**	5.00	(0.69)
36	2.12	(0.29)	2.18	(0.30)	1.21	(0.17)	2.20	(0.30)	2.21	(0.30)	3.00	(0.41)	**3.00**	**(0.41)**	3.00	(0.41)
38	-		1.61	(0.22)	2.96	(0.41)	-		-		-		-		-	
39	5.46	(0.75)	5.27	(0.73)	2.58	(0.36)	5.35	(0.74)	5.53	(0.76)	5.00	(0.69)	**5.00**	**(0.69)**	5.00	(0.69)
41	0.54	(0.07)	0.75	(0.10)	0.59	(0.08)	0.54	(0.07)	0.80	(0.11)	1.00	(0.14)	**1.00**	**(0.14)**	1.00	(0.14)
43	0.56	(0.08)	2.03	(0.28)	1.71	(0.24)	0.97	(0.13)	0.99	(0.14)	-		-		-	
47	-		0.02	(0.00)	0.86	(0.12)	-		-		-		-		-	
48	0.08	(0.01)	1.00	(0.14)	3.01	(0.41)	0.08	(0.01)	1.11	(0.15)	-		-		1.00	(0.14)
49	3.97	(0.55)	4.14	(0.57)	4.72	(0.65)	3.97	(0.55)	4.68	(0.64)	4.00	(0.55)	**4.00**	**(0.55)**	4.00	(0.55)
50	-		0.85	(0.12)	1.96	(0.27)	-		0.11	(0.02)	-		-		-	
53	11.06	(1.52)	11.99	(1.65)	14.08	(1.94)	10.96	(1.51)	11.33	(1.56)	14.00	(1.93)	**14.00**	**(1.93)**	14.00	(1.93)
56	3.45	(0.48)	4.52	(0.62)	5.77	(0.79)	3.39	(0.47)	4.60	(0.63)	5.00	(0.69)	**5.00**	**(0.69)**	6.00	(0.83)
58	-		0.71	(0.10)	1.65	(0.23)	-		-		-		-		-	
60	12.57	(1.73)	11.09	(1.53)	10.14	(1.40)	12.03	(1.66)	10.55	(1.45)	13.00	(1.79)	**11.00**	**(1.52)**	11.00	(1.52)
62	-		-		0.22	(0.03)	-		-		-		-		-	
63	-		-		0.25	(0.03)	-		-		-		-		-	
67	-		-		0.58	(0.08)	-		-		-		-		-	
75	-		-		0.02	(0.00)	-		-		-		-		-	
81	4.51	(0.62)	4.10	(0.56)	4.21	(0.58)	4.42	(0.61)	4.07	(0.56)	5.00	(0.69)	**5.00**	**(0.69)**	5.00	(0.69)
82	3.82	(0.53)	4.58	(0.63)	5.07	(0.70)	3.81	(0.52)	5.20	(0.72)	5.00	(0.69)	**5.00**	**(0.69)**	4.00	(0.55)
89	4.00	(0.55)	4.00	(0.55)	4.00	(0.55)	4.00	(0.55)	4.00	(0.55)	4.00	(0.55)	**4.00**	**(0.55)**	4.00	(0.55)
93	9.01	(1.24)	12.71	(1.75)	13.86	(1.91)	8.93	(1.23)	14.16	(1.95)	11.00	(1.52)	**11.00**	**(1.52)**	13.00	(1.79)
631	-		1.56	(0.21)	4.86	(0.67)	-		-		-		-		-	
646	-		0.01	(0.00)	0.21	(0.03)	-		-		-		-		-	
∞	111.48	(15.36)	111.48	(15.36)	120.62	(16.61)	108.33	(14.92)	106.99	(14.74)	78.00	(10.74)	**78.00**	**(10.74)**	78.00	(10.74)

Table A.23: Comparison of profiles for students in data set OS3

Rank	RSD		Pop-CHAT		ProB CHAT		MPO CHA		MPO CHAT		WOSMA	
1	577.85	(63.15%)	685.00	(74.86%)	**685.00**	**(74.86%)**	558.45	(61.03%)	611.00	(66.78%)	633.05	(69.19%)
2	54.15	(5.92%)	28.00	(3.06%)	**61.00**	**(6.67%)**	54.50	(5.96%)	50.00	(5.46%)	39.85	(4.36%)
3	52.20	(5.70%)	40.00	(4.37%)	**50.00**	**(5.46%)**	60.70	(6.63%)	57.00	(6.23%)	49.90	(5.45%)
4	36.00	(3.93%)	25.00	(2.73%)	**31.00**	**(3.39%)**	50.80	(5.55%)	49.00	(5.36%)	35.35	(3.86%)
5	39.20	(4.28%)	28.00	(3.06%)	**27.00**	**(2.95%)**	55.40	(6.05%)	53.00	(5.79%)	36.70	(4.01%)
6	42.60	(4.66%)	52.00	(5.68%)	**23.00**	**(2.51%)**	44.60	(4.87%)	38.00	(4.15%)	40.65	(4.44%)
7	19.30	(2.11%)	17.00	(1.86%)	**7.00**	**(0.77%)**	27.05	(2.96%)	15.00	(1.64%)	13.65	(1.49%)
8	11.60	(1.27%)	8.00	(0.87%)	**12.00**	**(1.31%)**	16.40	(1.79%)	8.00	(0.87%)	8.55	(0.93%)
9	8.05	(0.88%)	4.00	(0.44%)	**5.00**	**(0.55%)**	8.40	(0.92%)	7.00	(0.77%)	5.65	(0.62%)
10	7.55	(0.83%)	7.00	(0.77%)	**1.00**	**(0.11%)**	7.00	(0.77%)	6.00	(0.66%)	5.60	(0.61%)
11	4.70	(0.51%)	2.00	(0.22%)	**8.00**	**(0.87%)**	5.20	(0.57%)	2.00	(0.22%)	3.20	(0.35%)
12	4.60	(0.50%)	2.00	(0.22%)	**1.00**	**(0.11%)**	5.80	(0.63%)	4.00	(0.44%)	2.80	(0.31%)
13	2.60	(0.28%)	-		-		8.45	(0.92%)	6.00	(0.66%)	1.95	(0.21%)
14	1.95	(0.21%)	-		**2.00**	**(0.22%)**	4.15	(0.45%)	2.00	(0.22%)	1.90	(0.21%)
15	0.95	(0.10%)	-		-		1.90	(0.21%)	2.00	(0.22%)	0.60	(0.07%)
16	1.50	(0.16%)	-		-		2.20	(0.24%)	1.00	(0.11%)	1.00	(0.11%)
17	1.40	(0.15%)	-		**1.00**	**(0.11%)**	2.00	(0.22%)	2.00	(0.22%)	0.85	(0.09%)
18	0.50	(0.05%)	-		-		-		-		0.45	(0.05%)
19	0.70	(0.08%)	-		-		1.00	(0.11%)	1.00	(0.11%)	0.60	(0.07%)
20	0.10	(0.01%)	-		-		-		-		0.05	(0.01%)
21	0.15	(0.02%)	-		-		-		-		0.15	(0.02%)
22	0.20	(0.02%)	-		-		-		-		-	
∞	47.15	(5.15%)	17.00	(1.86%)	**1.00**	**(0.11%)**	1.00	(0.11%)	1.00	(0.11%)	32.50	(3.55%)