Too smart for their own good: Trading truthfulness for efficiency in the Israeli medical internship market

Ariel Rosenfeld*   Avinatan Hassidim[†]

Abstract

The two most fundamental notions in mechanism design are truthfulness and efficiency. In many market settings, such as the classic one-sided matching/assignment setting, these two properties partially conflict, creating a trade-off which is rarely examined in the real-world. In this article, we investigate this trade-off through the high-stakes Israeli medical internship market. This market used to employ a standard truthful yet sub-optimal mechanism and it has recently transitioned to an "almost" truthful, more efficient mechanism. Through this in-the-field study, spanning over two years, we study the interns' behavior using both official data and targeted surveys. We first identify that substantial strategic behaviors are exercised by the participants, virtually eliminating any efficiency gains from the transition. In order to mitigate the above, we performed an intervention in which conclusive evidence was provided showing that, for most of the interns, reporting truthfully was much better than what they actually did. Unfortunately, a re-examination of the market reveals that our intervention had only minor effects. These results combine to question the practical benefits of "almost" truthfulness in real-world market settings and shed new light on the typical truthfulness-efficiency trade-off.

Keywords: randomized assignment, truthfulness-efficiency tradeoff, "almost" truthful mechanisms

1 Introduction

The two most fundamental notions that underline the mechanism design field are *truthfulness* (also called incentive compatibility and strategyproofness) and *efficiency* (Nisan et al., 2007).

A mechanism is deemed *truthful* if it is a (weakly-)dominant strategy for each participating agent to reveal her true preferences. In other words, no agent can benefit from misrepresenting her preferences or by trying to manipulate the system, regardless of what the other agents do.

By and large, truthfulness does not guarantee *efficiency*. Namely, the fact that a mechanism is truthful does not assure us that it maximizes the social welfare in the market (e.g., (Anshelevich et al., 2013)). Generally speaking, participating agents are likely to be more interested in efficiency than in truthfulness, as the latter is the means, while the former is the end.

In this study, we investigate a possible trade-off between truthfulness and efficiency through the high-stakes, real-world Israeli Medical Internship Market (IMIM) in which the two properties indeed partially conflict. This market is a prime example of a classic many-to-one one-sided matching/assignment problem, which is at the core of computer science and operation research (Munkres, 1957). The market, which we have been investigating for the past two years, provides us with a rare and unique opportunity to examine the under-explored truthfulness-efficiency trade-off (Budish & Cantillon, 2012) for three main reasons:

1. Interns care a great deal about their assignment. Due to a wide variety of factors such as the hospitals' work terms, geographical location, teaching and professional guidance quality, etc., some hospitals are considered more desirable than others while different interns express different preferences. Let us consider Hadassah hospital[1] as an example. From official records published by the Israeli Ministry of Health[2] (MoH) and as depicted in Figure 1, 14% of the 2013 intern class[3] have ranked Hadassah as one of their top 3 choices (i.e., highly desirable), whereas 33% ranked it between the 9th and 11th positions (i.e., mid-ranking).

   In addition, interns are not allowed to trade their assigned internship for money. However, in the past, interns were caught illegally trading their placements with a market price for a good internship exceeding $2500. This, in turn, suggests that interns care a lot about their assignment.

*ORCID: 0000-0002-3230-3060. Information Science Department, Bar-Ilan University. Email: ariel.rosenfeld@biu.ac.il

†ORCID: 0000-0002-7034-7427. Computer Science Department, Bar-Ilan University.

[1]http://www.hadassah.org.il/

[2]www.moh.org.il

[3]We use the most recent data prior to 2014 as a transition from a truthful mechanism to a non-truthful one occurred during that year. Full details are provided in Section 2.2.
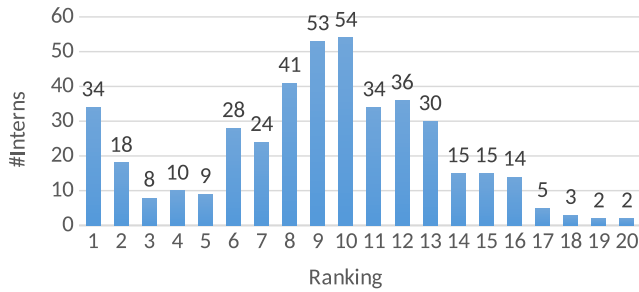
FIGURE 1: The number of interns in the class of 2013 who ranked Hadassah hospital in each ranking (a total of 20 hospitals were ranked that year).

2. The IMIM is the first (and only) market we are aware of that has transitioned from a standard truthful assignment mechanism (Random Serial Dictatorship (Abdulkadiroglu & Sonmez, 1998)) to a more efficient yet "almost truthful" variation thereof (defined more properly in Section 2.1), thus partially sacrificing the truthfulness property for efficiency.

3. We were given the opportunity to study the IMIM over the course of two years: 2018 and 2019. As such, we were able to first evaluate the market (2018 intern class) and reveal definitive evidence that the vast majority of interns should favor truthful reporting (see Section 3.1). We were then offered the opportunity to slightly *intervene* in the assignment process for the 2019 intern class and evaluate the effectiveness of our intervention. Specifically, we were allocated an hour-long presentation in the main internship national conference where the vast majority of would-be interns (80 − 90%) attend before submitting their preference lists, to learn about their duties, rights and opportunities during their internship period as well as to learn how the assignment mechanism works. We used this opportunity to better advise the 2019 intern class to adopt a truthful reporting strategy, in light of the results from the 2018 intern class (see Section 3.2).

It is self-evident that any intelligent system, let alone a deployed one, is only as good as its real-world performance. Therefore, the understanding of the interaction between real-world users (in our case, medical interns) and automated mechanisms (in our case, intelligent assignment mechanisms) is crucial (Rosenfeld & Kraus, 2018).

As the title of this article may have suggested, this work provides compelling evidence that *the transition from a truthful mechanism to a non-truthful yet more efficient one in the IMIM was not successful* considering the current market condition. Startlingly, virtually *all* interns (extensively) misrepresented their preferences before *and after* our presentation-based intervention in the market. This result is especially striking as the deployed mechanism is "almost" truthful

(i.e., only a small fraction of the interns benefit from manipulations in the current data) and our presentation in the 2019 intern class presented definitive empirical evidence that the vast majority of interns should favor truthful reporting. Moreover, the utility of the interns given that they misrepresent their preferences is similar to the utility they would have achieved if they had used the standard truthful mechanism deployed prior to the transition.[4] Taken jointly, these results suggest that "almost" truthful mechanisms may not be practically beneficial in the real-world and perhaps, as most theoretical work suggests, the truthfulness property is indeed binary with no "grey-scale" between the extremes.

The remainder of this article is structured as follows: In Section 2, we discuss the background and the literature related to the problem at hand (Section 2.1) as well as present the properties of the IMIM (Section 2.2). Section 3 presents our three-phase study: A) The investigation of the 2018 intern class (Section 3.1); B) Our intervention for the 2019 class (Section 3.2); and C) The following investigation of the 2019 intern class (Section 3.3). Then, in Section 4, we interpret the combined results and discuss their possible implications. Finally, we conclude the work and highlight future work directions in Section 5.

2   Background

2.1   Truthfulness and Efficiency of Mechanisms

A mechanism is said to be *truthful* if no agent has an incentive to misrepresent her true preferences. This property is considered highly desirable for mechanisms that are used in real-life markets. Indeed, many of the great success stories of market design employ truthful mechanisms, such as the second-price sealed-bid auction (Vickrey, 1961) (often deployed for online ad auctions), or Deferred Acceptance (DA) (Gale & Shapley, 1962) (often used for students' acceptance to schools). Truthfulness is also thought to reduce the efficiency costs of strategizing, to eliminate any possible advantages to more sophisticated agents, to provide robustness, to promote fairness by "leveling the playing field" (Basteck & Mantovani, 2018), to eliminate the costs associated with the collection of information on others, and to simplify the interpretation of reported preferences in terms of social welfare (Roth, 2008).

At the core of the attractiveness of these truthful mechanisms is the assumption that agents report their preferences truthfully. However, recent evidence has shown that this need not necessarily be the case. Rees-Jones and Skowronek (Rees-Jones & Skowronek, 2018) have examined survey data

---

[4]In order to obtain this result, one needs to put cardinal measures on the structure of the utility function. See Section 2.2.

from the American National Resident Matching Program[5] (NRMP) which deploys the truthful Deferred Acceptance (DA) mechanism, and they have found that 17% of medical graduates (futilely) try and manipulate the system. Similarly, Hassidim et al. (Hassidim et al., 2016) have reported a manipulation rate of at least 20% in the Israeli psychology matching market which also deploys a truthful mechanism (Hassidim et al., 2017b). Similar results were also reported for one-sided matching/assignment markets such as the Mexico City school assignment (Chen & Pereyra, 2019) which employs a special case of the DA mechanism called Serial Dictatorship (SD) (a randomized variant of SD is in the focus of this study, as discussed next). These and other empirical in-the-field studies (e.g., (Braun et al., 2014; Featherstone & Niederle, 2016)), as well as lab-based studies (e.g., (Chen & Sönmez, 2006; Li, 2017)), reveal that a small (yet non-negligible) portion of agents do misrepresent their preferences under truthful mechanisms, which in turn leads to significant losses in efficiency. Against this background, it may seem reasonable to sacrifice the truthfulness property, as it is already being violated by some agents, in order to achieve improved efficiency, as investigated in this article.

Non-truthful behaviors in truthful mechanisms are generally regarded as a (relatively minor) nuisance that can be (largely) explained by exogenous factors (low cognitive abilities, mistrust, complexity of the mechanism...) and are perhaps mitigated through targeted interventions and explicit information disclosure from the market designers (Hassidim et al., 2017a). On the other hand, non-truthful behaviors in non-truthful mechanisms can be significantly more troubling. For example, in voting settings (which are notorious for having no truthful mechanisms other than dictatorial ones when there are more than two alternatives (Gibbard et al., 1973; Satterthwaite, 1975)), strategic behavior seems inevitable. Indeed, many theoretical results show the wide range of possible manipulations in voting (Meir, 2018), which in turn are supported by empirical studies (e.g., (Tal et al., 2015)). In matching markets, (Artemov et al., 2017) have recently put forward evidence from an Australian student admission dataset suggesting that non-truthful behaviors in non-truthful matching mechanisms are largely payoff-irrelevant (i.e., they do not change the student's outcome and payoff). As such, these non-truthful behaviors may be effectively ignored. In contrast, our work provides strong evidence against this hypothesis, showing that, in the examined real-world market, a large portion of the participants are directly harmed by their non-truthful behavior.

In this work we examine agents' strategic behavior in an "almost" truthful mechanism where only a small portion of the agents can slightly improve their situation by misrepresenting their preferences. We refer to a mechanism as "almost truthful to an extent ϵ" if only a very small por-

tion of the agents (defined by ϵ) can manipulate it to their advantage in the resulting market. This practical definition is akin to the more theoretical one of *truthful with* $1 - \epsilon$ *probability* (Nisan et al., 2007, Ch. 13.4.6) where at most an $\epsilon > 0$ fraction of the agents can successfully gain by (some) misrepresentation of their preferences.[6] The almost truthful mechanism studied in this work is not the only almost truthful mechanism in the literature. In fact, Roth and Peranson (Roth & Peranson, 1999) have documented a (very) small portion of participants who can benefit through misrepresenting their preferences in the US residency match due to couples who participate in the match (Kojima et al., 2013; Ashlagi et al., 2014). As such, the deployed mechanism is almost truthful as well.

Two variants of the popular Random Serial Dictatorship (RSD) mechanism (Abdulkadiroglu & Sonmez, 1998) have been deployed in the IMIM thus far: 1) Standard RSD mechanism (denoted RSD henceforth); and 2) An extension of the RSD which we will refer to as RSD with Trading (or RSDT for short) (Bronfman et al., 2018, 2015).[7]

RSD works as follows: a random permutation of the agents is sampled from the uniform distribution followed by the agents taking turns in choosing their most preferred hospital placement among those hospitals which have yet to meet their capacity. It is easy to verify that RSD is a *truthful* mechanism: any agent has only one opportunity to pick a placement, so the dominant strategy for any agent is to pick the best available one. RSD is also *ex-post Pareto efficient*, meaning that no trade can happen right after the assignment or, in other words, no agent can get a better placement without having another agent getting a worse one. Intuitively, if Alice and Bob would like to swap places right after RSD is performed, and say Alice chose her hospital placement before Bob, then why didn't she choose Bob's placement when she had the chance?[8] Unfortunately, RSD is **not** *ex-ante Pareto efficient* under standard assumptions (i.e., agents having Von Neumann-Morgenstern utilities over random allocation, namely lotteries over the hospital placements). Namely, RSD provides each intern some probability of being assigned to any of the hospitals, depending on her place in the (random) ordering, on the decision of the interns who were drawn before her, and her own preferences. These probability vectors may be inefficient because there may exist other probability vectors that provide every agent a greater

---

[6]These two concepts should not be confused with that of an ϵ-Nash equilibrium (Nisan et al., 2007, Chapter 2.6.6) in which each agent has an incentive to misrepresent, where that incentive is bounded by ϵ.

[7]For ease of exposition, RSDT is introduced as an extension of RSD. It was also presented as such in prior literature. However, in order to be completely accurate, the RSDT mechanism can be theoretically used with initialization phases other than the RSD mechanism and thus can be considered in its own right.

[8]It is, however, possible that preferences change between the time the mechanism was employed and the actual assignment. In that case, trade may be beneficial for both sides.

expect utility (Bogomolnaia & Moulin, 2001). In fact, truthfulness and ex-ante Pareto efficiency are incompatible in the standard randomized assignment setting, requiring one to choose or balance the two in some reasonable way (Zhou, 1990; Liu & Pycia, 2016).

To mitigate the fact that the probability vectors induced by the RSD mechanism may be far from optimal, the RSDT mechanism was proposed: First, after each intern submits a ranked list of hospitals, a large number of RSD simulations is performed to approximate the true probability vector for each intern. Then, using a fitted utility function over probability vectors (learned through structured surveys), probabilities are automatically traded between the interns though a Linear Program (LP) which optimizes social welfare. The LP guarantees that each intern's expected utility (given the utility function) is no-worse than what she had before the trade. The full LP is given in 5. Then an assignment respecting the new probabilities is sampled.

In light of the theoretical incompatibility between efficiency and truthfulness as discussed before, it comes as no surprise that RSDT is not truthful. Namely, agents may benefit from misreporting their preferences due to the "trading" phase of the mechanism which is naturally manipulable (e.g., an intern may want to obtain probabilities for hospitals which others consider very desirable in the first phase of the RSDT, in order to trade them as a "bundle" with probabilities for undesired hospitals in the later phase). Still, using real-world data, we show that only minor benefits may be gained by only a very small fraction of the agents who have misrepresented their preferences, making RSDT almost truthful.[9] We replicate and extend this analysis in this study using newly collected data (which we believe to be more reliable than the old one) and additional utility functions which were not evaluated by the authors.

To the best of our knowledge, this is also the first study to examine a market which has already deployed a truthful mechanism, and have traded it in order to achieve a more efficient outcome.

2.2 The Israeli Medical Internship Market (IMIM)

The final step in getting an Israeli medical degree is performing a 12-month-long internship in one of 20–25 geographically-scattered hospitals in Israel.[10] Each hospital is allocated with a number of interns relative to the size of its associated patient population (with extra weight given to peripheral hospitals). This capacity is determined by the Israeli MoH and is published in the media. As a matter of

policy, interns are not assigned based on merits as talented interns should be allocated across the country. At the same time, interns have been shown to report different preferences for the hospitals. Specifically, interns differ in how they rank the hospitals from the most desired to the least. As mentioned before, taken together, the assignment of interns to hospitals in the IMIM is a special case of the classic many-to-one one-sided matching/assignment problem (Munkres, 1957). The assignment of tasks/goods to agents in the standard formulation is analogous to the assignment of interns to hospital placements, with the cost/profit associated with each pairing being analogous to the utility of the assigned intern from the paired hospital.

For about two and a half decades, until 2014, each intern was asked to submit her ranking of the hospitals relevant for her graduation class, and the assignment itself was decided by the RSD mechanism (with a few minor house rules aimed at providing special treatment for special intern groups such as PhD students and parents of young children), which has came to be known as the *Internship Lottery*. Up until a few years ago when the system was finally computerized, interns physically gathered in a large auditorium and ID numbers were drawn out of a hat. The RSDT mechanism has been deployed since 2014 in an attempt to increase efficiency (Bronfman et al., 2015) (see (Roth & Shorrer, 2015) for a review and discussion on the transition choice).

Interns we interviewed informally mentioned more than a few factors that affect the attractiveness of different hospitals. These range from personal preferences such as an intern's willingness to commute or move, to more objective factors such as the professional guidance provided by the hospitals. These, in turn, are individually weighted by the interns. That being said, according to the official data of the Israeli MoH, only small insignificant changes were observed until 2014 in the average rankings of hospitals. Specifically, the average rank of each hospital experienced very little change throughout the years, suggesting that the market was relatively stable until 2014.

The RSDT mechanism, as deployed in this market, assumes each intern uses an exponent-shaped utility function u. Specifically, let m denote the number of hospitals (in 2018 and 2019 cases, $m = 25$), p_j denote an intern's probability of receiving hospital j and $rank(j)$ denote the rank of hospital j within the intern's preference (1 being the most preferred). Then the utility of each intern is expressed as

$$u = \sum_j p_j (m - rank(j))^2$$

This utility function was fitted according to survey data collected by the RSDT developers in 2013.

In this study, we consider two additional utility functions as proposed by some interns we informally interviewed:

---

[9]For additional discussion and theoretical analysis of the RSDT mechanism, see (Bronfman et al., 2015, 2018).

[10]The exact number of hospitals in which an internship can be performed slightly changes over time, based on administrative needs determined by the Israeli MoH.
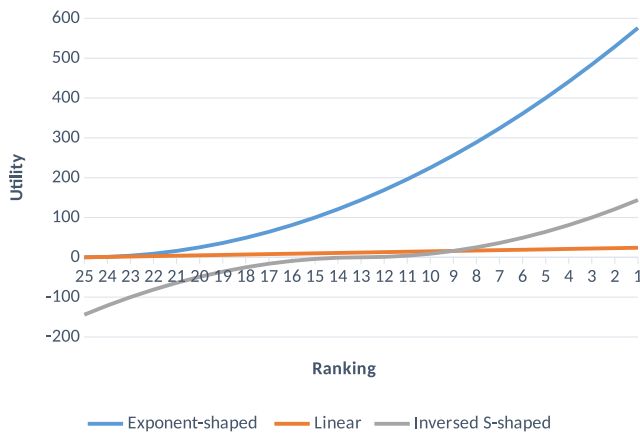
FIGURE 2: Marginal utility (y-axis) gained from each hospital as a function of its ranking (x-axis) and the assumed utility function (series).

1. Linear:

$$u = \sum_j p_j(m - rank(j))$$

2. Inversed S-shaped function:

$$u = \sum_j sign(rank(j) - \frac{m}{2})p_j(\lceil \frac{m}{2} \rceil - rank(j))^2$$

Figure 2 illustrates the marginal utility from each hospital as a function of its ranking.

In our analyses we also examine the case of utility *misidentification*. Namely, we examine the potential losses and room for strategic intern behavior in case the mechanism assumes one type of utility function but an intern has a different one.

The assignment mechanism is explained each year (both prior to 2014 as well as since then) to the students via an hour-long presentation at the annual intern conference where the vast majority of interns (80-90%) attend to learn about their duties, rights and opportunities during their internship period. While we do not have direct access to the presentations used prior to 2014, we were assured by the Israeli MoH officials that the fact that the mechanism (i.e., RSD) is truthful was highlighted to the students. The presentation used between 2014 and 2018 is available in Hebrew in the online appendix. The presentation used in 2019, presenting the results of this study for the 2018 intern class, is available as well. It is also important to note that the actual code used by the MoH was made available to the interns from 2014 onwards, promoting the credibility of the RSDT mechanism.

3   Method

This study consists of 3 main phases: First, we examined the state of the market through an in-depth inquiry of the 2018 intern class. This phase was conducted *after* the interns submitted their preferences and received their assignment so as to avoid the potential appearance of the inquiry having any effect on the assignment itself. Second, the results obtained through the first stage were used to design a presentation-based intervention for the intern class of 2019. The intervention took place about two weeks before the interns had to submit their preferences to the Israeli MoH. Finally, we repeated the inquiry from the first stage with the 2019 intern class to determine what if any impact our intervention had. As before, this phase was performed only after the interns' assignments were published and finalized.

3.1   Class of 2018

Following 4 years in which the RSDT mechanism was employed in the IMIM (2014–2017), we evaluated the mechanism through a user study.

3.1.1   Setup

The 2018 internship lottery took place on the 28th of January, 2018. The results were were published immediately after. A few weeks later, we approached the Israeli medical students council and asked for their help in distributing a specially designed questionnaire. The questionnaire consists of 4 simple parts:

1. *Simple demographics*: Age, gender and medical school.

2. *Real preferences*: The interns were asked to imagine a Utopian world in which they could be placed at whichever hospital they want regardless of what other interns want. Then, interns were asked to rank the hospitals from the most desired to the least without any other consideration other than their own true preferences. It is important to stress again that when the interns completed the survey, their assignment had already been already finalized. As such, interns have no incentive to lie or misrepresent their true preferences.

3. *Reported preferences:* The interns were asked to copy and paste the ranked list of hospitals as they had reported in the RSDT mechanism (the list was still accessible to them on the MoH website). Copying and pasting, while keeping the formatting, made it difficult to misrepresent the original list.

4. *Additional information*: We asked three questions in this part. First, we asked interns to what extent they understand how the assignment mechanism works. Second, following a few informal interviews with interns, we came up with a set of five possible "heuristics" aimed at manipulating the mechanism (e.g., placing a commonly less-desired hospital at the top of the ranking to increase its odds). Each intern was asked to select the reasons that apply to her, if any. We also
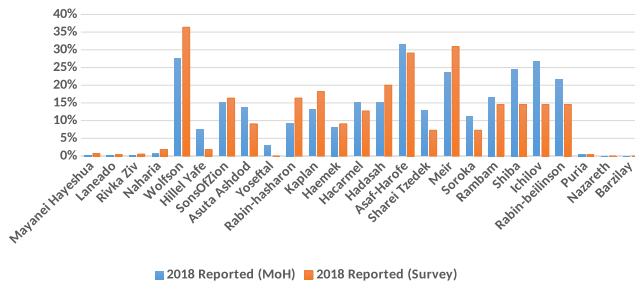
FIGURE 3: Percentage of interns who ranked each hospital in their top 3 choices.

FIGURE 4: Percentage of interns who ranked each of the hospitals in their top 3 choices: comparing real and reported preferences.

supported free-text reasons. Lastly, we asked interns whether they thought that the market should revert to the RSD mechanism.

The questionnaire is available in Appendix B.

Fortunately, the student council had agreed to officially distribute the questionnaire and encourage interns to complete it. The students' representatives from each of the 6 Israeli medical schools distributed the survey using both email and social media (i.e., designated Facebook groups) while, from the researchers' end, each intern was offered 20 NIS (about $7) for completing the survey.

The 2018 intern class consists of 640 interns, of which 239 interns (105 male, 132 female, 2 who wished not to be identified by gender) have completed the questionnaire online. The participating interns range in age between 23 and 39 (average of 28.5) and roughly represent the distribution of interns across the 6 Israeli medical schools (69 from the Hebrew University of Jerusalem, 48 from the Technion, 43 from Tel-Aviv University, 42 from Bar-Ilan University and 37 from Ben-Gurion University).

3.1.2 Results

We first compared the participating interns' *reported preferences* to those of all 640 interns (the latter were published by the Israeli MoH). To that end, we focused on the percentage of interns who ranked each of the hospitals in their top 3 choices. Using the Kolmogorov-Smirnov (KS) test (Goodman, 1954) we cannot reject the null hypothesis that our sample came from the same distribution as the data obtained through the Israeli MoH ($p = 0.91$)(see Figure 3).

Considering the interns' *real preferences* in Figure 4, one can see that, as expected, interns vary in the way they would rank the hospitals in a Utopian world. However, one can also see an apparent discordance between the real and reported preferences. For example, highly desired hospitals such as Sheba[11], Ichilov[12] and Rabin-Belinson[13] were ranked by sig-
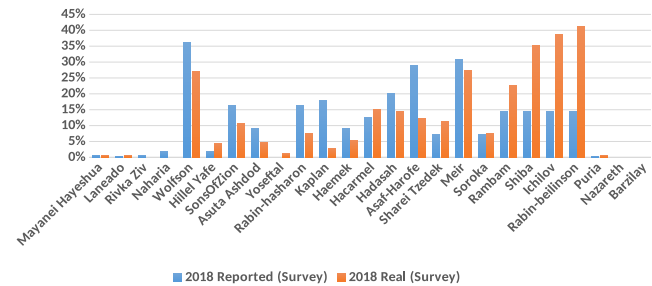
---

[11]https://www.sheba.co.il/
[12]https://www.tasmc.org.il/
[13]https://hospitals.clalit.co.il/rabin/

nificantly fewer interns in their top 3 hospitals compared to the interns' true preferences.

We further compare the interns' *real preferences* with the *reported preferences* of the 2013 intern class, which we assume to be *truthful*. Recall that the RSD mechanism, used in 2013, is truthful and the average ranking of each hospital in the IMIM remained constant leading to 2013 (see Section 2.2). Using the same technique as before, we consider the percentage of interns who ranked each of the hospitals in their top 3 choices. Using the KS test as before, we cannot reject the null hypothesis that the real preferences in our sample came from the same distribution as the reported preferences of the 2013 intern class ($p = 0.87$). As such, there is no significant difference in the real preferences of the 2018 intern class compared to the (assumed to be) latest available real preferences of the 2013 intern class.

However, for *all* 239 interns the real and reported preferences differ. In order to quantify the discordance between the real and reported preferences we use two common metrics: First, we calculate the mean absolute difference of ranks for each intern. Namely, we calculate the mean absolute shift in hospitals' location within the ranking between the two preference lists for each intern. The mean absolute difference was found to range between 1.12 and 9.2 with an average of 3.28 and a standard deviation of 1.38. Second, we calculate the Spearman's rank correlation coefficient (Ramsey, 1989) for each intern. The coefficients range between -0.39 and 0.77 with an average of 0.11 and a standard deviation of 0.23. Interestingly, for only 11% of the interns was the correlation across the 25 hospitals statistically significant, $p < 0.05$. These results strongly indicate that interns have heavily tried to manipulate the system, so much that on average there is an extremely weak (yet positive) correlation between the reported preferences and the truthful ones.

Against this background, an important question arises: how many of the interns have played sub-optimally? That is, how many of them could have done better by reporting their true preferences (and to what extent) given that all other interns continue to try to manipulate the system? To this end, using the original code used by the MoH, we re-evaluated the
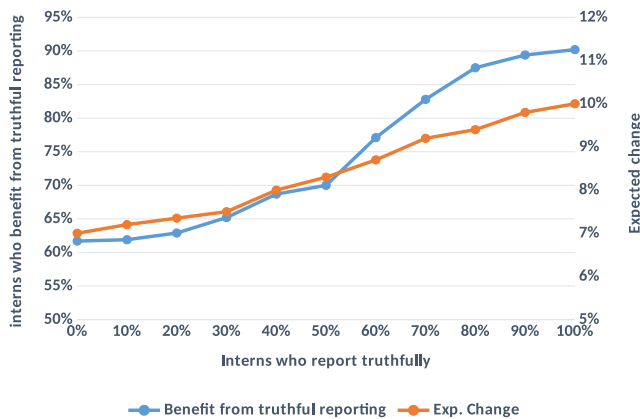
FIGURE 5: The percentage of interns who would benefit from truthful reporting (major y-axis) and the average change in expected utility (secondary y-axis) as a function of the portion of interns who report truthfully.

outcomes iteratively by "de-manipulating" each intern separately. That is, for each intern independently, we change the reported preferences to the true ones and run the mechanism from scratch. The results show that for 61.7% of the interns it would have been better to report their truthful preferences in the RSDT mechanism *while all other interns continue to misrepresent their preferences*. On average, by reporting truthfully, each intern could have improved her expected utility by 7% (assuming the exponent-shaped utility function used in the original code). This improvement is statistically significant using a pair-wise t-test, $p < 0.05$. That is to say, even when all other interns continue to try and manipulate the system, it is better for most interns to report truthfully. By varying the number of interns who report their truthful preferences we see that the number of interns who would have improved their expected utility increases from 61.7% (all others manipulate) to 90.2% (all others report truthfully) and the average expected utility increases from 7% to 10% (see Figure 5). In other words, the more interns who report truthfully, the more likely it is for every other intern to benefit from reporting truthfully and the larger the benefit from doing so.

We also examine the robustness of the above results by examining the two additional utility functions discussed in Section 2.2: a Linear and an Inversed S-shaped functions. We examine the case where the mechanism assumes one type of utility function while the interns use a different one (also known as utility misidentification). We evaluate the percentage of interns who would benefit from reporting truthfully (even when all other interns continue to manipulate the system) and the extent of the average change in interns' expected utility derived thereby. The results are presented in Table 1.

In light of the results discussed above, the empirical benefits of the RSDT mechanism may be questioned. To better address this concern, we compare the results of the RSDT

mechanism, despite the interns' "strategic" behavior, with those of a truthful RSD mechanism. Namely, we compare each intern's expected utility given her reported preferences under the RSDT mechanism to her expected utility given her real preferences under the RSD mechanism in which all interns report truthfully. The results show that 56.3% of the interns are better off using the RSDT mechanism, despite their misrepresentation of their preferences. However, the average expected benefit from the use of RSDT compared to a truthful RSD is very small (less than 1%) and it is not statistically significant. We further evaluate the expected change in the expected utility for each intern if she were to report truthfully in the RSDT mechanism and other interns were to continue to manipulate, compared with a truthful RSD where all interns are assumed to report truthfully. Surprisingly, **89%** of the interns are better off reporting truthfully in the RSDT mechanism (with all other interns misrepresenting their preferences) compared to the expected results of the (truthful) RSD mechanism. On average, each intern improves her expected utility by 7% compared to RSD. The improvement is found to be statistically significant using a pair-wise t-test, $p < 0.05$. This result is consistent when repeating the evaluation under different utility functions, as shown in Table 2.

Finally, we analyze the interns' answers to the additional questions we provided. 17% of the interns claimed that they *fully understand* the mechanism and half of them claimed that they *understand the main properties* thereof. Only 10% of the interns claimed that they *completely don't understand* the mechanism (the rest stated that they understand only the basics of the mechanism). Of the possible "manipulation heuristics" presented to the interns (in random order), most interns stated that they have used past statistics (available through the Israeli MoH website) to predict other interns' preferences (71% of the interns). Moreover, 59% of the interns stated that they have lowered their ranking of commonly low-ranked hospitals (as observed in previous years) and 53% have increased the ranking of some hospitals they deem "acceptable" and are generally ranked lower by others (in previous years). Only 10% of the interns reported that have changed their top-ranked hospital to a more "achievable" one, while 47% have stated the same for the second and third top-ranked hospitals. Finally, we asked the interns if they think that the market should revert to the RSD mechanism. 44% of the interns marked "yes", 32% of the interns marked "no" and the rest marked that they have no opinion on the matter.

3.1.3 Discussion

The above results reveal two orthogonal phenomena: First, in its existing condition, the market suffers from heavy manipulations which, in turn, make the transition from the RSD mechanism to the RSDT mechanism non-effective and per-

TABLE 1: Percentage of interns (class of 2018) who would benefit from reporting truthfully (given that all other interns continue to manipulate the system) under various utility conditions. In parentheses, the average change in the expected utility by reporting truthfully.

Mechanism/Interns	Exponent-shaped	Linear	Inversed S-shaped
Exponent-shaped	61.7% (7%)	62.5% (3%)	68.8% (21%)
Linear	59.4% (7%)	60.2% (3%)	67.2% (22%)
Inversed S-shaped	68.8% (6%)	64.1% (2%)	77.3% (19%)

TABLE 2: Percentage of interns (class of 2018) who would benefit from reporting truthfully under the RSDT mechanism (given that all other intern continue to manipulate the system) compared to a truthful RSD mechanism under various utility conditions. In parentheses, the average change in the expected utility.

Mechanism/Interns	Exponent-shaped	Linear	Inversed S-shaped
Exponent-shaped	83.6% (7%)	78.9% (3%)	89.1% (20%)
Linear	84.4% (7%)	83.6% (3%)	86.7% (22%)
Inversed S-shaped	86.5% (5%)	87.4% (2%)	90.6% (24%)

haps even slightly detrimental to the interns due to the indirect costs the interns endure by learning the somewhat complex RSDT mechanism, studying past statistics, coming up with manipulations, estimating the preferences of others, etc. Surprisingly, most interns claim that they understand how the mechanism works (or at least, its fundamental properties), yet they employ various heuristics that, as shown, are detrimental for the most part. Second, the results clearly show that, for the vast majority of interns, it is beneficial to behave truthfully under the RSDT mechanism *irrespective of what other interns do*. Substantial improvement in social welfare is expected in this case. In addition, even if all other interns continue to manipulate the system, 89% of the interns are better off reporting truthfully.

From a theoretical perspective, these two phenomena are orthogonal and cannot co-exist over time. It is expected that, in due course, (most) interns will realize that it is significantly better to report truthfully and the market will gradually shift towards a condition in which all (or at least the great majority of) interns report truthfully with little incentive to manipulate.

It is, however, possible that the market has yet to fully shift towards truthful behavior as the RSDT mechanism has "only" been deployed for four years. Be that as it may, this explanation seems to be unlikely as the almost truthfulness of the RSDT was highlighted to the interns again and again over the years. Furthermore, since the previously deployed RSD mechanism is truthful, it is highly abnormal that the interns would modify their behavior to such a drastic extent. The expectation that interns will remain truthful (for the most part) is due to the highly established Status quo bias (Kahneman et al., 1991) in which people tend to follow the same

decision-making procedures when the decision-making circumstances do not vary drastically. Another reason to reject the explanation is that the average ranking of previously highly desired hospitals (using the RSD mechanism, prior to 2014) has decreased gradually over the 5 years leading up to 2018 while mid-ranking hospitals have gradually increased in their ranking. For example, highly desired hospitals such as Ichilov and Rabin-Belinson with an average rank of 3.5 and 4 in 2013 have monotonically dropped over the last 5 years to an average rank of 5.1 and 5.3, respectively, in 2018. At the same time, commonly mid-ranked hospitals in 2013 such as HaCarmel and Wolfson have monotonically increased in average ranking from 11.2 and 9.4 (2013) to 8.4 and 6.7 (2018), respectively[14]. See Figure 6 for an illustration. We cannot conclusively exclude the possibility that interns' preferences have spontaneously changed over these years, yet this seems not to be the case given the comparison of the 2018 intern class's real preferences and those of the 2013 class (see above) and our discussions with intern representatives and MoH officials.

Since interns are highly-educated and highly-motivated individuals, the authors of this article have speculated that presenting the above results to the interns would bring about (at least some of) the desired change towards truthfulness. Specifically, since for the majority of the interns there is a very simple "better response" to other's decisions, one would hope that a simple intervention could change the market's condition. Unfortunately, this was not the case.

---

[14]As the number of hospitals varied (slightly) over the years, we have normalized the rankings to the number of hospitals available in 2013 (which was 20).
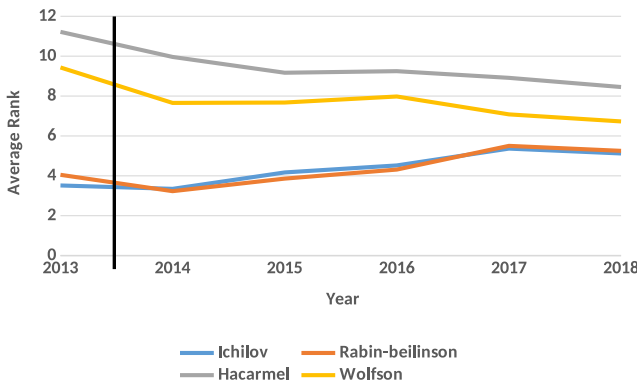
FIGURE 6: The average ranking (y-axis) of a few exemplary hospitals over time (x-axis). Rankings in 2013 were provided under the RSD mechanism while the following were provided under the RSDT mechanism.

In the next section, we discuss the design of our intervention and its administration followed by a re-evaluation of the 2019 intern class.

3.2 Intervention

Each year, interns are formally invited to the main internship national conference where they learn about their duties, rights and opportunities during their internship period as well as how the assignment mechanism works. We used this opportunity to present the main results detailed above (Section 3.1) in addition to highlighting, once again, the almost truthfulness of the RSDT mechanism.

The power-point presentation summarizing the results was presented by the second author on the 20th of December, 2018 at said conference. At the end of the presentation, the emails of both authors were given to the interns in order to accommodate any future questions.

3.3 Class of 2019

Following our intervention, we re-evaluated the market using the same procedure as discussed in Section 3.1.

3.3.1 Setup

The 2019 internship lottery took place on the 15th of January, 2019. The results were published immediately after. About 3 weeks after the lottery, the student council distributed the same questionnaire as used for the 2018 intern class (Appendix B) and encouraged interns to take part. As before, each intern was offered 20 NIS (about $7) for completing the survey.

The 2019 intern class consists of 692 interns, of which 116 interns (46 male, 69 female, 1 who wished not to be identified by gender) have completed the questionnaire online. The participating interns range in age between 23 and
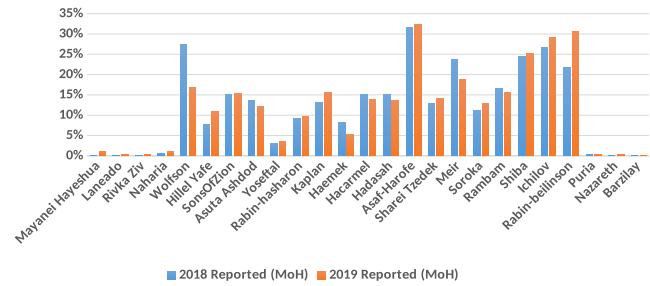

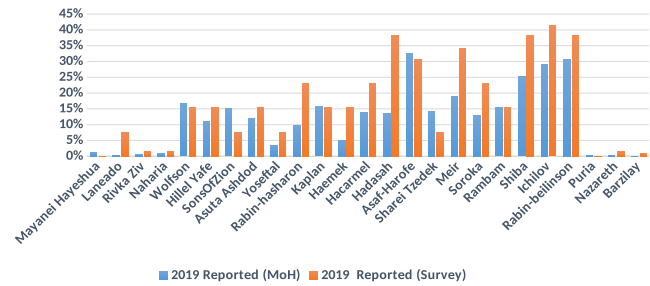FIGURE 7: Percentage of interns who ranked each of the hospitals in their top 3 choices.

FIGURE 8: Percentage of interns who ranked each of the hospitals in their top 3 choices: comparing real and reported preferences.

40 (average of 28.8) and span across the 5 Israeli medical schools (35 from the Hebrew University of Jerusalem, 12 from the Technion, 31 from Tel-Aviv University, 9 from Bar-Ilan University and 29 from Ben-Gurion University).

3.3.2 Results

We start our analysis by examining the participating interns' *reported preferences* compared to those of all 692 interns, which are published by the Israeli MoH. Similar to our 2018 analysis, we cannot reject the null hypothesis that our sample came from the same distribution using the KS test ($p = 0.57$). As can be observed in Figure 7, the reported preferences according to the Israeli MoH have not changed much between 2018 and 2019. We cannot reject the null hypothesis that the MoH data for 2018 and for 2019 came from the same distribution using the KS test ($p = 1$).

In the same vein as the above result, once again, an apparent discordance between the real and reported preferences is found, as shown in Figure 8.

As was the case for the 2018 intern class, we further compare the interns' *real preferences* with the reported preferences of the 2013 intern class and the real preferences of the 2018 intern class. Considering the percentage of interns who ranked each of the hospitals in their top 3 choices and using the KS test, we again see that we cannot reject the null hypotheses that the real preferences in the 2019 sample came from the same distribution as the reported preferences of the

2013 intern class ($p = 0.82$) or the same distribution as the reported preferences of the 2018 intern class ($p = 0.9$).

Examining each intern separately reveals that interns, again, heavily tried to manipulate the system. Except for only 4 interns, *all* 116 interns differ in their reported and real preferences. The mean absolute difference of indices was found to range between 0 and 8.56 with an average of 2.76 and a standard deviation of 2.06. The Spearman's rank correlation coefficients range between -0.43 and 1 with an average of 0.1 and a standard deviation of 0.29. For only 28% of the interns the correlation was found to be significant, $p < 0.05$. These results suggest that our intervention was not as effective as one could hope for with only a marginal shift towards truthful (or at least less "strategic") behavior.

Using the same analysis as before, we examine the percentage of interns who could have improved their expected utility by reporting truthfully under various utility function settings. The results are presented in Table 3.

As was the case for the intern class of 2018, most interns could have improved their expected utility by reporting truthfully (given that all continue to report as they did). Specifically, for the standard Exponent-shaped utility case (which is assumed in the deployed system), 67.4% of interns could have improved their expected utility by reporting truthfully. In this case, on average, each intern would increase her expected utility by 16%.

A comparison to the truthful RSD mechanism shows that, once more, a slight majority of interns (57.1%) are better off using the RSDT mechanism despite their misrepresention of their preferences. However, this time, the average change in expected utility from the RSDT compared to a truthful RSD is −1% (yet the difference is not statistically significant). If each intern were to report truthfully (and other interns would continue as they were), **85.7%** of them would be found to be better off using the RSDT mechanism compared to the truthful RSD mechanism. In this case, on average, each intern could improve her expected utility by 7% compared to RSD. This improvement is found to be statistically significant using a pair-wise *t*-test, $p < 0.05$. This result is consistent when repeating the evaluation under different utility functions, as shown in Table 4.

Lastly, the interns' answers to the last part of the questionnaire present an almost exact replica of the 2018 intern class. Specifically, 18% of the interns claim that they *fully understand* the mechanism, 57% claim that they understand the main properties thereof, and only 9% claim that they completely don't understand it. 47% of the interns stated that they used past statistics to predict other interns' preferences, 46% of the interns lowered their ranking of commonly low-ranked hospitals, and 38% increased the ranking of less-preferred hospitals which they deem "achievable" and are generally ranked lower by others. 15% of the interns misrepresented their top choice, while 36% reported the same for their second and third top-ranked hospitals. Finally, 41% of the

interns think it is better to revert to the RSD mechanism, 30% think it is better to stay with the RSDT mechanism, and the rest have no opinion on the matter.

During our presentation, the emails of both authors were given to the interns. Three emails were received by the first author in the days prior to the lottery (none by the second author). Two of the three emails were practically the same, describing (in words) some form of a utility function (e.g., "I don't want to be assigned to hospital X, what should I do?" and "It is most important to me to be allocated to one of the hospitals in region Y, what should I do?"). The third email related to a technical question regarding married couples in the mechanism (see (Hassidim et al., 2016)) and did not present any strategical considerations. The former emails were answered in the same way, reassuring the interns that a truthful representation of their preferences is (statistically) the best course of action. We cannot conclusively say that these two interns did not misrepresent their preferences in the system since our questionnaire was anonymous.

3.3.3 Discussion

The investigation of the 2019 intern class, following our intervention, presents very consistent patterns with those of the 2018 intern class. As was the case for the 2018 intern class, the 2019 intern class continues to heavily misrepresent their preferences and suffers from considerable utility losses for most of them as a consequence (compared to truthful reporting in RSDT). Specifically, the "strategic" behavior exercised by the interns is counter-effective for most interns (between 59.4% to 77.3%, depending on the utility setting). Once again, for the vast majority of interns (between 78.9% and 90.6%, depending on the utility setting), reporting truthfully in the RSDT condition (while the rest of the interns continue to behave as they did), would result in a substantial improvement in their expected utility over the RSD condition.

Regrettably, for both the 2018 and 2019 intern classes, interns' behavior practically eliminated the potential efficiency benefits of the RSDT mechanism that the market has hoped for.

4 Interpretation

The above results may be interpreted in one of two ways.

First, the most intuitive way to interpret the results is to consider them as "negative" results: the market has not capitalized on the possible efficiency benefits of the almost truthful mechanism (compared to the RSD mechanism) despite having what seems to be perfect conditions. Specifically, the RSD mechanism used prior to the transition is truthful, thus establishing a truthful starting point. The RSDT mechanism was presented and explained to the interns including

TABLE 3: Percentage of interns (class of 2019) who would benefit from reporting truthfully (given that all other interns continue to manipulate the system) under various utility conditions. In parentheses, the average change in expected utility by reporting truthfully.

Assumed/Real	Exponent-shaped	Linear	Inversed S-shaped
Exponent-shaped	67.4% (16%)	65.3% (8%)	73.5% (31%)
Linear	59.2% (9%)	55.1% (5%)	63.3% (25%)
Inversed S-shaped	61.2% (12%)	61.2% (4%)	69.4% (27%)

TABLE 4: Percentage of interns (class of 2019) who would benefit from reporting truthfully under the RSDT mechanism (given that all other interns continue to manipulate the system) compared to a truthful RSD mechanism under various utility conditions. In parentheses, the average change in expected utility.

Assumed/Real	Exponent-shaped	Linear	Inversed S-shaped
Exponent-shaped	85.7% (11%)	71.4% (8%)	89.8% (26%)
Linear	79.6% (8%)	81.6% (4%)	81.6% (25%)
Inversed S-shaped	82.1% (7%)	83.2% (3%)	85.7% (39%)

its almost truthfulness property (backed by real-world data). No other major changes have occurred in the market, the interns are highly-educated and highly-motivated to behave optimally, etc. The interns' "strategic" behavior continued in 2019, our intervention notwithstanding. Adopting this perspective would suggest that the IMIM should revert to the standard RSD mechanism, thereby reducing the RSDT-associated "strategic burden". More generally, this interpretation would also suggest that markets in which a truthful mechanism can be deployed should be very careful in future attempts to improve social welfare at the expense of truthfulness.

Second, perhaps more controversially, we could interpret the results as "positive"; not only did the interns not suffer any significant efficiency losses due to the transition (in terms of the expected utility according to the tested utility functions), but each intern can also significantly increase her expected outcome by reporting truthfully. Namely, while the interns did not capitalize on the possible benefits of the RSDT mechanism, they may still be able to do so in the future, perhaps by reading the results published in this article, which could have a more significant impact than our presentation-based intervention. In some respects, the interns *chose* not to utilize the potential benefits and they could easily do so, even at the individual level (in most of the cases). This interpretation is strengthened by the fact that the current behaviors are very different from a Nash equilibrium — most interns can easily respond better to the market condition, simply by being truthful. Adopting this interpretation would suggest that the IMIM should continue to use the current mechanism while trying to leverage additional intervention methods. More generally, this interpretation would suggest

that, as long as a non-truthful mechanism is not shown to significantly harm the participants and potential benefits are expected, it is worth considering.

Before one can derive conclusions from this study, it is worthwhile to consider two possible limitations: First, despite the many attempts to explain it to them, the interns may not have understand the RSDT mechanism or its almost-truthfulness property. It is often hypothesised that when individuals do not fully understand a mechanism they may resort to naïve theories of how the mechanism works. These theories are, in turn, translated into varying decision-making heuristics, which were found to be employed by the interns in the examined market. As such, some markets have chosen to trade efficiency for *transparency* (Pathak, 2017). As discussed before, the RSDT mechanism was clearly presented to the interns (its code was also provided), and its almost truthfulness was highlighted again and again (and backed by data). The interns' responses to the questionnaire further oppose this possible explanation with 67% (75%) of the 2018 (2019) interns reporting that they understand at least the main properties of the RSDT mechanism. The above combine to relieve some of this concern.

Second, since this study relies, for the most part, on self-reported data, the credibility of some of the results may be questioned. It is well-documented that human participants sometimes answer dishonestly in surveys (Stephens-Davidowitz & Pabon, 2017). In this study we use three types of data: 1) aggregative reported preferences taken from the Israeli MoH (which is as reliable as official data can be); 2) individual interns' reported preferences provided by the interns using copy-paste from the MoH website (which are fairly reliable since, as mentioned before, interns have no in-

centive to misreport and any such attempt would have been identified due to formatting issues); and 3) the interns' real preferences provided using a standard drag-and-drop tool. The latter is naturally subject to dishonest reporting, although it is important to stress that interns have no incentive to do so as their allocation has already been finalized. It is, however, possible that the real preferences of some interns have genuinely changed between the time of the lottery and the questionnaire. Furthermore, it is possible that some interns exhibit real indifference between two (or more) hospitals which may be broken in their reportings in a certain way purely by chance (neither the MoH system nor our questionnaire allowed for ties in the ranking). Since most interns have indicated at least one heuristic that they have used to adjust their true preferences, it is reasonable to attribute the larger part of the discordance between the reported and real preferences to strategic behavior.

5   Conclusions

In this article, we present and analyze a real-world transition from a "standard" randomized assignment mechanism to a novel, potentially more efficient yet almost truthful one in the IMIM. Following a two-year long study, including an intervention by the authors, the transition was found to bring about extensive manipulation attempts before *and after* our intervention in the market which, in turn, eliminated any of the potential efficiency gains due to the transition. At the same time, despite the extensive manipulation attempts, no significant utility losses were encountered for the interns.

Two general conclusions may be derived from this article: First, despite its potential benefits, sacrificing truthfulness for efficiency, even in the form of almost truthful mechanisms, may not bring about the utility gains one had hoped for, instead generating extensive manipulation attempts from the participating agents. In particular, the practicality of almost truthful mechanisms is questioned. Second, real-world user studies for evaluating the benefits and limitations of novel mechanisms and decision support systems are essential. Specifically, results obtained in offline evaluations using existing data may not be confirmed in the field.

Considering the IMIM, the authors believe that the decision either to continue using the RSDT mechanism or to revert to the RSD mechanism should be placed to a vote for the next intern class. The outcome of this vote is expected to be beneficial to the market no matter what the decision is: If the interns choose to continue using the RSDT mechanism after reading the results of this study, it will provide a strong signal to all interns that the majority of them believe the market would shift towards more truthful behavior which, in turn, increases each intern's incentive to report truthfully. If the interns choose to revert to the RSD mechanism, the transition will eliminate the "strategic burden" currently ex-perienced by virtually all interns and simplify the assignment process.

We plan to extend this study to additional real-world markets in which truthfulness is sacrificed to some extent. For example, in matching markets which must satisfy minimum and/or maximum quotas, a truthful mechanism does not exist under standard assumptions (e.g., (Gonczarowski et al., 2019)). For example, we plan to investigate other mechanisms which are expected to bring about more truthful behaviors in such market settings such as iterative mechanisms (Bó & Hakimov, 2020). This type of in-the-field investigation, assessing the strategic behavior from both sides of the matching market, if such exists, can enhance our understanding of non-truthful mechanisms and the way they impact participating agents in the real world.

References

Abdulkadiroglu, A. & Sonmez, T. (1998). Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, *66*(3), 689.

Anshelevich, E., Das, S., & Naamad, Y. (2013). Anarchy, stability, and utopia: Creating better matchings. *Autonomous Agents and Multi-Agent Systems*, *26*(1), 120–140.

Artemov, G., Che, Y.-K., & He, Y. (2017). *Strategic 'mistakes': Implications for market design research*. Technical report, Melbourne University.

Ashlagi, I., Braverman, M., & Hassidim, A. (2014). Stability in large matching markets with complementarities. *Operations Research*, *62*(4), 713–732.

Basteck, C. & Mantovani, M. (2018). Cognitive ability and games of school choice. *Games and economic behavior*, *109*, 156–183.

Bó, I. & Hakimov, R. (2020). Iterative versus standard deferred acceptance: Experimental evidence. *The Economic Journal*, *130*(626), 356–392.

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

Braun, S., Dwenger, N., Kübler, D., & Westkamp, A. (2014). Implementing quotas in university admissions: An experimental analysis. *Games and Economic Behavior*, *85*, 232–251.

Bronfman, S., Alon, N., Hassidim, A., & Romm, A. (2018). Redesigning the israeli medical internship match. *ACM Transactions on Economics and Computation (TEAC)*, *6*(3-4), 21.

Bronfman, S., Hassidim, A., Afek, A., Romm, A., Shreberk, R., Hassidim, A., & Massler, A. (2015). Assigning Israeli medical graduates to internships. *Israel journal of health policy research*, *4*(1), 6.

Budish, E. & Cantillon, E. (2012). The multi-unit assignment problem: Theory and evidence from course allocation at Harvard. *American Economic Review*, *102*(5), 2237–71.

Chen, L. & Pereyra, J. S. (2019). Self-selection in school choice. *Games and Economic Behavior*.

Chen, Y. & Sönmez, T. (2006). School choice: An experimental study. *Journal of Economic theory*, *127*(1), 202–231.

Featherstone, C. R. & Niederle, M. (2016). Boston versus deferred acceptance in an interim setting: An experimental investigation. *Games and Economic Behavior*, *100*, 353–375.

Gale, D. & Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, *69*(1), 9–15.

Gibbard, A. et al. (1973). Manipulation of voting schemes: A general result. *Econometrica*, *41*(4), 587–601.

Gonczarowski, Y. A., Kovalio, L., Nisan, N., & Romm, A. (2019). Matching for the Israeli" Mechinot" Gap-Year Programs: Handling Rich Diversity Requirements. *arXiv preprint arXiv:1905.00364*.

Goodman, L. A. (1954). Kolmogorov-Smirnov tests for psychological research. *Psychological bulletin*, *51*(2), 160.

Hassidim, A., Marciano, D., Romm, A., & Shorrer, R. I. (2017a). The mechanism is truthful, why aren't you? *American Economic Review*, *107*(5), 220–24.

Hassidim, A., Romm, A., & Shorrer, R. I. (2016). Strategic behavior in a strategy-proof environment. In *Proceedings of the 2016 ACM Conference on Economics and Computation* (pp. 763–764).: ACM.

Hassidim, A., Romm, A., & Shorrer, R. I. (2017b). Redesigning the Israeli Psychology Master's Match. *American Economic Review*, *107*(5), 205–09.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic perspectives*, *5*(1), 193–206.

Kojima, F., Pathak, P. A., & Roth, A. E. (2013). Matching with couples: Stability and incentives in large markets. *The Quarterly Journal of Economics*, *128*(4), 1585–1632.

Li, S. (2017). Obviously strategy-proof mechanisms. *American Economic Review*, *107*(11), 3257–87.

Liu, Q. & Pycia, M. (2016). Ordinal efficiency, fairness, and incentives in large markets. *Fairness, and Incentives in Large Markets (August 1, 2016)*.

Meir, R. (2018). Strategic Voting. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *13*(1), 1–167.

Munkres, J. (1957). Algorithms for the assignment and transportation problems. *Journal of the society for industrial and applied mathematics*, *5*(1), 32–38.

Nisan, N., Roughgarden, T., Tardos, E., & Vazirani, V. V. (2007). *Algorithmic game theory*. Cambridge university press.

Optimization, G. (2014). Inc.,"Gurobi optimizer reference manual".

Pathak, P. A. (2017). What really matters in designing school choice mechanisms. *Advances in Economics and Econometrics*, *1*, 176–214.

Ramsey, P. H. (1989). Critical values for Spearman's rank order correlation. *Journal of educational statistics*, *14*(3), 245–253.

Rees-Jones, A. & Skowronek, S. (2018). An experimental investigation of preference misrepresentation in the residency match. *Proceedings of the National Academy of Sciences*, *115*(45), 11471–11476.

Rosenfeld, A. & Kraus, S. (2018). Predicting human decision-making: From prediction to action. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *12*(1), 1–150.

Roth, A. E. (2008). What have we learned from market design? *Innovations: Technology, Governance, Globalization*, *3*(1), 119–147.

Roth, A. E. & Peranson, E. (1999). The redesign of the matching market for American physicians: Some engineering aspects of economic design. *American economic review*, *89*(4), 748–780.

Roth, A. E. & Shorrer, R. I. (2015). The redesign of the medical intern assignment mechanism in Israel. *Israel journal of health policy research*, *4*(1), 11.

Satterthwaite, M. A. (1975). Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of economic theory*, *10*(2), 187–217.

Stephens-Davidowitz, S. & Pabon, A. (2017). *Everybody lies: Big data, new data, and what the internet can tell us about who we really are*. HarperCollins New York, NY.

Tal, M., Meir, R., & Gal, Y. K. (2015). A study of human behavior in online voting. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems* (pp. 665–673).: International Foundation for Autonomous Agents and Multiagent Systems.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, *16*(1), 8–37.

Zhou, L. (1990). On a conjecture by Gale about one-sided matching problems. *Journal of Economic Theory*, *52*(1), 123–135.

Appendix A: The LP used by the RSDT Mechanism

Given a set of interns $\{i\}$, a set of hospitals $\{j\}$ and their associated capacity $\{c_j\}$, a utility function u, and a set of probability vectors $\{p_{i,j}\}$ (representing the probability that

intern i will be assigned to hospital j), the following mathematical program is solved:

$$\underset{p\hat{_{i,j}}}{Maximize} \sum_i \sum_j u(p\hat{_{i,j}})$$

$$\text{s.t} \quad \sum_j p\hat{_{i,j}} = 1 \quad \forall i$$

$$\sum_i p\hat{_{i,j}} = c_j \quad \forall j$$

$$\sum_j u(p\hat{_{i,j}}) \geq \sum_j u(p_{i,j}) \quad \forall i$$

$$p\hat{_{i,j}} \geq 0 \quad \forall i, j$$

The objective of the above is to find new probability vectors $p\hat{_{i,j}}$ that maximize the social welfare. This maximization is subject to the following constraints: First, each new vector is indeed a probability vector. Second, the capacity of each of the hospitals is respected. Third, the new probability vectors do not harm the interns w.r.t their expected utility under RSD. Lastly, all probabilities are naturally non-negative.

It is easy to verify that the above mathematical program is indeed an LP. Thus it can be easily solved through any off-the-shelf LP solver (for this work we used Gurobi (Optimization, 2014)).

Appendix B: The Questionnaire

The questionnaire was built using Qualtrics[15] and was administered in Hebrew. Below are the translated questions.

Part 1 – Demographics

1. Age [numerical answer only]

2. Gender [Male/Female/I wish not to say]

3. Medical school [Random order]

   - Tel-Aviv University
   - Bar-Ilan University
   - Hebrew University
   - Technion
   - Ben-Gurion University

Part 2 – Real Preferences

Imagine a Utopian world in which you can choose whichever hospital you want regardless of what other interns want. In that Utopian world, how would you rank the following hospitals [the names of the 25 hospitals were presented in a random order]

---

[15]www.qualtrics.com

Part 3 – Reported Preferences

Please access the MoH lottery website [link provided] and copy-paste the ranking you used in the following text field [illustrating figure provided]

Part 4 – Additional Information

1. How would you rate your level of understanding of the assignment mechanism used in the lottery

   - I don't understand it at all
   - I understand only the basics
   - I understand its general properties
   - I completely understand it

2. If your answers in parts 2 and 3 are different, which (if any) of the following statements would you say apply to your decision-making? [Random order, multiselect]

   - I used last year's statistics to predicts others' preferences
   - I changed my top choice to a more achievable one
   - I changed my 2nd and/or 3rd top choices to improve my odds
   - I lowered the ranking of hospitals that others deem less desirable in order to avoid them
   - I increased the ranking of commonly mid-ranking hospitals in order to avoid my less desired hospitals

3. Do you think that future interns would be better off using the "Hat" lottery (used prior to 2014)? [Yes/No/I don't know]