Online Supplementary Appendix to: Fair Matching under Constraints: Theory and Applications

Yuichiro Kamada and Fuhito Kojima*

August 24, 2018

Appendix C. Analysis of data from Bunkyo City

In this section, we report our simulations using data on daycare seat allocation from Bunkyo City, Japan. The numerical analysis we report here suggests that the main findings for Yamagata are robust to data features.

As explained in Section 5.2, Bunkyo City is one of the 23 special districts of Tokyo, with about 230,000 residents as of 2018. Bunkyo is much more urban than Yamagata. It has a population density about 30 times that of Yamagata, has a high concentration of educational institutions, and attracts many dual-income families investing heavily education and demanding childcare which, as we will see below, seems to make its daycare allocation problem more pressing than Yamagata's.[1] Part of our interest in studying Bunkyo's data is to investigate whether our numerical findings are robust to demographic features of different municipalities.

Our data involve the applicants (who are anonymized), usually parents, representing children who would begin attending the daycare in April of 2018. There were 2114 applicants aged between 0 to 5 as of April 1, 2018 on which they would begin attending the daycare. For each applicant, the data show her submitted preferences over the daycare centers and priority ranking (the priorities are common across daycare centers). Regarding submitted preferences, we note that the mechanism in Bunkyo is based on serial dictatorship but restricts applicants to list at most five daycare centers in their ranking.[2] Because of this restriction, Bunkyo's mechanism is not strategy-proof. This is one of the main reasons that we focus on Yamagata's data in our main analysis. Note that, however, there is a certain sense in which this mechanism is "less manipulable" than other

*Kamada: Haas School of Business, University of California, Berkeley, Berkeley, CA 94720, y.cam.24@gmail.com. Kojima: Department of Economics, Stanford University, Stanford, CA 94305, fkojima@stanford.edu.

[1]Bunkyo, whose literal translation would be "Literature Capital," is home to many higher education institutions such as University of Tokyo as well as prestigious elementary and secondary schools.

[2]There are a few additional differences between Bunkyo's mechanism and serial dictatorship, i.e., there are a few special rules, mainly regarding children with siblings. In our numerical analysis, however, this difference causes only a minor difference between the assignments from serial dictatorship (with limited length of preference lists) and the actual one.

mechanisms such as the Boston mechanism with the same length restriction (Pathak and Sönmez, 2013).

The priority order is based on the applicant characteristics such as parents' job status and the number of adults available for care at home (Bunkyo City, 2018). There are 63 daycare centers in our dataset. For each daycare center, the data show how many seats are supplied for each age.

In our simulation, we made several modeling choices given data limitation. First, as for Yamagata's data, Bunkyo's data we have involve ties although the actual priority order is strict. This is because our data lack information on some characteristics used by Bunkyo to determine the strict order, such as whether the child is currently in an alternative form of childcare and whether the family has a member with disability. As in our analysis of Yamagata's data, we randomly break ties using a single tie-breaking (that is, the tie-broken priorities are common across daycare centers) according to the uniform distribution. For each mechanism that we consider, we conducted 250 runs of simulations using such a tie-breaking rule.

The second limitation involves constraints. As is the case for Yamagata's dataset, for daycare centers, Bunkyo's dataset does not tell the entire family of feasible sets of children or the number of teachers corresponding to the flexible constraints. Instead, it only shows the number of advertised seats at each daycare center for each age, which is exactly enough to specify the rigid constraints. To overcome this limitation, we define m_s for each s in the daycare constraints (Equation (5.1)) by $m_s := \sum_{t \in T} r_t \cdot q^t$, where r_t and q^t are those in the data (recall that r_t is the teacher-child ratio under the national regulation, and q^t is the number of advertised seats for age t at daycare center s). That is, m_s is the minimum of the number of teachers such that the constraint implied by the number of advertised seats in data is a rigid constraint associated with our daycare constraint.[3] This method is identical to the one we used in our analysis of Yamagata's data.

We find that the effect of allowing flexibility in constraints is substantial in our data from Bunkyo, just as is the case of data from Yamagata: the average number of children who are matched with a strictly preferred daycare center in the flexible SOFM compared to the rigid SOFM is 1091.76, which amounts to 51.64% of all applicants (Table 3).[4] By contrast,

[3]We set m_s as the bare minimum that is consistent with the data on advertised seats so that we do not overstate our estimate of the gain from removing the rigid constraint. In a similar spirit, we allow for non-integral values of m_s although the number of teachers is an integer in practice. With an alternative specification setting m_s to be the integer rounded up from our present definition, for instance, our estimate of the gain from removing rigid constraints would be larger.

[4]This table as well as others also report simulations of other mechanisms we discuss below.

no applicant is made worse off, as implied by Proposition 2. The number of children who are unallocated changes from 1710.87 to 875.74, a 48.81% decrease (Figure 3). The average numbers of children who are matched to their first choice, first two choices, and the first three choices increase by 517.11%, 159.77% and 104.91%, respectively (Figure 4).[5] Our analysis suggests that substantial efficiency gain from utilizing the flexible nature of the constraints may be present not only in Yamagata City but more broadly.

From/To	rigid SOFM	flexible SOFM	actual allocation	flexible ETSD
rigid SOFM	0	1091.76 (51.64%)	913.72 (43.22%)	1156.84 (54.72%)
flexible SOFM	0	0	287.62 (13.61%)	102.84 (4.86%)
actual allocation	36.05 (1.71%)	534.94 (25.30%)	0	595.42 (28.17%)
flexible ETSD	0	0	266.95 (12.63%)	0

TABLE 3. The number of applicants who are made strictly better off by a change of a mechanism.

Next, we compare the rigid and flexible SOFMs with Bunkyo's actual assignment. Bunkyo's mechanism is based on rigid envy-tolerating serial dictatorship (rigid ETSD), just as is the case with Yamagata (with the limitation on the length of preference list).[6] This means that, among other things, there may remain justified envy between two children i and i' if they are of different ages, while by construction there is no justified envy between children of the same age. Bunkyo's assignment is expected to have some efficiency advantage over the rigid SOFM since justified envy is tolerated across different ages, while the comparison with the flexible SOFM is theoretically indeterminate because Bunkyo's assignment is based on the rigid constraint, which may or may not overwhelm the efficiency gain from tolerating justified envy across different ages.

We find that the flexible SOFM outperforms Bunkyo's assignment not only in terms of fairness but also in terms of efficiency. Regarding efficiency, all of our efficiency measures favor the flexible SOFM; the average fraction of unmatched children decreased by 11.18%, and 25.30% of children are matched with strictly preferred daycare under the flexible SOFM while only 13.61% are matched with strictly preferred daycare under the actual

[5]If an applicant lists k daycare centers in her submitted preferences and gets unassigned to any of them, then we list her as being assigned to her $(k+1)$st choice.

[6]Bunkyo's assignment mechanism has a few modifications to the serial dictatorship, most notably in the way it treats siblings. However, deviation of Bunkyo's assignment from pure serial dictatorship is very minor in magnitude. Detail is available upon request.

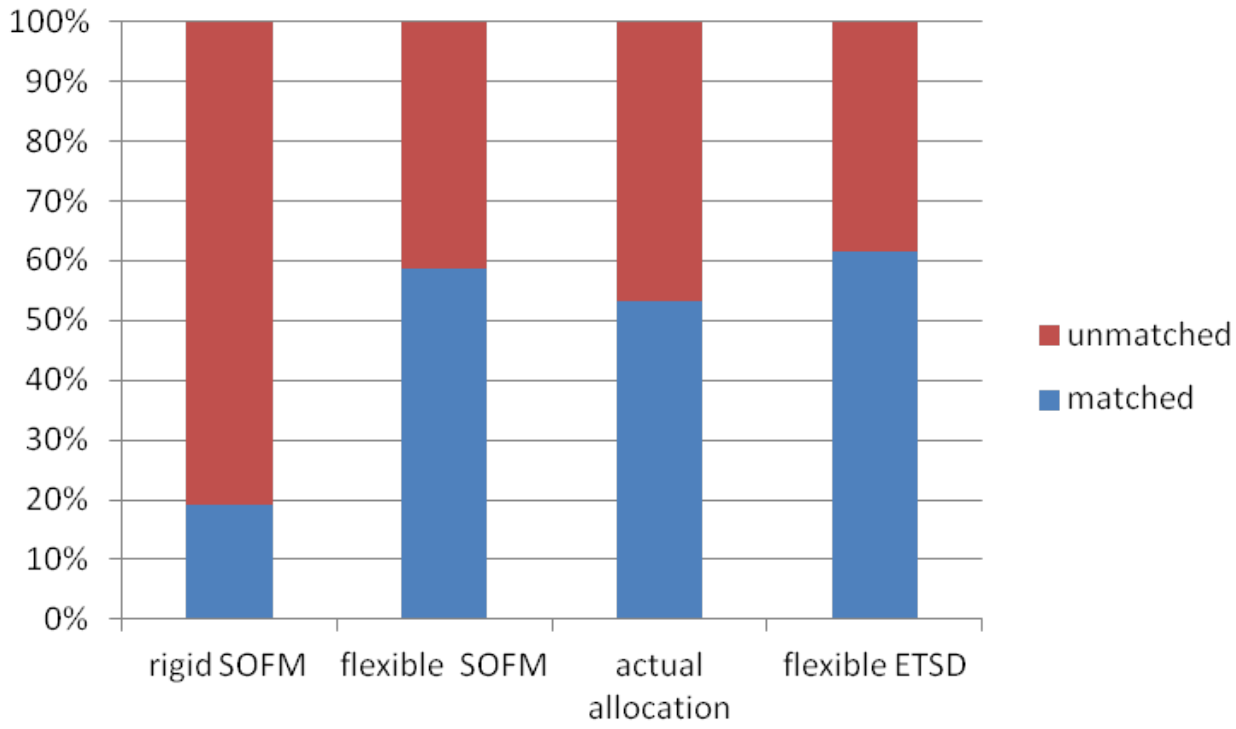

FIGURE 3. The fractions of matched applicants under different mechanisms.

allocation. Turning our focus to fairness, Table 2 provides several measures of envy for Bunkyo's assignment (note that all measures of envy are zero for the rigid and flexible SOFM). There are 1622 pairs (i, s) such that i has a justified envy toward some i' matched to s under the actual allocations. Also, students involved in at least one of such pairs and daycares involved are 40.82% and 96.83% of the respective total numbers. As for the analysis of Yamagata's data, the amount of envy for Bunkyo's actual assignment seems comparable to those in TTC on Boston and New Orleans data (Abdulkadiroglu et al., 2017).

	rigid SOFM	flexible SOFM	actual allocation	flexible ETSD
pairs with envy	0	0	1622 (18.79%)	923.00 (10.70%)
students with envy	0	0	863 (40.82%)	613.79 (29.03%)
daycares with envy	0	0	61 (96.83%)	44.35 (70.40%)

TABLE 4. Measures of justified envy under different mechanisms. The percentages for pairs with envies divide the numbers of pairs with envies by the numbers of pairs (i, s) such that s is acceptable to i.

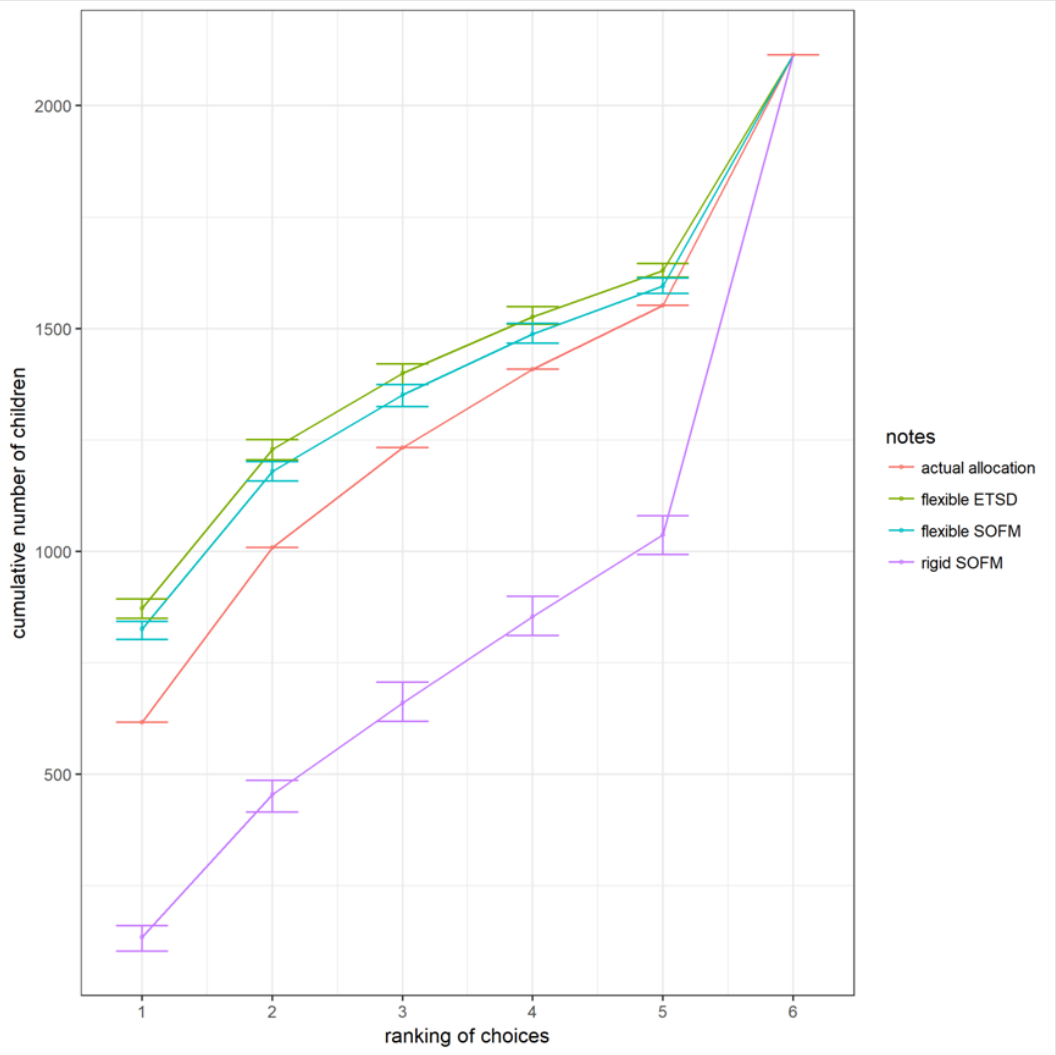

FIGURE 4. Rank distributions under different mechanisms: The graph reports the average cumulative number of children at each rank, as well as its range across all 250 simulation runs.

We also study what happens in serial dictatorship if the rigid constraint is removed so that it is only subject to the daycare constraint. In the induced mechanism, flexible envy-tolerating serial dictatorship (flexible ETSD), some envy is tolerated while the constraint is flexible in this mechanism. Thus, its efficiency is expected to be even higher than both Bunkyo's actual assignment and SOFM under daycare constraint. Somewhat surprisingly, however, the magnitude of the improvement of this mechanism over the flexible SOFM seems rather small; The average number of unmatched children decreases only by 63.83 (7.29%), and the average number of children who become strictly better off under the flexible ETSD is 102.84 (4.86%). This difference is smaller than improvement of the flexible

SOFM over Bunkyo's assignment, whose corresponding numbers are 110.26 (11.18%) and 534.94 (25.30%), respectively. Meanwhile, the measures of envy show similar magnitudes to those for Bunkyo's assignment. These numbers may suggest that the flexible SOFM may be a potentially useful mechanism in daycare allocation.

Overall, the numerical analysis we report here suggest that the main findings for Yamagata are robust to data features.