Supplementary Material

January 19, 2009

These notes consist of three sections. Section 1 defines and establishes the basic properties of the three algorithms in the large economy with a continuum of agents. Section 2 describes our simulation procedures in a greater detail than in the paper. Section 3 presents some omitted proofs of the paper.

1 Extensions of Algorithms to the Continuous Environment

It is convenient to explicitly model the randomizing device used to break the ties. For our purpose, it is sufficient to consider a vector $\omega = (\omega_1, ..., \omega_n) \in [0,1]^n =: \Omega$ of uniformly and independently generated numbers. (The vector of ω will be sufficiently rich enough to model the procedures we study.) Formally, we augment the type space by incorporating the random draw to $\mathcal{V} \times \Omega =: \Theta$, with its generic element denoted $\theta := (\mathbf{v}, \omega)$, and endow it with a product measure $\eta = \mu \times \xi_1 \times ... \times \xi_n$, where ξ_i is a uniform measure satisfying $\xi([0, \omega_i]) = \omega_i$ for each $\omega_i \in [0,1]$. This formalism avoids appealing to the law of large numbers (on the continuum of agents), by ensuring that a fraction ω_i of the student mass draws ω_i or less on each i-th random variable. A student of type $\theta = (\mathbf{v}, \omega)$ is then interpreted as having values $\mathbf{v}$ and drawing a vector ω. The student never observes ω, so her action required by the procedure will be measurable with respect to only $\mathbf{v}$; whereas (part or all of) ω component is "discovered" by the schools for their use in tie-breaking.

An *ex post allocation* is a measurable function $\psi := (\psi_1, ..., \psi_n) : \Theta \mapsto \Delta$ such that $\psi_i(\theta) \in \{0,1\}$ and that $\int \psi_i(\theta)\eta(d\theta) = 1$ for each $i \in S$. Namely, ψ assigns a student with $\mathbf{v}$ to school j upon drawing ω such that $\psi_j(\mathbf{v}, \omega) = 1$. Let $\mathcal{Y}$ be the set of all ex post allocations. Later, we shall describe how each procedure generates an ex post allocation. Some procedures may not use the entire vector of ω, so the ex post allocation they produce may be measurable with respect to only some components of ω.

We define the alternative DA procedures here.

Ordinal preferences. In any DA algorithm, every student submits a ranking of schools. Formally, students' ordinal preferences are represented by a measurable function $P : \Theta \to \Pi$, where $P(\mathbf{v}, \omega) \in \Pi$ is an ordered list of n schools (ordered not necessarily according to true preferences). Since the ω is unobserved by the students (at least at the time of submitting the ordinal preferences), we require that $P(\mathbf{v}, \omega) = P(\mathbf{v}', \omega')$ whenever $\mathbf{v} = \mathbf{v}'$. We say a DA algorithm is *ordinally strategy-proof* if it is a (weak) dominant strategy for each student with $\mathbf{v}$ to choose $P(\mathbf{v}, \omega) = \pi(\mathbf{v})$.

School priorities (tie-breaking rules). We introduce a tie-breaker function which determines the priority of each student for each school as a function of the random draw (as well as their auxiliary message in the case of CADA), in the event of a tie. Formally, *tie-breaker function for school* i is a bounded measurable function $F_i : \Theta \mapsto \mathbb{R}$, such that a student θ' is interpreted as having a higher priority than student θ if $F_i(\theta') < F_i(\theta)$. A *tie-breaker* is a profile $\mathcal{F} = \{F_i : i \in S\}$ of tie-breaker functions. Specifically, the tie-breakers for DA-STB, DA-MTB, and CADA are determined as follows:

- *DA-STB:* The STB rule uses the same tie-breaker function for all schools. This is modeled by a tie-breaker with

  $$F_i(\mathbf{v}, \omega_1, ..., \omega_n) = \omega_1,$$

  $\theta = (\mathbf{v}, \omega_1, ..., \omega_n)$, for *every* school $i \in S$. In other words, a draw's draw ω_1 serves as a priority number for all schools. Heuristically, a real number ω_1 is drawn randomly from an interval $[0, 1]$, for each student, which then serves as her priority score.[1]

- *DA-MTB:* The MTB rule produce a randomly and independently drawn priority list for each school. This is modeled by a tie-breaker, with

  $$F_i(\mathbf{v}, \omega_1, ..., \omega_n) = \omega_i,$$

  for $i \in S$ and for each $\theta = (\mathbf{v}, \omega_1, ..., \omega_n)$. In other words, for each student, a vector $(\omega_1, ..., \omega_n)$ of independent draws determines her priority scores at different schools.

- *CADA:* In CADA, each student sends an auxiliary message of a target school (in addition to their ordinal preferences over schools). Given a (measurable) strategy profile $s : \mathcal{V} \to S$ determining the auxiliary message for each intrinsic type $\mathbf{v}$, the tie-breaker function for school i is given by

  $$F_i(\mathbf{v}, \omega_1, ..., \omega_n) = \begin{cases} \omega_1 & \text{if } s(\mathbf{v}) = i \\ 1 + \omega_2 & \text{if } s(\mathbf{v}) \neq i \end{cases}$$

[1] This heuristics invokes a law of large numbers, but our formal method does not rely on it for we assume a well-behaved randomization device.

That is, under F_i, ties are broken first in favor of students who target i, within them according to the random draw ω_1, and then ties among the rest are broken according to a random draw $\omega_2 + 1$ (where 1 act as a "penalty score" 1). Clearly, F_i is a measurable function since ω_1 and s are measurable.

Definition of DA algorithms: Given ordinal preferences P and a tie-breaker $\mathcal{F} = \{F_i : i \in S\}$, a DA algorithm is defined as follows. First, we define a measurable function Ch_{F_i} over subsets of Θ as the set of best ranked students for school $i \in S$ according to F_i from a given set up to the capacity. Formally, for any measurable $X \subset \Theta$, let

$$Ch_{F_i}(X) := \sup\{Y \subset X | \eta(Y) \leq 1, F_i(\theta) < F_i(\theta'), \forall \theta \in Y, \theta' \in X \backslash Y\}$$

denote the set of students chosen from X such that the set does not exceed the capacity and that the chosen students have a higher priority than those not chosen.

Next, we define the $DA_{\mathcal{F}}$ (deferred acceptance) mapping. Consider first a mapping $Q : \Theta \to \Pi$, where $Q(\theta)$ is an ordered list of any $k \leq n$ schools. (Recall $P(\theta)$ is a special case involving the full set of schools.) The DA mapping, $Q' = DA_{\mathcal{F}}(Q) \in \Pi$ is determined as follows. Every student with θ applies to her most preferred school in $Q(\theta)$. Every school i (tentatively) admits from its applicants in the order of F_i. If all of its seats are assigned, it rejects the remaining applicants. If a student θ is rejected by i, $Q'(\theta)$ is obtained from $Q(\theta)$ by deleting i in $Q(\theta)$. If a student θ is not rejected, then $Q'(\theta) = Q(\theta)$. More formally, let $T_i(Q) = \{\theta \in \Theta : i \text{ is ranked first in } Q(\theta)\}$ be the set of students that rank i as first choice. Note that $T_i(Q)$ is measurable. Then each school i admits students in $Ch_{F_i}(T_i(Q))$ and rejects students in $T_i(Q) \backslash Ch_{F_i}(T_i(Q))$. If $\theta \in T_i(Q) \backslash Ch_{F_i}(T_i(Q))$ for some $i \in S$, then $Q'(\theta)$ is obtained from $Q(\theta)$ by deleting i from the top of $Q(\theta)$; otherwise $Q'(\theta) = Q(\theta)$. Since Q is a measurable function, Q' is also measurable.

Repeated application of the $DA_{\mathcal{F}}$ mapping gives us the DA algorithm. That is, given a problem $(P, \mathcal{F})$, let $Q^0 = P$ and define $Q^t = DA_{\mathcal{F}}(Q^{t-1})$ for $t > 0$. Then Q^t converges almost everywhere to some measurable Q^* (Theorem 0 below). The matching can be then found by assigning θ to its top choice of $Q^*(\theta)$. Formally, define a mapping $\psi^{(P,\mathcal{F})} : \Theta \mapsto \Delta$ such that $\psi_i^{(P,\mathcal{F})}(\theta) = 1$ if i is the top choice of $Q^*(\theta)$, and $\psi_i^{(P,\mathcal{F})}(\theta) = 0$ otherwise. Since the schools' capacities are respected in each round and also in the limit, the mapping must be an ex post allocation.

We present two main results:

Well-definedness of the Procedure. The existence of $\psi^{(P,\mathcal{F})}$ follows from the next theorem.

Theorem 0. *For every $(P, \mathcal{F})$, $DA^t_{\mathcal{F}}(P)$ converges almost everywhere to some measurable $Q^* : \Theta \to \Pi$.*

Proof: Define the set of rejected students as $R^t = \{\theta : \theta \in T_i(Q^t) \backslash Ch_{F_i}(T_i(Q^t))$ for some $i \in S\}$. Then $\eta(R^t)$ goes to zero as t goes to infinity. Otherwise, if $\eta(R^t) \geq \kappa > 0$ for all t, all the schools in every student's preference would be deleted in finite time because of finiteness of the number of schools, which in turn would imply that $\eta(R^t)$ goes to zero, a contradiction. Therefore, $DA^t_{\mathcal{F}}(P)$ converges almost everywhere to some Q^*. Since every $Q^t = DA^t_{\mathcal{F}}(P)$ is measurable, Q^* is also measurable. ▌

2 The Simulation Procedure

There are 5 schools each with a capacity of 20 seats and 100 students. Fix α. We independently draw 100 sets of vNM values for students. Let $\{\tilde{v}^s_{ij}\}$ denote a draw of vNM values, where superscript s denote the draw and $\tilde{v}^s_{ij}$ denotes student i's vNM value for school j. We normalize the level and the scale of each student's vNM utilities as follows:

$$v^s_{ij} = \frac{\tilde{v}^s_{ij} - \min_{j'} \tilde{v}^s_{ij'}}{\max_{j'} \tilde{v}^s_{ij'} - \min_{j'} \tilde{v}^s_{ij'}}$$

Given a normalized draw $\{v^s_{ij}\}$, fix the mechanism, define the following: p^s_{ij} is the probability that student i is assigned school j under the mechanism. $\pi^s_k(i)$ is the school that is ranked k-th in i's preference list. P^s is the set of popular schools. O^s is the set of oversubscribed schools in an equilibrium of CADA with no naive players.

A first best or utilitarian maximum solves

$$\overline{v}^s_{FB} = \frac{1}{100} \max_{\{\hat{p}^s_{ij}\}} \sum_i \sum_j \hat{p}^s_{ij} v^s_{ij}$$

Let $\{\overline{p}^s_{ij}\}$ denote a solution to the first best. There may be multiple solutions, we arbitrarily pick one.

Furthermore, we calculate

$$\delta^s_1 = \frac{1}{100} \sum_i \sum_j p^s_{ij} \cdot \mathbf{1}(\pi^s_1(i) = j)$$

where δ^s_1 is the average probability of assigning a student to her first choice.

In the CADA experiments with naive players, we divide the set of students into two: N is the set of naive players who always target their first choice, and S is the set of strategically sophisticated players who play their best response strategies given others' strategies. We calculate utilitarian welfare as before. We also compute the number of students targeting their k-th choice in equilibrium, which we denote by T^s_k, $k \in \{1, 2, 3, 4\}$.

Given a draw $\{v_{ij}^s\}$, the set P^s is determined trivially. Next we describe how the other numbers are computed.

A *single tie breaker* is a list of 100 randomly drawn lottery numbers, one for each student. Under DA-STB the ties at a school are broken according to students' single random numbers. In CADA, we draw two single tie breakers, one to be used to break ties at one's target school, the other to be used at one's other schools. A *multiple tie breaker* is a list of $100 \times 5 = 500$ randomly drawn numbers, one for each student at each school. Under DA-MTB, the ties at a school are broken according to students' tie breaker numbers at that school.

For each draw $\{v_{ij}^s\}$, we independently draw 2,000 single tie breakers for DA-STB, and an additional set of 2,000 single tie breakers for CADA, and 2,000 multiple tie breakers for DA-MTB. Then p_{ij}^s for a mechanism is computed by

$$\frac{\text{Number of tie breakers at which } i \text{ is assigned } j}{2,000}.$$

The equilibrium of CADA is computed with single tie breakers being fixed. Given the strategies of other students, a student's best response is found by computing that student's expected utility over those tie breakers. Then O^s, the set of oversubscribed schools, is found by using students's equilibrium target schools. In experiments with naive players, naive players' target schools are fixed at their first choice.

Note that we are approximating the equilibrium by drawing (two sets of) 2,000 independent tie-breakers. The exact numbers are computed by considering 100! single tie-breakers and $(100!)^5$ multiple tie breakers, which is beyond the capabilities of our computational resources. Any further increase in the number of tie breakers beyond 2,000 does not increase the precision of our computations significantly.

For each $z^s \in \{\overline{v}^s, \overline{v}^s_{FB}, \pi_1^s, T_1^s, T_2^s, T_3^s, T_4^s, |P^s|, |O^s|\}$, we compute the average of z^s by

$$z = \frac{1}{100} \sum_{s=1}^{100} z^s.$$

Note that we drop all "s" from a variable to denote its mean over 100 iterations of an experiment. We report $100\frac{\overline{v}}{\overline{v}_{FB}}$ in our welfare figures.

3 Omitted Proofs

3.1 Ordinal Strategy-proofness.

Fix arbitrary ordinal preferences P. Let $P_{-\mathbf{v}} : \mathcal{V} \backslash \{\mathbf{v}\} \to \Pi$ denote the ordinal preferences of all students but $\mathbf{v}$ determined by P. Recall that $\pi(\mathbf{v}) \in \Pi$ represents the truthful ordinal preference induced by $\mathbf{v}$, that is $\pi(\mathbf{v})$ lists i before j if and only if $v_i > v_j$. To simplify the notation, let $\psi^P := \psi^{(P,\mathcal{F})}$, with $\mathcal{F}$ suppressed, and let $\psi^* := \psi^{(\pi[\cdot],P_{-[\cdot]},\mathcal{F})}$ denote the matching outcome for any given type when it submits its ordinal preferences truthfully and the others report P. When students report P, a student with type $\mathbf{v}$ receives expected utility of

$$\mathbb{E}_\omega\left[v \cdot \psi^P(\mathbf{v},\omega)\right].$$

Theorem 1. *For every $(P,\mathcal{F})$, it is a (weak) dominant strategy for every student to submit her ordinal preferences truthfully to DA, that is, for all $\mathbf{v} \in \mathcal{V}$, P,*

$$\mathbb{E}_\omega\left[\mathbf{v} \cdot \psi^*(\mathbf{v},\omega)\right] \geq \mathbb{E}_\omega\left[\mathbf{v} \cdot \psi^P(\mathbf{v},\omega)\right].$$

Proof. It suffices to show that, for all $\theta = (\mathbf{v},\omega)$,

$$\mathbf{v} \cdot \psi^*(\mathbf{v},\omega) \geq \mathbf{v} \cdot \psi^P(\mathbf{v},\omega).$$

Suppose to the contrary that

$$\mathbf{v} \cdot \psi^*(\mathbf{v},\omega) < \mathbf{v} \cdot \psi^P(\mathbf{v},\omega), \tag{1}$$

for some $\theta = (\mathbf{v},\omega)$ and P. We show that there exists a finite many-to-one matching problem for which a DA algorithm fails strategy-proofness, which will then constitute a contradiction to the standard strategy-proofness result (Dubins and Friedman, 1981; Roth, 1982).

To begin, fix any $K \in \mathbb{N}_+$, and construct a discretization of (P,F) for θ as follows: For every $\mathbf{z} = (z_1,...,z_n)$ and $\mathbf{y} = (y_1,...,y_n)$ where $z_i, y_j \in \{0,...,K\}$, consider a set

$$\Theta_{\mathbf{z},\mathbf{y}} = \left\{(\tilde{\mathbf{v}},\tilde{\omega}) \in \Theta : \frac{z_i}{K} \leq \tilde{v}_i \leq \frac{z_i+1}{K}, \frac{ny_j}{K} \leq \tilde{\omega}_j \leq \frac{n(y_j+1)}{K}, i,j \in S\right\}.$$

Let $\eta_{K,\min} = \min\limits_{\Theta_{\mathbf{z},\mathbf{y}}} \eta(\Theta_{\mathbf{z},\mathbf{y}})$, and let $\#\Theta_{\mathbf{z},\mathbf{y}}$ be the integer part of $\frac{\eta(\Theta_{\mathbf{z},\mathbf{y}})}{\eta_{K,\min}}$.

Pick $\#\Theta_{\mathbf{z},\mathbf{y}}$ students in total from every set $\Theta_{\mathbf{z},\mathbf{y}}$ at random without repetition. Let $\{\theta^l\}$ denote the set of students that are picked. If $\frac{|\{\theta^l\}|}{n}$ is not an integer, pick additional students from the larger sets until obtaining an integer $\frac{|\{\theta^l\}|}{n}$. Note that the number of additional students to be picked this way is less than n and n is fixed, therefore this will be negligible in the limit as K goes to infinity. Now consider the problem in which the set $\{\theta^l\}$ of students are to be assigned to a set S of schools each with capacity $\frac{|\{\theta^l\}|}{n}$. Each student $\theta^l = (\mathbf{v}^l,\omega^l)$'s strict ordinal preference is given by $P(\theta^l)$. The schools' strict preferences are given by $\mathcal{F}$. Denote this problem by $(\{\theta^l\},S,P,\mathcal{F})_K$, and the associated ex post allocation ψ_K^P. As K goes to infinity, $(\{\theta^l\},S,P,\mathcal{F})_K$ approximate $(\Theta,S,P,\mathcal{F})$ arbitrarily closely.

Hence, $\psi_K^P \rightarrow_{a.e} \psi^P$ and $\psi_K^{\pi(\mathbf{v}),P_{-\mathbf{v}}} \rightarrow_{a.e.} \psi^*$ as $K \rightarrow \infty$. Hence, if (1) holds, then there exists K such that

$$\mathbf{v} \cdot \psi_K^*(\mathbf{v},\omega) < \mathbf{v} \cdot \psi_K^P(\mathbf{v},\omega).$$

This contradicts the fact that, in every finite problem, submitting true preferences to the student-proposing deferred acceptance mechanism is a dominant strategy for every student (Dubins and Friedman, 1981; Roth, 1982). ▌

3.2 Welfare Performances of CADA with Naive Students

We first prove the claim made in page 18 of the text that the characterization of Theorem 5-(ii) is tight in the sense that there is generally no bigger set that includes all oversubscribed schools *and some undersubscribed school* that supports Pareto efficiency. To see this, suppose there are four schools, $S = \{a, b, c, d\}$, and four types of students $\mathcal{V} = \{\mathbf{v}^1, \mathbf{v}^2, \mathbf{v}^3, \mathbf{v}^4\}$, with $\mu(\mathbf{v}^1) = \frac{3-\varepsilon}{2}$, $\mu(\mathbf{v}^2) = \frac{1+\varepsilon}{2}$, $\mu(\mathbf{v}^3) = \frac{3-\varepsilon}{2}$, and $\mu(\mathbf{v}^4) = \frac{1+\varepsilon}{2}$ where ε is a small number.

	v_j^1	v_j^2	v_j^3	v_j^4
$j = a$	10	10	20	20
$j = b$	3	5	9	8
$j = c$	1	4	8	1
$j = d$	0	0	0	0

In this case, type 1 and 3 students subscribe to school a, and type 2 and 4 students subscribe to school b. More specifically, the allocation ϕ^* has $\phi^*(\mathbf{v}^1) = \phi^*(\mathbf{v}^3) = (\frac{1}{3-\varepsilon}, 0, \frac{2-\varepsilon}{2(3-\varepsilon)}, \frac{2-\varepsilon}{2(3-\varepsilon)})$ and $\phi^*(\mathbf{v}^2) = \phi^*(\mathbf{v}^4) = (0, \frac{1}{1+\varepsilon}, \frac{\varepsilon}{2(1+\varepsilon)}, \frac{\varepsilon}{2(1+\varepsilon)})$. Although schools a and b are oversubscribed, this allocation is not PE within $\{a, b, c\}$ since type 1 students can trade probability shares of school 1 and 3 in exchange for probability share of b, with type 1 students. The allocation is not PE within $\{a, b, d\}$ either, since type 3 students can trade probability shares of school a and d in exchange for probability share of b, with type 4 students. Therefore $\{a, b\}$ is the largest set that supports Pareto efficiency.

Theorem 8. *In the presence of naive students, the equilibrium allocation of CADA satisfies the following properties: (i) The allocation is OE, and is thus pairwise PE. (ii) The allocation is PE within the set K of oversubscribed schools. (iii) If every student is naive, then the allocation is PE within $K \cup \{l\}$ for any undersubscribed school $l \in J := S \backslash K$.*

Proof: Part (i) is precisely the same as Part (i) of Theorem 5 and is a consequence of Part (ii) below and of Part (ii) of Lemma 5 (which does not depend on whether the students are naive or not). Hence it is omitted. To prove Part (ii), it is useful to establish the following lemma. As before, let ϕ^* denote the ex ante allocation arising from the CADA game and let K and $J = S\backslash K$ be respectively the sets of oversubscribed and undersubscribed schools in equilibrium.

Lemma N. *Any reassignment of* $\phi^*(\mathbf{v})$ *within* K *will make a naive student with* $\mathbf{v}$ *strictly worse off, for almost every* $\mathbf{v}$.

Proof: Consider a naive student with $\mathbf{v}$. Assume without loss of generality that she prefers i strictly over all other schools (i.e., $v_i > v_j$, $\forall j \neq i$). (This is without loss of generality since the values are distinct for almost every student type.) Since the student is naive, she subscribes to school i with probability 1. If school i is undersubscribed, then the result is trivial since $\phi^*_k(\mathbf{v}) = 0$ for all $k \in K$. Hence, suppose school i is oversubscribed. Then, any reassignment $\mathbf{x} \in \Delta^K_{\phi^*(\mathbf{v})}$ must satisfy

$$\sum_{j\in K} x_j = \phi^*_i(\mathbf{v})$$

Since $v_i > v_j \; \forall j \neq i$, for any $\mathbf{x} \in \Delta^K_{\phi^*(\mathbf{v})}$, $\mathbf{x} \neq \phi^*_i(\mathbf{v})$, we must have

$$\sum_{j\in K} x_j v_j < \sum_{j\in K} x_j v_i = \phi^*_i(\mathbf{v}) v_i,$$

which implies that the student must be strictly worse off from any such reassignment. ‖

We are now ready to prove Parts (ii) and (iii):

Part (ii): We make use of the proof of Theorem 5. By Lemma N, a type-$\mathbf{v}$ naive student's assignment from the CADA, $\phi^*(\mathbf{v})$, is a unique solution to $[P(\mathbf{v})]$, for a.e. $\mathbf{v}$, even without the constraint

$$\sum_{i\in K} p_i x_i \leq \sum_{i\in K} p_i \phi^*_i(\mathbf{v}). \tag{2}$$

Since $\phi^*(\mathbf{v})$ is feasible under (2), this must be a unique solution to $[P(\mathbf{v})]$.

For a non-naive student with a.e. $\mathbf{v}$, the proof of Theorem 5 follows directly, so $\phi^*(\mathbf{v})$, is also a unique solution of $[P(\mathbf{v})]$. Since the equilibrium assignment of both types solves $[P(\mathbf{v})]$, the rest of the argument in the proof of Theorem 5 applies, proving that we ϕ^* is PE within K. ‖

Part (iii): Again let ϕ^* be the ex ante allocation arising from CADA. Suppose to the contrary that there exists a within-$K \cup \{l\}$ reallocation $\tilde{\phi}$ of ϕ^* that Pareto dominates ϕ^*. By Part (ii), ϕ^* is PE within K, so $\tilde{\phi}_l(\mathbf{v}) \neq \phi^*_l(\mathbf{v})$ for a positive measure of $\mathbf{v}$, which in turn implies that there exists a set $A \subset \mathcal{V}$ with $\mu(A) > 0$ such that $\tilde{\phi}_l(\mathbf{v}) > \phi^*_l(\mathbf{v})$ for each $\mathbf{v} \in A$. Since $\tilde{\phi}(\mathbf{v}) \in \Delta^{K\cup\{l\}}_{\phi^*(\mathbf{v})}$, $\sum_{j\in K\cup\{l\}} \tilde{\phi}_j(\mathbf{v}) = \sum_{j\in K\cup\{l\}} \phi^*_j(\mathbf{v})$, so

$$\sum_{j\in K} \tilde{\phi}_j(\mathbf{v}) < \sum_{j\in K} \phi^*_j(\mathbf{v}) \;\; \text{for all } \mathbf{v} \in A.$$

Assume without loss that $\mathbf{v}$ satisfies $v_i > v_j$ for all $i \in K$ and for all $j \neq i$. ($i \neq l$ since $\tilde{\phi}_l(\mathbf{v}) > \phi^*_l(\mathbf{v})$ is impossible if $i = l$.) Then, the type-$\mathbf{v}$ student's expected payoff from $\tilde{\phi}$ is

$$\begin{aligned}
\sum_{j \in S} \tilde{\phi}_j(\mathbf{v}) v_j &= \sum_{j \in K} \tilde{\phi}_j(\mathbf{v}) v_j + \tilde{\phi}_l(\mathbf{v}) v_l + \sum_{j \in J \setminus \{l\}} \phi^*_j(\mathbf{v}) v_j \\
&< \sum_{j \in K} \tilde{\phi}_j(\mathbf{v}) v_i + \left(\tilde{\phi}_l(\mathbf{v}) - \phi^*_l(\mathbf{v}) \right) v_i + \phi^*_l(\mathbf{v}) v_l + \sum_{j \in J \setminus \{l\}} \phi^*_j(\mathbf{v}) v_j \\
&= \left(\sum_{j \in K \cup \{l\}} \tilde{\phi}_j(\mathbf{v}) - \phi^*_l(\mathbf{v}) \right) v_i + \sum_{j \in J} \phi^*_j(\mathbf{v}) v_j \\
&= \sum_{j \in S} \phi^*_j(\mathbf{v}) v_j.
\end{aligned}$$

Since this inequality holds for almost every $\mathbf{v} \in A$, and since $\mu(A) > 0$, $\tilde{\phi}$ cannot Pareto dominate ϕ^*. ▮