Posterior Stable Matching*

Archishman Chakraborty†
Baruch College, CUNY
New York

Alessandro Citanna‡
HEC
Paris

First Version: July, 2003. Revised: July 2006

Abstract

With interdependent values, assignments of students to schools may be challenged after the outcome of the assignment is known, prompting the question: are there assignments that are posterior stable and (weakly) implementable? When the entire allocation outcome is known to everyone after an assignment is selected, we show that there is no incentive compatible mechanism that implements posterior stable allocations in general. We then look at situations where each school knows only its own allocation outcome. In such environments, with one-sided incomplete information, we identify necessary and sufficient conditions on communication strategies available to blocking coalitions under which posterior stable and implementable allocations exist. We construct a 'modified serial dictatorship' direct mechanism that implements them. We also investigate efficiency and monotonicity properties of modified serial dictatorship.

Keywords: interdependent values, posterior stability, modified serial dictatorship.

JEL classification numbers: C78, D82.

---

*This research is sponsored by NSF grant #0617850. We wish to thank Atila Abdulkadiroglu, Dirk Bergemann, Ronnie Razin, Ilya Segal and Paolo Siconolfi for useful discussions. A preliminary version was circulated starting June 2003. The paper was completed while the second author was visiting the Graduate School of Business of Columbia University, New York.

†archishman_chakraborty@baruch.cuny.edu

‡citanna@hec.fr

1 Introduction

Consider a student assignment problem under incomplete information. Schools privately observe noisy signals of each student's quality before offering admission, where a student's quality may be common to all schools so that each school's signal is informationally relevant for the others. With such interdependent values, an assignment (equivalently termed an allocation, matching or admission) of students to schools may reveal valuable information to all parties. Consequently, a subset of schools or students may unanimously prefer an alternative assignment, given their *posterior beliefs* about the state of the world, updated after observing the allocation outcome, and based on their knowledge of the assignment mechanism. In this paper we characterize environments where the original mechanism is posterior stable, i.e., immune from renegotiation possibilities that arise at the posterior stage after everyone observes the outcome of the status quo allocation rule.

There are many other examples of matching problems that share important features of this student assignment problem. For instance, at entry level positions, employees may be assigned to work for divisions they rank unanimously from more to less prestigious. Each division can interview the employee and get information on their quality, over and beyond the publicly known prior. As another example, a decision-maker may have to delegate a decision to one of several experts, with each expert privately informed about the merits of taking on the case and interdependent values among the experts. As a final example, an author may have to submit an article to a peer-reviewed journal, when each journal only obtains a noisy signal of the paper's true quality and where the author may withdraw the paper after acceptance and submit it to another journal. Hereafter, we will stick to the student assignment problem as the baseline story underlying our terminology.

In addition, in our analysis of environments with one-sided incomplete information and interdependent values, we assume that students have no private information on their quality and that their preferences over schools are common knowledge and independent of the schools' private information. Furthermore, we focus on purely informational issues and so assume that schools have preferences that are separable over students, that they face no capacity constraints and that the student side of the matching market has homogeneous preferences. Finally, schools and students have a no-assignment option available, and monetary transfers are excluded.

Even in this simple set-up, a suitable definition of posterior stability has to address two fundamental issues. The first is observability, i.e., the extent to which the status quo outcome is publicly observed at the posterior stage when coalitions may raise objections. The second is communication, i.e., the restrictions if any that one imposes on information sharing among coalition members at the time they raise objections.

With respect to the observability issue, we consider two polar notions: strong and weak. Under strong observability, all schools and students observe the entire allocation outcome including parts that do not directly concern them. Under weak observability, each school (student) only observes the students (schools) assigned to it and does not directly observe the students assigned to other schools. If one thinks of the original mechanism as a centralized direct mechanism, then strong observability can be thought of as the case where a mechanism designer publicly announces the allocation outcome, whereas weak observability corresponds to the case where the designer privately communicates the relevant parts of the outcome to each party. We say that an allocation rule is strong (respectively, weak) posterior stable if it is posterior stable under strong (resp., weak) observability.

With respect to the communication issue, we allow coalitions to object to the decision made by the status quo mechanism via unanimous voting on an alternative decision. For a block to succeed, each coalition member must be strictly better off in expectation from accepting the alternative, conditional not only on the observed status quo outcome, but also on the acceptance of other coalition members. Such a notion allows communication to a limited degree, since acceptance strategies may depend on information that is private. However, the coalition cannot design a general communication mechanism. This last restriction is natural since we focus on situations where coalition members object to the status quo decision at the posterior stage by proposing an alternative decision, and do not necessarily object to the entire decision rule. It is also economically relevant since a general mechanism may be too costly to operate for a blocking coalition at the late stage when the status quo rule has already proposed an allocation. These costs may be direct or take the form of constraints in designing complex mechanisms that require precise specification of randomization schemes and message games (for a similar point see Forges, 1994). However, quite apart from how reasonable these economic constraints are, we also show that such restrictions on communication are not arbitrary but actually *necessary* in order to obtain any general existence result.

We first establish some preliminary properties of posterior stable allocation rules, in particular their existence and their relation to other already established interim and ex-post notions of stability. Since we have in mind environments where information is private and, in particular, cannot be verified ex-post (by a court of law, say), it is natural to think of posterior stability in conjunction with the question of eliciting truthful information revelation. Accordingly, we look next at posterior stable allocation rules that are also weakly implementable, i.e., incentive compatible when written as a direct mechanism.

We show that strong posterior stable allocations are generally not incentive compatible. We provide a nonexistence example, which obtains even under strong affiliation and symmetry conditions on the environment, and does not depend on the restrictions on communication possibilities that we impose. This result shows that under strong observability, too much information may be contained in the outcome of any incentive compatible rule for it to be posterior stable.

On the other hand, we show that under weak observability, the set of posterior stable and incentive compatible allocation rules is always nonempty. We provide a direct mechanism that can be used to construct such rules as a function of the signal and student distribution, that we call *modified serial dictatorship*. With such a mechanism, schools are ordered in a descending sequence according to the student's preferences. Each school is then assigned the student based on the school's evaluation of the student, conditional on the school's own private signal and on the signals of schools above (but not below) them in the student's rankings. With such an allocation, and under weak observability, no subset of schools or students can infer from the observed outcome enough information that allows it to profit via a blocking coalition.

In contrast to the negative result under strong observability, the positive result on the existence of weak posterior stable incentive compatible rules does depend on the restrictions on communication that we impose. Specifically, we assume that the alternative decision voted on by a blocking coalition (i) must make each coalition member strictly better off relative to the status quo, (ii) cannot involve lotteries and other randomization schemes and (iii) cannot depend on messages sent by coalition members. Absent restriction (i), coalition members may sell 'free' information to other coalition members by participating in a block even though they do not directly gain from it, precluding the existence of stable rules. On the other hand, absent restriction (ii), coalition members may 'stochastically reward' other members for information via the use of lotteries, also precluding the existence of stable rules. Similarly, absent restriction (iii), coalition members may artificially create randomization schemes by making allocations depend on payoff-irrelevant messages. In this precise sense, as we show below, all of the restrictions (i)–(iii) are necessary to obtain the existence of weak posterior stable and incentive compatible rules, giving our definitions and results a normative content.

We also consider what happens when 'renegotiation is anticipated' by the schools when they submit their signals. We modify the notion of blocks by allowing schools to object to a given allocation also after lying to the mechanism. Allocations obtained through modified serial dictatorship are incentive compatible and stable even under this wider notion of blocks.

Modified serial dictatorship is posterior efficient and gives rise to ex-post efficient allocations when students rank higher the schools with higher opportunity costs of admitting a student. In such situations, it is also natural to expect that for a stable rule, the average quality of the student at a better school must be greater than at a worse school. While this is true for strong posterior stable allocations, we show that weak posterior stable allocations may not have this property, even in affiliated environments. In this sense the matching pattern may not be monotonic or assortative in incomplete information settings.

The literature on matching has usually considered private value environments. Roth and Sotomayor (1990) provide a survey of many earlier results, including those of Roth (1989). In simultaneous and independent work, Ostrovsky (2005) considers the question of posterior stability in a one-to-one matching problem under common values. He provides an impossibility result under strong observability, similar in spirit to our Example 2. He assumes that schools have identical preferences over students, requires schools to block in a pre-specified order and obtains nonexistence in the presence of at least three students. Instead, we consider a many-to-one problem, allow general preferences for schools, use a notion of blocks similar to definitions of the core under incomplete information, and are able to show nonexistence under strong observability with only one student. We also provide necessary and sufficient conditions for the existence of incentive compatible and stable rules under weak observability, a question not addressed by Ostrovsky (2005). Other recent papers that include some notion of reallocation proofness, such as Abdulkadiroglu and Sonmez (1999, 2003) and Papai (2000), study school choice, as opposed to student assignments, and are not concerned with the common value aspect of the problem.

Our work is closely related to notions of the core under incomplete information considered in the literature on cooperative games (e.g., Wilson (1978), Vohra (1999), Forges, Mertens and Vohra (2001); for a survey, see Forges, Minelli and Vohra (2001)). An important issue in this literature that is shared by the present paper concerns the question of communication, i.e., endogenous sharing of information and the associated incentive compatibility and self-selection constraints that arise when coalitions raise objections under incomplete information (Volij (2000), Dutta and Vohra (2001)). As shown in Example 1, our focus on objections at the posterior stage distinguishes it from interim notions such as the coarse core, since we use different measurability restrictions on blocks and different information to let the agents evaluate them. In particular, the information available to coalitions at the time they may raise objections is endogenous in our set up, since the allocation outcome itself (an endogenous function of private signals) is used by agents to evaluate the allocation.

Forges (1994) also considers situations where the allocation outcome is used to evaluate an allocation —a notion of strong posterior efficiency. As a difference, we insist on asking for more than just efficiency (i.e., stability to objections by the grand coalition). Our Example 2 shows that, once stability is taken into account, the positive results of Forges (1994, Properties 1, 4) are reversed.

In the implementation literature, our work is related to the idea of durable mechanisms in Holmstrom and Myerson (1983), in the sense of allowing endogenous information sharing via unanimous voting on alternatives. As in Maskin and Moore (1999), we look at implementation and renegotiation, but again we ask for stability of the outcome, and do not require ex-post efficiency.

The rest of the paper proceeds as follows. In Section 2 we set up our environment. In Section 3 we define our notions of posterior stability and compare them with interim core notions. We demonstrate the non-existence of incentive compatible strong posterior stable rules, provide a constructive proof of existence of weak posterior stable ones via modified serial dictatorship, and establish the necessity of restricting communication among blocks to obtain existence. In Section 4 we discuss renegotiation and the efficiency and monotonicity properties of modified serial dictatorship. Section 5 concludes while the Appendix contains all the proofs and the details of some examples.

2 The environment

The agents are students and schools. The set of students is denoted by $\mathbf{S} = \{0, 1, ..., S\}$, with $S \geq 1$ and typical element s, whereas the set of schools is denoted by $\mathbf{I} = \{0, 1, .., I\}$, with $I > 1$ and typical element i. We denote by $i = 0$ a fictitious null school (to capture the possibility that a student does not go to any school) and let $\mathbf{I}_+ = \mathbf{I} \setminus \{0\}$ denote the set of real schools. An analogous role is played by $s = 0$, with $\mathbf{S}_+ = \mathbf{S}\backslash\{0\}$ denoting the set of real students. For some analysis, to each real student we assign an observed type $\lambda \in [0,1]$ and let $\mathbf{S}_\lambda$ be the set of students assigned to type λ, with element s_λ. Let Λ be the set of possible types with $\sum_{\lambda\in\mathbf{\Lambda}} \#(\mathbf{S}_\lambda) = S$.

Students $s \in \mathbf{S}_+$ also differ in their unobserved quality or ability $q_s \in \mathbf{Q}$, a finite set. Neither schools nor students know the quality q_s of any student s, but each school $i \in \mathbf{I}_+$ receives a private signal $x_{i,s}$. This can be thought of as the outcome of a privately observed informative test for student s. We assume that $x_{i,s} \in \mathbf{X} = \{x_1, ..., x_n\}$. Let $x_s = (x_{i,s})_{i\in\mathbf{I}_+}$ be the vector of signals associated with each student s , and $x_i = (x_{i,s})_{s\in\mathbf{S}_+}$ be the vector of signals received by each school $i \in \mathbf{I}_+$, so that $x = (x_s)_{s\in\mathbf{S}_+}$ $=(x_i)_{i\in\mathbf{I}_+}$ represents the private information available on the quality of all students $s \in \mathbf{S}_+$ by all schools $i \in \mathbf{I}_+$. The null school $i = 0$ has no private information. We let Pr be the joint probability distribution over qualities and signals, with full support on $\mathbf{X}^{SI} \times \mathbf{Q}^S$.

The payoff to school i from accepting student s of quality q_s when the signal is x_s is $w_i(x_s, q_s)$, whereas the payoff from not getting a student is 0. In general we will put no further restrictions on w_i, except in the context of some examples, as discussed below. Students derive utility from going to school i, represented by a payoff $v_s(i, x)$. Let $w = \{w_i\}_{i\in\mathbf{I}}$ and $v = \{v_s\}_{s\in\mathbf{S}_+}$.

An environment is identified by an array $(\mathbf{I}, \mathbf{S}, \mathbf{\Lambda}, \mathbf{Q}, \mathbf{X}, \Pr, w, v)$. The problem we face in these environments is to allocate students to schools. Let $\mu_{i,s}(x) \geq 0$ be a real number for $i \in \mathbf{I}$, $s \in \mathbf{S}$ and $x \in \mathbf{X}^{SI}$ that specifies the fraction of time student s is going to school i, given the information x.[1] An *allocation given* $x \in \mathbf{X}^{SI}$ is $\mu(x) = \{\mu_i(x)\}_{i\in\mathbf{I}}$, where $\mu_i(x) = \{\mu_{i,s}(x)\}_{s\in\mathbf{S}}$ is the allocation pertaining to school i, while that pertaining to student s is $\mu_s(x) = \{\mu_{i,s}(x)\}_{i\in\mathbf{I}}$, and $\sum_{i\in\mathbf{I}} \mu_{i,s}(x) = 1$ for all $s \in \mathbf{S}$ and $x \in \mathbf{X}^{SI}$. An *allocation (rule)* is $\mu = \{\mu(x)\}_{x\in\mathbf{X}^{SI}}$.

In general, school i does not know the realized vector x and computes its expected payoff from an allocation given its information. We denote the information of school i by $\mathcal{F}_i$, a partition of $\mathbf{X}^{SI}$. Let $\mathbf{S}' \subset \mathbf{S}$ be a subset of students. If $\mathbf{F}_i \in \mathcal{F}_i$, we assume that the conditional expected payoff to school i from an allocation μ relative to students $\mathbf{S}'$ and given $\mathbf{F}_i$ is additive across students and linear in probabilities, and write it as

$$U_i(\mu, \mathbf{S}'|\mathbf{F}_i) = \sum_{s\in\mathbf{S}'} \sum_{x\in\mathbf{X}^{SI}} u_i(s|x)\mu_{i,s}(x) \Pr[x|\mathbf{F}_i]$$

whenever this is well–defined, i.e., $\Pr(\mathbf{F}_i) > 0$, with the expected payoff being arbitrarily defined otherwise. Here, $u_i(s|x) = \sum_{q_s} w_i(x_s, q_s) \Pr(x_s, q_s|x)$ is the expected payoff to school i from matching with student s when signals are x. For notational simplicity, when $\mathbf{S}'$ is the entire set of students $\mathbf{S}$ we write $U_i(\mu, \mathbf{S}'|\mathbf{F}_i) = U_i(\mu|\mathbf{F}_i)$, whereas when $\mathbf{S}' = \{0, s\}$ and $\mu_{i,s}(x) = 1$ for all $x \in \mathbf{F}_i$, we write $U_i(\mu, \mathbf{S}'|\mathbf{F}_i) = u_i(s|\mathbf{F}_i)$.

Similarly, the information student s has can generally be represented by a partition $\mathcal{F}_s$. The expected payoff for student s from μ conditional on her information $\mathbf{F}_s \subset \mathbf{X}^{SI}$ is

[1] We are allowing a student to attend multiple schools simultaneously. This is a convenient analytical representation, although multiple attendance will not arise at stable allocations. As usual, $\mu_{i,s}(x)$ can also be interpreted as a probability that student s is going to school i at x, without altering anything to follow.

$$U_s(\mu|\mathbf{F}_s) = \sum_{i\in\mathbf{I}} \sum_{x\in\mathbf{X}^{SI}} v_s(i;x)\mu_{i,s}(x)\Pr(x|\mathbf{F}_s)$$

Since an allocation may depend on information that is private to the schools, an allocation μ is informationally feasible if it is incentive compatible for the schools to reveal their private information. Taking recourse to the revelation principle, we think of an allocation $\mu(\widehat{x})$ as a function of direct messages $\widehat{x}$ from the schools. The definition of incentive compatibility is then the usual one. An allocation μ is *incentive compatible* if for each $i \in \mathbf{I}_+$ and each x_i

$$x_i \in \arg\max_{\widehat{x}_i} \sum_{s\in\mathbf{S}} \sum_{x\in\mathbf{X}^{SI}} u_i(s|x_s)\mu_{i,s}(\widehat{x}_i, x_{-i})\Pr[x|x_i] \tag{IC}$$

We impose the following assumptions on our environments. First, qualities and signals are independent across students, and for any student s_λ, the distribution $\Pr_{s_\lambda}$ depends only on the type λ and not on the identity s of the student:

A1 For any $s, s' \in \mathbf{S}_+$, $s \neq s'$, $\Pr(x_s, q_s, x_{s'}, q_{s'}) = \Pr_s(x_s, q_s)\Pr_{s'}(x_{s'}, q_{s'})$; and for each $s_\lambda \in \mathbf{S}_\lambda$, $\lambda \in \mathbf{\Lambda}$, $\Pr_{s_\lambda}(x_{s_\lambda}, q_{s_\lambda}) = \Pr(x_{s_\lambda}, q_{s_\lambda}; \lambda)$.

Assumption A1 helps eliminate informational spillovers across students. Since, aside from λ, $\Pr(x_{s_\lambda}, q_{s_\lambda}; \lambda)$ is identical across students, hereafter we refer to it as $\Pr(x_{s_\lambda}, q_{s_\lambda})$ for notational simplicity, at times omitting altogether reference to type λ if convenient. Next, we suppose that

A2 Schools have no capacity constraints.

Assumption A2 allows us to highlight the purely informational aspects of the assignment problem that are independent of capacity constraints.[2] Moreover, the students' (common knowledge) preferences over schools are set by assuming that:

A3 $v_s(i, x) = v_{i,s}$, with $v_{i,s} \neq v_{j,s}$ all $i, j \in \mathbf{I}$, $i \neq j$, and $v_{0,s} < v_{i,s}$ for all $i \in \mathbf{I}_+$.

The assumption that student preferences are independent of private signals observed by schools, while natural in the context of the applications we have in mind, is important for our positive results. Furthermore, we also assume that preferences are homogeneous on the student side of the market, i.e.,

A3′ $v_{i,s} = v_i$ for all $s \in \mathbf{S}_+$ and v_i is increasing in i.

Assumption A3$'$ is not crucial for our main results, although it simplifies some proofs. As discussed in Section 3.2.2, for all our results through Section 3, A3$'$ can be replaced by assuming instead that schools' preferences are homogeneous, i.e.,

[2] Our negative (non–existence) results do not depend on A2 and the main positive (existence) result extends to the case with binding capacity constraints, under assumption A3′ on student preferences.

A3″ $w_i(x_s, q_s)$ is identical across i up to affine transformations, for all $s \in \mathbf{S}$.

Finally, for ease of exposition we will carry out the analysis in the paper for the generic environments in which the following condition holds:

A4 Pr and $w_i(x_s, q_s)$ are such that $u_i(s|\mathbf{F}_i) \neq 0$ for all partitions $\mathcal{F}_i$, all $i \in \mathbf{I}_+$, $s \in \mathbf{S}_+$.

If $\mathbf{Q}$, $\mathbf{X}$ are totally ordered and $\Pr(x_{s_\lambda}, q_{s_\lambda})$ is log–supermodular in x_{s_λ}, q_{s_λ} and λ, then the signals are affiliated —for any student s_λ a higher signal obtained by school i makes it more likely that the signals obtained by other schools are also likely to be higher. Log–supermodularity together with the monotonicity of $w_i(x_{s_\lambda}, q_{s_\lambda})$ in its arguments imply that $\mathbb{E}_{s_\lambda}[w_i(x_{s_\lambda}, q_{s_\lambda})|..., a_i \leq x_{i,s_\lambda} \leq b_i, ...]$ is increasing in $a_i, b_i \in \mathbf{X}$, all i, and in λ (see, e.g., Milgrom and Weber (1982, Theorems 23, 24)). For brevity, we call such environments *affiliated*. In what follows, affiliation will be used to construct some examples. In these examples we frequently use the special case of an environment where the set of unobserved qualities is $\mathbf{Q} = \{0, 1\}$, the observed type λ is the prior probability of $q_{s_\lambda} = 1$, and $w_i(x_{s_\lambda}, q_{s_\lambda}) = q_{s_\lambda} - c_i$, where the preference parameter $c_i \in (0, 1)$ is identified as an opportunity cost of admitting a student.

Finally, we introduce the following notation that will be useful later. Given an allocation rule μ, we let $\mathbf{M}(\mu)$, $\mathbf{M}_i(\mu)$ and $\mathbf{M}_s(\mu)$ be the images of $\mu(\cdot)$, $\mu_i(\cdot)$ and $\mu_s(\cdot)$, respectively. With $\mathbf{M}_s(\mu) \subset \mathbb{R}^I$, we let m_s^k denote the k-th canonical basis vector of $\mathbb{R}^I$. We define $\iota_\mu : \mathbf{S} \times \mathbf{X}^{SI} \rightrightarrows \mathbf{I}$ as the schools student s is assigned to attend at an allocation μ when the information is x. That is, $i \in \iota_\mu(s; x)$ if $\mu_{i,s}(x) > 0$; generally ι_μ is a correspondence, unless $\mu_{i,s}(x) = 1$, some i. When $\mu_{i,s}(x) \in \{0, 1\}$ for all i, s and x we say that μ is *degenerate* while if $\mu(x)$ does not depend on x we say that it is *constant*. We turn now to our notions of posterior stable allocations.

3 Posterior stable allocations

As mentioned in the Introduction, we consider two notions of posterior stability, strong and weak. These differ in terms of what is observed by agents at the blocking stage. Strong posterior stability relates to the case each school and student observes the entire allocation outcome $\mu(x)$ at each state of the world x, whereas weak posterior stability relates to the case where school i only observes $\mu_i(x)$ and student s only $\mu_s(x)$.

3.1 Definitions

With incomplete information, expected payoffs of schools and students from an alternative allocation μ', that is feasible for a coalition of schools and students, depend on the ability of schools and students to make inferences from the proposed allocations and to communicate whatever information is left private at that stage. We model a block as a 'unanimous voting game' over an alternative allocation outcome such that members of the blocking coalition can infer which signals others have received from their acceptance of μ'. In other words, each member of a coalition evaluates the proposal μ' at a (Bayesian) Nash equilibrium of this game, knowing the equilibrium acceptance rule of the other members. We first define the notion of feasibility of an alternative allocation for a coalition $\mathbf{C}$ of agents. Then, we define acceptance strategies for members of $\mathbf{C}$, considering strong posterior stability first.

We consider coalitions $\mathbf{C} = (\mathbf{I}',\mathbf{S}')$, with $\mathbf{I}' \subset \mathbf{I}$ and $\mathbf{S}' \subset \mathbf{S}$, such that either $\mathbf{I}'_+ = \mathbf{I}' \cap \mathbf{I}_+ \neq \varnothing$ or $\mathbf{S}'_+ = \mathbf{S}' \cap \mathbf{S}_+ \neq \varnothing$. The allocation μ' is *feasible* for $\mathbf{C}$ relative to μ (or, simply, feasible for $\mathbf{C}$) if for every $s \in \mathbf{S}'$, $\sum_{i \in \mathbf{I}'} \mu'_{i,s}(x) = 1$ and $\mu'_{i,s}(x) = 0$ for each $i \notin \mathbf{I}'$, and for each $i \in \mathbf{I}'_+$, each $s \notin \mathbf{S}'_+$, either $\mu'_{i,s}(x) = \mu_{i,s}(x)$, or $\mu'_{i,s}(x) = 0$, all x. Therefore, μ' is an allocation only for the students and schools in $\mathbf{C}$. For instance, if a school blocks alone (i.e., $\mathbf{S}'_+ = \varnothing$), it can only drop some students it was assigned to under μ.

For $i \in \mathbf{I}'_+$, let $\alpha_i(x'_i, m) \in \{0, 1\}$ be i's acceptance ($\alpha_i = 1$) or rejection ($\alpha_i = 0$) of μ' as a function of his private signal $x'_i \in \mathbf{X}^S$ given that $m \in \mathbf{M}(\mu)$ has been observed.[3] Similarly, for $s \in \mathbf{S}'_+$ let $\alpha_s(m) \in \{0, 1\}$ be the student's acceptance given the observed outcome $m \in \mathbf{M}(\mu)$. Let $\alpha_{\mathbf{C}} = \{\alpha_k\}_{k \in \mathbf{C}}$. Throughout what follows we assume that α_k for the fictitious school or student is equal to 1. Let $piv_k(\alpha_{\mathbf{C}}) = \{x' \in \mathbf{X}^{SI} \mid \alpha_{k'} = 1,\ k' \neq k\}$ be the event that $k \in \mathbf{C}$ is pivotal in the acceptance game for the alternative proposal μ', i.e., the event that all other members of $\mathbf{C}$ have accepted the alternative μ'. The following is our notion of a strong posterior stable allocation.

Definition 1 *An allocation μ is strongly blocked at $x \in \mathbf{X}^{SI}$ if there exist $\mathbf{C} = (\mathbf{I}', \mathbf{S}')$, a constant and degenerate μ' that is feasible for $\mathbf{C}$ and acceptance strategies $\alpha_{\mathbf{C}}$ such that for all $i \in \mathbf{I}'_+$ and $s \in \mathbf{S}'_+$, and for all $x'_i \in \mathbf{X}^S$, $m \in \mathbf{M}(\mu)$,*

$$\alpha_i(x'_i, m) = 1 \text{ iff } U_i(\mu' - \mu | x'_i, m, piv_i(\alpha_{\mathbf{C}})) > 0, \tag{1}$$

$$\alpha_s(m) = 1 \text{ iff } U_s(\mu' - \mu | m, piv_s(\alpha_{\mathbf{C}})) > 0 \tag{2}$$

and $\alpha_i(x_i, \mu(x)) = \alpha_s(\mu(x)) = 1$ for all $i \in \mathbf{I}'$ and $s \in \mathbf{S}'$. An allocation μ is strong posterior stable if there is no x where it is strongly blocked.

Notice from Definition 1 that all members of the coalition must strictly prefer to accept the blocking allocation μ', and that μ' must be degenerate and constant. This is identical to requiring that each member of a blocking coalition strictly prefers to the status quo a feasible and non-random alternative allocation, as opposed to an allocation rule. These are restrictions on the ability of blocking coalitions to "effectively negotiate" (see Myerson, 1991, Ch. 9-10). Especially when $\mu(x)$ does not fully reveal the state of the world, one could imagine that a blocking coalition may be able to design a communication game and thereby implement a general incentive compatible alternative allocation rule μ', neither constant nor degenerate (see, e.g., Definition 3.2 in Forges, 1994). In contrast, Definition 1 allows communication only to the extent of what can be inferred from the

[3] In general $\mu(x)$ is nondegenerate since the allocation rule may prescribe assigning a student s to multiple schools, given x. As an alternative, one could assume that schools observe only which schools got what students, and not the time they spend in each school (part-time vs. full time, say). With our observability assumption however, posterior stable allocations turn out to be degenerate.

equilibrium of the unanimous voting game, i.e., from conditioning on the event $piv_k(\alpha_{\mathbf{C}})$. As we show below however, even with these restrictions on blocks, strong posterior stable allocations that are also required to be incentive compatible do not exist in general —given any incentive compatible μ, there may be too much information contained in $\mu(x)$ at some x for μ to be strong posterior stable. This is the sense in which our analysis is free of renegotiation protocol details.

This also motivates our definition of weak posterior stability, a notion under which less than full information about the status–quo allocation outcome $\mu(x)$ is made available to the agents. Specifically, at each x, school i only observes $\mu_i(x)$ and student s only observes $\mu_s(x)$. Notice that this is the minimum possible information that will be observed by the schools and students once the allocation outcome $\mu(x)$ is selected. The only ensuing difference for the acceptance game is that now a strategy α_i is a function of school i's private information $x_i' \in \mathbf{X}$ and its observed part of the allocation, $\mu_i(x') \in \mathbf{M}_i(\mu)$, whereas the acceptance strategy α_s for student s is a function of $\mu_s(x') \in \mathbf{M}_s(\mu)$.

Definition 2 *An allocation μ is weakly blocked at $x \in \mathbf{X}^{SI}$ if there exist $\mathbf{C} = (\mathbf{I}', \mathbf{S}')$, a constant and degenerate μ' that is feasible for $\mathbf{C}$, and acceptance strategies $\alpha_{\mathbf{C}}$ such that for all $i \in \mathbf{I}'_+$ and $s \in \mathbf{S}'_+$, and for all $x_i' \in \mathbf{X}^S$, $m_i \in \mathbf{M}_i(\mu)$, $m_s \in \mathbf{M}_s(\mu)$,*

$$\alpha_i(x_i', m_i) = 1 \textit{ iff } U_i(\mu' - \mu | x_i', m_i, piv_i(\alpha_{\mathbf{C}})) > 0, \tag{3}$$

$$\alpha_s(m_s) = 1 \textit{ iff } U_s(\mu' - \mu | m_s, piv_s(\alpha_{\mathbf{C}})) > 0, \tag{4}$$

and $\alpha_i(x_i, \mu_i(x)) = \alpha_s(\mu_s(x)) = 1$ for all $i \in \mathbf{I}'$ and $s \in \mathbf{S}'$. An allocation μ is weak posterior stable if there is no x where it is weakly blocked.

Our first result shows that, absent incentive compatibility considerations, both weak and strong posterior stable allocations always exist since, in particular, ex–post stable allocations are posterior stable (see the proof for the definition of ex–post stability).

Proposition 1 *Posterior stable allocations exist and are degenerate.*

Proof. See the Appendix. □

In Example 1 below, we characterize the full set of weak and strong posterior stable rules, in order to gain intuition. We also use the example to compare our notion of posterior stability to the seminal notion of interim stability provided by Wilson (1978), i.e., the coarse core, and its incentive compatible version (see Vohra (1999)). In the coarse core, blocks occur at the interim stage, allow signal-contingent alternatives, and students evaluate allocations in expected terms prior to knowing the allocation outcome. No inference can be made from acceptance of a block, since each member of a coalition must be strictly better off from accepting in all states that are consistent with what is commonly known by the coalition. The example shows that the sets of posterior stable and (incentive compatible) coarse core allocations may not even intersect.

Example 1 *The sets of posterior stable and (incentive compatible) coarse core allocations do not intersect.*

We take an affiliated environment with $\mathbf{I}_+ = \{1, 2\}$, one student s and binary signals for each school, i.e., $\mathbf{X} = \{L, H\}$ with $L < H$. The preferences are summarized by the following:
(i) $u_i(s|x) > 0$ iff $x \neq (L, L)$, $i = 1, 2$; (ii) $u_1(s|x_1 = L) > 0$;
(iii) $u_2(s|x_2 = L) < 0$, (iv) $v_1 > v_2 \Pr[x \neq (L, L)]$.
Let μ^1, μ^2, μ^3 be three different degenerate rules defined respectively by[4]

$$\iota_{\mu^1}(s; x) = \left\{ \begin{array}{cc} 2 & 2 \\ 0 & 2 \end{array} \right\}, \iota_{\mu^2}(s; x) = \left\{ \begin{array}{cc} 2 & 2 \\ 0 & 1 \end{array} \right\}, \iota_{\mu^3}(s; x) = \left\{ \begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array} \right\}$$

Observe first that the set of weak posterior stable allocations is $\{\mu^1, \mu^2\}$. For at any candidate weak stable allocation μ, $\mu_{2,s}(x) = 1$ if $x_2 = H$, or school 2 will block in a coalition with the student, by condition (i). Next, $\mu_{1,s}(L, L) = 0$, or school 1 will block alone, again by (i); and $\mu_{2,s}(L, L) = 0$, or school 2 will block alone, by either (i) or (iii). Finally, $\mu_{0,s}(H, L) = 0$, or school 1 blocks with the student, by (i). Then, μ^1, μ^2 are the only remaining possibilities. Indeed, μ^1 is strong posterior stable but μ^2 is not, since under strong stability school 2 can identify the state of the world $x = (H, L)$ upon observing the allocation corresponding to the the bottom right element of the matrix ι_{μ^2}, and consequently will block with the student, by (i).

However, neither μ^1 nor μ^2 is in the coarse core. Using μ^3 as a feasible alternative, school 1 and the student have a coarse objection to them: for school 1 by (i), (ii); for the student, the utility from μ^1 is $v_2 \Pr[x \neq (L, L)]$, the utility of μ^2 is $v_2 \Pr[x_2 = H] + v_1 \Pr[(H, L)]$, and v_1 is the utility under μ^3, which is higher by (iv). Indeed, μ^1, μ^2 are not even in the incentive compatible coarse core, since μ^3 is an incentive compatible coarse objection. For completeness, observe that μ^3 is in the (incentive compatible) coarse core, as can be readily verified.■

A consequence of Example 1 is that posterior stable allocations may not be in the fine core either (see Wilson (1978) for a definition), since this last is contained in the coarse core. A similar no-inclusion relation holds for the credible core (Dutta and Vohra, 2001), which is contained in the incentive compatible coarse core. Example 1 then also shows that the requirement of stability at *both* the interim and the posterior stages may not be feasible.

3.2 Incentive compatibility

3.2.1 Strong posterior stability: impossibility

Proposition 1 establishes the existence of posterior stable allocations, strong and weak, absent incentive compatibility requirements. The next example shows that there exist robust environments where strong posterior stable allocation rules are bound to fail incentive compatibility.

[4] In all examples, we represent ι_μ as a matrix, with the columns denoting the signal of school 1, increasing as one moves right, and the rows denoting the signal of school 2, increasing as one moves up. In later examples with three schools, the signals of school 3 are denoted by different matrices.

Example 2 *Non–existence of strong posterior stable and incentive compatible allocation rules.*

We take a symmetric affiliated environment with $\mathbf{I}_+ = \{1, 2\}$, one student s and three signals for each school, i.e., $\mathbf{X} = \{L, M, H\}$ with $L < M < H$. The preference parameters are as follows:
(i) $u_1(s|x) > 0$ iff $\min_i x_i \geq M$ or $\max_i x_i \geq H$; (ii) $u_1(s|x_1 = M, x_2 < H) > 0$;
(iii) $u_2(s|x) > 0$ iff $\max_i x_i \geq H$; (iv) $u_2(s|x_1 > L, x_2 = L) < 0$; (v) $u_2(s|x_1 > L, x_2 = M) > 0$;
(vi) $u_2(s|x_1 \neq M, x_2 = M) < 0$.

Observe first that any strongly stable allocation is degenerate, by Proposition 1. Moreover, $\mu_{2,s}(x_1, H) = 1$, for all x_1 or $i = 2$ and the student would (strongly) block at (x_1, H), by (iii). Second, $\mu_{i,s}(L, x_2) = 0$ for $i \in \mathbf{I}_+$ for $x_2 = L, M$, or i would block (using (i) for $i = 1$ and (iii), (iv), (vi) for $i = 2$). Finally, $\mu_{2,s}(M, L) = 0$ or else school 2 will block at (M, L), by (iii) and (iv).

We now consider two cases: (a) $\mu_{2,s}(M, M) \neq \mu_{2,s}(H, M)$ and (b) $\mu_{2,s}(M, M) = \mu_{2,s}(H, M)$.

Case (a): $\mu_{2,s}(M, M) \neq \mu_{2,s}(H, M)$

Since μ is degenerate, in this case we either have $\mu_{2,s}(M, M) = 1$ or $\mu_{2,s}(H, M) = 1$, but not both. Since in the former case 2 would block at (M, M) by (iii), we see $\mu_{2,s}(H, M) = 1$ and so $\mu_{2,s}(M, M) = 0$. Next, $\mu_{1,s}(M, M) = 1$. For if not then school 1 and the student will block at (M, M), by (i) and (ii). Further, if $\mu_{2,s}(H, L) = 0$, then $\mu_{1,s}(H, L) = 1$ or else 1 and the student will block at (H, L), by (i). This leaves three possible allocation rules with corresponding ι_μ's given by

$$\left\{\begin{matrix} 2 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{matrix}\right\}, \quad \left\{\begin{matrix} 2 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{matrix}\right\}, \quad \left\{\begin{matrix} 2 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 1 & 1 \end{matrix}\right\}$$

None of these are incentive compatible as in each, school 1 has an incentive to announce M when $x_1 = H$.

Case (b): $\mu_{2,s}(M, M) = \mu_{2,s}(H, M)$

In this case, $\mu_{2,s}(M, M) = \mu_{2,s}(H, M) = 1$. For if $\mu_{2,s}(M, M) = \mu_{2,s}(H, M) = 0$, then $\mu_{1,s}(x_1, M) = 1$ for $x_1 = M, H$ as otherwise 1 and the student would block at (x_1, M) by (i) and (ii), implying in turn, since 2 observes the entire outcome $\mu(x)$, that 2 and the student would block at (H, M), by (v). It follows that $\mu_{1,s}(M, L) = 0$ as otherwise 1 will block at (M, L) (by (i)). Further, $\mu_{2,s}(H, L) = 1$. For if $\mu_{2,s}(H, L) = 0$, then we must have $\mu_{1,s}(H, L) = 1$ or else 1 and the student will block at (H, L) by (i), implying in turn, since 2 observes the entire outcome $\mu(x)$, that 2 and the student would block at (H, L), by (iii). Thus, the only possible μ in this case has ι_μ given by

$$\left\{\begin{matrix} 2 & 2 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{matrix}\right\}$$

This is not incentive compatible as 2 has an incentive to announce L when $x_2 = M$, by (iii). This concludes the example.■

When private information must be elicited in an incentive compatible manner, strong posterior stability may be impossible to obtain. Notice in this respect that in the arguments of Example 2, we only used blocking coalitions of either singleton schools or school–student pairs. In all such blocks, the privately informed school in a blocking coalition did not infer anything from the acceptance decisions of other coalition members (if any) that it did not already know from observing the allocation outcome $\mu(x)$. Thus, the nonexistence result of Example 2 does not depend on the restrictions on communication possibilities between members of a blocking coalition. Since there is only one student, it does not depend on the homogeneity of the students' preferences, or on the absence of capacity constraints.

3.2.2 Weak posterior stability: modified serial dictatorship

Our next result is positive, and establishes the existence of weakly stable incentive compatible allocations. The proof uses a *modified serial dictatorship* algorithm to deliver allocations with such properties. The algorithm works as follows. For each student, the student's value to each school i is gauged using information relative to that student only. In particular, this value is measured using school i's own signal and the signals possessed by all schools above i in the student's rankings. Then, the mechanism assigns the student to school i if this value is positive and the student is available, i.e., given that the student was not assigned to any of the higher schools in the student's rankings. If school i is not assigned the student, then the next school is considered, and so on, starting from the top school and moving down the student's rankings. Given A3$'$, each student s is then allocated to school $i \in \mathbf{I}$ if and only if $u_i(s|x_{i,s}, x_{i+1,s}, ..., x_{I,s}) \geq 0 > u_j(s|x_{j,s}, x_{j+1,s}, ..., x_{I,s})$ for all $j > i$. We denote the resulting allocation as μ^{MSD}.

Theorem 1 *The allocation rule μ^{MSD} obtained via modified serial dictatorship is incentive compatible and weak posterior stable in every environment.*

Proof. See the Appendix. □

By the properties of modified serial dictatorship, each school in a coalition can effectively condition its decision on the private information held by other schools in the coalition that are higher in the student's ranking. In fact, it follows from the way the allocation μ^{MSD} is constructed that μ^{MSD} cannot be weak blocked if, in addition to $\mu_i(x)$, each school is also allowed to observe the actual signals of all schools above it in the student rankings at the blocking stage. Under A3$'$, this implies that the lowest school in any coalition cannot learn anything from other schools in the coalition, and so eliminates the need for considering blocks involving multiple schools and students —any time such a block exists, another one involving only one school (specifically, the lowest) and one student also does.[5] However, a block involving a student and a single school

[5]Multi–school coalitions can also be ruled out if one replaces A3$'$ by A3$''$, although for a different "no trade" argument. Since multi-school coalitions can all gain only if they trade students amongst themselves, the preference homogeneity among schools imposed by A3$''$ eliminates such blocks due to arguments similar to the "no trade theorem" of Milgrom and Stokey (1982). Details are available upon request.

cannot obtain, since the acceptance of the block by a student who was not initially allocated to that school conveys to the school only that the student was admitted to a school below it in the student's ranking. By the properties of modified serial dictatorship, the school cannot be better off from accepting such a student conditional only on this information. Since it is immediate that μ^{MSD} is interim individually rational for each school and student, this establishes weak stability. Furthermore, μ^{MSD} is incentive compatible for each school i, not only in interim terms, but also conditional on the actual signals received by all schools $j > i$, given truth-telling by schools $j' < i$. Thus, it is ex-post incentive compatible for school 1 and, given truth-telling by school 1, incentive compatible for school 2 even given the actual signals of all schools above 2, and so on. We will utilize these properties of μ^{MSD} further in Section 4.1.[6]

Modified serial dictatorship delivers posterior stability under weak observability because, by construction, no school is able to infer anything about the information held by schools below it in the student's ranking from its own allocation. Instead, under strong observability, by observing the entire allocation a school may infer valuable information about the signal obtained by schools below it in the student's ranking, and so profitably block μ^{MSD}. For instance, in the environment of Example 2 the allocation rule μ^{MSD} is given by:

$$\iota_{\mu^{MSD}}(s;x) = \left\{ \begin{array}{ccc} 2 & 2 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{array} \right\}$$

This rule is not strong posterior stable. When $x_2 = M$, upon observing that the student has been allocated to school 1 (i.e., $\mu_{1,s}(x) = 1$), school 2 will infer that the signal of school 1, $x_1 \in \{M, H\}$, and so it will be willing to admit the student via a blocking coalition, even though the student will always accept such an offer.[7] However, μ^{MSD} is posterior stable under weak observability since when $x_2 = M$ school 2 will only observe $\mu_{2,s}(x_1, M) = 0$, for each $x_1 \in \{L, M, H\}$. It will therefore be unable to infer anything about x_1 from its own allocation. Neither will it infer anything by conditioning on the fact that the student accepts a block since the student will always accept a block that gains him admission to school 2, regardless of which school he has been admitted to under the status quo allocation. Such a block therefore cannot obtain.

We conclude this section by comparing modified serial dictatorship with a similar algorithm, *simple* serial dictatorship. As with modified serial dictatorship, under simple serial dictatorship the student is allocated to a school in the descending order given by the student's preferences. However, a student's value to school i is not gauged with the actual signals obtained by higher schools,

[6] It can be shown that an amended version of MSD also applies when schools have known capacity constraints, provided students have identical preferences, i.e, A3$'$ applies. In the amended version each school (starting from the top) is allocated as many students as it is willing to have subject to its capacity, the availability of the student, and given its own signals and signals of schools above it.

[7] As shown in Example 2 this ability to deduce the information held by lower schools in the context of strong observability leads to the nonexistence of any incentive compatible strongly stable allocation rule. Exactly the same intuition underlies the non–existence of incentive compatible ex–post stable allocation rules.

but using only school i's own signal, and the fact that the student is still available. In affiliated environments (i.e., where Pr displays affiliation and w is a collection of increasing functions), simple serial dictatorship also delivers allocations that are weak posterior stable and incentive compatible. This is because affiliation implies that conditional expectations are monotonically increasing, linking a student's value across schools. This is enough to exclude blocks among multiple schools (Steps 1 and 2 in the proof of Theorem 1) through a 'no trade' argument: acceptance of a student's move to a higher school in a block reveals to the lower school that the student is worth keeping.[8] Outside of affiliated environments, simple serial dictatorship may not deliver weak posterior stability, as Example 3 demonstrates. Instead, with modified serial dictatorship information on the higher school's signal is embedded in the construction of the allocation rule, and as a result the mechanism operates in more general environments, i.e., with arbitrary Pr and w.

Example 3 *Simple serial dictatorship may not be weak posterior stable.*

We take an environment with $\mathbf{I}_+ = \{1, 2\}$, one student s and binary signals for each school. The preference parameters are as follows: (i) $u_1(s|x) > 0$ iff $x = (H, H)$; (ii) $u_1(s|x_1 = H) > 0$; (iii) $u_2(s|x) > 0$ iff $x = (H, L)$; (iv) $u_2(s|x_2 = L) < 0$.

Consider the allocations μ^1 and μ^2 below:

$$\iota_{\mu^1} = \left\{ \begin{array}{cc} 0 & 1 \\ 0 & 1 \end{array} \right\}, \quad \iota_{\mu^2} = \left\{ \begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array} \right\}$$

It is straightforward to verify that μ^1 is obtained through simple serial dictatorship and μ^2 through modified serial dictatorship, and that μ^1 is not weak posterior stable: at (H, L) the grand coalition of schools 1 and 2 and the student will block with $\mu'_{2,s} = 1$ using the acceptance rules $\alpha_s = 1$ for all m_s, $\alpha_1 = 1$ iff $x_1 = H$ and $\alpha_2 = 1$ iff $x_2 = L$, since at (H, L) all parties prefer that the student goes to school 2 instead of 1. Instead, by Theorem 1, μ^2 is weak posterior stable and incentive compatible.■

3.2.3 The necessity of restrictions on blocks

Apart from the partial observability that defines weak stability, there are a number of other restrictions on the ability of blocking coalitions to effectively negotiate that we commented upon before. These are (i) all members of such a coalition must be *strictly* better off from accepting the block, (ii) the blocking rule must be *degenerate,* and (iii) it must be a *constant.* In particular, we do not allow a blocking coalition to construct a general direct mechanism to implement a possibly non–degenerate and non–constant alternative allocation. This can be justified on the grounds of a general mechanism being too costly to run at the renegotiation stage when the original mechanism has already picked an allocation. In view of the positive result of Theorem 1 it is still natural to ask however if these restrictions are also necessary for existence of posterior stable incentive compatible

[8] Details available upon request.

rules. We answer this question now by considering alternative notions of weak stability, relaxing each of the restrictions (i)–(iii) individually.

We begin by relaxing (i) above, i.e., we allow participation in blocks even if an agent is only weakly better off from doing so. One can then imagine a situation where two schools form a coalition, communicate the state and agree not to admit the student to either school. Such a possibility is ruled out by condition (3) in Definition 2 since at least one school cannot strictly gain from such a block vis-a-vis a degenerate status quo μ. But consider a weakening conditions (3) and (4) in Definition 2 to, respectively,

$$\begin{aligned} \alpha_i(x'_i, m_i) = 1 \text{ only if } U_i(\mu' - \mu | x'_i, m_i, piv_i(\alpha_{\mathbf{C}})) \geq 0, \\ \alpha_s(m_s) = 1 \text{ only if } U_s(\mu' - \mu | m_s, piv_s(\alpha_{\mathbf{C}})) \geq 0, \end{aligned} \tag{5}$$

with one strict inequality. We call such a notion of stability *weak* posterior stability.* Our next example concerns incentive compatible weak* stable rules.

Example 4 *Non–existence of weak* posterior stable and incentive compatible allocation rules.*

Consider the same environment as in Example 2, with $\mathbf{I}_+ = \{1, 2\}$, one student s and three signals for each school, $\mathbf{X} = \{L, M, H\}$ with $L < M < H$. The preference parameters are as follows:

(i) $u_1(s|x) > 0$ iff $\min_i x_i \geq M$ or $\max_i x_i \geq H$; (ii) $u_1(s|x_1 = M, x_2 < H) < 0$;
(iii) $u_2(s|x) > 0$ iff $\max_i x_i \geq H$; (iv) $u_2(s|x_1 > L, x_2 = L) < 0$; (v) $u_2(s|x_2 = M) > 0$.

We show in the Appendix that no weak* posterior stable and incentive compatible allocations exist in this environment. To appreciate the force of weak* blocks, we show here instead how the modified serial dictatorship outcome can be weak* blocked. In this example, the allocation μ^{MSD} can be written as

$$\iota_{\mu^{MSD}}(s; x) = \left\{ \begin{array}{ccc} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 0 & 0 & 1 \end{array} \right\}.$$

By Theorem 1, μ^{MSD} is incentive compatible and weak posterior stable. However it can be weak* blocked at $x = (L, M)$ by the coalition $\mathbf{I}' = \{0, 1, 2\}$ and $\mathbf{S}' = \{0\}$ with the proposal $\mu'_{i,s} = 0$ for all $i \in \mathbf{I}_+$ and the acceptance rules $\alpha_1(x'_1, m_1) = 1$ iff $x'{}_1 = L$, all $m_1 \in \mathbf{M}_1(\mu)$, $\alpha_2(x'_2, m_2) = 1$ iff $x'_2 = M$, all $m_2 \in \mathbf{M}_2(\mu)$. Intuitively, school 1 participates in this block and sells information to school 2 at price zero, making school 2 strictly better off and school 1 indifferent.■

Next, we relax (ii) above, i.e., consider the non-degeneracy of the alternative proposal μ'. We show that if in defining weak stability we allow for proposals μ' to be non–degenerate, again we get nonexistence. To see this, consider Definition 2 of weak stability, but now allow μ' to be a possibly random feasible allocation for a coalition $\mathbf{C}$ (i.e., a lottery), although still constant across $\mathbf{X}$. Call the resulting stable allocations *weak** posterior stable*.

Example 5 *Non-existence of weak** posterior stable and incentive compatible allocation rules.*

Suppose that there is one student s, three schools $\mathbf{I}_+ = \{1,2,3\}$, each with binary signals, $\mathbf{X} = \{L, H\}$, and that the environment is affiliated. Suppose further that the following conditions on preferences and distributions hold:

(i) $u_1(s|x) > 0$ for all $x \in \mathbf{X}^I$; (ii) $\frac{u_1(s|H,H,L)}{u_1(s|H,L,L)} > \frac{\Pr(x_2=L|x_1=H,,x_3=L)}{\Pr(x_2=H|x_1=H,x_3=L)}$;

(iii) $u_2(s|x) < 0$ iff $x \in \{(L,L,L), (H,L,L), (L,H,L), (L,L,H)\}$; (iv) $u_2(s|\ x_2 = H, x_i = L) > 0$ for $i = 1,3$; (v) $u_2(s|x_2 = L) > 0$;

(vi) $u_3(s|x) < 0$ iff $x \in \{(L,L,L), (H,L,L), (L,H,L), (L,L,H)\}$; (vii) $u_3(s|\ (x_1,x_2) \neq (L,L), x_3 = L) < 0$; (viii) $u_3(s|\ x_1 = H, x_3 = L) < 0$; (ix) $u_3(s|x_2 = H, x_3 = L) > 0$, but (x) $u_3(s|\ (x_1,x_2) \neq (H,L), x_3 = H) < 0$, $u_3(s|\ (x_1,x_2) \neq (L,H), x_3 = H) < 0$; (xi) $u_3(s|\ x_3 = H) > 0$.

Once again, in the Appendix we show how the possibility of sharing the student via lotteries precludes the existence of any weak** stable and incentive compatible rule in this environment. To illustrate the power of weak** blocks, we show here how the allocation rule obtained via modified serial dictatorship can be weak** blocked. In this example, the modified serial dictatorship allocation μ^{MSD} can be written as

$$\iota_{\mu^{MSD}}(s; x_{-3}, L) = \left\{ \begin{array}{cc} 2 & 2 \\ 1 & 1 \end{array} \right\} \quad \iota_{\mu^{MSD}}(s; x_{-3}, H) = \left\{ \begin{array}{cc} 3 & 3 \\ 3 & 3 \end{array} \right\}$$

where we use matrices for school 3's signal x_3.

By Theorem 1, μ^{MSD} is weak posterior stable and incentive compatible. However, it can be weak** blocked at $x = (H,H,L)$ by the coalition $\mathbf{C}$ with $\mathbf{I}'_+ = \{1,3\}$ and $\mathbf{S}'_+ = \{s\}$ with the proposal $\mu'_{1,s} = p = 1 - \mu'_{3,s}$ where $p \in (0,1)$ is such that $pv_1 + (1-p)v_3 > v_2$. In this block, each school accepts if and only if it has not already been allocated the student, i.e., $\alpha_i(x_i, m_i) = 1$ iff $m_i \neq 1$, and the student accepts if and only if he has not already been assigned to school 3. When school 3 does not receive the student under μ^{MSD}, the acceptance strategy of school 1 effectively allows it to communicate to school 3 that the student has been assigned to school 2, revealing school 2's signal and making the block profitable for school 3. School 1 also gains since it obtains the student a proportion p of the time, while the student gains since he prefers the lottery μ' over schools 1 and 3 to being in school 2 for sure.■

Examples 4 and 5 together show the necessity of each of the two restrictions (i) and (ii) to obtain existence, namely that coalition members must be strictly better off and that the alternative μ' must be degenerate as a function of x. We now turn to restriction (iii), namely that μ' is constant.

If μ' is allowed to depend nontrivially on signals and privately known allocation outcomes, we need to explicitly take into account incentive compatibility and self-selection constraints that apply to the communication game played by the blocking coalition. However, since Example 5 tells us that μ' must be degenerate in order to obtain existence, the revelation principle does not apply at the blocking stage. This is because in a Bayesian Nash equilibrium of a general mechanism that applies at the blocking stage agents may play mixed strategies, a possibility that cannot be captured if one restricts μ' to be an incentive compatible direct mechanism that is degenerate as a function of reported types. Accordingly, we turn now to a notion of a block in terms of a general communication game.

For a coalition $\mathbf{C}$, let $\mathbf{T}_k = \mathbf{X}^S \times \mathbf{M}_k(\mu)$ be the type space of agent $k \in \mathbf{C}$, with $t_k = (x_k, m_k)$, and $x_k \equiv 0$ if $k \in \mathbf{S}'$. Thus, t_k consists of the privately observed signal x_k of agent k (if any) as well as the privately observed allocation outcome m_k. As in Definition 2, we suppose that each agent $k \in \mathbf{C}$ can either accept ($\alpha_k(t_k) = 1$) or reject ($\alpha_k(t_k) = 0$) the block, with the block succeeding only if all coalition members accept. In contrast to Definition 2 however, we now allow each agent $k \in \mathbf{C}$ to also send a message $\xi_k \in \mathbf{\Xi}_k \supset \mathbf{T}_k$ when he accepts. Let $\Xi = \times_{k \in \mathbf{C}} \Xi_k$. An admissible alternative allocation rule $\mu' : \mathbf{\Xi} \to \mathbb{R}^{SI}$ for the coalition $\mathbf{C}$ chooses a degenerate allocation $\mu'(\xi)$ that is feasible for $\mathbf{C}$, as a function of messages $\xi \in \mathbf{\Xi}$, given that all agents have accepted.

For each $k \in \mathbf{C}$, let $\mathbf{A}_k = \{t_k \in \mathbf{T}_k | \alpha_k(t_k) = 1\}$ be the set of types where k accepts the block and let $\mathbf{A} = \times_{k \in \mathbf{C}} \mathbf{A}_k$ and $\mathbf{A}_{-k} = \times_{k' \in \mathbf{C}, k' \neq k} \mathbf{A}_{k'}$. Further, given acceptance, let $\sigma_k(\xi_k | t_k)$ be the probability with which k announces message ξ_k given his type $t_k \in \mathbf{A}_k$ and let $\sigma_{-k}(\xi_{-k} | t_{-k})$ be the probability of message ξ_{-k} for agents other than k when their type is t_{-k}. Finally, for $t_k \in \mathbf{A}_k$, let $\Xi_k(\sigma_k, t_k) = \{\xi_k | \ \sigma_k(\xi_k | t_k) > 0\}$ be the support of σ_k given $t_k \in \mathbf{A}_k$.

Given the alternative rule μ', for any school i of type t_i in the coalition $\mathbf{C}$, we write the expected payoff from accepting the block and announcing a message ξ_i given that other coalition members have also accepted the block and sent messages according to σ_{-i}, as follows:

$$U_i(\mu', \xi_i, \sigma_{-i} | t_i, \mathbf{A}_{-i}) = \sum_{s \in \mathbf{S}'} \sum_{t_\mathbf{C} \in \mathbf{T}_\mathbf{C}} \sum_{\xi_{-i}} u_i(s | t_\mathbf{C}) \mu'_{i,s}(\xi_i, \xi_{-i}) \sigma_{-i}(\xi_{-i} | t_{-i}) \Pr[t_\mathbf{C} | t_i, \mathbf{A}_{-i}]$$

Similarly, for a student s in the coalition $\mathbf{C}$ we write the corresponding expression as

$$U_s(\mu', \xi_s, \sigma_{-s} | t_s, \mathbf{A}_{-s}) = \sum_{i \in \mathbf{I}} \sum_{t_\mathbf{C} \in \mathbf{T}_\mathbf{C}} \sum_{\xi_{-s}} v_{i,s} \mu'_{i,s}(\widehat{\xi}_s, \xi_{-s}) \sigma_{-s}(\xi_{-s} | t_{-s}) \Pr[t_\mathbf{C} | t_s, \mathbf{A}_{-s}].$$

We say that the allocation rule μ is *weak*** posterior blocked* at x if there exists a coalition $\mathbf{C}$, an admissible μ', acceptance sets $\{\mathbf{A}_k\}_{k \in \mathbf{C}}$ and announcement strategies $\{\sigma_k\}_{k \in \mathbf{C}}$ such that, for all $k \in \mathbf{C}$,

$$\max_{\xi_k \in \Xi_k} U_k(\mu', \xi_k, \sigma_{-k} | t'_k, \mathbf{A}_{-k}) > U_k(\mu | t'_k, \mathbf{A}_{-k}) \text{ iff } t'_k \in \mathbf{A}_k, \tag{6}$$

and

$$U_k(\mu', \xi_k, \sigma_{-k} | t_k, \mathbf{A}_{-k}) \geq U_k(\mu', \widehat{\xi}_k, \sigma_{-k} | t_k, \mathbf{A}_{-k}) \text{ for all } t_k \in \mathbf{A}_k, \xi_k \in \Xi_k(\sigma_k, t_k), \widehat{\xi}_k \in \Xi_k \tag{7}$$

and, furthermore $(x_k, \mu_k(x)) \in \mathbf{A}_k$. The rule μ is *weak*** posterior stable* if it cannot be weak*** posterior blocked at any x.

In words, condition (6) represents a self-selection constraint. It states that a coalition member k strictly prefers the alternative μ' if and only if his type t_k is in his acceptance set $\mathbf{A}_k$, given that $t_{-k} \in \mathbf{A}_{-k}$. Condition (7) states that it is a best response for k to announce according to σ_k when

$t_k \in \mathbf{A}_k$, given that other coalition members have also accepted and sent messages according to σ_{-k}. Finally, the condition $(x_k, \mu_k(x)) \in \mathbf{A}_k$ states that the acceptance sets are non-empty.

Observe that this definition is similar to known core concepts at the interim stage (for instance, see Dutta and Vohra (2001)). It boils down to Definition 2 when $\mathbf{\Xi}_k = \mathbf{T}_k$ and μ' is constant —the acceptance region $\mathbf{A}_{-k}$, when projected onto $\mathbf{X}^{SI}$ using μ and truth-telling, corresponds to the set $piv_k(\alpha_{\mathbf{C}})$ for $k \in \mathbf{C}$; condition (7) is trivially satisfied if μ' is constant, and condition (6) is equivalent to (3) and (4). We show in Example 6 that there exist environments where weak*** posterior stable and incentive compatible allocations fail to exist.

Example 6 *Non-existence of weak*** posterior stable and incentive compatible allocation rules.*

Suppose that there are two students, s_1 and s_2 and three schools $\mathbf{I}_+ = \{1, 2, 3\}$. Students are ex ante identical and payoffs and probabilities for each student are given by conditions (i) – (xi) of Example 5. In the Appendix we demonstrate the non-existence of any weak*** stable and incentive compatible rule in this environment. The intuition is as follows. Since the alternative μ' can depend non-trivially on messages ξ, in a coalition with both students one can effectively make the allocation of student s_i to be, from his perspective, a lottery whose outcomes vary with the (random) messages sent by the other student s_j, $i \neq j$. The logic of Example 5 then applies to rule out the existence of any weak*** stable and incentive compatible rules.■

This concludes the demonstration of the necessity of restrictions (i)-(iii) in establishing the existence of stable and incentive compatible rules under weak observability.

4 Further properties of modified serial dictatorship

4.1 Renegotiation

With modified serial dictatorship we have found weak posterior stable allocations that are incentive compatible. As a result, on the equilibrium path of the direct mechanism students and schools have no interest to block the allocation suggested by the mechanism. Hence, no renegotiation would occur on the equilibrium path. In this section we study whether such allocations would still be implementable if schools and students could always renegotiate, and block also 'off the equilibrium path', i.e., after lying to the mechanism, in the spirit of Maskin and Moore (1999).

We first have to modify condition (IC). In terms of expected payoff comparison, we need to make sure that 'truth telling and no blocking' dominates not only 'truth telling and blocking' (as guaranteed by posterior stability) and 'lying and no blocking' (as guaranteed so far by (IC)), but also 'lying and blocking'. Informally, this is simply requiring that truth telling and no blocking must be the outcome on the path of play of a (perfect Bayesian) equilibrium of a two-stage game of information revelation and blocking.

Formally, let $m_i(\widehat{x}_i, x_{-i}) \in \mathbf{M}_i(\mu)$ be the allocation outcome from μ for school i when submitting announcement $\widehat{x}_i \in \mathbf{X}^S$ and other schools have truthfully announced x_{-i}. Let $\widehat{\beta}_i \equiv$

$\widehat{\beta}_i(x_i, m_i(\widehat{x}_i, x_{-i}))$ be a blocking strategy conditional on the information contained in $x_i, m_i(\widehat{x}_i, x_{-i})$, including the possibility of not blocking. That is, $\widehat{\beta}_i$ captures the possibility that school i may profitably form a blocking coalition in the sense of Definition 2, even when μ is weak posterior stable, by exploiting the fact that it has lied at the announcement stage. Notice in this respect that we assume that renegotiation takes place with no other information for school i —the actual messages sent by all other schools are still not known when blocking. Moreover, a block is described as before, where communication occurs through a unanimous acceptance game. Finally, members of a coalition believe that all other agents within and outside the coalition have made truthful reports at the announcement stage.

Let $\mu_{i,s}(\widehat{\beta}_i(x_i, m_i(\widehat{x}_i, x_{-i})))$ be the induced allocation outcome for school i, assigning student s according to $\widehat{\beta}_i(x_i, m_i(\widehat{x}_i, x_{-i}))$ conditional on $x_i, m_i(\widehat{x}_i, x_{-i})$. In particular, it equals $\mu(\widehat{x}_i, x_{-i})$ when i does not block conditional on $x_i, m_i(\widehat{x}_i, x_{-i})$. We say that an allocation μ is *immune from renegotiation* if, for each $i \in \mathbf{I}_+$ and each x_i, (x_i, β_i) solves

$$\max_{\widehat{x}_i, \widehat{\beta}_i} \sum_{s \in \mathbf{S},} \sum_{x \in \mathbf{X}^{SI}} u_i(s|x_s) \mu_{i,s}(\widehat{\beta}_i(x_i, m_i(\widehat{x}_i, x_{-i}))) \Pr[x|x_i] \tag{IC-R}$$

where β_i is such that $\mu_{i,s}(\beta_i(x_i, m_i(\widehat{x}_i, x_{-i}))) = \mu_{i,s}(\widehat{x}_i, x_{-i})$ for all $x_i, \widehat{x}_i$. Notice how, if μ is weak posterior stable, $\mu_{i,s}(\beta_i(x_i, m_i(x_i, x_{-i})) = \mu_{i,s}(x_i, x_{-i})$, since we already know that when everyone tells the truth μ cannot be blocked successfully. If μ is immune from renegotiation then no school can profit even after lying to the mechanism and then blocking. If μ is also incentive compatible in the sense of (IC), it then follows that no school can do better than reporting truthfully to the mechanism and then not blocking. We have the following result.

Proposition 2 *The allocations obtained via modified serial dictatorship are immune from renegotiation.*

Proof. See the Appendix. □

The intuition for this result is similar to that of Theorem 1. By the properties of μ^{MSD}, the allocation of a student to any given school depends only on that school's signals, as well as on the signals of all schools above (but not below) it in the student's ranking. This together with the observability restrictions in the definition of weak stability implies that no school can profit from lying at the announcement stage, in order to obtain more information from the observed outcome, and subsequently forming a blocking coalition.

4.2 Efficiency

We ask whether modified serial dictatorship gives rise to allocations satisfying efficiency properties other than weak posterior stability against the grand coalition. As with posterior stability, where it matters whether acceptance decisions are allowed to be weak or strict, one needs care in defining dominance when considering posterior efficiency. Along the lines of what Example 4 has shown, if dominance can be weak for those agents who share information used in constructing an alternative allocation, agents can be actively part of a coalition by selling information at zero price to other coalition members, and allocations obtained through modified serial dictatorship are not guaranteed to be efficient in this sense. If dominance must be strict for every agent, and the alternative is still restricted to be constant and degenerate, the grand coalition cannot generally be effective in blocking any degenerate allocation— for instance, if there is a single student. This motivates the following definitions of dominance and efficiency. First, we require that an agent who shares information within the grand coalition has to strictly profit from the alternative allocation. In addition, no agent can be hurt by the counterproposal, which is still assumed to be constant and degenerate.

Formally, we say that μ is *strong posterior w-dominated* at $x \in \mathbf{X}^{SI}$ if it can be strongly blocked by some coalition $\mathbf{C} = (\mathbf{I}', \mathbf{S}')$ according to Definition 1, without reducing the expected payoff conditional on $x_i, \mu(x)$ for any $i \in \mathbf{I}$, $s \in \mathbf{S}$; μ is *strong posterior w-efficient* if it cannot be strong posterior w-dominated at any $x \in \mathbf{X}^{SI}$.

We will also measure modified serial dictatorship against ex-post efficiency: μ is *ex-post dominated* at $x \in \mathbf{X}^{SI}$ if there exists a degenerate μ' that is feasible for $\mathbf{C} = (\mathbf{I}, \mathbf{S})$ such that for all $i \in \mathbf{I}_+$ and $s \in \mathbf{S}_+$, $U_i(\mu' - \mu|x) \geq 0$, $U_s(\mu' - \mu|x) \geq 0$, with one strict inequality; and μ is *ex-post efficient* if it cannot be ex-post dominated at any $x \in \mathbf{X}^{SI}$. Note that for ex-post efficiency we allow improvements not to be strict. Also, because of the restrictions on alternative proposals (specifically, constancy), strong posterior w-efficiency does not automatically imply ex-post efficiency. The allocations obtained with modified serial dictatorship have the following efficiency properties.

Proposition 3 *Let μ be obtained through modified serial dictatorship: (i) if the environment is affiliated, then μ is strong posterior w-efficient; (ii) if $w_i(x_s, q_s)$ is decreasing in i, then μ is ex-post efficient.*

Proof. See the Appendix. □

Notice first that (ii) does not need affiliation. Notice next from Example 3 that, the case where payoffs w_i are decreasing with i is also necessary for (ii) to obtain. This case represents environments where students rank unanimously as top schools those that have a higher opportunity cost of admitting them, and therefore where it is more difficult to get in. A leading example is provided by $w_i(x_s, q_s) = q_s - c_i$, with c_i increasing in i. In this case, if simple serial dictatorship also delivers weak stable and incentive compatible allocations, these allocations are not guaranteed to be ex-post efficient, as the following example shows.

Example 7 *Simple serial dictatorship does not guarantee ex-post efficiency.*

Consider the same environment as in Example 2, but where (ii) is reversed in sign, and the allocations μ^1, μ^2 with ι_μ's given by

$$\iota_{\mu^1} = \left\{ \begin{array}{ccc} 2 & 2 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{array} \right\}, \iota_{\mu^2} = \left\{ \begin{array}{ccc} 2 & 2 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{array} \right\}$$

One can easily check that μ^1 is weakly stable and it is obtained through simple serial dictatorship. However, it is not ex-post efficient, as it is dominated by μ^2 obtained via modified serial dictatorship. Not only does modified serial dictatorship work where simple serial dictatorship may not, but when both work, it has better efficiency properties.■

4.3 Matching patterns

Little can be said in general about the average quality of students assigned to a school and the school's ex-ante position in the students' rankings. However, in the leading example where $w_i(x_{s_\lambda}, q_{s_\lambda}) = q_{s_\lambda} - c_i$, and c_i is increasing in i, it seems natural to expect that the average quality of admitted students is increasing with the quality of the school, (i.e., with i), especially when the environment is affiliated. We consider this question now.

Let

$$\mathcal{I}_\mu(i) = \left\{(\lambda, x_{s_\lambda}) \in \mathbf{\Lambda} \times \mathbf{X}^I \mid \iota_\mu(s_\lambda; x_{s_\lambda}) = i\right\}$$

be the information extracted from an allocation μ about student quality only given that a student goes to school $i \in \mathbf{I}$ (but not given λ). We evaluate the probability of this set based on the induced measure $N(\lambda) = \#\mathbf{S}_\lambda / S$, and we denote it as $\Pr_N[\mathcal{I}_\mu(i)]$. If λ denotes the generic student prior quality, a random variable, when $\Pr_N[\mathcal{I}_\mu(i)] > 0$, $\mathbb{E}_N[q_{s_\lambda} \mid \mathcal{I}_\mu(i)]$ is the average or expected quality of the students who go to school i under μ. Strict monotonicity of students' preferences and monotonicity of posteriors in λ imply the following immediate result.

Proposition 4 *Suppose $w_i(x_{s_\lambda}, q_{s_\lambda}) = q_{s_\lambda} - c_i$, with c_i is increasing in i. The expected quality of students at school i is non–decreasing in i, for all i with $\Pr_N[\mathcal{I}_\mu(i)] > 0$, for all strongly stable allocations μ.*

Proof. See the Appendix. □

Notice that the last result obtains even when the environment is not affiliated. On the other hand, weak posterior stable allocations do not necessarily display this kind of monotonicity, even in affiliated environments with c_i increasing in i, as the next example demonstrates.

Example 8 *The average quality of a student admitted at $i + 1$ may be lower than that at i, for weak posterior stable allocations.*

Consider affiliated environments with $\mathbf{I}_+ = \{1, 2\}$, $\mathbf{Q} = \{0, 1\}$, binary, symmetric signals, conditionally independent given q, where $\Pr(x_{s_\lambda}, q_{s_\lambda}; \lambda)$ is continuous in λ, and two types of students λ_a and λ_b For student λ_a, (i) $u_i(s_{\lambda_a}|x) > 0$ iff $x_{s_{\lambda_a}} \neq (L, L)$, $i = 1, 2$, (ii) $u_2(s_{\lambda_a}|x_{2,s_{\lambda_a}} = L) < 0$, whereas for student λ_b, (iii) $u_i(s_{\lambda_b}|x) > 0$ iff $x_{s_{\lambda_b}} = (H, H)$, $i = 1, 2$, (iv) $u_2(s_{\lambda_a}|x_{2,s_{\lambda_b}} = H) > 0$.

We also assume that $\Pr(x_{i,s_\lambda} = H|q_{s_\lambda} = 1) = \Pr(x_{i,s_\lambda} = L|q_{s_\lambda} = 0) = \beta$, and $\frac{1}{2} < \beta < 1$. Note that here $\mathbb{E}[q_{s_\lambda}] = \lambda$: students of type λ_a are of better average quality, ex-ante and ex-post for the same x_s.

For $k = 0, 1, 2,\ \ j = 1, 2$ and $i \in \mathbf{I}_+$, let $L_i^k = \min\{\lambda \in [0,1] \mid \mathbb{E}[q_{s_\lambda} - c_i)|x_{i,s_\lambda} = H$ for k schools$] \geq 0\}$, and $l_i^j = \min\{\lambda \in [0,1] \mid u_i(s_\lambda|x_{i,s_\lambda} = x_j) \geq 0\}$, which are well-defined by continuity of Pr in λ. Note that $\lambda_a \in (L_2^1, l_2^1)$, $\lambda_b \in (l_2^2, L_1^1)$, and $L_i^2 < l_i^2 < L_i^1 < l_i^1 < L_i^0$, $L_1^1 < L_2^1$. Consider the allocation μ^{MSD} :

$$\iota_\mu(s_{\lambda_a}; x_{s_{\lambda_a}}) = \left\{ \begin{matrix} 2 & 2 \\ 0 & 1 \end{matrix} \right\} \text{ and } \iota_\mu(s_{\lambda_b}; x_{s_{\lambda_b}}) = \left\{ \begin{matrix} 2 & 2 \\ 0 & 0 \end{matrix} \right\}$$

and choose λ_b, λ_a and $N(.)$ as follows: $\Pr_N(\lambda_b|\mathcal{I}_\mu(2))/\Pr_N(\lambda_a|\mathcal{I}_\mu(2)) \to 1$; $\lambda_b = l_2^2 + \varepsilon < L_1^1,\ \ \varepsilon > 0$ such that $\mathbb{E}\ [q_{s_{\lambda_b}}|\ x_{2,s_{\lambda_b}}{=}H] = \beta\lambda_b/[\beta\lambda_b + (1-\beta)(1-\lambda_b)] = c_2 + \delta$, $\delta > 0$, and $\lambda_a = c_2 + 2\delta < l_2^1$. This can be done since conditional probabilities are increasing and continuous in λ. Then, school 2's students are not better than school 1's: $\mathbb{E}_N[q_{s_{\tilde{\lambda}}}|\ \mathcal{I}_\mu(1)] = \lambda_a > \mathbb{E}_N[q_{s_{\tilde{\lambda}}}|\ \mathcal{I}_\mu(2)]$. ■

Since both types of students get into school 2 with positive probability, while only type λ_a may be assigned to school 1, and since most students are of type λ_b, expected qualities are decreasing in i for λ_b large enough. While the expected quality of a student in school 1 is greater than school 2's cutoff c_2, school 2 cannot block such an assignment without knowing which, if any, student has been admitted to school 1.

5 Conclusion

We have considered student assignment problems with interdependent values, introducing a notion of stability against renegotiation after all or parts of allocation outcome is observed by the participants. We have shown that in general posterior stability and incentive compatibility are mutually inconsistent when the allocation outcome is publicly observed. In contrast, when each agent privately observes its own parts of the allocation outcome, modified serial dictatorship delivers posterior stability and incentive compatibility, provided communication in the blocking stage is restricted. Restrictions on communication among blocking coalitions are necessary to obtain existence. We also analyze efficiency and monotonicity properties of modified serial dictatorship.

6 Appendix

Proof of Proposition 1. Step 1: μ posterior stable $\Rightarrow$ μ degenerate. We show this for weak stable allocations; the proof working also for strong stable allocations, they must also be degenerate. The allocation μ is degenerate iff the set $\mathbf{P}_{s,x} = \{k \in \mathbf{I} \mid \mu_{k,s}(x) > 0\}$ is a singleton for each x, all $s \in \mathbf{S}$. Suppose not and let $i, j \in \mathbf{P}_{s,x}$ with $v_{i,s} > v_{j,s}$, some s, x. Without loss of generality, assume that $v_{i,s}$ is increasing in i following the natural order.

Start with $i = I$. First, $U_I(\mu, s|x_I, \mu_I(x)) \geq 0$, or else the coalition $\mathbf{C} = \{\mathbf{I}', \mathbf{S}'\}$ with $\mathbf{I}' = \{0, I\}$ and $\mathbf{S}' = \{0\}$ would block μ at x: here $\mu'_{I,s} = 0$, and all we need to verify is that the strategy $\alpha_I(x'_I, m_I)$ constructed according to Definition 2 has the property $\alpha_I(x_I, \mu_I(x)) = 1$. Now, notice that at $x'_I = x_I$ and $m_I = \mu_I(x)$, $U_I(\mu, s|x'_I, m_I, piv_I(\alpha_{\mathbf{C}})) = U_I(\mu, s|x_I, \mu_I(x)) < 0$, while of course $U_I(\mu', s|x'_I, m_I, piv_I(\alpha_{\mathbf{C}})) = 0$, verifying our claim that I would block.

By A4, $U_I(\mu, s|x_I, \mu_I(x)) = u_I(s|x_I, \mu_I(x))\mu_{I,s}(x) > 0$. If so, student s and school I would form a blocking coalition with $\mathbf{S}' = \{0, s\}$ at x using $\mu'_{I,s} = 1$ (and $\mu'_s = \mu_s$, $s \notin \mathbf{S}'$). By A3 and $\mu_{I,s}(x) < 1$, $\alpha_s(\mu_s(x)) = 1$. At $x'_I = x_I$ and $m_I = \mu_I(x), U_I(\mu|x'_I, m_I, piv_I(\alpha_{\mathbf{C}})) = U_I(\mu|x_I, \mu_I(x))$ and $U_I(\mu', \mathbf{S}'|x'_I, m_I, piv_I(\alpha_{\mathbf{C}})) = u_I(s|x_I, \mu_I(x)) > U_I(\mu, \mathbf{S}'|x_I, \mu_I(x))$. Hence, if the set $\mathbf{P}_{s,x}$ is not a singleton, it cannot contain I. Now proceed by induction on i to show that if $\mathbf{P}_{s,x}$ is not a singleton and cannot contain $i' > i$, it cannot contain i. This leads to a contradiction, i.e., $\mathbf{P}_{s,x}$ is a singleton. As a consequence, to find weakly stable allocations, we can restrict attention to degenerate μ's: $\mu_{i,s}(x) > 0$ implies $\mu_{i,s}(x) = 1$.

Step 2: ex-post stable allocations exist. An allocation μ is ex-post blocked at $x \in \mathbf{X}^{SI}$ if there exist $\mathbf{C} = (\mathbf{I}', \mathbf{S}') \subset \mathbf{I} \times \mathbf{S}$, a degenerate μ' that is feasible for $\mathbf{C}$, and acceptance strategies $\alpha_{\mathbf{C}}$ such that for all $i \in \mathbf{I}'$ and $s \in \mathbf{S}'$, $\alpha_i(x) = 1$ iff $U_i(\mu' - \mu|x) \geq 0$, $\alpha_s(x) = 1$ iff $U_s(\mu' - \mu|x) \geq 0$, with one strict inequality. An allocation μ is ex-post stable if there is no x where it is ex-post blocked.

Notice that, given A1, A2, an allocation μ is ex-post blocked at x if and only if it is blocked by coalitions $\mathbf{C}$ with $\#\mathbf{I}'_+ \leq 1$: the only role of multiple school blocks is to share information among schools, but this is useless ex-post. On the other hand, given A1, A2 and separability of schools' preferences over students if coalitions with $\#\mathbf{S}'_+ > 1$ can block, then so can coalitions with the same school and one student at a time. It is then immediate to show that an allocation μ is ex-post blocked at $x \in \mathbf{X}^{SI}$ if and only if either $\mu_{i,s}(x) > 0$ and $u_i(s|x) < 0$, or $\mu_{i,s}(x) > 0$ and there is an $i' > i$ with $u_{i'}(s|x) \geq 0$. Then, μ is ex-post stable if for all $s \in \mathbf{S}_+$ and all $x \in \mathbf{X}$, $\mu_s(x)$ solves

$$\max_{\mu \in \mathbb{R}^I_+} \sum_{i \in \mathbf{I}} \mu_{i,s}(x) v_{i,s} \quad \text{s.t.} \quad \sum_{i \in \mathbf{I}} \mu_{i,s}(x) = 1 \text{ and } \mu_{i,s}(x) u_i(s|x) \geq 0 \qquad \text{(EP-stable)}$$

Indeed, μ solves (EP-stable) $\Rightarrow$ μ ex-post stable. For suppose μ is not ex-post stable, but it solves (EP-stable). Clearly the second constraint in (EP-stable) rules out $\mu_{i,s}(x) > 0$ and $u_i(s|x) < 0$. For the remaining blocking possibility, suppose there is $s \in \mathbf{S}_+$, $x \in \mathbf{X}^{SI}$ and $i \in \mathbf{I}$ with $u_i(s|x) \geq 0$, $v_i > v_j$ and $\mu_{j,s}(x) > 0$. Since $u_j(s|x) \geq 0$, we can find $\mu"_s(x) = \mu_s(x) + \Delta\mu$, where $\Delta\mu \in \mathbb{R}^I$ is such that $\Delta\mu_i = -\Delta\mu_j > 0$, $\Delta\mu_k = 0$ if $k \neq i, j$ and $[\mu_{j,s}(x) + \Delta\mu_j]u_j(s|x) \geq 0$. Then, we have $\sum_{k \in \mathbf{I}} \mu"_{k,s}(x) v_{k,s} > \sum_{k \in \mathbf{I}} \mu_{k,s}(x) v_{k,s}$, while $\mu"_{i,s}(x) u_i(s|x) \geq 0$, $\mu"_{j,s}(x) u_j(s|x) \geq 0$, and $\sum_{k \in \mathbf{I}} \mu"_{k,s}(x) = 1$. Hence, μ does not solve (EP-stable), a contradiction. Since the set of $\mu_s(x) \in \mathbb{R}^I_+$ with $\sum_{i \in \mathbf{I}} \mu_{i,s}(x) = 1$ and $\mu_{i,s}(x) u_i(s|x) \geq 0$ is nonempty, compact and the objective function is continuous, a solution to (EP-stable) exists. Existence of ex-post stable allocations now follows.

Step 3: μ ex-post stable $\Rightarrow$ μ posterior stable. We prove this for strong stable allocations, again the proof working also for weak stable allocations, they must also include ex-post stable ones. If μ is not strong stable, then suppose there exists an $x \in \mathbf{X}^{SI}$, a $\mathbf{C} = (\mathbf{I}', \mathbf{S}')$ and a constant and degenerate μ', feasible for $\mathbf{C}$, such that $\mathbf{C}$ blocks μ at x with μ'. Suppose that

$\mathbf{S}'_+ = \varnothing$. Then, it must be that for all $i \in \mathbf{I}'_+$, $U_i(\mu' - \mu|x_i, \mu(x)) > 0$. Since $U_i(\mu' - \mu|x_i, \mu(x)) = \sum_{x_{-i} \in \mathbf{X}^{S(I-1)}} U_i(\mu' - \mu|x_i, x_{-i}) \Pr(x_{-i}|x_i, \mu(x)) > 0$, there exists $\widehat{x}_{-i}$ such that $\Pr(\widehat{x}_{-i}|x_i, \mu(x)) > 0$ and $U_i(\mu' - \mu|x_i, \widehat{x}_{-i}) > 0$. Then i will ex-post block alone at $\widehat{x} = (x_i, \widehat{x}_{-i})$, with μ'_i. Suppose instead that $\mathbf{S}'_+ \neq \varnothing$. Then, for each $i \in \mathbf{I}'_+$,

$$0 < U_i(\mu' - \mu|x_i, \mu(x), piv_i(\alpha_{\mathbf{C}})) = \sum_{x_{-i} \in \mathbf{X}^{I-1}} U_i(\mu' - \mu|x_i, x_{-i}) \Pr(x_{-i}|x_i, \mu(x), piv_i(\alpha_{\mathbf{C}})).$$

and there exists an $\widehat{x}_{-i} \in \mathbf{X}^{S(I-1)}$ with $\Pr(\widehat{x}_{-i}|x_i, \mu(x), piv_i(\alpha_{\mathbf{C}})) > 0$ and $U_i(\mu' - \mu|x_i, \widehat{x}_{-i}) > 0$, and there exists $s \in \mathbf{S}_+$ with $U_i(\mu' - \mu, s|x_i, \widehat{x}_{-i}) > 0$. Let $\mathbf{S}_i = \{s \in \mathbf{S}_+ | U_i(\mu' - \mu, s|x_i, \widehat{x}_{-i}) > 0\}$, and consider the set $\mathbf{S}_i(\mu') = \{s \in \mathbf{S}_i |\ \mu'_{i,s} = 1\}$. If $\mathbf{S}_i(\mu') \subset \mathbf{S}_i$ for at least an $i \in \mathbf{I}'_+$, then again school i will ex-post block alone at $\widehat{x}$. If $\mathbf{S}_i(\mu') = \mathbf{S}_i$ for all $i \in \mathbf{I}'_+$, then $\mathbf{S}_i \subset \mathbf{S}'_+$ and observe that for each i, $\mu_{i,s}(x) = \mu_{i,s}(\widehat{x}) < 1 = \mu'_{i,s}$ for all $s \in \mathbf{S}_i$. On the other hand, using A3, $U_s(\mu' - \mu|m, piv_s(\alpha_{\mathbf{C}})) > 0$ implies that $\sum_j v_{j,s}(\mu'_{j,s} - m_{j,s}) > 0$ where $m = \mu(x)$ and $\mu_{j,s}(x) = \mu_{j,s}(\widehat{x}) = m_{j,s}$. Since μ' is degenerate, for each s, $(\mu'_{j,s} - m_{j,s}) > 0$ for at most one $j \in \mathbf{I}'_+$, and one student $s \in \mathbf{S}_i$ strictly accepts to block with one school $i \in \mathbf{I}'_+$ ex-post at $\widehat{x}$.■

Proof of Theorem 1. We break the proof up into a few steps.

Step 0: Construction of modified serial dictatorship (MSD)

For each $s \in \mathbf{S}_+$, and any $i, j \in \mathbf{I}$, we write $i >_s j$ if $v_{i,s} > v_{j,s}$. Let $y_{i,s}$ be the vector of signals received by all $j >_s i$, while ${}_{i,s}y$ be the vector of signals received by all $j <_s i$, at x. Notice that when $y_{i,s}$ is fixed, so are the $y_{j,s}$ for $j >_s i$.

Let a student $s \in \mathbf{S}_+$ and an $x_s \in \mathbf{X}^I$ be given, with $I_s = \arg\max_i v_{i,s}$. Construct the (unique and degenerate) MSD allocation μ: starting from school I_s and going down to school 0, for each school i allocate the student entirely to that school if and only if $u_i(s|x_{i,s}, y_{i,s}) \geq 0$ and if the student has not yet been allocated to a school $j >_s i$. Then, for $s \in \mathbf{S}_+$, using A4,

$$\mu_{i,s}(x_s) = 1 \text{ iff } u_i(s|x_{i,s}, y_{i,s}) > 0 > u_j(s|x_{j,s}, y_{j,s}) \text{ for all } j \text{ s.t. } j >_s i,$$

Notice also that if $x_s = ({}_{i,s}y, x_{i,s}, y_{i,s})$ and $x'_s = ({}_{i,s}y', x_{i,s}, y_{i,s})$ with ${}_{i,s}y' \neq {}_{i,s}y$, then $\mu_{i,s}(x_s) = \mu_{i,s}(x'_s)$.

Step 1: If a coalition $\mathbf{C} = \{\mathbf{I}', \mathbf{S}'\}$ weak blocks the allocation of Step 1 at some x, then $\#\mathbf{S}'_+ \leq 1$ without loss of generality.

Suppose that a coalition $\mathbf{C}$ blocks μ at x. If $\mathbf{S}'_+ = \varnothing$, there is nothing to prove, so suppose that $\mathbf{S}'_+ \neq \varnothing$. Then $\mathbf{I}'_+$ is nonempty, by strict acceptance for the students. Pick $i \in \mathbf{I}'_+$ and, using A1, the separability of U_i over the students and the dependence of $\mu_s(x)$ only on x_s, observe from (3) that there exists $s \in \mathbf{S}_+$ (but not necessarily in $\mathbf{S}'_+$) such that

$$\sum_{x'_s \in \mathbf{X}^I} u_i(s|x'_s)[\mu'_{i,s} - \mu_{i,s}(x_s)] \Pr(x'_s|x_{i,s}, \mu_{i,s}(x_s), piv_i(\alpha_{\mathbf{C}})) > 0 \tag{8}$$

First, note that $s \in \mathbf{S}'_+$ without loss of generality. For if not, by feasibility, $\mu'_{j,s} = 0$ for all $j \in \mathbf{I}'_+$ and so, the degeneracy of μ implies that $\mu_{i,s}(x_s) = 1$. It follows that every other school $j \neq i$,

$j \in \mathbf{I}'_+$ neither gains nor loses from student s at x, implying by (3) that $\alpha_j(x'_j, \mu_j(x))$ does not vary with $x'_{j,s}$. But then the event $\{x_{i,s}, \mu_{i,s}(x_s), piv_i(\alpha_{\mathbf{C}})\}$ contains no additional information about x_s for i compared to the event $\{x_{i,s}, \mu_{i,s}(x_s)\}$ allowing us to conclude that the singleton coalition of school i can also block at such x. Next notice that if i is the only element of $\mathbf{I}'_+$ then, by strict acceptance of s, we must have $\mu'_{i,s} = 1$ and so from (8) it is then immediate that the coalition of i and s can also block at x. Consequently, we can focus on the case $\#\mathbf{I}'_+ \geq 2$ in what follows.

We show now that $\mu'_{i,s} = 1$. Suppose not, i.e., $\mu'_{i,s} = 0$. It follows from (8) that $\mu_{i,s}(x_s) = 1$ and, by strict acceptance for s, that $\mu'_{j,s} = 1$ for some $j >_s i$, with $j \in \mathbf{I}'_+$. Moreover, every other school $k \in \mathbf{I}'_+$, $k \neq i, j$, neither gains nor loses from student s at x so that from (3) it follows that $\alpha_k(x'_k, \mu_k(x))$ does not vary with $x'_{k,s}$. But then the event $\{x_{i,s}, \mu_{i,s}(x_s), piv_i(\alpha_{\mathbf{C}})\}$ contains no additional information for i about x_s compared to the event $\{x_{i,s}, \mu_{i,s}(x_s), \alpha_j = 1\}$. From (8) we then obtain

$$U_i(\mu' - \mu, s | x_{i,s}, \mu_{i,s}(x_s), \alpha_j = 1) > 0$$

which by breaking down the expectations can be written as

$$\sum_{y_{i,s}} U_i(\mu' - \mu, s | x_{i,s}, y_{i,s}, \mu_{i,s}(x_s), \alpha_j = 1) \Pr(y_{i,s} | x_{i,s}, \mu_{i,s}(x_s), \alpha_j = 1) > 0 \tag{9}$$

Recall from Step 0 that $\Pr(\mu_{i,s}(x_s) = 1 | x) = \Pr(\mu_{i,s}(x_s) = 1 | x_{i,s}, y_{i,s})$. Further, if $\widehat{x} = ({}_i\widehat{y}, x_i, y_i)$ differs from $x = ({}_iy, x_i, y_i)$ only in ${}_{i,s}\widehat{y} \neq {}_{i,s}y$, and $\mu_{i,s}(x_s) = 1$, then $\alpha_j(\widehat{x}_j, \mu_j(\widehat{x})) = 1$ whenever $\alpha_j(x_j, \mu_j(x)) = 1$. In other words, $\Pr(\alpha_j = 1 | {}_{i,s}y, x_{i,s}, y_{i,s}, \mu_{i,s}(x_s)) = \Pr(\alpha_j = 1 | x_{i,s}, y_{i,s}, \mu_{i,s}(x_s))$. Using Bayes' rule and basic properties of conditional expectations, it follows that

$$\Pr({}_{i,s}y | x_{i,s}, y_{i,s}, \mu_{i,s}(x_s), \alpha_j = 1) = \Pr({}_{i,s}y | x_{i,s}, y_{i,s})$$

implying in turn that we can write (9) as

$$\sum_{y_{i,s}} U_i(\mu' - \mu, s | x_{i,s}, y_{i,s}) \Pr(y_{i,s} | x_{i,s}, \mu_{i,s}(x_s), \alpha_j = 1) > 0$$

Now, if $\Pr(y_{i,s} | x_{i,s}, \mu_{i,s}(x_s), \alpha_j = 1) > 0$, then $\Pr(y_{i,s} | x_{i,s}, \mu_{i,s}(x_s)) > 0$. But $\mu_{i,s}(x_s) = 1$ implies by construction of μ that $y_{i,s}$ is such that $U_i(\mu' - \mu, s | x_{i,s}, y_{i,s}) | x_{i,s}, y_{i,s}) \leq 0$, a contradiction. Thus, $\mu'_{i,s} = 1$.

Now invoke A3$'$. Since $i \in \mathbf{I}'_+$ was arbitrary, we can let $i = \inf \mathbf{I}'_+$. Further, since $\mu'_{i,s} = 1$, by strict acceptance for s, we conclude that $\mu_{i,s}(x_s) = 0$. It follows from (8) that

$$U_i(\mu' - \mu, s | x_{i,s}, \mu_{i,s}(x_s), piv_i(\alpha_{\mathbf{C}})) > 0.$$

Given the students' acceptance strategy, $\mu_{j,s}(x_s) = 0$ for all $j > i$. It follows that every other school $j \in \mathbf{I}'_+$, $j > i$, neither gains nor loses from student s so that from (3) it follows that $\alpha_j(x'_j, \mu_j(x))$ does not vary with $x'_{j,s}$, i.e., the event piv_j does not convey any information known by schools $j \in \mathbf{I}'_+$ with $j > i$, to i about x_s. Since $i = \inf \mathbf{I}'_+$, a coalition $\mathbf{C}'' = \{\mathbf{I}'', \mathbf{S}''\}$ with $\mathbf{I}''_+ = \{i\}$ and $\mathbf{S}''_+ = \{s\}$ can then also block μ at x.

Given Step 1, we do not need to check stability against blocks of coalitions with $\#\mathbf{S}'_+ > 1$. By separability of schools' payoffs over students, construction of μ and A1, if μ is incentive compatible student by student, then it is incentive compatible overall. Hence, without loss of generality hereafter we assume that $\mathbf{S} = \{0, s\}$ and we drop the subscript s from the signals.

Step 2: If a coalition weak blocks the allocation of Step 1, then $\#\mathbf{I}'_+ \leq 1$ without loss of generality.

No coalition $\mathbf{C}$ with $\mathbf{S}'_+ = \varnothing$ and $\#\mathbf{I}'_+ > 1$ can weak block μ, by strict acceptance and feasibility of μ'. So suppose that $\mathbf{S}'_+ \neq \varnothing$.

First, it must be that there is $i \in \mathbf{I}'_+$ with $\mu_{i,s}(x) = 1$. If not, and $\mu_{i,s}(x) = 0$ for all $i \in \mathbf{I}'_+$, since μ' must be degenerate and constant, there is $j \in \mathbf{I}'_+$ which cannot strictly accept the block, a contradiction. Then, $\mu_{j,s}(x) = 0$ for $j \neq i$, $j \in \mathbf{I}'_+$. Next, $\mu'_{j,s} = 1$ for some $j \in \mathbf{I}'_+$, $j \neq i$. If not, and $\mu'_{j,s} = 0$, then j cannot strictly accept. In fact, $\#\mathbf{I}'_+ = 2$. Further, since $\mu'_{j,s} = 1$, then $j >_s i$, or the student will not accept, by A3. Since i accepts strictly, $U_i(\mu' - \mu | x_i, \mu_i(x), piv_i(\alpha_{\mathbf{C}})) > 0$. Now, we apply the same reasoning as in Step 1 to get a contradiction with school i's payoffs at any y_i with $\Pr(y_i | x_i, \mu_i(x), piv_i(\alpha_{\mathbf{C}})) > 0$, using $\mu_i(x) = 1$, and the construction of μ.

Step 3: The allocation of Step 0 is weakly stable.

Consider first a block by i alone at x so that $\mu_{i,s}(x) = 1$ at x. By Step 0 this depends only on x_i, y_i. We can then write $U_i(\mu' - \mu | x_i, \mu_i(x)) > 0$ as

$$\begin{aligned} &= \sum_{y_i} U_i(\mu' - \mu | x_i, y_i, \mu_i(x)) \Pr(y_i | x_i, \mu_i(x)) \\ &= \sum_{y_i} U_i(\mu' - \mu | x_i, y_i) \Pr(y_i | x_i, \mu_i(x)) \end{aligned}$$

where the second equality holds since once we know y_i, we also know whatever is revealed by $\mu_{i,s}(x) = 1$, from Step 0. But each $U_i(\mu' - \mu | x_i, y_i) \leq 0$ if $\Pr(y_i | x_i, \mu_i(x)) > 0$ by construction of μ, and so is the average, and i does not want to block alone.

As for the remaining blocking possibility of a school i and a student s at x, we must have $\mu_{i,s}(x) = 0$ and $\mu'_{i,s} = 1$ and

$$U_i(\mu' - \mu, s | x_i, \mu_i(x), piv_i(\alpha_{\mathbf{C}})) > 0. \tag{10}$$

As in Step 1, using the properties of μ derived in Step 0 and Bayes' rule, it follows that $\Pr({}_i y | x_i, y_i, \mu_{i,s}(x), piv_i(\alpha_{\mathbf{C}})) = \Pr({}_i y | x_i, y_i)$ implying in turn that we can write (10) as

$$\sum_{y_i} u_i(s | x_i, y_i) \Pr(y_i | x_i, \mu_{i,s}(x), piv_i(\alpha_{\mathbf{C}})) > 0$$

Now, if $\Pr(y_i | x_i, \mu_{i,s}(x), piv_i(\alpha_{\mathbf{C}})) > 0$, then $\Pr(y_i | x_i, \mu_{i,s}(x)) > 0$. But $\mu_{i,s}(x) = 0$ implies by construction of μ that y_i is such that $u_i(s | x_i, y_i) < 0$, a contradiction.

Step 4: The allocation of Step 0 is incentive compatible.

For any i, let x_i be the true signal for i, and $\widehat{x}_i \neq x_i$ be i's false report; (IC) can be written as

$$\sum_{y_i}\{\sum_{iy} u_i(s|x)[\mu_{i,s}(x) - \mu_{i,s}(\widehat{x}_i, x_{-i})] \Pr[_iy|x_i, y_i]\} \Pr[y_i|x_i] \geq 0 \tag{D}$$

Fix y_i and consider the corresponding term within the curly braces. Suppose first that x_i, y_i is such that $u_i(s|x_i, y_i) < 0$, so that $\mu_{i,s}(x) = 0$. Then lying is going to affect the outcome μ only if at $\widehat{x}_i$, $u_i(s|\widehat{x}_i, y_i) \geq 0$ and y_i is such that $u_j(s|x_j, y_j) < 0$ for all $j >_s i$. In this case, $\mu_{i,s}(\widehat{x}_i, x_{-i}) = 1$, so that in (D) the term in the first summation corresponding to y_i becomes

$$-\sum_{iy} u_i(s|x) \Pr[_iy|x_i, y_i] = -u_i(s|x_i, y_i) > 0.$$

Suppose next that x_i, y_i is such that $u_i(s|x_i, y_i) \geq 0$ so that $\mu_{i,s}(x) = 1$. Again, lying is going to affect the outcome μ only if at $\widehat{x}_i$, $u_i(s|\widehat{x}_i, y_i) < 0$. Then, $\mu_{i,s}(\widehat{x}_i, x_{-i}) = 0$, and so in (D) the term in the first summation corresponding to y_i becomes

$$\sum_{iy} u_i(s|x) \Pr[_iy|x_i, y_i] = u_i(s|x_i, y_i) \geq 0$$

In every other case, lying is not going to affect the outcome. It follows that (IC) is satisfied for all i contingent on y_i, and so it is satisfied in interim expected terms (i.e., contingent on x_i only).■

Proof of Proposition 2. Let μ be obtained through modified serial dictatorship. We prove a stronger version of the result, i.e., that (IC-R) must hold for all i conditional not only on x_i but also on y_i, where $y_i = (y_{i,s})_{s \in \mathbf{S}_+}$ and $y_{i,s}$ is defined in the proof of Theorem 1, Step 0. Furthermore, we suppose that, apart from i's signal x_i, message $\widehat{x}_i$ and observed outcome $m_i(\widehat{x}_i, x_{-i})$, the best blocking strategy of i can also depend on y_i and we write this as $\widehat{\beta}_i(x_i, y_i, m_i(\widehat{x}_i, x_{-i}))$. The stated result then immediately obtains if $\widehat{\beta}_i$ cannot be made contingent on y_i and when (IC-R) must hold only in expected terms over y_i given x_i.

Notice first that for each $i \in \mathbf{I}_+$, given x_i, y_i, and any message $\widehat{x}_i$ the observed outcome $m_i(\widehat{x}_i, x_{-i})$ contains no information for i that is not already contained in x_i, y_i. This follows from the properties of modified serial dictatorship (see the proof of Theorem 1, Step 0). In other words, $\Pr[_iy|x_i, y_i, \widehat{x}_i, m_i(\widehat{x}_i, x_{-i})] = \Pr[_iy|x_i, y_i]$ where $_iy = (x_j)_{j<i}$ is the only information about the state of the world unknown to i given x_i, y_i.

Suppose now that for some $i \in \mathbf{I}_+$, and given x_i, y_i, (IC-R) does not hold, i.e., i is strictly better off from an alternative message $\widehat{x}_i \neq x_i$ and subsequently using the blocking strategy $\widehat{\beta}_i(x_i, y_i, m_i(\widehat{x}_i, y_i))$. Since μ is incentive compatible, from Theorem 1 it follows that $\widehat{\beta}_i(x_i, y_i, m_i(\widehat{x}_i, y_i))$ must be a non-trivial block. We show now that in fact the blocking coalition $\mathbf{C}' = \{\mathbf{I}', \mathbf{S}'\}$ associated with $\widehat{\beta}_i(x_i, y_i, m_i(\widehat{x}_i, y_i))$ must contain at least one other school $j \in \mathbf{I}_+$ with $j < i$.

For suppose not. Consider first the case where $\mathbf{I}'_+ = \{i\}$. With fixed y_i, the set of students who are assigned to schools $k > i$ does not depend on i's message, by properties of MSD. Hence, by announcing $\widehat{x}_i$ instead of x_i school i can only manage to lose some students s with $u_i(s|x_i, y_i) > 0$ to schools $j < i$ and also possibly gain some students s' with $u_i(s'|x_i, y_i) < 0$ from schools $j' < i$. Subsequently, a blocking strategy where i is the singleton school in the coalition cannot attract any students from schools $k > i$ and can only attract students lost to schools lost to $j < i$ and drop those gained from $j' < i$ as a result of the message $\widehat{x}_i$. Further, it will not convey any information (via the acceptance strategies of students in the coalition) that i already does not know from x_i, y_i. Thus, the best such block cannot make i better off than announcing x_i truthfully in the first place. It follows that if (IC-R) does not hold, the ensuing blocking strategy must involve a school $j \neq i$ from whom i obtains information at the blocking stage via acceptance strategies. Moreover, since i can condition on y_i it is evident that $j < i$, since otherwise j does not have any information that i does not. Thus, there must exist $j \in \mathbf{I}'_+$ with $j < i$.

Now consider $j = \min \mathbf{I}'_+$. Since j participates in the block under the presumption that i tells the truth, arguments identical to that contained in Theorem 1, Step 1, lead us to conclude that a coalition $\mathbf{C}'' = \{\mathbf{I}'', \mathbf{S}''\}$ with $\mathbf{I}''_+ = \{j\}$ and $\mathbf{S}''_+ = \{s\}$, some $s \in \mathbf{S}_+$, can weak block μ at $({}_iy, \widehat{x}_i, y_i)$ for some ${}_iy$, a contradiction with Theorem 1. Thus, (IC-R) must hold for all i and x_i.■

Proof of Proposition 3. (i) Suppose μ^{MSD} is not strong posterior w-efficient. Let $\mathbf{C} = (\mathbf{I}', \mathbf{S}')$ be the blocking coalition. Necessarily, $\mathbf{S}'_+ \neq \varnothing$, or students will be worse off. First consider the case where $\mu'_{i,s} = 0$ all $s \in \mathbf{S}'_+$, for some $i \in \mathbf{I}'_+$. Then, the set $\mathbf{S}_i(\mu) = \{s \in \mathbf{S}_+ \mid \mu_{i,s}(x_s) = 1\}$ is nonempty, and $\mathbf{S}_i(\mu) \subset \mathbf{S}'_+$, or again students would be worse off. Then, for each $s \in \mathbf{S}_i(\mu)$, $\mu'(j, s) = 1$ for some $j >_s i$, or the student would not be better off, and effectively the blocking coalition $\mathbf{C} = (\mathbf{I}', \mathbf{S}')$ has $\#\mathbf{I}'_+ \geq 2$. Since i's acceptance is strict, $U_i(\mu' - \mu, \mathbf{S}_i(\mu)|x_i, \mu(x), piv_i(\alpha_{\mathbf{C}})) > 0$, and using A1, construction of μ (Theorem 1, Step 0), and separability of schools' payoffs across students, this can be written as (9) for some $s \in \mathbf{S}_i(\mu)$, conditioning on $\mu_s(x_s)$ — or equivalently on $\mu_{i,s}(x_s) = 1$. As in Theorem 1, Step 1, given $x_{i,s}, y_{i,s}$, there is nothing else that is revealed by $\mu_s(x_s)$ once $x_{i,s}, y_{i,s}$ is known, while by construction of μ, $\Pr({}_{i,s}y|x_{i,s}, y_{i,s}, \mu_s(x_s), piv_i(\alpha_{\mathbf{C}})) = \Pr({}_{i,s}y|x_{i,s}, y_{i,s})$. So $U_i(\mu' - \mu, s|x_{i,s}, y_{i,s}, \mu_s(x_s), piv_i(\alpha_{\mathbf{C}})) = U_i(\mu' - \mu, s|x_{i,s,}y_{i,s})$, and $\mu_{i,s}(x_s) = 1$ implies by construction of μ that $y_{i,s}$ is such that $U_i(\mu' - \mu, s|x_{i,s,}y_{i,s}) < 0$, a contradiction. If instead $\mu'_{i,s} = 1$ for some $s \in \mathbf{S}'_+$, for all $i \in \mathbf{I}'_+$, then by A3$'$, $\mu_{0,s}(x_s) = 1$. Indeed, $\mu_{k,s}(x_s) = 0$ for $k \geq i = \inf \mathbf{I}'_+$, by A3$'$ and because the student must be strictly better off; while if $\mu_{j,s}(x_s) = 1$ some $j < i$, $j \notin \mathbf{C}$, then $U_j(\mu' - \mu|x_j, \mu(x)) = U_j(\mu' - \mu|x_j, \mu_j(x)) \geq 0$, with strict inequality by A4, which implies j had students it did not want, a contradiction to weak stability of μ. Let $\mathbf{S}_i$ be the set of students whose payoff from $\mu' - \mu$ is positive for i. Again $\mathbf{S}_i \subset \mathbf{S}'_+$ or the students would be worse off. If $s \in \mathbf{S}_i$ implies $\mu'_{i,s} = 0$, then we get a contradiction as in the previous case. If $s \in \mathbf{S}_i$ and $\mu'_{i,s} = 1$, observe that by construction of μ (Theorem 1, Step 0) $u_i(s|x_{i,s}, y_{i,s}) < 0$, and by affiliation this holds iff $x_{i,s} < \overline{x}_{i,s}(y_{i,s})$, where $\overline{x}_{i,s}(y_{i,s}) = \inf\{z \in \mathbf{X}|u_i(s|z, y_{i,s}) \geq 0\}$ for $i < I$, and $\overline{x}_{I,s} = \inf\{z \in \mathbf{X}|u_I(s|z) \geq 0\}$. Hence, the set $\{x' \in \mathbf{X}^{SI}|\mu_{j,s}(x') = 0 \text{ for all } j > 0\}$ is identical to the set $\{x' \in \mathbf{X}^{SI}|x_{j,s} < \overline{x}_{j,s}(y_{j,s}) \text{ for all } j > 0\}$. Therefore, if $\Pr(x_{-i,s}|x_{i,s}, \mu(x), piv_i(\alpha_{\mathbf{C}})) > 0$, then by affiliation $U_i(\mu' - \mu, s|x_{i,s}, x_{-i,s}) \leq U_i(\mu' - \mu, s|x_{i,s}, y_{i,s})$, and knowing further the signals for $j < i$ reduces the expected payoff for school i,

$$
\begin{aligned}
& U_i(\mu' - \mu, s|x_i, \mu(x), piv_i(\alpha_{\mathbf{C}})) \\
= & \sum_{x_{-i,s}} U_i(\mu' - \mu, s|x_{i,s}, x_{-i,s}, \mu(x), piv_i(\alpha_{\mathbf{C}})) \Pr(x_{-i,s}|x_i, \mu(x), piv_i(\alpha_{\mathbf{C}})) \leq U_i(\mu' - \mu, s|x_{i,s}, y_{i,s})
\end{aligned}
$$

Hence school i is made worse off by μ', and μ cannot be strongly dominated. This completes the proof of (i).

(ii) Suppose w_i is decreasing in i, but μ^{MSD} is not ex-post efficient. Then at some $x \in \mathbf{X}^{SI}$ there exists a degenerate μ' that is feasible for $\mathbf{C} = (\mathbf{I}, \mathbf{S})$ such that for all $i \in \mathbf{I}_+$ and $s \in \mathbf{S}_+, U_i(\mu' - \mu|x) \geq 0$, $U_s(\mu' - \mu|x) \geq 0$, with one strict inequality. Observe first that there must exist $i \in \mathbf{I}_+$ who gets a student s_i under μ' and s_i was assigned to $i' < i$ under μ, for if not, either $\mu' = \mu(x)$ or one student is worse off, a contradiction with ex-post domination. Next, for at least one such $i \in \mathbf{I}_+$ and corresponding s_i we must have $u_i(s_i|x_{s_i}) > 0$. For suppose not. Let i be the highest school who is assigned a student under μ' not assigned to it under μ. If $u_i(s_i|x_{s_i}) < 0$ for every such student s_i then for i to be better off under μ', school i must lose some student s' assigned to i under μ, to some $j \neq i$. Since s' is also better off under μ' we must have $j > i$. But then i cannot be the highest school who is assigned a student under μ' not assigned to it under μ, a contradiction.

Since there exists $i \in \mathbf{I}_+$ who gets a student s_i under μ' with $u_i(s_i|x_{s_i}) > 0$ and since s_i is better off under μ' we conclude that s_i was assigned to $i' < i$ under μ. Since $w_i < w_{i'}$ we see that $u_{i'}(s_i|x_s) > 0$. Since i' is also better off under μ' this implies that i' must get some other student $s_{i'}$ under μ' with $u_{i'}(s_{i'}|x_{s_{i'}}) > 0$ from some $i'' < i'$ with $u_{i''}(s_{i'}|x_{s_{i'}}) > 0$, and so on. Continuing this logic, we obtain a school i^n and a student s_{i^n} obtained by i^n under μ' with $u_{i^n}(s_{i^n}|x_{s_{i^n}}) > 0$ where s_{i^n} comes from school 0. But since $w_{i^n} \leq w_1$, $u_{i^n}(s_{i^n}|x_{s_{i^n}}) \leq u_1(s_{i^n}|x_{s_{i^n}}) \leq 0$, a contradiction, where the last inequality follows from the properties of MSD for all s_{i^n} assigned to school 0 under μ.■

Proof of Proposition 4. Fix N and suppose μ is strong posterior stable. When $\Pr_N[\mathcal{I}_\mu(i)] > 0$, we can write (dropping the subscript N),

$$\mathbb{E}[q_{s_\lambda}|\mathcal{I}_\mu(i)] = \sum_{\lambda \in \mathbf{\Lambda}} \sum_{x_i \in \mathbf{X}^S} \mathbb{E}[q_{s_\lambda}|\lambda, x_i, \mathcal{I}_\mu(i)] \Pr[\lambda, x_i|\mathcal{I}_\mu(i)]$$

Since λ is observed by all schools, using A1, we see that $\mathbb{E}[q_{s_\lambda}|\lambda, x_i, \mathcal{I}_\mu(i)]$ must be equal to the expected value of student s_λ computed by school i when it receives the student and observes the signal x_i for every λ, x_i with $\Pr[\lambda, x_i|\mathcal{I}_\mu(i)] > 0$ (i.e., there exists x_{-i} such that $\mu_{i,s_\lambda}(x_i, x_{-i}) = 1$). Since μ is stable, we see that $\mathbb{E}[q_{s_\lambda}|\lambda, x_i, \mathcal{I}_\mu(i)] \geq c_i$ for all such λ, x_i as otherwise i would block μ at $x = (x_i, x_{-i})$ for some x_{-i} with $\mu_{i,s_\lambda}(x) = 1$. It follows that $\mathbb{E}[q_{s_\lambda}|\mathcal{I}_\mu(i)] \geq c_i$. Now if $\mathbb{E}[q_{s_\lambda}| \mathcal{I}_\mu(i-1)] > \mathbb{E}[q_{s_\lambda}| \mathcal{I}_\mu(i)]$ for some $i \in \mathbf{I}_+$, then we see that

$$\mathbb{E}[q_{s_\lambda}|\mathcal{I}_\mu(i-1)] = \sum_{\lambda \in \mathbf{\Lambda}} \sum_{x_i \in \mathbf{X}^{SI}} \mathbb{E}[q_{s_\lambda}|\lambda, x_i, \mathcal{I}_\mu(i-1)] \Pr[\lambda, x_i|\mathcal{I}_\mu(i-1)] > c_i$$

But then there must exist λ', x'_i with $\Pr[\lambda', x'_i|\mathcal{I}_\mu(i-1)] > 0$ such that $\mathbb{E}[q_{s_{\lambda'}}|\lambda', x'_i, \mathcal{I}_\mu(i-1)] > c_i$, implying that school i and student $s_{\lambda'}$ will block μ at some $x' = (x'_i, x'_{-i})$ where $\mu_{i-1,s_{\lambda'}}(x') = 1$, a contradiction with the strong stability of μ.■

Construction of Example 4. It is easy to see that any weak* (posterior) stable allocation

μ must be degenerate (this follows the logic of the first step in the proof of Proposition 1). At any weak* stable allocation $\mu_{2,s}(x_1, H) = 1$ for all $x_1 \in \mathbf{X}$. If not, it is easy to see that school 2 will block with the student at any such $x = (x_1, H)$.

Next, $\mu_{i,s}(L, L) = 0$ for $i \in \mathbf{I}_+$. For suppose not, and let $\mu_{1,s}(L, L) = 1$. Then, the coalition $\mathbf{I}' = \{0, 1, 2\}$ and $\mathbf{S}' = \{0\}$ will block at $x = (L, L)$ with the proposal $\mu'_{i,s} = 0$ for all $i \in \mathbf{I}_+$, and $\alpha_1(x'_1, m_1) = 1$ iff $x'_1 = L$, all $m_1 \in \mathbf{M}_1(\mu)$, while $\alpha_2(x'_2, m_2) = 1$ iff $x'_2 = L$, all $m_2 \in \mathbf{M}_2(\mu)$, or $x'_2 = M$, $m_2 = 1$. To verify that $\alpha_{\mathbf{C}}$ satisfies (5), we only need to consider combinations x_i, m_i which are consistent with μ, $piv_i(\alpha_{\mathbf{C}})$, since otherwise payoffs are arbitrary and α_i is obviously consistent with (5). First, let $i = 1$, and $m_1 = 1$ –hence $m_2 = 0$. Given α_2, $x' \in piv_1(\alpha_{\mathbf{C}})$ and $m_1 = 1$ imply $x'_2 = L$. If $x'_1 = L$, then $U_1(\mu' - \mu | x'_1, m_1, piv_1(\alpha_{\mathbf{C}})) = -u_1(s; L, L) > 0$ by (i), and $\alpha_1(x_1, \mu_1(x)) = 1$, with strict acceptance. If $x'_1 = H$, $U_1(\mu' - \mu | x'_1, m_1, piv_1(\alpha_{\mathbf{C}}))$ is either arbitrary or $-u_1(s; H, L) < 0$, by (i), consistent with $\alpha_1(H, m_1) = 0$. Now let $m_1 = 0$. Given α_2, $x' \in piv_1(\alpha_{\mathbf{C}})$ and $m_1 = 0$ imply $x'_2 < H$. If $x'_1 = L$, then $x'_2 = M$. Now $U_1(\mu' - \mu | x'_1, m_1, piv_1(\alpha_{\mathbf{C}})) = 0$, so $\alpha_1(L, m_1) = 1$ is consistent with (5). If $x'_1 > L$, $U_1(\mu' - \mu | x'_1, m_1, piv_1(\alpha_{\mathbf{C}})) = 0$, or it is arbitrary, and this is consistent with $\alpha_1(x'_1, m_1) = 0$. For $i = 2$, given α_1, $x' \in piv_2(\alpha_{\mathbf{C}})$ implies $x'_1 = L$. Let $m_2 = 0$. Then, when $x'_2 = L$, $U_2(\mu' - \mu | x'_2, m_2, piv_2(\alpha_{\mathbf{C}})) = 0$, and $\alpha_2(x_2, \mu_2(x)) = 1$. When $x'_2 > L$, school 2 is either indifferent between dropping the student (when $\mu_{2,s}(L, x'_2) = 0$), or strictly prefers to reject it given μ, such as at $x = (L, H)$, by (iii). Finally, let $m_2 = 1$. Then, when $x'_2 = H$, $U_2(\mu' - \mu | x'_2, m_2, piv_2(\alpha_{\mathbf{C}})) = -u_2(s; L, H) < 0$ by (iii); if $x'_2 = M$, $U_2(\mu' - \mu | x'_2, m_2, piv_2(\alpha_{\mathbf{C}})) = -u_2(s; L, M) > 0$, and $\alpha_2(x'_2, m_2) = 1$ is consistent with (5). So $\alpha_{\mathbf{C}}$ satisfies (5), and μ is weak* blocked. A similar block can be successful if $\mu_{2,s}(L, L) = 1$.

Observe now that $\mu_{i,s}(L, M) = 0$ for $i \in \mathbf{I}_+$. Otherwise, suppose that $\mu_{1,s}(L, M) = 1$. Then, $i = 1$ will block alone by (i). And if $\mu_{2,s}(L, M) = 1$, it can be verified that the coalition $\mathbf{I}' = \{0, 1, 2\}$ and $\mathbf{S}' = \{0\}$ will block at $x = (L, M)$ with the proposal $\mu'_{i,s} = 0$ for all $i \in \mathbf{I}_+$, and the acceptance rule: $\alpha_1(x'_1, m_1) = 1$ iff $x'_1 = L$, all $m_1 \in \mathbf{M}_1(\mu)$, $\alpha_2(x'_2, m_2) = 1$ iff $x'_2 = M$, all $m_2 \in \mathbf{M}_2(\mu)$.

Note next that $\mu_{2,s}(M, L) = 1$ is blocked by $i = 2$ alone, either by (iii) (if $\mu_{2,s}(H, L) = 0$), or by (iv) (if $\mu_{2,s}(H, L) = 1$), so that $\mu_{0,s}(M, L) = 1$. Also notice that $\mu_{2,s}(H, M) = 1$, or otherwise school 2 will block alone, by (iii) if $\mu_{2,s}(M, M) = 1$, and by (v) if $\mu_{2,s}(M, M) = 0$. Furthermore, we claim that $\mu_{0,s}(M, M) = 0$. If not, one can check that the coalition $\mathbf{I}' = \{0, 1, 2\}$ and $\mathbf{S}' = \{0, s\}$ will block at $x = (M, M)$ with the proposal $\mu'_{1,s} = 1$, and $\alpha_1(x'_1, m_1) = 1$ iff $x'_1 > L$, $\alpha_2(x'_2, m_2) = 1$ iff $x'_2 = M$ and $m_2 = 0$, and $\alpha_s(m_s) = 1$ iff $m_s = m_s^0$.

Finally, suppose that $\mu_{2,s}(M, M) = 1$. Then, it must be that $\mu_{2,s}(H, M) = 1$, or again by (iii) $i = 2$ will block with the student. Then, the coalition $\mathbf{I}' = \{0, 1, 2\}$ and $\mathbf{S}' = \{0\}$ will block at $x = (M, M)$ with the proposal $\mu'_{i,s} = 0$ for all $i \in \mathbf{I}_+$, and $\alpha_{\mathbf{C}}$ such that $\alpha_i(x'_i, m_i) = 1$ iff $x_i = M$, $i \in \mathbf{I}'_+$. As earlier, it can be checked that this block will satisfy (5). Then, $\mu_{1,s}(M, M) = 1$ at a weak* stable allocation, and we are left with the allocations whose ι_μ are

$$\left\{\begin{matrix} 2 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{matrix}\right\}, \quad \left\{\begin{matrix} 2 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{matrix}\right\}, \quad \left\{\begin{matrix} 2 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{matrix}\right\} \tag{11}$$

However none of these three rules are incentive compatible: in the first and the second, school 2 has an incentive to lie at $x_2 = L$ and report $x_2 = M$; while in the last one, school 1 lies at $x_1 = H$, reporting instead $x_1 = M$. ■

Construction of Example 5. Again weak** stable allocations must be degenerate. We first establish some preliminary facts.

Fact 0. At any weak** stable allocation, $\mu_{0,s}(x) = 0$. Otherwise, school 1 and the student will block.

Fact 1. $\mu_{3,s}(x_1, x_2, L) = 1$ iff $\{x_2 = H, \text{ all } x_1\}$, or $\{(x_1, x_2) = (H, H)\}$.

Indeed, it cannot be that: $\mu_{3,s}(x_1, x_2, L) = 1$ all $(x_1, x_2) \in \mathbf{X}^2$, by (vii); $\mu_{3,s}(x_1, x_2, L) = 0$ for a unique $(x_1, x_2) \in \mathbf{X}^2$, $(x_1, x_2) \neq (H, H)$, by (vii).

It also cannot be that $\mu_{3,s}(x_1, L, L) = 1$ all x_1, or $\mu_{3,s}(L, x_2, L) = 1$, all x_2, by (vi). It cannot be that $\mu_{3,s}(H, x_2, L) = 1$ all x_2, by (viii), and any other combination of three's which is dominated by affiliation of the environment. Finally, it cannot be that $\mu_{3,s}(x_1, x_2, L) = 1$ for a unique $(x_1, x_2) \in \mathbf{X}^2$, $(x_1, x_2) \neq (H, H)$, by (vi). All these allocations will be blocked by 3 alone.

Fact 2. It cannot be that $\mu_{1,s}(x_1, H, L) = 1$ all x_1, or school 2 will block with the student, by (iv).

Fact 3. It cannot be that $\mu_{2,s}(L, H, L) = 1$ and $\mu_{2,s}(H, H, L) = \mu_{2,s}(x_1, H, H) = 0$ all x_1, or 2 blocks alone at (L, H, L), by (iii).

Fact 4. It cannot be that $\mu_{2,s}(x_1', L, L) = 1$ any x_1', and $\mu_{2,s}(x_1, L, H) = 0$ all x_1, or 2 blocks alone at (x_1', L, L), by (iii).

Fact 5. If $\mu_{3,s}(H, L, H) = 0$, then $\mu_{2,s}(H, L, H) = 1$. For if not, either $\mu_{2,s}(x_1, L, x_3) = 0$ for all x_1, x_3, but using Fact 1 this will be blocked by 2 with the student, by (v); or $\mu_{2,s}(x_1, L, x_3) = 1$ for some x_1, x_3, but this μ will be blocked by 2 alone, from (iii).

We distinguish several cases, depending on the assignment of the student to school 3 when $x_3 = H$. To highlight the role played by lotteries, and for future use, we also first eliminate assignments based on incentive compatibility and blocks which do not use lotteries. In a second step, we use lottery blocks to eliminate the remaining possible configurations.

Step 1

Case (a): $\mu_{3,s}(x_1, x_2, H) = 1$, all $(x_1, x_2) \in \mathbf{X}^2$ —notice how this is possible given (xi). Then, $\mu_{1,s}(x_1, L, L) = 1$ for all x_1. Suppose not, then $\mu_{3,s}(x_1, L, L) = 0$ by Fact 1. It cannot be that $\mu_{2,s}(x_1, L, L) = 1$, by Fact 4. It cannot be that $\mu_{1,s}(x_1, H, L) = 1$ for all x_1, by Fact 2. Invoking Fact 3, the remaining possibilities for $x_3 = L$ aside for $\mu_{2,s}(x_1, H, L) = 1$ for all x_1 —which will be treated in Step 2— are

$$\left\{ \begin{array}{cc} 1 & 2 \\ 1 & 1 \end{array} \right\}, \left\{ \begin{array}{cc} 1 & 3 \\ 1 & 1 \end{array} \right\}, \left\{ \begin{array}{cc} 3 & 3 \\ 1 & 1 \end{array} \right\}. \tag{12}$$

But the first two are not incentive compatible for $i = 1$ at $x_1 = H$, and the third for $i = 3$ at $x_3 = H$, by (x) and affiliation of the environment.

Case (b): $\mu_{3,s}(x_1, x_2, H) = 0$ for a unique $(x_1, x_2) \in \mathbf{X}^2$. Using (x) and affiliation of the environment, it must be that $\mu_{3,s}(L, L, H) = 0$. Suppose $\mu_{2,s}(L, L, H) = 1$. Then school 2 will block by (iii). Suppose $\mu_{1,s}(L,L,H)=1$. It cannot be that $\mu_{2,s}(x_1,L,L)=1$ for any x_1, by Fact 4. Hence, by Fact 1, $\mu_{1,s}(x_1,L,L)=1$ for all x_1. Fact 2 excludes $\mu_{1,s}(x_1,H,L)=1$ for all x_1. Aside for the allocation where $\mu_{2,s}(x_1,H,L)=1$ for all x_1 —which again will be treated in Step 2— and invoking Facts 1 and 3, the remaining possibilities for $x_3=L$ are the same as (12).

Case (c): $\mu_{3,s}(x_1,x_2,H)=1$ for only a pair of (x_1,x_2), $(x_1,x_2)'\in\mathbf{X}^2$.

Then, either: c.1) $\mu_{3,s}(x_1,H,H)=1$ all x_1; or c.2) $\mu_{3,s}(H,x_2,H)=1$ all x_2. If c.1), $\mu_{3,s}(x_1,L,L)=0$ all x_1, from Fact 1. Then, it cannot be that $\mu_{2,s}(L,L,H)=1$ while $\mu_{2,s}(H,L,H)=0$, by (iii). If $\mu_{1,s}(x_1,L,H)=1$ all x_1, then $\mu_{1,s}(x_1,L,L)=1$ all x_1, by Facts 1 and 4. It must that $\mu(2,s;H,L,H)=1$, by Fact 5. Then, the remaining possibilities are of the form

$$\left\{\begin{array}{cc}(1,2,3) & (1,2,3)\\ (1,2) & (1,2)\end{array}\right\}\left\{\begin{array}{cc}3 & 3\\ (1,2) & 2\end{array}\right\} \tag{13}$$

where we list ι_μ as pairs of matrices $\{\iota_\mu(s;x_{-3},L)\},\{\iota_\mu(s;x_{-3},H)\}$, and multiple entries in parenthesis $(\cdot)$ mean that at this stage μ can be constructed using any of those entries in that cell, independently of those in others. Now, $\mu_{3,s}(H,H,L)=1$, or μ is not incentive compatible for 3 at $x_3=L$, by (ix). Then, $\mu_{2,s}(L,H,L)=0$, or 2 will block alone (see Fact 3). But then μ is not incentive compatible for 2 at $x_2=H$.

If c.2), since again by Fact 1 $\mu(3,s;x_1,L,x_3)=0$ all $(x_1,x_3)\neq(H,H)$, then $\mu_{1,s}(x_1,L,x_3)=1$ all $(x_1,x_3)\neq(H,H)$, by Fact 4. Suppose that $\mu_{2,s}(L,H,L)=\mu_{2,s}(L,H,H)=1$, and $\mu_{1,s}(H,H,L)=1$. Then, 2 will block with the student at $x=(H,H,L)$, by (iii). Therefore the remaining possibilities are of the form

$$\left\{\begin{array}{cc}(1,2,3) & (1,2,3)\\ 1 & 1\end{array}\right\}\left\{\begin{array}{cc}(1,2) & 3\\ 1 & 3\end{array}\right\} \tag{14}$$

and these are not incentive compatible for 1 at $x_1=H$.

Case (d): $\mu_{3,s}(x_1,x_2,H)=1$ for a unique $(x_1,x_2)\in\mathbf{X}^2$. Then, it must be $\mu_{3,s}(L,L,H)=0$, by (i). Three are the possibilities: d.1) $\mu_{3,s}(H,H,H)=1$; d.2) $\mu_{3,s}(H,L,H)=1$; d.3) $\mu_{3,s}(L,H,H)=1$.

d.1) Again, it must that $\mu_{2,s}(H,L,H)=1$, by Fact 5, and $\mu_{3,s}(x_1,L,L)=0$ all x_1, from Fact 1. Also, $\mu_{2,s}(L,H,H)=1$, or 2 will block with the student, by (iv). Now, $\mu_{3,s}(H,H,L)=1$, or μ is not incentive compatible for 3. Then, $\mu_{3,s}(L,H,L)=0$, or again μ is not incentive compatible for 3. Assuming that $\mu_{2,s}(L,L,H)=1$ —which will be proved in Step 2— what is left is of the form

$$\left\{\begin{array}{cc}(1,2) & 3\\ (1,2) & (1,2)\end{array}\right\}\left\{\begin{array}{cc}2 & 3\\ 2 & 2\end{array}\right\} \tag{15}$$

These are not incentive compatible for 2 at $x_2=H$, because they violate (iv) and affiliation of the environment in the least favorable case for lying, when $\mu_{2,s}(L,L,L)=1$, and $\mu_{2,s}(L,H,L)=\mu_{2,s}(H,L,L)=0$.

d.2) Using Fact 1, and (iii), it must be that $\mu_{1,s}(x_1,L,x_3)=1$ all $(x_1,x_3)\neq(H,H)$, or 2 will block alone. Then again, $\mu_{2,s}(x_1,H,H)=1$ all x_1, or 2 will block with the student, by (iv) and affiliation of the environment. The remaining possibilities are of the form

$$\left\{ \begin{array}{cc} (1,2,3) & (1,2,3) \\ 1 & 1 \end{array} \right\} \left\{ \begin{array}{cc} 2 & 2 \\ 1 & 3 \end{array} \right\} \tag{16}$$

which are not incentive compatible for 2 at $x_2 = L$.

d.3) From Fact 5, $\mu_{2,s}(H, L, H) = 1$, and $\mu_{2,s}(H, H, H) = 1$, or 2 will block with the student by (iv) and affiliation of the environment. Next, $\mu_{3,s}(x_1, H, L) = 0$ for some x_1, or μ will not be incentive compatible for 3 at $x_3 = H$. By Fact 1, $\mu_{3,s}(x_1, x_2, L) = 1$ implies $(x_1, x_2) = (H, H)$. Then, assuming $\mu_{2,s}(L, L, H) = 1$ —which will be proved in Step 2— we look at μ when $x_3 = L$, whether $\mu_{3,s}(H, H, L) = 0$ or not.

Suppose that $\mu_{3,s}(H, H, L) = 0$. Then, $\mu_{2,s}(H, H, L) = 1$. For if $\mu_{1,s}(H, H, L) = 1$, then 2 will block with the student at $x = (H, H, L)$. Now, it must be that $\#\{x_2|\mu_{1,s}(H, x_2, L) = 1\} = \#\{x_2|\mu_{1,s}(L, x_2, L) = 1\}$. If not, by the previous step we know that $\#\{x_2|\mu_{1,s}(H, x_2, L) = 1\} < \#\{x_2|\mu_{1,s}(L, x_2, L) = 1\}$, but then school 1 will lie, using (i). As a result, if $\mu_{1,s}(H, L, L) = 1$, then $\mu_{1,s}(L, L, L) = 1$ and $\mu_{2,s}(L, H, L) = 1$, by (ii) —which cannot occur, as shown below in Step 2. Otherwise, $\mu_{2,s}(H, L, L) = 1$, and then $\mu_{2,s}(x_1, x_2, L) = 1$ for all x_1, x_2, and ι_μ is

$$\left\{ \begin{array}{cc} 2 & 2 \\ 2 & 2 \end{array} \right\} \left\{ \begin{array}{cc} 3 & 2 \\ 2 & 2 \end{array} \right\} \tag{17}$$

which is not incentive compatible since school 2 will lie at $x_2 = H$, by (iii).

Suppose finally that $\mu_{3,s}(H, H, L) = 1$. Still, $\#\{x_2|\mu_{1,s}(H, x_2, L) = 1\} = \#\{x_2|\mu_{1,s}(L, x_2, L) = 1\}$. If $\mu_{1,s}(H, L, L) = 1$, then $\mu_{1,s}(L, L, L) = 1$ and $\mu_{2,s}(L, H, L) = 1$, by (ii), and ι_μ is

$$\left\{ \begin{array}{cc} 2 & 3 \\ 1 & 1 \end{array} \right\} \left\{ \begin{array}{cc} 3 & 2 \\ 2 & 2 \end{array} \right\} \tag{18}$$

which is not incentive compatible since school 2 will lie at $x_2 = H$, by (iii). While if $\mu_{2,s}(H, L, L) = 1$, then $\mu_{2,s}(x_1, x_2, L) = 1$ for all $(x_1, x_2) \neq (H, H)$, so that ι_μ is

$$\left\{ \begin{array}{cc} 2 & 3 \\ 2 & 2 \end{array} \right\} \left\{ \begin{array}{cc} 3 & 2 \\ 2 & 2 \end{array} \right\} \tag{19}$$

which is also not incentive compatible since school 2 will lie by (iii).

Step 2

When $\mu_{3,s}(x_1, x_2, H) = 1$, all $(x_1, x_2) \in \mathbf{X}^2$, and when $\mu_{3,s}(x_1, x_2, H) = 0$ for a unique $(x_1, x_2) \in \mathbf{X}^2$ —and then $\mu_{1,s}(L, L, H) = 1$— we need to rule out the allocation which has $\mu_{2,s}(x_1, H, L) = 1$ for all x_1. The coalition with $\mathbf{I}' = \{0, 1, 3\}$ and $\mathbf{S}' = \{0, s\}$ will block this μ at $x = (x_1, H, L)$ with $\mu'_{1,s} = p = 1 - \mu'_{3,s}$ where p is such that $pv_1 + (1-p)v_3 > v_2$, and $\alpha_s(m_s) = 1$ all $m_s \in \mathbf{M}_s(\mu)$ with $m_s \neq m_s^3$, $\alpha_1(x'_1, m_1) = 1$ iff $m_1 = 0$, and $\alpha_3(x'_3, m_3) = 1$ iff $x'_3 = L$, $m_3 = 0$. To verify (3) for $i = 1$, observe that $x' \in piv_1(\alpha_\mathbf{C})$ implies $x'_3 = L$, $m_3 = 0$. At any x'_1, $U_1(\mu' - \mu|x'_1, m_1, piv_1(\alpha_\mathbf{C})) = u_1(s|x'_3 = L)(p - m_1) > 0$ iff $m_1 = 0$, therefore α_1 satisfies (3), and $\alpha_1(x_1, \mu_1(x)) = 1$. For $i = 3$,

$x' \in piv_3(\alpha_{\mathbf{C}})$, $m_3 = 0$ and $x'_3 = L$ is equivalent to $(x'_2, x'_3) = (H, L)$. So at $x'_3 = L$, $m_3 = 0$, $U_3(\mu' - \mu | x'_3, m_3, piv_3(\alpha_{\mathbf{C}})) = u_3(s | x'_2 = H, x'_3 = L)(1-p) > 0$ by (ix), and α_3 verifies (3), and $\alpha_3(x_3, \mu_3(x)) = 1$. By choice of p, α_s immediately verifies (4).

Next, suppose that $\mu_{3,s}(H,H,H) = 1$ is the unique (x_1, x_2) where school 3 gets the student when $x_3 = H$ —and then, as argued in Step 1, $\mu_{3,s}(L,L,H) = 0$, and $\mu_{2,s}(H,L,H) = \mu_{2,s}(L,H,H) = 1$— but that $\mu_{2,s}(L,L,H) = 0$. Then the coalition with $\mathbf{I}' = \{0,1,3\}$ and the student will block at $x = (H,L,H)$ with $\mu'_{1,s} = p = 1 - \mu'_{3,s}$ where p is as above, and $\alpha_s(m_s) = 1$ all $m_s \in \mathbf{M}_s(\mu)$ with $m_s \neq m_s^3$, $\alpha_1(x'_1, m_1) = 1$ iff $m_1 = 0$, and $\alpha_3(x'_3, m_3) = 1$ iff $x'_3 = H$, $m_3 = 0$. It can be readily verified that such an $\alpha_{\mathbf{C}}$ satisfies (3) and (4).

If instead $\mu_{3,s}(L,H,H) = 1$ is the unique (x_1, x_2) where school 3 gets the student when $x_3 = H$, and as we know $\mu_{3,s}(L,L,H) = 0$, while $\mu_{2,s}(H,L,H) = \mu_{2,s}(H,H,H) = 1$, we need to show first that $\mu_{2,s}(L,L,H) = 1$. Suppose not; then again the coalition $\mathbf{C}$ with $\mathbf{I}' = \{0,1,3\}$ will block at $x = (H,L,H)$ with $\mu'_{1,s} = p = 1 - \mu'_{3,s}$ where p has already been defined, and $\alpha_s(m_s) = 1$ all $m_s \in \mathbf{M}_s(\mu)$ with $m_s \neq m_s^3$, $\alpha_1(x'_1, m_1) = 1$ iff $m_1 = 0$, and $\alpha_3(x'_3, m_3) = 1$ iff $x'_3 = H$, $m_3 = 0$ (if $\mu_{3,s}(H,H,L) = 1$) or $x'_3 = L$, $m_3 = 0$ (if $\mu_{3,s}(H,H,L) = 0$), as can be verified. Finally, we need to show that if $\mu_{2,s}(H,H,L) = 1$, we cannot have $\mu_{1,s}(H,L,L) = 1$, and then $\mu_{1,s}(L,L,L) = \mu_{2,s}(L,H,L) = 1$. Indeed, this is not stable, as again the same coalition with $\mathbf{I}' = \{0,1,3\}$ and the student will block at $x = (H,L,H)$ with $\mu'_{1,s} = p = 1 - \mu'_{3,s}$ where p is as above, and appropriate strategies $\alpha_{\mathbf{C}}$.■

Construction of Example 6. For each student s_h, $h = 1,2$ we can apply Step 1 in the construction of Example 5 to rule out the corresponding allocations as being weak*** posterior stable and incentive compatible.

For the remaining allocations excluded in Step 2, consider all possible pairs (μ_{s_1}, μ_{s_2}) resulting from combinations of such allocations for students s_1, s_2. Observe that there exist signals x_{s_1} and x_{s_2} such that, at $x = (x_{s_1}, x_{s_2})$ the coalition of $\mathbf{I}' = \{0,1,3\}$ and $\mathbf{S}' = \{0, s_1, s_2\}$ can block each pair of allocations as follows. Let (μ_{s_1}, μ_{s_2}) be the candidate pair of allocations for weak*** stability and incentive compatibility. Let $\mathbf{\Xi}_i = \mathbf{T}_i$ for $i = 1,3$, and $\mathbf{\Xi}_{s_h} = \{\xi^1_{s_h}, \xi^2_{s_h}\} \cup \mathbf{T}_{s_k}$ where $\xi^1_{s_h}, \xi^2_{s_h}$ are arbitrary messages.

The alternative allocation μ'_{s_h} is set as follows: when student s_h, $h = 1,2$, announces message $\xi^1_{s_h}$, the other student $s_{h'}$, $h' \neq h$, is allocated to school 1 whereas when student s_h announces message $\xi^2_{s_h}$ the other student $s_{h'}$ is allocated to school 3, regardless of the message of other coalition members; for any other message sent by student s_h, student $s_{h'}$ is allocated to school 0. Notice that μ' does not depend on the messages sent by the schools and that the allocation for any student does not depend on the messages sent by that student but only on the messages sent by the other student.

For student s_h, $h = 1,2$, let $\mathbf{A}_{s_h}$ correspond to states when s_h has not been allocated by μ to school 3. For $t_{s_h} \in \mathbf{A}_{s_h}$, let student s_h choose a mixed announcement strategy $\sigma_{s_h}(\xi^1_{s_h} | t_{s_h}) = p = 1 - \sigma_{s_h}(\xi^2_{s_h} | t_{s_h})$ where $pv_1 + (1-p)v_3 > v_2$ and otherwise let the announcement strategies for all parties be arbitrary. For $i = 1,3$, let $\mathbf{A}_i$ correspond to the set of types for school i for which

$\alpha_i = 1$, in each of the cases considered in Step 2, Example 5. Observe that with such a choice of μ', $\{\mathbf{A}_k\}_{k \in \mathbf{C}}$ and $\{\boldsymbol{\sigma}_k\}_{k \in \mathbf{C}}$, from the perspective of each student s_h we have recreated the lotteries used Step 2, Example 5 by using the random messages sent by the other student. It then follows that conditions (6) and (7) are satisfied, due to arguments identical to those used in that example. We conclude that no allocation $\mu = (\mu_{s_1}, \mu_{s_2})$ is both weak*** stable and incentive compatible.■

References

[1] Abdulkadiroglu, A., and Sonmez, T., (1999) House Allocation with Existing Tenants, Journal of Economic Theory, v. 88, 233-60.

[2] Abdulkadiroglu, A., and Sonmez, T., (2003) School Choice: A Mechanism Design Approach, American Economic Review, v. 93, 729-47.

[3] Dutta, B., and Vohra, R., (2001) Incomplete Information, Credibility and the Core, W.P. 2001-02, Brown University. Forthcoming, Mathematical Social Sciences.

[4] Forges, F., (1994) Posterior Efficiency, Games and Economic Behavior, v. 6, 238-61.

[5] Forges, F., Mertens, J.F., and Vohra, R., (2001) The Ex Ante Incentive Compatible Core in the Absence of Wealth Effects, Econometrica, v. 70, 1865-92.

[6] Forges, F., Minelli, E., and Vohra, R., (2001) Incentives and the Core of an Exchange Economy: A Survey, Journal of Mathematical Economics, v. 38, 1-41.

[7] Holmstrom, B., and Myerson, R., (1983) Efficient and Durable Decision Rules with Incomplete Information, Econometrica, v.51, 1799-1820.

[8] Maskin, E., and Moore, J., (1999) Implementation and Renegotiation, Review of Economic Studies, v. 66, 39-56.

[9] Milgrom, P., and Stokey, N., (1982) Information, Trade and Common Knowledge, Journal of Economic Theory, v. 26, 17-27.

[10] Milgrom, P., and Weber, R., (1982) A Theory of Auctions and Competitive Bidding, Econometrica, v. 50, 1089-1122.

[11] Myerson, R., (1991) Game Theory, Harvard University Press, Cambridge, MA.

[12] Ostrovsky, M., (2005) Two-sided matching with common values, mimeo, Stanford University.

[13] Papai, S., (2000) Strategy-Proof Assignment by Hierarchical Exchange, Econometrica, v. 68, 1403-22.

[14] Roth, A., (1989), Two-Sided Matching with Incomplete Information about Others' Preferences, Games and Economic Behavior, v. 1, 191-209.

[15] Roth, A., and Sotomayor, M., (1990) Two-Sided Matching, Cambridge University Press, Cambridge, MA.

[16] Vohra, R., (1999) Incomplete Information, Incentive Compatibility, and the Core, Journal of Economic Theory, v. 86, 123-47.

[17] Volij, O., (2000) Communication, Credible Improvements, and the Core of an Economy With Asymmetric Information, International Journal of Game Theory, v. 29, 63-79.

[18] Wilson, R., (1978) Information, Efficiency, and the Core of an Economy, Econometrica, v. 46, 807-16.