Matching Markets
under (In)complete Information[*]

Lars Ehlers[†]   Jordi Massó[‡]

November 2014

[†]Département de Sciences Économiques and CIREQ, Université de Montréal, C.P. 6128 Succursale Centre-Ville, Montréal, Québec H3C 3J7, Canada; e-mail: `lars.ehlers@umontreal.ca`; Telephone: +1-514-3437532.

[‡]Universitat Autònoma de Barcelona and Barcelona GSE. Departament d'Economia i Història Econòmica, Edifici B. Campus UAB. 08193 Cerdanyola del Vallès (Barcelona), Spain; e-mail: `jordi.masso@uab.es`

Abstract

We introduce incomplete information to centralized many-to-one matching markets. This is important because in real life markets (i) any agent is uncertain about the other agents' true preferences and (ii) most entry-level matching is many-to-one (and not one-to-one). We show that given a common prior, a strategy profile is an ordinal Bayesian Nash equilibrium under incomplete information in a stable mechanism if and only if, for any true profile in the support of the common prior, the submitted profile is a Nash equilibrium under complete information in the direct preference revelation game induced by the stable mechanism.

1   Introduction

Centralized many-to-one matching markets operate as follows to match the agents from two sides, the firms (colleges, hospitals, schools, etc.) and the workers (students, medical interns, children, etc.): a centralized clearinghouse collects for each participant a ranked list of potential partners and matches via a mechanism firms and workers on the basis of the submitted ranked lists. In applications many of the successful mechanisms are stable.[1][2] The literature has considered stability of a matching (in the sense that all agents have to be matched to acceptable partners and no unmatched pair of a firm and a worker prefer each other rather than the proposed partners) to be its main characteristic in order to survive.[3] This is puzzling because there exists no stable mechanism which makes truth-telling a dominant strategy for all agents (Roth, 1982). Therefore, an agent's (submitted) ranked lists of potential partners are not necessarily his true ones and the implemented matching may not be stable for the true profile. The literature has studied intensively Nash equilibria of direct preference revelation games induced by different stable mechanisms under complete information.[4]

We use the (ordinally) Bayesian approach in many-to-one matching markets by assuming that nature selects a preference profile according to a commonly known probability distribution on the set of profiles (a common prior).[5] Since matching

---

[1]See Roth (1984a), Roth and Peranson (1999), and Roth (2002) for a careful description and analysis of the American entry-level medical market. Roth (1991), Kesten (2005), Ünver (2005), and Ehlers (2008) describe and analyze the equivalent UK markets.

[2]Chen and Sönmez (2006), Ergin and Sönmez (2006), and Abdulkadiroğlu, Che, and Yosuda (2011) study the case of public schools in Boston, Abdulkadiroğlu and Sönmez (2003) studies the cases of public schools in Boston, Lee County (Florida), Minneapolis, and Seattle, and Abdulka-diroğlu, Pathak, and Roth (2005, 2009) study the case of public high schools in New York City.

[3]See, for instance, Roth (1984a) and Niederle and Roth (2003).

[4]See Dubins and Freedman (1981), Roth (1982, 1984b, 1985), Gale and Sotomayor (1985), Shin and Suh (1996), Sönmez (1997), Ma (1995, 2002), and Alcalde (1996).

[5]Roth (1989) is the first paper studying strategic incentives generated by stable mechanisms

markets require to report ranked lists and not their specific utility representations, we stick to the ordinal setting and assume that probability distributions are evaluated according to the first-order stochastic dominance criterion. Then, a strategy profile is an ordinal Bayesian Nash equilibrium (OBNE) if, for every von Neumann Morgenstern (vNM)-utility function of an agent's preference ordering (his type), the submitted ranked list maximizes his expected utility in the direct preference revelation game induced by the common prior and the mechanism.[6] For direct preference revelation games under incomplete information induced by a stable mechanism, our main result shows a link between Nash equilibria under complete information and OBNE under incomplete information. More precisely, Theorem 1 states that, given a common prior, a strategy profile is an OBNE under incomplete information in a stable mechanism if and only if for any profile in the support of the common prior, the submitted profile is a Nash equilibrium under complete information at the true profile in the direct preference revelation game induced by the stable mechanism.

The paper is organized as follows. Section 2 describes the many-to-one matching market with responsive preferences and introduces incomplete information and the notion of ordinal Bayesian Nash equilibrium. Section 3 states our main result, Theorem 1, and its applications. Section 4 discusses variations of our main result and the Appendix contains all proofs.

---

under incomplete information. He shows that for particular vNM-utility representations of the ordinal preferences, Bayesian Nash equilibria under incomplete information may not satisfy appealing properties of Nash equilibria under complete information. Chakraborty, Citanna, and Ostrovsky (2010) study two-sided matching markets with interdependent preferences.

[6]This notion was introduced by d'Aspremont and Peleg (1988) who call it "ordinal Bayesian incentive-compatibility". Majumdar and Sen (2004) use it to relax strategy-proofness in the Gibbard-Satterthwaite Theorem. Majumdar (2003), Pais (2005), and Ehlers and Massó (2007) have already used this ordinal equilibrium notion in one-to-one matching markets.

2 Many-To-One Matching Markets

Let W denote the set of workers, F denote the set of firms, and $V \equiv F \cup W$ denote the set of agents. For each firm f, there is a maximum number $q_f \geq 1$ of workers that f may hire, f's *quota*. Let $q = (q_f)_{f \in F}$ denote the vector of quotas. Each worker w has a strict preference ordering P_w over $F \cup \{\emptyset\}$, where $\emptyset$ stands for being unmatched. Each firm f has a strict preference ordering P_f over $W \cup \{\emptyset\}$, where $\emptyset$ stands for leaving a position unfilled. A *profile* $P = (P_v)_{v \in V}$ is a list of preference orderings. Given $S \subseteq V$, we sometimes write (P_S, P_{-S}) instead of P. Let $\mathcal{P}_v$ be the set of all preference orderings of agent v. Let $\mathcal{P} = \times_{v \in V} \mathcal{P}_v$ be the set of all profiles and let $\mathcal{P}_{-v} = \times_{v' \in V \setminus \{v\}} \mathcal{P}_{v'}$. Let R_v denote the weak preference associated with P_v. Given $w \in W$, $P_w \in \mathcal{P}_w$, and $v \in F \cup \{\emptyset\}$, let $B(v, P_w)$ denote the *weak upper contour set* of P_w at v; *i.e.*, $B(v, P_w) = \{v' \in F \cup \{\emptyset\} \mid v' R_w v\}$. Let $A(P_w)$ be the set of *acceptable* firms for w according to P_w; *i.e.*, $A(P_w) = \{f \in F \mid f P_w \emptyset\}$. Given a subset $S \subseteq F \cup \{\emptyset\}$, let $P_w|S$ denote the restriction of P_w to S. Similarly, given $P_f \in \mathcal{P}_f$, $v \in W \cup \{\emptyset\}$, and $S \subseteq W \cup \{\emptyset\}$, we define $B(v, P_f)$, $A(P_f)$, and $P_f|S$. A *many-to-one matching market (or college admissions problem)* is a quadruple (F, W, q, P). Because F, W and q remain fixed, a problem is simply a profile $P \in \mathcal{P}$. If $q_f = 1$ for all $f \in F$, (F, W, q, P) is called a *one-to-one matching market*.

A *matching* is a function $\mu : V \to 2^V$ satisfying the following: (m1) for all $w \in W$, $\mu(w) \subseteq F$ and $|\mu(w)| \leq 1$; (m2) for all $f \in F$, $\mu(f) \subseteq W$ and $|\mu(f)| \leq q_f$; and (m3) $\mu(w) = \{f\}$ if and only if $w \in \mu(f)$. We will often write $\mu(w) = f$ instead of $\mu(w) = \{f\}$. If $\mu(w) = \emptyset$, we say that w is *unmatched* at μ. If $|\mu(f)| < q_f$, we say that f has $q_f - |\mu(f)|$ *unfilled* positions at μ. Let $\mathcal{M}$ denote the set of all matchings. Given $P \in \mathcal{P}$ and $\mu \in \mathcal{P}$, μ is *stable (at P)* if (s1) for all $v \in V$, $\mu(v) \subseteq A(P_v)$ (*individual rationality*); and (s2) there exists no pair $(w, f) \in W \times F$ such that $f P_w \mu(w)$ and either $[w P_f \emptyset$ and $|\mu(f)| < q_f]$ or $[w P_f w'$ for some $w' \in \mu(f)]$ (*pairwise stability*). Let $C(P)$ denote the set of stable matchings at P (or the core of P). A (direct) mechanism is a function $\varphi : \mathcal{P} \to \mathcal{M}$. A mechanism φ is *stable* if for all $P \in \mathcal{P}$, $\varphi[P]$ is stable at

P. The most popular stable mechanisms are the *deferred-acceptance algorithms (DA-algorithms)* (Gale and Shapley, 1962): the firm-proposing DA-algorithm is denoted by DA_F and the worker-proposing DA-algorithm is denoted by DA_W.

A mechanism matches each firm f to a *set* of workers, taking into account only f's preference ordering P_f over individual workers. To study firms' incentives, preference orderings of firms over individual workers have to be extended to preference orderings over subsets of workers. The preference extension P_f^* over 2^W is *responsive* to $P_f \in \mathcal{P}_f$ if for all $S \in 2^W$, all $w \in S$, and all $w' \notin S$: (r1) $S \cup \{w'\} P_f^* S \Leftrightarrow |S| < q_f$ and $w' P_f \emptyset$; and (r2) $(S \backslash \{w\}) \cup \{w'\} P_f^* S \Leftrightarrow w' P_f w$. Let R_f^* denote the weak preference associated with P_f^* and $resp(P_f)$ denote the set of all responsive extensions of P_f. Moreover, given $S \in 2^W$, let $B(S, P_f^*)$ be the *weak upper contour set of P_f^* at S*; *i.e.*, $B(S, P_f^*) = \{S' \in 2^W \mid S' R_f^* S\}$.

Any mechanism and any true profile define a direct (ordinal) preference revelation game under complete information for which we can define the natural (ordinal) notion of Nash equilibrium. We denote such a game by $(\mathcal{P}, \varphi, P)$ where P is the true profile, φ is a mechanism and any agent v's set of strategies is $\mathcal{P}_v$. Given a mechanism φ and $P, P' \in \mathcal{P}$, P' is a *Nash equilibrium (NE) in the mechanism φ under complete information P* if (n1) for all $w \in W$, $\varphi[P'](w) R_w \varphi[\hat{P}_w, P'_{-w}](w)$ for all $\hat{P}_w \in \mathcal{P}_w$; and (n2) for all $f \in F$ and all $P_f^* \in resp(P_f)$, $\varphi[P'](f) R_f^* \varphi[\hat{P}_f, P'_{-f}](f)$ for all $\hat{P}_f \in \mathcal{P}_f$. Truth-telling is a NE in φ under P if P is a NE in φ under P.

A *common prior* is a probability distribution $\tilde{P}$ over $\mathcal{P}$. Given $P \in \mathcal{P}$, let $\Pr\{\tilde{P} = P\}$ denote the probability that $\tilde{P}$ assigns to P. Given $v \in V$, let $\tilde{P}_v$ denote the marginal distribution of $\tilde{P}$ over $\mathcal{P}_v$. Given a common prior $\tilde{P}$ and $P_v \in \mathcal{P}_v$, let $\tilde{P}_{-v}|_{P_v}$ denote the probability distribution which $\tilde{P}$ induces over $\mathcal{P}_{-v}$ conditional on P_v. It describes agent v's (Bayesian) uncertainty about the preferences of the other agents, given that his preference ordering is P_v.[7]

---

[7] This formulation does not require symmetry nor independence of priors; conditional priors might be very correlated if agents use similar sources to form them (*i.e.*, rankings, grades, recommendation letters, etc.).

A random matching $\tilde{\eta}$ is a probability distribution over the set of matchings $\mathcal{M}$. Given $\mu \in \mathcal{M}$, let $\Pr\{\tilde{\eta} = \mu\}$ denote the probability that $\tilde{\eta}$ assigns to μ. Let $\tilde{\eta}(w)$ denote the distribution which $\tilde{\eta}$ induces over w's set of potential partners $F \cup \{\emptyset\}$, and let $\tilde{\eta}(f)$ denote the distribution which $\tilde{\eta}$ induces over f's set of potential partners 2^W. Given two random matchings $\tilde{\eta}$ and $\tilde{\eta}'$, (fo1) for $w \in W$ and $P_w \in \mathcal{P}_w$ we say that $\tilde{\eta}(w)$ *first-order stochastically* P_w-*dominates* $\tilde{\eta}'(w)$, denoted by $\tilde{\eta}(w) P_w^{sd} \tilde{\eta}'(w)$, if for all $v \in F \cup \{\emptyset\}$, $\sum_{v' \in F \cup \{\emptyset\}: v' R_w v} \Pr\{\tilde{\eta}(w) = v'\} \geq \sum_{v' \in F \cup \{\emptyset\}: v' R_w v} \Pr\{\tilde{\eta}'(w) = v'\}$; and (fo2) for $f \in F$ and $P_f \in \mathcal{P}_f$, $\tilde{\eta}(f)$ *first-order stochastically* P_f-*dominates* $\tilde{\eta}'(f)$, denoted by $\tilde{\eta}(f) P_f^{sd} \tilde{\eta}'(f)$, if for all $P_f^* \in resp(P_f)$ and all $S \in 2^W$,[8] $\sum_{S' \in 2^W: S' R_f^* S} \Pr\{\tilde{\eta}(f) = S'\} \geq \sum_{S' \in 2^W: S' R_f^* S} \Pr\{\tilde{\eta}'(f) = S'\}$.[9]

A mechanism φ and a common prior $\tilde{P}$ define a direct (ordinal) preference revelation game under incomplete information. A strategy of agent v is a function $s_v : \mathcal{P}_v \to \mathcal{P}_v$ specifying for each type P_v of v a list that v submits to the mechanism, $s_v(P_v)$. We restrict our analysis to pure strategies in the main text. The Appendix generalizes our main result to mixed strategies and random mechanisms. A *strategy profile* is a list $s = (s_v)_{v \in V}$ of strategies specifying for each true profile P a submitted profile $s(P)$. Given a mechanism $\varphi : \mathcal{P} \to \mathcal{M}$ and a common prior $\tilde{P}$ over $\mathcal{P}$, a strategy profile $s : \mathcal{P} \to \mathcal{P}$ induces a random matching $\varphi[s(\tilde{P})]$ in the following way: for all $\mu \in \mathcal{M}$,

$$\Pr\{\varphi[s(\tilde{P})] = \mu\} = \sum_{P \in \mathcal{P}: \varphi[s(P)] = \mu} \Pr\{\tilde{P} = P\}.$$

Using Bayesian updating, the relevant random matching for agent v, given his type P_v and a strategy profile s, is $\varphi[s_v(P_v), s_{-v}(\tilde{P}_{-v}|P_v)]$ (where $s_{-v}(\tilde{P}_{-v}|P_v)$ is the probability distribution over $\mathcal{P}_{-v}$ which s_{-v} and $\tilde{P}$ induce conditional on P_v).

---

[8] Observe that this definition requires that $\tilde{\eta}$ first-order stochastically dominates $\tilde{\eta}'$ according to *all* responsive extensions of P_f.

[9] It is well-known that (fo1) is equivalent to that for any vNM-representation of P_w the expected utility of $\tilde{\eta}$ is greater than or equal to the expected utility of $\tilde{\eta}'$ (and similarly for (fo2) and all vNM-representations of any responsive extension of P_f). See for instance, Theorem 3.11 in d'Aspremont and Peleg (1988).

Definition 4 (Ordinal Bayesian Nash Equilibrium) Let $\tilde{P}$ be a common prior. A strategy profile s is an *ordinal Bayesian Nash equilibrium* (*OBNE*) in the mechanism φ under incomplete information $\tilde{P}$ if and only if for all $v \in V$ and all $P_v \in \mathcal{P}_v$ such that $\Pr\{\tilde{P}_v = P_v\} > 0$,[10]

$$\varphi[s_v(P_v), s_{-v}(\tilde{P}_{-v}|_{P_v})](v) P_v^{sd} \varphi[P'_v, s_{-v}(\tilde{P}_{-v}|_{P_v})](v) \quad \text{for all } P'_v \in \mathcal{P}_v. \tag{1}$$

Truth-telling is an OBNE in the mechanism φ under incomplete information $\tilde{P}$ if and only if for all $v \in V$ and all $P_v \in \mathcal{P}_v$ such that $\Pr\{\tilde{P}_v = P_v\} > 0$,

$$\varphi[P_v, \tilde{P}_{-v}|_{P_v}](v) P_v^{sd} \varphi[P'_v, \tilde{P}_{-v}|_{P_v}](v) \quad \text{for all } P'_v \in \mathcal{P}_v. \tag{2}$$

In general for arbitrary mechanisms there is no connection between NE under complete information and OBNE under incomplete information. For instance, suppose that the common prior $\tilde{P}^u$ is uniform in the sense that it puts equal probability on all preference profiles. Furthermore, suppose that the mechanism φ matches a worker and a firm if and only if they rank each other as their most preferred choice (and φ leaves all other positions unfilled and all other workers unmatched). Then it is easy to verify that truth-telling is an OBNE in the mechanism φ under the uniform prior $\tilde{P}^u$.[11] However, truth-telling is not always a NE in the mechanism φ under complete information since for some profiles, a firm may rank a worker first and a worker the firm second, and if the worker is unmatched, then she profitably manipulates by moving the firm to the first position of her submitted ranking. Our main result will show that such a disconnection between NE and OBNE is only possible for unstable mechanisms.

---

[10]In the definition of OBNE optimal behavior of agent v is only required for the preferences of v which arise with positive probability under $\tilde{P}$. If $P_v \in \mathcal{P}_v$ is such that $\Pr\{\tilde{P}_v = P_v\} = 0$, then the conditional prior $\tilde{P}_{-v}|_{P_v}$ cannot be derived from $\tilde{P}$. However, we could complete the prior of v in the following way: let $\tilde{P}_{-v}|_{P_v}$ put probability one on a profile where all other agents submit lists which do not contain v.

[11]For any agent v and any $P_v \in \mathcal{P}_v$, $\tilde{P}^u_{-v}|_{P_v}$ is uniform over $\mathcal{P}_{-v}$. For all agents belonging to the opposite side of the market, the probability that she ranks v first is identical. Hence, v cannot do better than submitting the true preference relation.

3   The Main Result and Its Applications

The support of a common prior $\tilde{P}$ is the set of profiles on which $\tilde{P}$ puts positive probability: $P \in \mathcal{P}$ belongs to the support of $\tilde{P}$ if and only if $\Pr\{\tilde{P} = P\} > 0$.

Theorem 1 *Let $\tilde{P}$ be a common prior, s be a strategy profile, and φ be a stable mechanism. Then, s is an OBNE in φ under incomplete information $\tilde{P}$ if and only if for any profile P in the support of $\tilde{P}$, $s(P)$ is a Nash equilibrium in φ under complete information P.*

Theorem 1 has several consequences and applications. One immediate consequence is that for determining whether a strategy profile is an OBNE, we only need to check whether for each realization of the common prior the submitted preference orderings constitute a Nash equilibrium under complete information. This means that the uniquely relevant information for an OBNE is the support of the common prior and no calculations of probabilities are necessary. This consequence is very important for applications because we need to check equilibrium play only for the realized (or observed) profiles. Furthermore, by Theorem 1, we can use properties of NE (under complete information) to deduce characteristics of OBNE.

Observe that, given a common belief $\tilde{P}$, the set of OBNE in a stable mechanism is non-empty. For instance, imagine that the workers and the firms are divided into "local" matching markets as follows: let $(W_f)_{f \in F}$ be a partition of the set of workers (allowing $W_f = \emptyset$ for some firms f) where W_f denotes the set of workers belonging to the "local" market of f. In words, under the following strategy profile, if worker w belongs to the local market of firm f, then w ranks f uniquely acceptable if f is preferred to being unmatched and otherwise w ranks no firm acceptable. Any firm f ranks as acceptable (and in the true order) all workers which both belong to its local market and are acceptable according to its true preference relation. Let the strategy profile s be defined in the following way: (i) for any $w \in W$ and any $P_w \in \mathcal{P}_w$, $A(s_w(P_w)) = \{f\}$ if $f \in A(P_w)$ and $w \in W_f$, and $A(s_w(P_w)) = \emptyset$ otherwise; and (ii)

for all $f \in F$ and all $P_f \in \mathcal{P}_f$, let $A(s_f(P_f)) = A(P_f) \cap W_f$ and $s_f(P_f)|A(P_f) \cap W_f = P_f|A(P_f) \cap W_f$. Then for any stable mechanism and any profile P, $s(P)$ is a NE under complete information. Hence, s is an OBNE in any stable mechanism under any common belief. In the special case where $W_f = W$ for some firm f, firm f has a monopolistic market.

Restricting ourselves to truth-telling (where $s(P) = P$ for all $P \in \mathcal{P}$), Theorem 1 shows that truth-telling is an OBNE in a stable mechanism φ if and only if for any profile in the support of the common belief, truth-telling is a NE in φ under complete information for any profile in the support of the common prior. In other words, in many-to-one matching *a stable mechanism φ is ordinally Bayesian incentive compatible under the common prior $\tilde{P}$ if and only if φ restricted to the support of $\tilde{P}$ is incentive compatible.* Furthermore, it can be easily seen that P is a NE in the stable mechanism φ under complete information P if and only if P is a NE in any stable mechanism under complete information P. All what matters is the stability of the mechanism (and not which specific stable matching is chosen).

While the proof of Theorem 1's (If)-part is straightforward, its (Only if)-part proceeds roughly as follows. If for some profile P in the support of $\tilde{P}$, $s(P)$ does not constitute a NE, then some agent v has a profitably deviation from $s(P)$ under complete information P. Using this deviation we then construct another profitable deviation for v from $s(P)$ under complete information P such that agent v strictly increases the probability of the weak upper contour set (at his type P_v) of assigned partner(s) of the deviation under P. This implies that strategy profile s cannot be an OBNE. The construction and the proof use repeatedly the following peculiarities of stable matchings in many-to-one matching markets (see Appendix A.1): (1) invariance of unmatched agents and unfilled positions: the set of unmatched agents and any firm's number of unfilled positions are the same for all stable matchings; and (2) comparative statics: starting from any many-to-one matching market and its workers-optimal matching, when new workers become available all firms weakly

8

prefer any matching, which is stable for the enlarged market, to the workers-optimal stable matching of the smaller market.

Below we turn to the applications of Theorem 1.

3.1 Application I: Structure of OBNE

By Theorem 1, a strategy profile is an OBNE if and only if the agents play a Nash equilibrium for any profile in the support of the common prior. Therefore, (a) the set of OBNE is identical for any two common priors with equal support and (b) the set of OBNE shrinks if the support of the common prior becomes larger. We state these two facts as Corollary 1 below.

Corollary 1 (Invariance) *Let s be a strategy profile and φ be a stable mechanism.*

(a) *Let $\tilde{P}$ and $\tilde{P}'$ be two common priors with equal support. Then, s is an OBNE in the stable mechanism φ under $\tilde{P}$ if and only if s is an OBNE in the stable mechanism φ under $\tilde{P}'$.*

(b) *Let $\tilde{P}$ and $\tilde{P}'$ be two common priors such that the support of $\tilde{P}'$ is contained in the support of $\tilde{P}$. If s is an OBNE in the stable mechanism φ under $\tilde{P}$, then s is an OBNE in the stable mechanism φ under $\tilde{P}'$.*

Now by (a) of Corollary 1, for stable mechanisms any OBNE is robust to perturbations of the common prior which leave its support unchanged. Therefore, any OBNE remains an equilibrium if agents have different priors with equal support, *i.e.* each agent v may have a private prior $\tilde{P}^v$ but all private priors have identical (or common) support.[12] This consequence is especially important for applications since for many of them, the common prior assumption might be too strong.

---

[12]Then in Definition 4 of OBNE the common prior $\tilde{P}$ is replaced for each agent v by his private prior $\tilde{P}^v$. Theorem 1 and its proof show that, a strategy profile s is an OBNE in a stable mechanism φ under private priors $(\tilde{P}^v)_{v \in V}$ if and only if for all $v \in V$ and any profile P in the support of $\tilde{P}^v$, $s_v(P_v)$ is a best response to $s_{-v}(P_{-v})$ in φ under complete information P. In other words,

By (b) of Corollary 1, the set of OBNE with full support (*i.e.* all common priors which put positive probability on all profiles) is contained in the set of OBNE of any arbitrary common prior (or support). Therefore, any OBNE for a common prior with full support is an OBNE for any arbitrary prior. Hence, such OBNE are invariant with respect to the common prior and remain OBNE if the agents' priors are not necessarily derived from the same common prior (and the "local" markets example is an OBNE in any stable mechanism under any priors).

3.2 Application II: Truth-Telling under Correlated Preferences

In empirical applications the preferences of one side of the market are often perfectly correlated. For example, each firm may rank all workers according to an objective criterion such as their degree of qualifications or each college may rank all students according to their grades. Furthermore, it is common in labor economics or search theory to often assume that all workers have identical preferences over firms.[13] One-sided perfect correlation is an extreme case of interdependence of preferences where an agent's preference may depend on the preferences of the other agents on his side.

We say that a common prior $\tilde{P}$ is *F-correlated* if for any profile P in the support of $\tilde{P}$, all firms have identical preferences.[14] Similarly we say that a prior $\tilde{P}$ is *W-correlated* if for any profile P in the support of $\tilde{P}$, all workers have identical preferences. Theorem 1 also helps us to prove the following result under F-correlated or

---

each agent's strategy s_v chooses a best response to the other reported preferences for any profile belonging to the support of his private prior. If all private priors have equal support, then it follows that a strategy profile s is an OBNE with private priors (with common support) if and only if for any profile P in the common support, $s(P)$ is a Nash equilibrium in the mechanism φ under complete information P.

[13]For instance, Shi (2002) provides a long list of papers on directed search models in labor markets where at least one side of the market is homogenous.

[14]Formally this means for all $f, f' \in F$, $A(P_f) = A(P_{f'})$ and $P_f|W = P_{f'}|W$.

W-correlated common priors.

Proposition 1 *Let $\tilde{P}$ be a common prior.*

(a) *If $\tilde{P}$ is F-correlated or W-correlated, then truth-telling is an OBNE in any stable mechanism under incomplete information $\tilde{P}$.*

(b) *Let s be a strategy profile such that $s_w(P_w) = P_w$ for all $w \in W$ and all $P_w \in \mathcal{P}_w$. If $\tilde{P}$ is W-correlated and s is an OBNE in the stable mechanism DA_W under incomplete information $\tilde{P}$, then for all profiles P in the support of $\tilde{P}$, $DA_W[s(P)]$ is stable with respect to P. The analogous statement is true for the stable mechanism DA_F.*

Although Proposition 1 focuses on completely correlated priors, it is easy to extend it in the following direction. Suppose that each worker has a certain qualification and each firm only offers positions having the same job-specific qualification. Let all firms, which are interested in the same qualification, have identical preferences over all workers possessing this qualification for any realization in the common prior. Then the qualifications segregate the matching market and the conclusions of Proposition 1 apply. For example, each firm may represent a certain department in a hospital and they would like to fill their positions with physicians who studied the medical specialty of their department.

4 Variations

Recall for truth-telling to be an OBNE for a common belief, it must be that for any firm and any of its realized preference over firms, truth-telling first order stochastically dominates submitting any other ranking *for all* responsive extensions of the true ranking. It is natural to ask whether Theorem 1 breaks down when we restrict the set of responsive extensions firms may have.

First, it is easy to see that the proof of Theorem 1 remains true if firms responsive extensions are additive, i.e. where a firm has a numerical value for each worker and the value of a set of workers is the sum of the values of the hired workers.[15]

Second, we show that Theorem 1 depends on firms having responsive extensions which are not monotonic: given $P_f \in \mathcal{P}_f$ and $P_f^* \in resp(P_f)$, P_f^* is *monotonic* if for all $S, S' \in 2^W$ such that $|S'| < |S| \leq q_f$ and $S \subseteq A(P_f)$, we have $S P_f^* S'$. Let $mresp(P_f)$ denote the set of all monotonic responsive extensions of P_f. We will call a strategy profile a *monotonic OBNE* in a mechanism under incomplete information if (fo2) holds for all monotonic responsive extensions of any firm's preference relation over individual workers. In the example below we show that truth-telling is a monotonic OBNE in DA_W, while, for some preference profile P in the support of the common belief, truth-telling is not a Nash equilibrium under complete information P in the direct preference revelation game induced by DA_W.

Example 1 Consider a many-to-one matching market with three firms $F = \{f_1, f_2, f_3\}$ and four workers $W = \{w_1, w_2, w_3, w_4\}$. Firm f_1 has capacity $q_{f_1} = 2$ and firms f_2 and f_3 have capacity $q_{f_2} = q_{f_3} = 1$. Consider the common belief $\tilde{P}$ with $\Pr\{\tilde{P} = P\} = p$ and $\Pr\{\tilde{P} = \bar{P}\} = 1 - p$, where $0 < p < 1/2$, and P and $\bar{P}$ are the following profiles:

P_{f_1}	P_{f_2}	P_{f_3}	P_{w_1}	P_{w_2}	P_{w_3}	P_{w_4}	$\bar{P}_{f_1}$	$\bar{P}_{f_2}$	$\bar{P}_{f_3}$	$\bar{P}_{w_1}$	$\bar{P}_{w_2}$	$\bar{P}_{w_3}$	$\bar{P}_{w_4}$
w_1	w_1	w_3	f_3	f_2	f_1	f_1	w_1	w_2	w_4	f_1	f_2	f_1	f_3
w_2	w_2	w_1	f_1	f_1	f_3	f_2	w_2						
w_3	w_3	w_2	f_2	f_3	f_2	f_3	w_3						
w_4	w_4	w_4					w_4						

Note that $P_{f_1} = \bar{P}_{f_1}$. It is straightforward to verify that both profiles have a singleton

[15]A responsive preference ordering P_f^* is additive if there exists an injective function $g : W \to \mathbb{R}\backslash\{0\}$ such that for all $S, S' \in 2^W$ with $|S| \leq q_f$ and $|S'| \leq q_f$, we have $S P_f^* S' \Leftrightarrow \sum_{w \in S} g(w) > \sum_{w' \in S'} g(w')$. In footnotes we show that any responsive extension in the proof of Theorem 1 can be chosen to be additive.

core and $C(P) = \{\mu\}$ and $C(\bar{P}) = \{\bar{\mu}\}$, where

$$\mu = \begin{pmatrix} f_1 & f_2 & f_3 \\ \{w_3, w_4\} & w_2 & w_1 \end{pmatrix} \quad \text{and} \quad \bar{\mu} = \begin{pmatrix} f_1 & f_2 & f_3 \\ \{w_1, w_3\} & w_2 & w_4 \end{pmatrix}.$$

Let φ be a stable mechanism. Thus, by stability of φ, $\varphi[P] = \mu$ and $\varphi[\bar{P}] = \bar{\mu}$.

Claim 1: Truth-telling is *not* a Nash equilibrium under complete information P in the direct preference revelation game induced by DA_W.

Proof of Claim 1: Let $P'_{f_1} \in \mathcal{P}_{f_1}$ be such that $P'_{f_1} : w_1 w_2 w_4 \emptyset w_3$. Then $C(P'_{f_1}, P_{-f_1}) = \{\mu'\}$ where

$$\mu' = \begin{pmatrix} f_1 & f_2 & f_3 \\ \{w_1, w_4\} & w_2 & w_3 \end{pmatrix}.$$

Hence, by stability of φ, $\varphi[P'_{f_1}, P_{-f_1}] = \mu'$. Obviously, for all responsive extensions $P^*_{f_1}$ of P_f we have $\{w_1, w_4\} P^*_{f_1} \{w_3, w_4\}$, which is equivalent to $\varphi[P'_{f_1}, P_{-f_1}](f_1) P^*_{f_1} \varphi[P](f_1)$. Therefore, truth-telling is not a Nash equilibrium in any stable mechanism φ under complete information P (and profile P belongs to the support of the common belief $\tilde{P}$). $\square$

Claim 2: *Truth-telling is a monotonic OBNE in the stable mechanism DA_W under incomplete information $\tilde{P}$.*

Proof of Claim 2: Note that for all $v \in V \backslash \{f_1\}$, if v observes his preference relation, then v knows whether P was realized or $\bar{P}$ was realized. Since at both of P and $\bar{P}$ the core is a singleton and firms f_2 and f_3 have quota one, it follows from the proof of Theorem 1 in Ehlers and Massó (2007) that v cannot gain by deviating.

Next we consider firm f_1. All arguments except for the last one apply to any stable mechanism φ. Observe that $P_{f_1} = \bar{P}_{f_1}$ and the random matching $\varphi[P_{f_1}, \tilde{P}_{-f_1}|_{P_{f_1}}]$ assigns to f_1 the set $\{w_3, w_4\}$ with probability p and the set $\{w_1, w_3\}$ with probability $1 - p$. Let $P^*_{f_1}$ be a monotonic responsive extension of P_{f_1} and $P''_{f_1} \in \mathcal{P}_{f_1}$. We show that

$$\varphi[P_{f_1}, \tilde{P}_{-f_1}|_{P_{f_1}}](f_1) P^{*sd}_{f_1} \varphi[P''_{f_1}, \tilde{P}_{-f_1}|_{P_{f_1}}](f_1). \tag{3}$$

We distinguish two cases.

Case 1: $|\varphi[P''_{f_1}, P_{-f_1}](f_1)| = 1$ or $|\varphi[P''_{f_1}, \bar{P}_{-f_1}](f_1)| = 1$.

Now if (3) does not hold, then by monotonicity of $P^*_{f_1}$ and the fact that when submitting P_{f_1}, f_1 is assigned the set $\{w_3, w_4\}$ with probability p and the set $\{w_1, w_3\}$ with probability $1 - p$ (where $1 - p > 1/2$), we have that $\varphi[P''_{f_1}, P_{-f_1}](f_1) P^*_{f_1} \{w_1, w_3\}$ or $\varphi[P''_{f_1}, \bar{P}_{-f_1}](f_1) P^*_{f_1} \{w_1, w_3\}$. Obviously, from the definition of $\bar{P}_{-f_1}$, the last is impossible. Thus, $\varphi[P''_{f_1}, P_{-f_1}](f_1) P^*_{f_1} \{w_1, w_3\}$ and, by responsiveness of $P^*_{f_1}$, we have $\varphi[P''_{f_1}, P_{-f_1}](f_1) = \{w_1, w_2\}$. But then, without loss of generality, we would have $DA_F[P''_{f_1}, P_{-f_1}](f_1) = \{w_1, w_2\}$ (because DA_F chooses the most preferred stable matching from the firms' point of view).[16] Since we have $C(P) = \{\mu\}$ and $DA_F[P_{f_1}, P_{-f_1}](f_1) = \{w_3, w_4\}$, this would imply that in the corresponding one-to-one matching problem DA_F is group manipulable by the two copies of f_1 (with each copy gaining strictly), a contradiction to the result of Dubins and Freedman (1981).[17]

Case 2: $|\varphi[P''_{f_1}, P_{-f_1}](f_1)| = 2$ and $|\varphi[P''_{f_1}, \bar{P}_{-f_1}](f_1)| = 2$.

Then by definition of $\bar{P}_{-f_1}$ and $|\varphi[P''_{f_1}, \bar{P}_{-f_1}](f_1)| = 2$, we have $\varphi[P''_{f_1}, \bar{P}_{-f_1}](f_1) = \{w_1, w_3\}$. Thus, by stability of φ, $\{w_1, w_3\} \subseteq A(P''_{f_1})$. We consider two subcases.

Subcase 2.1: For all $\mu'' \in C(P''_{f_1}, P_{-f_1})$, $\mu''(w_4) = \emptyset$.

Then $w_4 \notin A(P''_{f_1})$ and by definition of P_{-f_1} and $w_3 \in A(P''_{f_1})$, $DA_W[P''_{f_1}, P_{-f_1}](f_1) = \{w_3\}$. Then f_1 does not fill all its positions at the workers-optimal matching and by Roth and Sotomayor (1990), f_1 is matched to the same set of workers at all stable matchings. Hence, $\varphi[P''_{f_1}, P_{-f_1}](f_1) = \{w_3\}$ and (3) holds (because when submitting P''_{f_1}, firm f_1 is matched with probability p to w_3 and with probability $1 - p$ to $\{w_1, w_3\}$).

---

[16]If $DA_F[P''_{f_1}, P_{-f_1}](f_1) \neq \{w_1, w_2\}$, then choose P'''_{f_1} such that $A(P'''_{f_1}) = \{w_1, w_2\}$. Then we obtain $DA_F[P'''_{f_1}, P_{-f_1}](f_1) = \{w_1, w_2\}$.

[17]Their result says that in a one-to-one matching market no group of firms can profitably manipulate DA_F at the true profile under complete information (with strict preference holding for all firms belonging to the group).

Subcase 2.2: For some $\mu'' \in C(P''_{f_1}, P_{-f_1})$, $\mu''(w_4) \neq \emptyset$.

Then by definition of P_{-f_1}, $\mu''(w_4) = f_1$; otherwise the pair (w_2, f_2) would block μ'' at (P''_{f_1}, P_{-f_1}) if $\mu''(w_4) = f_2$ and the pair (w_1, f_3) would block μ'' at (P''_{f_1}, P_{-f_1}) if $\mu''(w_4) = f_3$. Thus, by $\{w_3, w_4\} \subseteq A(P''_{f_1})$, $\mu \in C(P''_{f_1}, P_{-f_1})$ and $DA_W[P''_{f_1}, P_{-f_1}] = \mu$. Hence, (3) holds for the stable mechanism DA_W.[18] $\qquad\qquad\qquad\square$

The above example has another important implication: suppose that firms submit preference orderings over sets of workers instead of submitting preference orderings over individual workers only and the common belief is a distribution over profiles where firms' preference orderings are over sets of workers. Now if the common belief puts only positive probability on profiles where all firms' preference orderings are responsive and monotonic, then the above example shows that truth-telling can be an OBNE while not necessarily at all profiles in the support truth-telling is a NE under complete information.

This is partly due to the fact our main result is a statement for any common belief. Once we put certain conditions on the common belief, our main result continues to hold even if firms submit preference orderings over sets of workers. Without going into details, let $\mathcal{P}^*_f$ denote the set of all responsive preference orderings of f over 2^W and $\mathcal{P}^* = (\times_{f \in F} \mathcal{P}^*_f) \times (\times_{w \in W} \mathcal{P}_w)$. Let the common belief $\tilde{P}^*$ on $\mathcal{P}^*$ be such that for all $P^*_{-f} \in \mathcal{P}^*_{-f}$ and all $P_f, P'_f \in \mathcal{P}^*_f$ such that $P_f|_{W \cup \{\emptyset\}} = P'_f|_{W \cup \{\emptyset\}}$, $\Pr\{\tilde{P}^* = (P_f, P^*_{-f})\} > 0 \Leftrightarrow \Pr\{\tilde{P}^* = (P'_f, P^*_{-f})\} > 0$. In words, if whenever the common belief puts positive probability on some profile, then for any firm and any other preference ordering which is responsive to the same ordering, the belief also puts positive probability on the profile where the firm's preference is replaced by this preference ordering (the support of the common belief does not distinguish preference orderings which are

---

[18]Note that Example 1 does not contradict Theorem 1. When considering the non-monotonic extension $P^*_{f_1}$ such that $\{w_1\} P^*_{f_1} \{w_3, w_4\}$, then firm w_1 can gain by submitting the list $\hat{P}_{f_1}$ where worker w_1 is the unique acceptable worker (*i.e.* $A(\hat{P}_{f_1}) = \{w_1\}$). Then we have both $\varphi[\hat{P}_{f_1}, P_{-f_1}](f_1) = \{w_1\}$ and $\varphi[\hat{P}_{f_1}, \bar{P}_{-f_1}](f_1) = \{w_1\}$, which means that truth-telling is not an OBNE in any stable mechanism φ under incomplete information $\tilde{P}$.

responsive to the same ordering over individual workers). The proof of Theorem 1 then shows that truth-telling is an OBNE in the stable mechanism φ if and only if for any profile in the support of $\tilde{P}^*$, truth-telling is a NE in the stable mechanism under complete information.

It would be interesting to identify other economic environments where a similar link between BNE under incomplete information and NE under complete information holds. In those environments the strategic analysis under complete information is essential to undertake the corresponding analysis under incomplete information.

References

[1] A. Abdulkadiroğlu, Y.-K. Che, Y. Yosuda, Resolving conflicting preferences in school choice: the Boston mechanism reconsidered, Amer. Econ. Rev. 101 (2011) 399-410.

[2] A. Abdulkadiroğlu, P. Pathak, A. Roth, The New York City high school match, Amer. Econ. Rev. (Papers and Proceedings) 95 (2005) 364-367.

[3] A. Abdulkadiroğlu, P. Pathak, A. Roth, Strategy-proofness versus efficiency in matching with indifferences: redefining the NYC high school match, Amer. Econ. Rev. 99 (2009) 1954-1978 (2009).

[4] A. Abdulkadiroğlu, T. Sönmez, School choice: a mechanism design approach, Amer. Econ. Rev. 93 (2003) 729-747.

[5] C. d'Aspremont, B. Peleg, Ordinal bayesian incentive compatible representation of committees, Soc. Choice Welfare 5 (1988) 261-280.

[6] J. Alcalde, Implementation of stable solutions to marriage problems, J. Econ. Theory 69 (1996) 240-254.

[7] A. Chakraborty, A. Citanna, M. Ostrovsky, Two-sided matching with interdependent values, J. Econ. Theory 145 (2010) 85-105.

[8] Y. Chen, T. Sönmez, School choice: an experimental study, J. Econ. Theory 127 (2006) 202-231.

[9] V.P. Crawford, Comparative statics in matching markets, J. Econ. Theory 54 (1991) 389-400.

[10] L. Dubins, D. Freedman, Machiavelli and the Gale-Shapley algorithm, Amer. Math. Monthly 88 (1981) 485-494.

[11] L. Ehlers, Truncation strategies in matching markets, Math. Oper. Research 33 (2008) 327-335.

[12] L. Ehlers, J. Massó, Incomplete information and singleton cores in matching markets, J. Econ. Theory 136 (2007) 587-600.

[13] H. Ergin, T. Sönmez, Games of school choice under the Boston mechanism, J. Public Economics 90 (2006) 215-237.

[14] D. Gale, L. Shapley, College admissions and the stability of marriage, Amer. Math. Monthly 69 (1962) 9-15.

[15] D. Gale, M. Sotomayor, Ms. Machiavelli and the stable matching problem, Amer. Math. Monthly 92 (1985) 261-268.

[16] O. Kesten, An advice to the organizers of entry-level labor markets in the United Kingdom, Working Paper (2005).

[17] J. Ma, Stable matchings and rematching-proof equilibria in a two-sided matching market, J. Econ. Theory 66 (1995) 352–369.

[18] J. Ma, Stable matchings and the small core in Nash equilibrium in the college admissions problem, Rev. Econ. Design 7 (2002) 117-134.

[19] D. Majumdar, Ordinally bayesian incentive-compatible stable matchings, Working Paper (2003).

[20] D. Majumdar, A. Sen, Ordinally bayesian incentive-compatible voting rules, Econometrica 72 (2004) 523-540.

[21] M. Niederle, A.E. Roth, Unraveling reduces mobility in a labor market: gastroenterology with and without a centralized match, J. Political Economy 111 (2003) 1342-1352.

[22] J. Pais, Incentives in Random Matching Markets, Ph.D. Thesis, Universitat Autònoma de Barcelona, 2005.

[23] J. Pais, Random matching in the college admissions problem, Econ. Theory 35 (2008) 99-116.

[24] A.E. Roth, The economics of matching: stability and incentives, Math. Oper. Research 7 (1982) 617-628.

[25] A.E. Roth, The evolution of the labor market for medical interns and residents: a case study in game theory, J. Political Economy 92 (1984a) 991-1016.

[26] A.E. Roth, Misrepresentation and stability in the marriage problem, J. Econ. Theory 34 (1984b) 383-387.

[27] A.E. Roth, The college admissions problem is not equivalent to the marriage problem, J. Econ. Theory 36 (1985) 277-288.

[28] A.E. Roth, Two-sided matching with incomplete information about others' preferences, Games Econ. Behav. 1 (1989) 191-209.

[29] A.E. Roth, A natural experiment in the organization of entry level labor markets: regional markets for new physicians and surgeons in the U.K., Amer. Econ. Rev. 81 (1991) 415-440.

[30] A.E. Roth, The economist as engineer: game theory, experimentation, and computation as tools for design economics, Econometrica 70 (2002) 1341-1378.

[31] A.E. Roth, E. Peranson, The redesign of the matching market for American physicians: some engineering aspects of economic design, Amer. Econ. Rev. 89 (1999) 748-780.

[32] A.E. Roth, M. Sotomayor, The college admissions problem revisited, Econometrica 57 (1989) 559-570.

[33] A.E. Roth, M. Sotomayor, Two-sided Matching: A Study in Game-Theoretic Modelling and Analysis, Cambridge University Press, Cambridge, England, 1990.

[34] S. Shi, Directed search model of inequality with heterogenous skills and skill-biased technology, Rev. Econ. Stud. 69 (2002) 467-491.

[35] S. Shin, S.-C. Suh, A mechanism implementing the stable rule in marriage problems, Econ. Letters 51 (1996) 185-189.

[36] T. Sönmez, Games of manipulation in marriage problems, Games Econ. Behav. 20 (1997) 169-176.

[37] U. Ünver, On the survival of some unstable two-sided matching mechanisms, Int. J. Game Theory 33 (2005) 239-254.

Before we prove Theorem 1, we recall the following properties of the core of a many-to-one matching market. These properties will be used frequently in the proof. It will be convenient to write $(F, W, P; q)$ for any many-to-one matching market (F, W, q, P) in which $q_f = 1$ for all $f \in F$.

A.1 Properties of the Core

The core of a many-to-one matching market has a special structure. The following well-known properties will be useful in the sequel:[19]

(P1) For each profile $P \in \mathcal{P}$, $C(P)$ contains two stable matchings, the firms-optimal stable matching μ_F and the workers-optimal stable matching μ_W, with the property that for all $\mu \in C(P)$, $\mu_W(w) R_w \mu(w) R_w \mu_F(w)$ for all $w \in W$, and for all $f \in F$, $\mu_F(f) R_f^* \mu(f) R_f^* \mu_W(f)$ for all $P_f^* \in resp(P_f)$. The *deferred-acceptance algorithms (DA-algorithms)* (Gale and Shapley, 1962) are denoted by $DA_F : \mathcal{P} \to \mathcal{M}$ and $DA_W : \mathcal{P} \to \mathcal{M}$: for all $P \in \mathcal{P}$, $DA_F[P] = \mu_F$ and $DA_W[P] = \mu_W$.

(P2) For each profile $P \in \mathcal{P}$ and any responsive extensions $P_F^* = (P_f^*)_{f \in F}$ of $P_F = (P_f)_{f \in F}$, $C(P)$ coincides with the set of group stable matchings at (P_W, P_F^*), where group stability corresponds to the usual cooperative game theoretical notion of weak blocking.[20] This means that the set of group stable matchings (relative to P) is invariant with respect to any specific responsive extensions of P_F.

(P3) For each $P \in \mathcal{P}$, the set of unmatched agents is the same for all stable matchings (see Roth and Sotomayor, 1990, Theorems 5.12 and 5.13): for all $\mu, \mu' \in C(P)$, and for all $w \in W$ and $f \in F$, (i) if $\mu(w) = \emptyset$, then $\mu'(w) = \emptyset$; (ii) $|\mu(f)| = |\mu'(f)|$; and (iii) if $|\mu(f)| < q_f$, then $\mu(f) = \mu'(f)$.

---

[19]See Roth and Sotomayor (1990) for a detailed presentation of these properties.

[20]A matching μ is weakly blocked by coalition $S \subseteq V$ under (P_W, P_F^*) if there exists a matching μ' such that (b1) for all $v \in S$, $\mu'(v) \subseteq S$, (b2) for all $w \in W \cap S$, $\mu'(w) R_w \mu(w)$, and (b3) for all $f \in F \cap S$, $\mu'(f) R_f^* \mu(f)$, with strict preference holding for at least one $v \in S$.

(P4) Consider a one-to-one matching market $(F, W, P; q)$ and suppose that new workers enter the market. Let $(F, W', P'; q)$ be this new one-to-one matching market where $W \subseteq W'$ and P' agrees with P over F and W. Let $DA_W[P] = \mu_W$. Then, for all $f \in F$, $\mu'(f) R'_f \mu_W(f)$ for all $\mu' \in C(F, W', P'; q)$ (Gale and Sotomayor, 1985; Crawford, 1991).

(P5) Given (F, W, q, P), split each firm f into q_f identical copies of itself (all having the same preference ordering P_f) and let F' be this new set of $\sum_{f \in F} q_f$ split firms. Set $q'_{f'} = 1$ for all $f' \in F'$ and replace f by its copies in F' (always in the same order) in each worker's preference relation P_w. Then, $(F', W, P; q')$ is a one-to-one matching market for which we can uniquely identify its matchings with the matchings of the original many-to-one matching market (F, W, q, P), and vice versa (Roth and Sotomayor, 1990, Lemma 5.6). Then, and using this identification, we write $C(F, W, q, P) = C(F', W, P; q')$.

A.2 Proof of Theorem 1

Below we extend our result to random stable mechanisms[21] and mixed strategies.

Let $\Delta(\mathcal{M})$ denote the set of all probability distributions over $\mathcal{M}$. A random mechanism is a function $\tilde{\varphi} : \mathcal{P} \to \Delta(\mathcal{M})$ choosing for each profile $P \in \mathcal{P}$ a distribution $\tilde{\varphi}[P]$ over $\mathcal{M}$. The random mechanism $\tilde{\varphi}$ is stable if for all $P \in \mathcal{P}$, the support of $\tilde{\varphi}[P]$ is contained in $C(P)$. Given $v \in V$, let $\Delta(\mathcal{P}_v)$ denote the set of all probability distributions over $\mathcal{P}_v$. A mixed strategy of agent v is a function $m_v : \mathcal{P}_v \to \Delta(\mathcal{P}_v)$ specifying for each type P_v of v a distribution $m_v(P_v)$ over $\mathcal{P}_v$. A (mixed) strategy profile is a list $m = (m_v)_{v \in V}$.

Given a random mechanism $\tilde{\varphi}$, $P \in \mathcal{P}$, and m, $m(P)$ is a NE (in mixed strategies) in $\tilde{\varphi}$ under complete information P (i.e. $m(P)$ is a NE in $(\mathcal{P}, \tilde{\varphi}, P)$) if for all $v \in V$ and all $P'_v \in \mathcal{P}_v$, $\tilde{\varphi}[m(P)](v) P^{sd}_v \tilde{\varphi}[P'_v, m_{-v}(P_{-v})](v)$.[22]

---

[21]Pais (2008) provides a strategic analysis of random stable mechanisms under complete information.

[22]Note that implicitly here we use the fact that under complete information a mixed strategy is a

Definition 5 (OBNE in mixed strategies) Let $\tilde{P}$ be a common prior. A mixed strategy profile m is an *ordinal Bayesian Nash equilibrium* (*OBNE*) in the random mechanism $\tilde{\varphi}$ under incomplete information $\tilde{P}$ if and only if for all $v \in V$ and all $P_v \in \mathcal{P}_v$ such that $\Pr\{\tilde{P}_v = P_v\} > 0$,

$$\tilde{\varphi}[m_v(P_v), m_{-v}(\tilde{P}_{-v}|_{P_v})](v) P_v^{sd} \tilde{\varphi}[u_v, m_{-v}(\tilde{P}_{-v}|_{P_v})](v) \quad \text{for all } u_v \in \Delta(\mathcal{P}_v). \quad (4)$$

As usual, if s is an OBNE in pure strategies in the (deterministic) mechanism φ under $\tilde{P}$, then s is an OBNE in mixed strategies in the mechanism φ (where φ is a random mechanism putting probability one on a unique matching for each profile).

Theorem 2 *Let $\tilde{P}$ be a common prior, m be a mixed strategy profile, and $\tilde{\varphi}$ be a random stable mechanism. Then, m is an OBNE in $\tilde{\varphi}$ under incomplete information $\tilde{P}$ if and only if for any profile P in the support of $\tilde{P}$, $m(P)$ is a Nash equilibrium in $\tilde{\varphi}$ under complete information P.*

Proof. Let $\tilde{P}$ be a common prior, m be a mixed strategy profile and $\tilde{\varphi}$ be a random stable mechanism. For any probability distribution D we denote by $supp(D)$ its support (and e.g., $supp(\tilde{P})$ is the support of $\tilde{P}$).

($\Leftarrow$) Suppose that for any profile P in the support of $\tilde{P}$, $m(P)$ is a Nash equilibrium in the mechanism $\tilde{\varphi}$ under complete information P. Let $v \in V$ and $P_v \in \mathcal{P}_v$ be such that $\Pr\{\tilde{P}_v = P_v\} > 0$. By the previous fact, then we have for all $P_v' \in \mathcal{P}_v$ and all $P_{-v} \in \mathcal{P}_{-v}$ such that $\Pr\{\tilde{P}_{-v}|_{P_v} = P_{-v}\} > 0$, $\tilde{\varphi}[m(P)](v) P_v^{sd} \tilde{\varphi}[P_v', m_{-v}(P_{-v})](v)$ for all $P_v' \in \mathcal{P}_v$. Hence,

$$\tilde{\varphi}[m_v(P_v), m_{-v}(\tilde{P}_{-v}|_{P_v})](v) P_v^{sd} \tilde{\varphi}[P_v', m_{-v}(\tilde{P}_{-v}|_{P_v})](v),$$

and for any $u_v \in \Delta(\mathcal{P}_v)$,

$$\tilde{\varphi}[m_v(P_v), m_{-v}(\tilde{P}_{-v}|_{P_v})](v) P_v^{sd} \tilde{\varphi}[u_v, m_{-v}(\tilde{P}_{-v}|_{P_v})](v).$$

---

best response (in the set of all mixed strategies) if and only if the mixed strategy is weakly better than any pure strategy $P_v' \in \mathcal{P}_v$.

Hence, m is an OBNE in mixed strategies in $\tilde{\varphi}$ under $\tilde{P}$, the desired conclusion.

($\Rightarrow$) Let m be an OBNE in mixed strategies in the random stable mechanism $\tilde{\varphi}$ under $\tilde{P}$.

First we show that for all $P \in \mathcal{P}$ such that $\Pr\{\tilde{P} = P\} > 0$,

$$\mu(v) \subseteq A(P_v) \text{ for all } v \in V \text{ and all } \mu \in supp(\tilde{\varphi}[m(P)]). \tag{5}$$

If for some P in the support of $\tilde{P}$, for some $v \in V$, and some $\mu \in supp(\tilde{\varphi}[m(P)])$, $\mu(v) \not\subseteq A(P_v)$, then let agent v choose P_v instead of $m_v(P_v)$. By stability of $\tilde{\varphi}$, we have $\mu'(v) \subseteq A(P_v)$ for all $P'_{-v} \in \mathcal{P}_{-v}$ and all $\mu' \in supp(\tilde{\varphi}[P_v, m_{-v}(P'_{-v})])$. Let $v \in F$ (the case $v \in W$ is analogous and easier). We choose a responsive extension P_v^* of P_v such that for all $W' \in 2^W$ with $|W'| \le q_v$, $W' R_v^* \emptyset$ if and only if $W' \subseteq A(P_v)$.[23] Hence, by $\mu(v) \not\subseteq A(P_v)$, $\mu \in supp(\tilde{\varphi}[m(P)])$, and $\Pr\{\tilde{P}_{-v}|_{P_v} = P_{-v}\} > 0$, it follows that

$$\Pr\{\tilde{\varphi}[P_v, m_{-v}(\tilde{P}_{-v}|_{P_v})](v) \in B(\emptyset, P_v^*)\} = 1 > \Pr\{\tilde{\varphi}[m_v(P_v), m_{-v}(\tilde{P}_{-v}|_{P_v})](v) \in B(\emptyset, P_v^*)\},$$

which means that m is not an OBNE in the random stable mechanism $\tilde{\varphi}$ under $\tilde{P}$, a contradiction. Hence, (5) holds.

Second, suppose that for some $P \in supp(\tilde{P})$, $m(P)$ is not a NE in $\tilde{\varphi}$ under P. Then, without loss of generality, there exist $f \in F$ and $P'_f \in \mathcal{P}_f$ such that

$$\tilde{\varphi}[m(P)](f) \; \not\!P_f^{sd} \; \tilde{\varphi}[P'_f, m_{-f}(P_{-f})](f).$$

Then there exists a responsive extension P_f^* of P_f and $P'_{-f} \in supp(m_{-f}(P_{-f}))$ such that for some $\mu' \in supp(\tilde{\varphi}[P'_f, P'_{-f}])$ and $\mu \in supp(\tilde{\varphi}[m_f(P_f), P'_{-f}])$ we have

$$\mu'(f) P_f^* \mu(f). \tag{6}$$

---

[23]The responsive extension P_v^* of P_v can be chosen to be additive by selecting $g : W \to \mathbb{R}\backslash\{0\}$ such that (i) for all $w, w' \in W$, $w P_v w' \Leftrightarrow g(w) > g(w')$, (ii) for all $w \in A(P_v)$, $g(w) \in (0, 1)$, and (iii) for all $w \in W\backslash A(P_v)$, $g(w) < -|W|$. It is easy to see that for all $W' \in 2^W$ with $|W'| \le q_v$, $\sum_{w' \in W'} g(w') \ge 0 \Leftrightarrow W' \subseteq A(P_v)$.

The case where a worker has a profitable deviation is analogous to the case where a firm with quota one has a profitable deviation.

Let $\mu'(f) = \{w'_1, w'_2, \ldots, w'_{|\mu'(f)|}\}$ where $w'_1 P_f w'_2 P_f \cdots P_f w'_{|\mu'(f)|}$ and $\mu(f) = \{w_1, w_2, \ldots, w_{|\mu(f)|}\}$ where $w_1 P_f w_2 P_f \cdots P_f w_{|\mu(f)|}$. We now construct from P'_f another deviation P''_f and from $\mu'(f)$ both a responsive extension P^{**}_f of P_f and a subset of workers W^*, and prove that the random matching $\tilde{\varphi}[m_f(P_f), m_{-f}(\tilde{P}_{-f}|_{P_f})]$ does not first-order stochastically P_f-dominate the random matching $\tilde{\varphi}[P''_f, m_{-f}(\tilde{P}_{-f}|_{P_f})]$ since $\Pr\{\tilde{\varphi}[P''_f, m_{-f}(\tilde{P}_{-f}|_{P_f})](f) \in B(W^*, P^{**}_f)\} > \Pr\{\tilde{\varphi}[m_f(P_f), m_{-f}(\tilde{P}_{-f}|_{P_f})](f) \in B(W^*, P^{**}_f)\}$. We proceed by distinguishing between two mutually exclusive cases.

Case 1: There exists $k \in \{1, \ldots, |\mu'(f)|\}$ such that $w'_k P_f w_k$ and $w_l R_f w'_l$ for all $l \in \{1, \ldots, k-1\}$.

Note that $w'_k \in A(P_f)$ because $w'_k P_f w_k$ and by (5), $w_k \in \mu(f) \subseteq A(P_f)$. Let $P''_f \in \mathcal{P}_f$ be such that $A(P''_f) = B(w'_k, P_f)$ and $P''_f|A(P''_f) = P'_f|A(P''_f)$.

First we show that for all $\mu'' \in supp(\tilde{\varphi}[P''_f, P'_{-f}])$, $\mu''(f)$ contains at least k workers. Note that any profile implicitly specifies the set of agents of the matching problem. For the time being, below we specify both the profile and the quota of the matching problem.

Because $\tilde{\varphi}$ is stable and $\mu' \in supp(\tilde{\varphi}[P'_f, P'_{-f}])$, $\mu' \in C(P'_f, P'_{-f}; q)$. Let μ'' be the matching for the problem $(F, W \setminus \{w'_{k+1}, \ldots, w'_{|\mu'(f)|}\}, (k, q_{-f}), (P'_f, P'_{-\{f\} \cup \{w'_{k+1}, \ldots, w'_{|\mu'(f)|}\}}))$ such that $\mu''(f) = \{w'_1, \ldots, w'_k\}$ and $\mu''(f') = \mu'(f')$ for all $f' \in F \setminus \{f\}$. Then from $\mu' \in C(P'_f, P'_{-f}; q)$ it follows that

$$\mu'' \in C(P'_f, P'_{-\{f\} \cup \{w'_{k+1}, \ldots, w'_{|\mu'(f)|}\}}; k, q_{-f}). \tag{7}$$

By our choice of P''_f, we have $\mu''(f) \subseteq A(P''_f)$ and $P''_f|A(P''_f) = P'_f|A(P''_f)$. Hence, we also have by (7),

$$\mu'' \in C(P''_f, P'_{-\{f\} \cup \{w'_{k+1}, \ldots, w'_{|\mu'(f)|}\}}; k, q_{-f}). \tag{8}$$

Thus, by $\mu''(f) = \{w'_1, \ldots, w'_k\}$ and the fact that any firm is matched to the same number of workers under all stable matchings, firm f is matched to k workers for all

matchings belonging to $C(P''_f, P'_{-\{f\} \cup \{w'_{k+1}, \ldots, w'_{|\mu'(f)|}\}}; k, q_{-f})$. Now if firm f is matched to fewer than k workers in some matching belonging to $C(P''_f, P'_{-\{f\} \cup \{w'_{k+1}, \ldots, w'_{|\mu'(f)|}\}}; q)$, then this matching is also stable for the problem $(P''_f, P'_{-\{f\} \cup \{w'_{k+1}, \ldots, w'_{|\mu'(f)|}\}}; k, q_{-f})$, a contradiction to the previous fact. Hence, f is matched to at least k workers in any stable matching belonging to $C(P''_f, P'_{-\{f\} \cup \{w'_{k+1}, \ldots, w'_{|\mu'(f)|}\}}; q)$. Now when considering the workers-optimal matching in this core, we may split firm f into q_f copies (all having the same preference P''_f) and each copy of firm f weakly prefers according to P''_f any matching in $C(P''_f, P'_{-f}; q)$ to this matching. Since at least k copies of f are matched to a worker under the workers-optimal matching in $C(P''_f, P'_{-\{f\} \cup \{w'_{k+1}, \ldots, w'_{|\mu'(f)|}\}}; q)$, at least k copies of f must be also matched to a worker under any stable matching in $C(P''_f, P'_{-f}; q)$. Therefore, by $supp(\tilde{\varphi}[P''_f, P'_{-f}]) \subseteq C(P''_f, P'_{-f}; q)$, for all $\mu'' \in supp(\tilde{\varphi}[P''_f, P'_{-f}])$, $\mu''(f)$ contains at least k workers.

Second we choose a responsive extension P^{**}_f of P_f. Let $W^* \subseteq B(w'_k, P_f)$ be such that W^* consists of the k lowest ranked workers (according to P_f) in the set $B(w'_k, P_f)$, i.e. $|W^*| = k$ and for all $w \in B(w'_k, P_f) \backslash W^*$ and all $w^* \in W^*$, $w P_f w^*$. Let P^{**}_f be the responsive extension of P_f be such that for all $W'' \in 2^W$, $W'' R^{**}_f W^*$ if and only if the following three conditions hold: (i) $W'' \subseteq A(P_f)$, (ii) $k \le |W''| \le q_f$, and (iii) if $W'' = \{w''_1, w''_2, \ldots, w''_{|W''|}\}$ where $w''_1 P_f \cdots P_f w''_{|W''|}$ and $W^* = \{w^*_1, \ldots, w^*_k\}$ where $w^*_1 P_f \cdots P_f w^*_k$, then $w''_l R_f w^*_l$ for all $l \in \{1, \ldots, k\}$.[24] Since for all $\mu'' \in supp(\tilde{\varphi}[P''_f, P'_{-f}])$, $\mu''(f)$ contains at least k workers and $A(P''_f) = B(w'_k, P_f)$, our construction implies that for all $\mu'' \in supp(\tilde{\varphi}[P''_f, P'_{-f}])$, $\mu''(f) P^{**}_f \mu(f)$. More precisely, for Case 1 the set $\mu(f)$ violates (iii) and our choice of P^{**}_f and W^* yields for all $\mu'' \in supp(\tilde{\varphi}[P''_f, P'_{-f}])$,

$$\mu''(f) R^{**}_f W^* P^{**}_f \mu(f) \tag{9}$$

---

[24]The responsive extension P^{**}_f of P_f can be chosen to be additive by selecting $g : W \to \mathbb{R} \backslash \{0\}$ such that (a) for all $w, w' \in W$, $w P_f w' \Leftrightarrow g(w) > g(w')$, (b) for all $w \in B(w'_k, P_f)$, $g(w) \in [|W|, |W| + 1]$, (c) for all $w \in A(P_f) \backslash B(w'_k, P_f)$, $g(w) \in (0, 1)$, and (d) for all $w \in W \backslash A(P_f)$, $g(w) < -|W|^2$. It is easy to see that for all $W'' \in 2^W$ with $|W''| \le q_f$, $\sum_{w' \in W''} g(w') \ge \sum_{w \in W^*} g(w) \Leftrightarrow$ (i)-(iii) hold for W''.

and $\mu \in supp(\tilde{\varphi}[m_f(P_f), P'_{-f}])$.

The following claim will be the key to the proof. We show that for any profile, if some stable matching is weakly preferred to W^* under P_f^{**}, then *all* matchings, which are stable under the profile where f's preference ordering is replaced by P_f'', are weakly preferred to W^* under P_f^{**}.

Claim: Let $\hat{P} \in \mathcal{P}$. If for some $\hat{\mu} \in C(\hat{P})$, $\hat{\mu}(f)R_f^{**}W^*$, then for all $\bar{\mu} \in C(P_f'', \hat{P}_{-f})$, $\bar{\mu}(f)R_f^{**}W^*$.

Proof of Claim: By $\hat{\mu}(f)R_f^{**}W^*$ and our choice of P_f^{**},

$$\hat{\mu}(f) \cap B(w'_k, P_f) \text{ contains at least } k \text{ workers.} \tag{10}$$

If for some $\bar{\mu}' \in C(P_f'', \hat{P}_{-f})$, $\bar{\mu}'(f)$ contains at least k workers, then all these workers belong to $B(w'_k, P_f)$ because $A(P_f'') = B(w'_k, P_f)$. Thus, by our choice of P_f^{**} and W^*, $\bar{\mu}'(f)R_f^{**}W^*$. Because by (P3), f is matched to at least k workers in any stable matching belonging to $C(P_f'', \hat{P}_{-f})$ and $A(P_f'') = B(w'_k, P_f)$, it follows from our construction that $\bar{\mu}(f)R_f^{**}W^*$ for all $\bar{\mu} \in C(P_f'', \hat{P}_{-f})$, the desired conclusion.

Suppose that for all $\bar{\mu} \in C(P_f'', \hat{P}_{-f})$, $\bar{\mu}(f)$ contains fewer than k workers. Let $\hat{\mu}(f) = \{\hat{w}_1, \ldots, \hat{w}_{|\hat{\mu}(f)|}\}$ where $\hat{w}_1 P_f \cdots P_f \hat{w}_{|\hat{\mu}(f)|}$. By (10), $\hat{\mu}(f) \cap B(w'_k, P_f)$ contains at least k workers. Thus, $k \leq |\hat{\mu}(f)|$. For the time being, below we specify both the profile and the quota of the matching problem. Then we have $\hat{\mu} \in C(\hat{P}; q)$. Let $\hat{\mu}'$ be the matching for the problem $(F, W \backslash \{\hat{w}_{k+1}, \ldots, \hat{w}_{|\hat{\mu}(f)|}\}, (k, q_{-f}), (\hat{P}_f, \hat{P}_{-\{f\} \cup \{\hat{w}_{k+1}, \ldots, \hat{w}_{|\hat{\mu}(f)|}\}}))$ such that $\hat{\mu}'(f) = \{\hat{w}_1, \ldots, \hat{w}_k\}$ and $\hat{\mu}'(f') = \hat{\mu}(f')$ for all $f' \in F \backslash \{f\}$. Then, from $\hat{\mu} \in C(\hat{P}_f, \hat{P}_{-f}; q)$ it follows that

$$\hat{\mu}' \in C(\hat{P}_f, \hat{P}_{-\{f\} \cup \{\hat{w}_{k+1}, \ldots, \hat{w}_{|\hat{\mu}(f)|}\}}; k, q_{-f}). \tag{11}$$

Let $\hat{w} \in \hat{\mu}'(f)$ be such that $\hat{\mu}'(f) \subseteq B(\hat{w}, \hat{P}_f)$ (in other words, $\hat{w}$ is the worker who is least preferred in $\hat{\mu}'(f)$ according to $\hat{P}_f$). Let $\dot{P}_f \in \mathcal{P}_f$ be such that $A(\dot{P}_f) = B(\hat{w}_k, P_f) \cap B(\hat{w}, \hat{P}_f)$ and $\dot{P}_f | A(\dot{P}_f) = P_f'' | A(\dot{P}_f)$. Note that $\hat{\mu}'(f) \subseteq A(\dot{P}_f)$. Then we have $\hat{\mu}' \in C(\dot{P}_f, \hat{P}_{-\{f\} \cup \{\hat{w}_{k+1}, \ldots, \hat{w}_{|\hat{\mu}(f)|}\}}; k, q_{-f})$: first, note that $\hat{\mu}'$ is individually rational

because both $\hat{\mu}'(f) \subseteq B(\hat{w}_k, P_f)$ and $\hat{\mu}'(f) \subseteq B(\hat{w}, \hat{P}_f)$ (by our choice of $\hat{w}$); second, if there would exist a blocking pair for $\hat{\mu}'$, then by (11) and the fact that only firm f's preference changed from $\hat{P}_f$ to $\dot{P}_f$, firm f needs to be part of this blocking pair; third, if (w, f) blocks $\hat{\mu}'$ under $(\dot{P}_f, \hat{P}_{-\{f\} \cup \{\hat{w}_{k+1}, ..., \hat{w}_{|\hat{\mu}(f)|}\}}; k, q_{-f})$, then $w \notin \hat{\mu}'(f)$; now this implies $w \neq \hat{w} \in \hat{\mu}'(f)$, and $w \in A(\dot{P}_f) = B(\hat{w}_k, P_f) \cap B(\hat{w}, \hat{P}_f)$; but then we have $w \hat{P}_f \hat{w}$ (because $\hat{w}$ is the least preferred worker in $\hat{\mu}'(f)$ and $A(\dot{P}_f) \subseteq B(\hat{w}, \hat{P}_f)$) and (w, f) also blocks $\hat{\mu}'$ under $(\hat{P}_f, \hat{P}_{-\{f\} \cup \{\hat{w}_{k+1}, ..., \hat{w}_{|\hat{\mu}(f)|}\}}; k, q_{-f})$, a contradiction to (11).

Thus, since $|\hat{\mu}'(f)| = k$, firm f is matched to k workers for all matchings belonging to $C(\dot{P}_f, \hat{P}_{-\{f\} \cup \{\hat{w}_{k+1}, ..., \hat{w}_{|\hat{\mu}(f)|}\}}; k, q_{-f})$. Now if firm f is matched to fewer than k workers for some $\tilde{\mu} \in C(\dot{P}_f, \hat{P}_{-\{f\} \cup \{\hat{w}_{k+1}, ..., \hat{w}_{|\hat{\mu}(f)|}\}}; q)$, then $\tilde{\mu}$ is also stable under $(\dot{P}_f, \hat{P}_{-\{f\} \cup \{\hat{w}_{k+1}, ..., \hat{w}_{|\hat{\mu}(f)|}\}}; k, q_{-f})$, a contradiction to the previous fact. Hence, f is matched to at least k workers in any stable matching belonging to $C(\dot{P}_f, \hat{P}_{-\{f\} \cup \{\hat{w}_{k+1}, ..., \hat{w}_{|\hat{\mu}(f)|}\}}; q)$. Now when considering the workers-optimal matching in this core, we may split firm f into q_f copies (all having the same preference $\dot{P}_f$) and each copy of firm f weakly prefers according to $\dot{P}_f$ any matching in $C(\dot{P}_f, \hat{P}_{-f}; q)$ to this matching. Since at least k copies of f are matched to a worker under the workers-optimal matching in $C(\dot{P}_f, \hat{P}_{-\{f\} \cup \{\hat{w}_{k+1}, ..., \hat{w}_{|\hat{\mu}(f)|}\}}; q)$, on the one hand

at least k copies of f are matched to a worker in any matching in $C(\dot{P}_f, \hat{P}_{-f}; q)$.

$$(12)$$

On the other hand, for all $\bar{\mu} \in C(P_f'', \hat{P}_{-f})$, $\bar{\mu}(f)$ contains fewer than k workers. Let $\tilde{\mu} \in C(P_f'', \hat{P}_{-f})$. Let $\tilde{\mu}'$ be the matching for the problem $(F, W \backslash (\tilde{\mu}(f) \backslash A(\dot{P}_f)), q, (P_f'', \hat{P}_{-\{f\} \cup (\tilde{\mu}(f) \backslash A(\dot{P}_f))}))$ such that $\tilde{\mu}'(f) = \tilde{\mu}(f) \cap A(\dot{P}_f)$ and $\tilde{\mu}'(f') = \tilde{\mu}(f')$ for all $f' \in F \backslash \{f\}$. Since $\tilde{\mu} \in C(P_f'', \hat{P}_{-f}; q)$ and $\tilde{\mu}(f)$ contains fewer than q_f workers, we have $\tilde{\mu}' \in C(P_f'', \hat{P}_{-\{f\} \cup (\tilde{\mu}(f) \backslash A(\dot{P}_f))}; q)$. Thus, by $\tilde{\mu}'(f) \subseteq A(\dot{P}_f)$ and $\dot{P}_f | A(\dot{P}_f) = P_f'' | A(\dot{P}_f)$, we also obtain $\tilde{\mu}' \in C(\dot{P}_f, \hat{P}_{-\{f\} \cup (\tilde{\mu}(f) \backslash A(\dot{P}_f))}; q)$. Hence, in any matching belonging to this core firm f is matched to $|\tilde{\mu}'(f)| = |\tilde{\mu}(f) \cap A(\dot{P}_f)|$ workers. Now when considering the workers-optimal matching in this core, we may split each firm $f' \in F \backslash \{f\}$ into $q_{f'}$ copies (all having the same preference $\hat{P}_{f'}$) and each copy of firm f' weakly prefers

27

according to $\hat{P}_{f'}$ any matching in $C(\dot{P}_f, \hat{P}_{-f}; q)$ to this matching. Thus, in total all the copies of all firms $f' \in F \backslash \{f\}$ receive at least the same number of workers in $C(\dot{P}_f, \hat{P}_{-f}; q)$ as they did previously. Since exactly $|\tilde{\mu}(f) \backslash A(\dot{P}_f)|$ new workers are available and f was matched to $|\tilde{\mu}'(f)| = |\tilde{\mu}(f) \cap A(\dot{P}_f)|$ workers before, firm f can be matched to at most $|\tilde{\mu}(f)|$ workers under any stable matching in $C(\dot{P}_f, \hat{P}_{-f}; q)$. Since $|\tilde{\mu}(f)| < k$, this contradicts (12) and the fact that under responsive preferences, firm f is matched to the same number of workers for any two matchings in $C(\dot{P}_f, \hat{P}_{-f}; q)$. Hence, for $\bar{\mu} \in C(P_f'', \hat{P}_{-f})$, $\bar{\mu}(f)$ cannot contain fewer than k workers. $\blacklozenge$

Because $\tilde{\varphi}$ is stable, the Claim implies that for all $\hat{P} \in \mathcal{P}$,

$$\Pr\{\tilde{\varphi}[\hat{P}_f, \hat{P}_{-f}](f) \in B(W^*, P_f^*)\} \leq \Pr\{\tilde{\varphi}[P_f'', \hat{P}_{-f}](f) \in B(W^*, P_f^*)\}.$$

Thus,

$$\Pr\{\tilde{\varphi}[\hat{P}_f, m_{-f}(\tilde{P}_{-f}|_{P_f})](f) \in B(W^*, P_f^*)\} \leq \Pr\{\tilde{\varphi}[P_f'', m_{-f}(\tilde{P}_{-f}|_{P_f})](f) \in B(W^*, P_f^*)\}.$$

By (9), there exists $\dot{P}_f \in supp(m_f(P_f))$ such that $\mu \in supp(\tilde{\varphi}[\dot{P}_f, P'_{-f}])$. Thus, by $\mu'(f) R_f^{**} W^* P_f^{**} \mu(f)$ and $\mu' \in C(P_f'', P'_{-f})$, (9) implies

$$\Pr\{\tilde{\varphi}[\dot{P}_f, P'_{-f}](f) \in B(W^*, P_f^*)\} < \Pr\{\tilde{\varphi}[P_f'', P'_{-f}](f) \in B(W^*, P_f^*)\}.$$

Hence, by $m_f(P_f) \in \Delta(\mathcal{P}_f)$ and both $\dot{P}_f \in supp(m_f(P_f))$ and $P'_{-f} \in supp(m_{-f}(\tilde{P}_{-f}|_{P_f}))$,

$$\Pr\{\tilde{\varphi}[m_f(P_f), m_{-f}(\tilde{P}_{-f}|_{P_f})](f) \in B(W^*, P_f^*)\} < \Pr\{\tilde{\varphi}[P_f'', m_{-f}(\tilde{P}_{-f}|_{P_f})](f) \in B(W^*, P_f^*)\},$$

which means that m is not an OBNE in $\tilde{\varphi}$ under $\tilde{P}$.

Case 2: Otherwise.

Then we have $w_l R_f w_l'$ for all $l \in \{1, \ldots, \min\{|\mu(f)|, |\mu'(f)|\}\}$. Let $k = |\mu(f)|$. If $|\mu'(f)| \leq |\mu(f)|$, then by responsiveness of P_f^* and $\mu(f) \subseteq A(P_f)$, we have $\mu(f) R_f^* \mu'(f)$, which contradicts (6). Hence, we have $|\mu'(f)| > |\mu(f)| = k$, $q_f > k$, and $w'_{k+1} \in A(P_f)$. Let $P_f'' \in \mathcal{P}_f$ be such that $A(P_f'') = B(w'_{k+1}, P_f)$ and $P_f''|A(P_f'') = P_f'|A(P_f'')$.

Since $\mu(f) \subseteq B(w'_{k+1}, P_f) = A(P''_f)$ and $\mu(f)$ does not fill the quota of firm f, we have $\mu \in C(P''_f, P'_{-f}; q)$. Hence, on the one hand

$$\text{firm } f \text{ is matched to } k \text{ workers under any matching in } C(P''_f, P'_{-f}; q). \qquad (13)$$

On the other hand, let μ'' be the matching for the problem $(F, W \setminus \{w'_{k+2}, \ldots, w'_{|\mu'(f)|}\}, (k+1, q_{-f}), (P''_f, P'_{-\{f\} \cup \{w'_{k+2}, \ldots, w'_{|\mu'(f)|}\}}))$ such that $\mu''(f) = \{w'_1, \ldots, w'_{k+1}\}$ and $\mu''(f') = \mu'(f')$ for all $f' \in F \setminus \{f\}$. Then from $\mu' \in C(P'_f, P'_{-f}; q)$ it follows that $\mu'' \in C(P'_f, P'_{-\{f\} \cup \{w'_{k+2}, \ldots, w'_{|\mu'(f)|}\}}; k+1, q_{-f})$. Thus, by $\mu''(f) \subseteq B(w'_{k+1}, P_f) = A(P''_f)$ and $P''_f | A(P''_f) = P'_f | A(P''_f)$, $\mu'' \in C(P''_f, P'_{-\{f\} \cup \{w'_{k+2}, \ldots, w'_{|\mu'(f)|}\}}; k+1, q_{-f})$. Now if firm f is matched to fewer than $k+1$ workers in some matching belonging to $C(P''_f, P'_{-\{f\} \cup \{w'_{k+2}, \ldots, w'_{|\mu'(f)|}\}}; q)$, then this matching is also stable for the problem $(P''_f, P'_{-\{f\} \cup \{w'_{k+2}, \ldots, w'_{|\mu'(f)|}\}}; k+1, q_{-f})$, a contradiction to the previous fact. Hence, f is matched to at least $k+1$ workers in any stable matching belonging to $C(P''_f, P'_{-\{f\} \cup \{w'_{k+2}, \ldots, w'_{|\mu'(f)|}\}}; q)$. Now when considering the workers-optimal matching in this core, we may split firm f into q_f copies (all having the same preference P''_f) and each copy of firm f weakly prefers according to P''_f any matching in $C(P''_f, P'_{-f}; q)$ to this matching. Since at least $k+1$ copies of f are matched to a worker under the workers-optimal matching in $C(P''_f, P'_{-\{f\} \cup \{w'_{k+2}, \ldots, w'_{|\mu'(f)|}\}}; q)$, at least $k+1$ copies of f must be also matched to a worker under any matching in $C(P''_f, P'_{-f}; q)$, which contradicts (13) and the fact that firm f is matched to the same number of workers under any matching in $C(P''_f, P'_{-f}; q)$. Hence, Case 2 cannot occur. $\qquad \square$

Theorem 1 is a corollary of Theorem 2 by restricting Theorem 2 and its proof to pure strategies and deterministic mechanisms.

A.3 Proof of Proposition 1

Proposition 1 *Let $\tilde{P}$ be a common prior.*

(a) *If $\tilde{P}$ is F-correlated or W-correlated, then truth-telling is an OBNE in any stable mechanism under incomplete information $\tilde{P}$.*

(b) Let s be a strategy profile such that $s_w(P_w) = P_w$ for all $w \in W$ and all $P_w \in$
   $\mathcal{P}_w$. If $\tilde{P}$ is W-correlated and s is an OBNE in the stable mechanism DA_W
   under incomplete information $\tilde{P}$, then for all profiles P in the support of $\tilde{P}$,
   $DA_W[s(P)]$ is stable with respect to P. The analogous statement is true for the
   stable mechanism DA_F.

Proof. (a) Let φ be a stable mechanism and $\tilde{P}$ be a common prior. Without loss
of generality, let $\tilde{P}$ be F-correlated. The case where $\tilde{P}$ is W-correlated is analogous
to the case where $\tilde{P}$ is F-correlated and all firms have quota 1. Let P be in the
support of $\tilde{P}$. Because all firms' preferences are identical at P, we have $|C(P)| = 1$,
say $C(P) = \{\mu\}$. By stability of φ, $\varphi[P] = \mu$. By Theorem 1, it suffices to show that
P is a NE in the mechanism φ under complete information P.

Because all firms have identical preferences, say $P_f : w_1 w_2 \cdots w_k \emptyset w_{k+1} \cdots$ for all
$f \in F$, $\mu(w_1)$ is w_1's most preferred firm (if any) under P_{w_1}. Then $\mu(w_2)$ is w_2's most
preferred firm (if any) from $F \backslash \{\mu(w_1)\}$ under P_{w_2}, and in general for $i = 1, \ldots, |W|$,
$\mu(w_i)$ is w_i's most preferred firm (if any) from $F \backslash \{\mu(w_1), \ldots, \mu(w_{i-1})\}$ under P_{w_i}. By
stability of φ, obviously no worker can profitably manipulate.

Let $f \in F$, $P'_f \in \mathcal{P}_f$, and $\mu' = \varphi[P'_f, P_{-f}]$. Suppose that for some $P^*_f \in resp(P_f)$ we
have $\mu'(f) P^*_f \mu(f)$. Hence, $\mu'(f) \neq \mu(f)$. By stability of φ, without loss of generality
we may suppose $A(P'_f) = \mu'(f)$ and because any firm's set of acceptable workers is
$\{w_1, \ldots, w_k\}$, $A(P'_f) \subseteq A(P_f)$. Again, by stability of φ, $\mu(f) \subseteq A(P_f)$. First, suppose
that $|\mu'(f)| > |\mu(f)|$. Note that $\mu \in C(P)$ and $\mu' \in C(P'_f, P_{-f})$. Then, without loss
of generality, we may suppose $|\mu'(f)| = q_f$.[25] Let $P''_f \in \mathcal{P}_f$ be such that $A(P''_f) = A(P_f)$ and $P''_f | A(P''_f) = P'_f | A(P''_f)$. By $|\mu(f)| < q_f$, $A(P''_f) = A(P_f)$ and (P3), we
obtain $\mu \in C(P''_f, P_{-f})$. By $|\mu(f)| \neq |\mu'(f)|$ and (P3), we have $\mu' \notin C(P''_f, P_{-f})$.
Thus, μ' is blocked by some pair (w', f') under (P''_f, P_{-f}). By $\mu' \in C(P'_f, P_{-f})$ and

---

[25]If $|\mu'(f)| < q_f$, then set $q'_f = |\mu'(f)|$. From (P3) (where we specify both the profile and the
quotas), $\mu \in C(P; q)$ and $|\mu(f)| < q'_f$ imply $\mu \in C(P; q'_f, q_{-f})$, and similarly $\mu' \in C(P'_f, P_{-f}; q)$ and
$\mu'(f) = A(P'_f)$ imply $\mu' \in C(P'_f, P_{-f}; q'_f, q_{-f})$.

$A(P'_f) \subseteq A(P''_f)$, we have $f' = f$ and $w' \notin \mu'(f)$. But then by $A(P'_f) = \mu'(f)$ and $P''_f|A(P''_f) = P'_f|A(P''_f)$, we have $wP_f w'$ for all $w \in \mu'(f)$. Now f has an unfilled slot under μ' and $|\mu'(f)| < q_f$, a contradiction.

Hence, $|\mu'(f)| \leq |\mu(f)|$. Let $\underline{w}$ be the least P_f-preferred worker in $\mu(f)$, i.e. $wR_f\underline{w}$ for all $w \in \mu(f)$. If for all $w \in \mu'(f)\backslash\mu(f)$ we have $\underline{w}P_f w$, then by $|\mu'(f)| \leq |\mu(f)|$ and responsiveness of P^*_f, we have $\mu(f)R^*_f\mu'(f)$, a contradiction. Let $w_l \in \mu'(f)\backslash\mu(f)$ be such that $w_l P_f\underline{w}$. We show that there exists an index $i(l) < l$ such that $w_{i(l)} \in \mu(f)\backslash\mu'(f)$. Let $\mu(w_l) = f_l$. Note that by $f_l \neq f$, $w_l P_f\underline{w}$ and $\mu \in C(P)$ we have $f_l P_{w_l} f$. Thus, by $\mu' \in C(P'_f, P_{-f})$, we have $|\mu'(f_l)| = q_{f_l}$ and $wP_{f_l}w_l$ for all $w \in \mu'(f_l)$. But then w_l's position at firm f_l is filled with some new worker w', i.e. $w' \in \mu'(f_l)\backslash\mu(f_l)$ and $w'P_{f_l}w_l$. Now by $\mu \in C(P)$ and $\mu(w_l) = f_l$, $\mu(w')P_{w'}f_l$. If $\mu(w') = f$, then w' has an index $i(l) < l$ such that $w' = w_{i(l)}$ and $w_{i(l)} \in \mu(f)\backslash\mu'(f)$. Otherwise, let $\mu(w') = f' \neq f$. Then again as above from $f'P_{w'}f_l$ we have $|\mu'(f')| = q_{f'}$ and $wP_{f'}w'$ for all $w \in \mu'(f')$, and w''s position at firm f' is filled with some new worker w'', i.e. $w'' \in \mu'(f')\backslash\mu(f')$ and $w''P_{f'}w'$. By $P'_{-f} = P_{-f}$ and the finiteness of W and F, in the end for $w_l \in \mu'(f)\backslash\mu(f)$ there exists $w_{i(l)} \in \mu(f)\backslash\mu'(f)$ with $i(l) < l$. Furthermore, from the above arguments, we can choose $i(l) \neq i(l')$ for all $l \neq l'$ such that $w_l, w_{l'} \in \mu'(f)\backslash\mu(f)$. Since $\mu(f) \subseteq A(P_f)$ and $|\mu'(f)| \leq |\mu(f)|$, responsiveness of P^*_f implies $\mu(f)R^*_f\mu'(f)$, a contradiction.

(b) Let P be in the support of $\tilde{P}$. Since $s_w(P_w) = P_w$ for all w and $\tilde{P}$ is W-correlated, we have by (a) that no worker can gain by manipulation. Furthermore, by Theorem 1, $s(P)$ must be a NE in DA_W under P. Because $\tilde{P}$ is W-correlated, all workers have identical preferences, say $P_w : f_1 f_2 \cdots f_l \emptyset f_{l+1} \cdots$ for all $w \in W$. Suppose that $DA_W[s(P)]$ is not stable with respect to P. Since $s(P)$ is a NE in DA_W under P, no agent is matched to any partner under $DA_W[s(P)]$ which is unacceptable according to its true preference relation. Suppose that some unmatched worker-firm pair (w, f) blocks $DA_W[s(P)]$. Then $f \in A(P_w)$ and by $s_w(P_w) = P_w$, $fP_w DA_W[s(P)](w)$. But then, along the DA_W-algorithm which produces $DA_W[s(P)]$, worker w proposed

to f before proposing to $DA_W[s(P)](w)$ and because all workers' submitted lists are identical, at that step all unmatched workers proposed to f (and the set of unmatched workers shrinks from one step to the next one). Let w' be the least preferred worker according to P_f in $DA_W[s(P)](f)$. But now f profitably deviates from $s(P)$ in DA_W by submitting a list P'_f where $A(P'_f) = (DA_W[s(P)](f) \cup \{w\})\backslash\{w'\}$. When in $DA_W[P'_f, s_{-f}(P_{-f})]$ worker w proposes to f, all unmatched workers propose to f in that step because all workers' submitted lists are identical. Firm f accepts $(DA_W[s(P)](f) \cup \{w\})\backslash\{w'\}$ which is strictly preferred to $DA_W[s(P)](f)$ under any responsive extension P_f^* of P_f. Hence, $s(P)$ is not a NE in DA_W under P, a contradiction. $\qquad\square$