Abstract: We study strategic behavior of agents participating in centralized stable mechanisms in two-sided many-to-one matching markets when each agent has incomplete information about the preferences of the other agents. It is known that in any stable mechanism there is a strong link between ordinal Bayesian Nash equilibria under incomplete information and Nash equilibria under complete information: given a common belief, a strategy is an ordinal Bayesian Nash equilibrium in a stable mechanism if and only if at each profile in the support of the belief the strategy prescribes a Nash equilibrium in the stable
mechanism under complete information at the profile. In this paper we show that in the workers-optimal stable mechanism this link does not hold for truth-telling if firms’ preferences over families of subsets of workers are monotonic responsive extensions of their rankings over individual workers. Firms have monotonic responsive extensions if having all positions filled with acceptable workers is always strictly preferred to having some positions vacant. However, we show that in the firms-optimal stable mechanism the link holds for truth-telling: given a common belief, truth-telling is a monotonic ordinal Bayesian Nash equilibrium in the firms’ optimal stable mechanism under incomplete information if and only if truth-telling is a monotonic Nash equilibrium under complete information at each profile in the support of the common belief.

*JEL classification number*: C78; D81; J44.

*Keywords*: Many-to-one matching market; Stability; Incomplete information; Monotonic responsive extensions.

1 Introduction

Both empirical and theoretical studies of two-sided many-to-one matching markets have been useful in applications. Many such markets have developed centralized market clearing mechanisms (in response to various failures of the decentralized market) to match the agents from the two sides: the institutions (*firms*, colleges, hospitals, schools, etc.) and the individuals (*workers*, students, medical interns, children, etc.).

The National Resident Matching Program is the most well-studied example of this kind of two-sided matching markets. Each year around 20,000 medical students look for a four-years position in American hospital programs to undertake their medical

---

1Roth and Sotomayor (1990) gives a masterful overview of two-sided matching markets.
internships. In many countries, each year thousands of students seek for positions in colleges, six years old children have to be assigned to public schools or 8th graders high school students to high schools, as well as civil servants to similar jobs in public positions scattered in different cities across a country.

All of these entry-level many-to-one matching markets share two specific features. First, workers enter the market by cohorts (often once per year) and, second, agents are matched through a centralized procedure: an institution (clearinghouse) collects, for each participant, a ranked list of potential partners and proposes, after processing the profile of submitted ranked lists, a final matching between firms and workers.

Yet, and in order to survive, the proposed matching has to be stable (relative to the true preference profile) 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. Stability constitutes a minimal requirement that a matching has to fulfill if the assignment is voluntary rather than compulsory. The literature has considered stability of a matching to be its main characteristic in order to survive. Indeed, many of the successful mechanisms are stable despite the fact that 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

---


5 See, for instance, Roth (1984a) and Niederle and Roth (2003).
stable for the true profile. As a consequence, the literature has studied intensively Nash equilibria (NE) of direct preference revelation games induced by different stable mechanisms for a given true preference profile.⁶

Nevertheless all this strategic analysis presupposes that the true profile of preferences is both certain and common knowledge among all agents; the very definition of Nash equilibrium under complete information requires it. Indeed, participants in these markets perceive the outcome of the mechanism as being uncertain because the submitted preferences of the other participants are unknown. To model this uncertainty and to overcome the limitation of the complete information set up, we follow the Bayesian approach by assuming that participants share a common belief (or common prior); namely, nature selects a preference profile according to a commonly known probability distribution on the set of profiles. Each agent then becomes aware of his preference ordering (his type) and submits to the mechanism a preference ordering according to his strategy (which specifies a preference ordering for each type). Then, given a mechanism, a strategy profile (a list of individual strategies) translates the uncertainty that each agent faces about preference profiles into probability distributions on matchings and hence, on the set of potential partners. Thus, each type, when evaluating the consequences of submitting alternative ranked lists has to evaluate probability distributions (conditional on his type) on the set of potential partners.⁷

Since matching 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 that represents an agent’s preference ordering (his type), the submitted ranked list maximizes his expected utility in the


⁷Roth (1989) is the first to study Bayesian Nash equilibria (for a fixed utility representation of agents’ preferences) in one-to-one matching markets under incomplete information.
direct preference revelation game induced by the common belief and the mechanism.\footnote{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, 2008), and Ehlers and Massó (2007, 2015) have already used, among others, this ordinal equilibrium notion in two-sided matching markets.}

In two earlier papers we have studied strategic incentives (\textit{i.e.,} OBNE) induced by centralized stable matching mechanisms under incomplete information. Specifically Theorem 1 in Ehlers and Massó (2015) shows that, for many-to-one matching markets, there is a strong link between NE under complete information and OBNE under incomplete information. Namely, given a common belief, 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 belief, the submitted profile is a NE under complete information at the true profile in the direct preference revelation game induced by the stable mechanism. This result implies the former Theorem 1 in Ehlers and Massó (2007) which says that, for one-to-one matching markets, truth-telling is an OBNE in the Bayesian direct revelation game induced by a common belief and a stable mechanism if and only if the support of the common belief is contained in the set of profiles with a unique stable matching.

However, the proofs of the two results use responsive extensions of firms’ preferences that are not monotonic. And we think that, in many applications (for instance, those described at the beginning of this Introduction), monotonicity is the norm: firms consider that having all positions filled with acceptable workers is always strictly preferred to having some positions vacant. In particular, in settings where, although workers may have different individual qualities for a firm, those differences are not sufficiently important to compensate the loss of two acceptable workers with only one worker.

In this paper we first define monotonic NE and monotonic OBNE by requiring that when firms compare sets of workers (for NE) or probability distributions by means of the first order stochastic dominance criterion (for OBNE), they only use
monotonic responsive extensions. Secondly, we restrict attention only to truth-telling. In particular, in Theorem 1 here, we show that truth-telling is a monotonic OBNE in a stable mechanism under incomplete information only if the support of the common belief is contained in the set of profiles with a unique stable matching. We then present Example 1 showing that the link between complete and incomplete information of Theorem 1 in Ehlers and Massó (2015) is broken when firms evaluate subsets of workers only with monotonic responsive extensions of their rankings on individual workers. Specifically, in the many-to-one matching market of Example 1, truth-telling is a monotonic OBNE in the workers-optimal stable mechanism, and yet there is a preference profile in the support of the common belief at which truth-telling is not a monotonic NE in any stable mechanism (in particular, in the workers-optimal stable mechanism) under complete information at this profile. Hence, to hold, Theorem 1 in Ehlers and Massó (2015) requires that firms ought to have non-monotonic responsive extensions. Nevertheless, we show that in Example 1 truth-telling is not a monotonic OBNE in the firms-optimal stable mechanism. This fact raises the question of whether or not the link between incomplete and complete information holds for this stable mechanism, even when firms responsive preferences are restricted to be monotonic. Theorem 2 answers this question positively: given a common belief, truth-telling is a monotonic OBNE in the firms’ optimal stable mechanism under incomplete information if and only if the support of the common belief is contained in the set of all profiles where truth-telling is a monotonic NE in the firms’ optimal stable mechanism under complete information.

The paper is organized as follows. Section 2 describes the many-to-one matching market with responsive preferences. Section 3 introduces incomplete information and the notion of OBNE (and its monotonic variant), states and proves our two results (Theorems 2 and 3) and presents Example 1.
2 Many-To-One Matching Markets

2.1 Agents, Quotas, and Preferences

The agents of a many-to-one matching market consist of two disjoint sets, the set of firms \( F \) and the set of workers \( W \). A generic firm will be denoted by \( f \), a generic worker by \( w \), and a generic agent by \( v \in V \equiv F \cup W \). Both the set of workers \( W \) and the set of firms \( F \) are assumed to be finite. While workers can only work for at most one firm, firms may hire different numbers of workers. For each firm \( f \), there is a maximum number \( q_f \geq 1 \) of workers that \( f \) may hire, \( f \)'s quota. Let \( q = (q_f)_{f \in F} \) be the vector of quotas. Each worker \( w \) has a strict preference ordering \( P_w \) over \( F \cup \{\emptyset\} \), where \( \emptyset \) means the prospect of not being hired by any firm. Each firm \( f \) has a strict preference ordering \( P_f \) over \( W \cup \{\emptyset\} \), where \( \emptyset \) stands for leaving a position unfilled.\(^9\) A profile \( P = (P_v)_{v \in V} \) is a list of preference orderings, one for each agent.

To emphasize the preference orderings of a subset of agents \( S \subseteq V \) we often denote a profile \( P \) by \( (P_S, P_{-S}) \). 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 \( \mathcal{P}_{-v} \) denote the set \( \times_{v' \in V \setminus \{v\}} \mathcal{P}_{v'} \).

Since agent \( v \) might have to compare potentially the same partner, we denote by \( R_v \) the weak preference ordering corresponding to \( P_v \); namely, for \( v', v'' \in V \cup \{\emptyset\} \), \( v'R_v v'' \) means either \( v' = v'' \) or \( v'P_v v'' \). Momentarily fix a worker \( w \) and his preference ordering \( P_w \). Given \( v \in F \cup \{\emptyset\} \), let \( B(v, P_w) \) be 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., \( A(P_w) = \{f \in F \mid fP_w \emptyset\} \). Given a subset \( S \subseteq F \cup \{\emptyset\} \), let \( P_w | S \) denote the restriction of \( P_w \) to \( S \). Similarly, given \( P_f \in \mathcal{P}_f \), \( v \in W \cup \{\emptyset\} \),

\(^9\)Observe that firm \( f \)'s preference ordering is only defined on the set \( W \cup \{\emptyset\} \) in spite of the fact that firm \( f \) will be matched to a subset in \( 2^W \). Hence, \( f \) should be able to compare also subsets of workers with cardinality larger than one. For this reason, we will later extend any preference ordering \( P_f \) on \( W \cup \{\emptyset\} \) to a responsive preference \( P^*_f \) on \( 2^W \) (to be defined in Subsection 2.4). Although each preference \( P_f \) on \( W \cup \{\emptyset\} \) admits many responsive extensions on \( 2^W \), the set of stable matchings is invariant to the particular chosen responsive extensions.
and $S \subseteq W \cup \{\emptyset\}$, we define $B(v, P_f), A(P_f)$, and $P_f|S$. Given $P_f$ and $w, w' \in W$, let $P_f^{w \leftrightarrow w'}$ stand for the ordering where $w$ and $w'$ switch positions in $P_f$.\footnote{Formally, (i) $P_f^{w \leftrightarrow w'}|W \cup \{\emptyset\}\{w, w'\} = P_f|W \cup \{\emptyset\}\{w, w'\}$, (ii) $wP_fw'$ iff $w'P_f^{w \leftrightarrow w'}$, (iii) for all $v \in W \cup \{\emptyset\}\{w, w'\}$, $vP_fw$ iff $vP_f^{w \leftrightarrow w'}w'$, and (iv) for all $v \in W \cup \{\emptyset\}\{w, w'\}$, $vP_fw'$ iff $vP_f^{w \leftrightarrow w'}w$.} For example, if $P_f : w_1w_2w_3w_4\emptyset$, then $P_f^{w_1 \leftrightarrow w_3} : w_3w_2w_1w_4\emptyset$.\footnote{We will use the convention that $P_f : w_1w_2w_3w_4\emptyset$ means $w_1P_fw_2P_fw_3P_fw_4P_f\emptyset$.}

A many-to-one matching market (also known as a college admissions problem) is a quadruple $(F, W, q, P)$. Because $F$, $W$ and $q$ will often remain fixed, a problem will simply be 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 (also known as a marriage market).

### 2.2 Stable Matchings

The assignment problem consists of matching workers with firms keeping the bilateral nature of their relationship, complying with firms’ capacities given by their quotas, and allowing for the possibility that both workers and firms may remain unmatched. Formally, given a college admissions problem $(F, W, q, P)$, a matching $\mu$ is a mapping from the set $V$ to the set of all subsets of $V$ such that:

1. \(m_1\) for all $w \in W$, either $|\mu(w)| = 1$ and $\mu(w) \subseteq F$ or else $\mu(w) = \emptyset$;
2. \(m_2\) for all $f \in F$, $\mu(f) \subseteq W$ and $|\mu(f)| \leq q_f$; and
3. \(m_3\) $\mu(w) = \{f\}$ if and only if $w \in \mu(f)$.

Abusing notation, we will often write $\mu(w) = f$ instead of $\mu(w) = \{f\}$. If $\mu(w) = \emptyset$ we say that $w$ is unmatched at $\mu$ and if $|\mu(f)| < q_f$ we say that $f$ has $q_f - |\mu(f)|$ unfilled positions at $\mu$; $f$ is unmatched at $\mu$ when it has $q_f$ unfilled positions at $\mu$. Let $\mathcal{M}$ denote the set of all matchings.

Not all matchings are equally likely. Stability of a matching is considered to be its main characteristic in order to survive. A matching is stable if no agent is matched to
an unacceptable partner (individual rationality) and no unmatched worker-firm pair mutually prefers each other to (one of) their current assignments (pair-wise stability). That is, given a many-to-one matching market \( P \in \mathcal{P} \), a matching \( \mu \in \mathcal{M} \) is stable (at \( P \)) if

(s1) for all \( w \in W \), \( \mu(w) \mathrel{R}_w \emptyset \);

(s2) for all \( f \in F \) and all \( w \in \mu(f) \), \( w \mathrel{P}_f \emptyset \); and

(s3) there is no pair \((w, f) \in W \times F\) such that \( w \notin \mu(f) \), \( f \mathrel{P}_w \mu(w) \), and either \( w \mathrel{P}_f w' \)
   
   for some \( w' \in \mu(f) \) or \( w \mathrel{P}_f \emptyset \) if \( |\mu(f)| < q_f \).

Notice that this definition declares a matching to be stable if it is not blocked (in the sense of the core) by either individual agents or unmatched pairs. Gale and Shapley (1962) establishes that all college admissions problems have a non-empty set of stable matchings and Roth and Sotomayor (Theorem 3.3, 1990) states that larger coalitions do not have additional (weak) blocking power because the set of stable matchings coincides with the core. We denote by \( C(P) \) the non-empty set of stable matchings at \( P \) (or the core of \( P \)).

### 2.3 Stable Matching Mechanisms

Whether or not a matching is stable depends on the preference orderings of agents, and since they are private information, agents have to be asked about them. A mechanism requires each agent \( v \) to report some preference ordering \( P_v \) and associates a matching with any reported profile \( P \). Namely, a (direct revelation) mechanism is a function \( \varphi : \mathcal{P} \rightarrow \mathcal{M} \) mapping each preference profile \( P \in \mathcal{P} \) to a matching \( \varphi[P] \in \mathcal{M} \). Then \( \varphi[P](v) \) is the match of agent \( v \) at preference profile \( P \) under mechanism \( \varphi \).

Note that, for all \( w \in W \), \( \varphi[P](w) \in F \cup \{\emptyset\} \) and, for all \( f \in F \), \( \varphi[P](f) \in 2^W \). A mechanism \( \varphi \) is stable if for all \( P \in \mathcal{P} \), \( \varphi[P] \in C(P) \).

The most popular stable mechanisms are the two associated to the two deferred-acceptance algorithms (DA-algorithms) (Gale and Shapley, 1962): the firms-proposing
DA-algorithm, as a mechanism denoted by $DA_F$, and the workers-proposing DA-algorithm, as a mechanism denoted by $DA_W$. Fix a profile $P \in \mathcal{P}$. At any step of the algorithm in which workers make offers, each worker $w$ proposes to the most-preferred firm among the set of firms that have not already rejected $w$ during previous steps, while a firm $f$ accepts the $q_f$-most preferred workers among the set of current offers plus the set of workers provisionally matched to $f$ in the previous step (if any). The algorithm stops at the step when either all offers are accepted or workers have no more acceptable firms to whom they want to make an offer; the provisional matching becomes then definite and is denoted by $DA_W[P]$. At any step of the algorithm in which firms make offers, each firm $f$ proposes itself to set of $q_f$-most preferred workers who have not already rejected $f$ during the previous steps, while each worker $w$ accepts the offer of the best firm among the set of current offers plus the one made by the firm provisionally matched in the previous step (if any). The algorithm stops at the step at which all offers are accepted; the provisional matching becomes then definite and is denoted by $DA_F[P]$.

2.4 Responsive Extensions

The notion of a mechanism in which firms (like workers) only submit rankings on individual agents fits with most of the mechanisms used in real life centralized matching markets. But a mechanism matches each firm $f$ to a set of workers, taking into account only $f$‘s preference ordering $P_f$ over individual workers. Thus, to study firms’ incentives in direct revelation mechanisms, preference orderings of firms over individual workers have to be extended to preference orderings over subsets of workers. But

\[\text{It is well-known that for each profile } P \in \mathcal{P}, \ (i) \ DA_W[P] \text{ matches each worker } w \text{ with } w\text{'s best partner, among all partners that } w \text{ is matched to across all stable matchings, and (ii) } DA_F[P] \text{ matches each firm } f \text{ with } f\text{'s best set of partners—according to any responsive extension of } P_f \text{ (see Subsection 2.4 below for the definition of responsiveness)—, among all partners that } f \text{ is matched to across all stable matchings. This is why they are named workers-optimal and firms-optimal stable matchings, respectively.}\]
a firm $f$ may have different rankings over subsets of workers respecting its quota $q_f$ and the ranking $P_f$ over individual workers. For instance, let $W = \{w_1, w_2, w_3, w_4\}$ be the set of workers and let $P_f$ be such that $P_f : w_1 w_2 w_3 w_4 \emptyset$ and $q_f = 2$. While it is reasonable to assume that, under the absence of very strong complementarities among workers, the set $\{w_1, w_2\}$ is preferred by $f$ to the set $\{w_3, w_4\}$ or to the set $\{w_1, w_3\}$, firm $f$’s preference between the sets $\{w_1, w_4\}$ and $\{w_2, w_3\}$ is ambiguous since $P_f$ does not convey this information. Following the literature, we will only require these extensions to be responsive in the sense that replacing a worker in a set (or an unfilled position) by a better worker (or an acceptable worker) makes a set more preferred; for example, in all responsive extensions $\{w_1, w_2\}$ is preferred to $\{w_1\}$, to $\{w_3, w_4\}$ and to $\{w_1, w_3\}$ but for some responsive extensions $\{w_1, w_4\}$ is preferred to $\{w_2, w_3\}$ while for other responsive extensions $\{w_2, w_3\}$ is preferred to $\{w_1, w_4\}$. For later purposes we also introduce the notion of monotonic responsive extensions under which having all positions filled with acceptable workers is always strictly preferred to having some positions vacant.

**Definition 1 (Responsive Extensions)**

(a) The preference extension $P_f^*$ over $2^W$ is responsive to the preference ordering $P_f$ over $W \cup \{\emptyset\}$ if for all $S \in 2^W$, all $w \in S$, and all $w' \notin S$:

(r1) $|S| > q_f$ implies $\emptyset P_f^* S$.

(r2) $S \cup \{w'\} P_f^* S$ if and only if $|S| < q_f$ and $w' P_f \emptyset$.

(r3) $S \cup \{w'\} P_f^* S \setminus \{w\}$ if and only if $w' P_f w$.

(b) We say that a responsive extension $P_f^*$ of $P_f$ 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 $SP_f^* S'$.

Observe that responsive extensions are not necessarily monotonic. For instance, let $W = \{w_1, w_2, w_3\}$ be the set of workers and consider a firm $f$ with $q_f = 2$ and

---

13See for instance, Roth and Sotomayor (1990).
$P_f : w_1w_2w_3\emptyset$. Responsiveness does not rule out that worker $w_1$ may be more desirable than workers $w_2$ and $w_3$ together. For instance, the ordering

$$\{w_1, w_2\}P_f^*\{w_1, w_3\}P_f^*\{w_2, w_3\}P_f^*\{w_3\}P_f^*\emptyset$$

is responsive to $P_f$ but not monotonic since $\{w_1\}P_f^*\{w_2, w_3\}$ and $\{w_2, w_3\} \subseteq A(P_f)$. For some applications monotonicity is meaningful; for instance, in matching problems where leaving positions unfilled is very costly.

Given a responsive extension $P_f^*$ of $P_f$, let $R_f^*$ denote its corresponding weak preference ordering on $2^W$. Moreover, given $S \in 2^W$, let $B(S, P_f^*)$ be the weak upper contour set of $P_f^*$ at $S$; i.e., $B(S, P_f^*) = \{S' \in 2^W \mid S'R_f^*S\}$. Given $P_f \in \mathcal{P}_f$, we denote by $\text{resp}(P_f)$ and $\text{mresp}(P_f)$ the sets of responsive and monotonic responsive extensions of $P_f$, respectively.

### 2.5 Nash Equilibrium under Complete Information

Clearly any mechanism $\varphi : \mathcal{P} \rightarrow \mathcal{M}$ and any true profile $P \in \mathcal{P}$ define a direct (ordinal) preference revelation game under complete information for which we can define the natural (ordinal) notions of Nash equilibrium and monotonic Nash equilibrium.

**Definition 2 (Nash Equilibrium)**

(a) Let $P \in \mathcal{P}$ be a profile. A profile $P'$ is a Nash equilibrium (NE) under complete information $P$ in the direct preference revelation game induced by the mechanism $\varphi$ if 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 for all $f \in F$ and all $P_f^* \in \text{resp}(P_f)$, $\varphi[P'_f](f)R_f^*\varphi[\hat{P}_f, P'_{-f}](f)$ for all $\hat{P}_f \in \mathcal{P}_f$.

(b) Let $P \in \mathcal{P}$ be a profile. A profile $P'$ is a monotonic Nash equilibrium (monotonic NE) under complete information $P$ in the direct preference revelation game induced by the mechanism $\varphi$ if 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 for all $f \in F$ and all $P_f^* \in \text{mresp}(P_f)$, $\varphi[P'_f](f)R_f^*\varphi[\hat{P}_f, P'_{-f}](f)$ for all $\hat{P}_f \in \mathcal{P}_f$. 

12
A large literature on matching studies Nash equilibrium under complete information in direct preference revelation games induced by stable mechanisms; in particular, for the mechanisms $DA_W$ and $DA_F$. However, for many applications the assumption that the true profile is common knowledge is extremely unrealistic. In the next section we present an ordinal way of studying Bayesian Nash Equilibria under incomplete information, which has already been used (see Ehlers and Massó (2007, 2015), for instance).

3 Incomplete Information

3.1 Preliminaries and OBNE

We depart from the very strong assumption that the true profile is common knowledge and consider the Bayesian direct preference revelation game induced by a mechanism and a belief about the true profile, which is shared among all agents. A common belief is a probability distribution $\tilde{P}$ over $\mathcal{P}$. Given a profile $P$ and the common belief $\tilde{P}$, $\Pr\{\tilde{P} = P\}$ is the probability that $\tilde{P}$ assigns to the event that the true profile is $P$.

Given $v \in V$, let $\tilde{P}_v$ denote the marginal distribution of $\tilde{P}$ over $\mathcal{P}_v$. Observe that, following the Bayesian approach, the common belief $\tilde{P}$ describes agents’ uncertainty about the true profile before agents learn their types. Now, given a common belief $\tilde{P}$ and a preference ordering $P_v$ (agent $v$’s type), 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 uncertainty about the preferences of the other agents, given that his preference ordering is $P_v$. This formulation does not require symmetry nor independence of beliefs; conditional beliefs might be very correlated if agents use similar sources to form them (i.e., rankings, grades, recommendation letters, etc.).

An agent with incomplete information about the others’ preference orderings

---

14Strictly speaking $\tilde{P}$ cannot be set equal to $P$ because $\tilde{P}$ is not a random variable but a probability distribution on $\mathcal{P}$. However, for convenience we use this notation as if $\tilde{P}$ were a random variable.
(more importantly, about their submitted lists) will perceive the outcome of a mechanism as being uncertain. A random matching \(\tilde{\eta}\) is a probability distribution over the set of matchings \(\mathcal{M}\). Given a matching \(\mu\) and the random matching \(\tilde{\eta}\), \(\Pr\{\tilde{\eta} = \mu\}\) is the probability that \(\tilde{\eta}\) assigns to matching \(\mu\). But the uncertainty important for agent \(v\) is not over matchings but over \(v\)'s set of potential partners. Let \(\tilde{\eta}(w)\) denote the probability distribution which \(\tilde{\eta}\) induces over worker \(w\)'s set of potential partners \(F \cup \{\emptyset\}\) and let \(\tilde{\eta}(f)\) denote the probability distribution which \(\tilde{\eta}\) induces over firm \(f\)'s set of potential partners \(2^W\). Namely, for \(w \in W\) and all \(v \in F \cup \{\emptyset\}\),

\[
\Pr\{\tilde{\eta}(w) = v\} = \sum_{\{\mu \in \mathcal{M}|\mu(w) = v\}} \Pr\{\tilde{\eta} = \mu\}
\]

and for \(f \in F\) and all \(S \in 2^W\),

\[
\Pr\{\tilde{\eta}(f) = S\} = \sum_{\{\mu \in \mathcal{M}|\mu(f) = S\}} \Pr\{\tilde{\eta} = \mu\}.
\]

A mechanism \(\varphi\) and a common belief \(\tilde{P}\) define a direct (ordinal) preference revelation game under incomplete information as follows. Before submitting a list to the mechanism, agents learn their types. Thus, a strategy of agent \(v\) is a function \(s_v : \mathcal{P}_v \to \mathcal{P}_v\) specifying for each type of agent \(v\), \(P_v\), a list that \(v\) submits to the mechanism, \(s_v(P_v)\).\(^{15}\) 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 belief \(\tilde{P}\) 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 each \(\mu \in \mathcal{M}\),

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

is the probability of matching \(\mu\). But after learning their types, agents also update their beliefs. Hence, the relevant random matching for agent \(v\), given his type \(P_v\) and

\(^{15}\)Ehlers and Massó (2015) contains the analysis when agents may use mixed strategies, and shows that the addition of the randomness coming from mixed strategies does not change the results. Hence, we restrict here our analysis to pure strategies.
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 $P_v$). But again, the relevant uncertainty that agent $v$ faces is given by $\varphi[s_v(P_v), s_{-v}(\tilde{P}_{-v}|P_v)](v)$, the probability distribution which the random matching $\varphi[s_v(P_v), s_{-v}(\tilde{P}_{-v}|P_v)]$ induces over $v$’s set of potential partners. Observe that this uncertainty is held by each agent $v$, given his type $P_v$, before the mechanism proposes a matching. Then, after collecting the preference profile $s(P)$, the mechanism $\varphi$ proposes a unique matching $\varphi[s(P)]$, the one that agents have to carry out. Our stability condition on the mechanism applies to this proposed matching (i.e., it is the standard notion of stability under complete information). This is why we do not need an ex ante notion of stability relative to the incomplete information environment when the partners are still uncertain and no particular matching has been proposed yet.

To evaluate the consequences of their declared preference orderings agents will compare their induced probability distributions on the set of their potential partners according to the first-order stochastic dominance criterion.

**Definition 3 (First-Order Stochastic Dominance)**

(fo1) A random matching $\tilde{\eta}$ first-order stochastically $P_w$—dominates a random matching $\tilde{\eta}'$, denoted by $\tilde{\eta}(w) \succ_P w \tilde{\eta}'(w)$, if for all $v \in F \cup \{\emptyset\}$,

$$\sum_{\{v' \in F \cup \{\emptyset\} | v' \sim R_{w,v}\}} \Pr\{\tilde{\eta}(w) = v'\} \geq \sum_{\{v' \in F \cup \{\emptyset\} | v' \sim R_{w,v}\}} \Pr\{\tilde{\eta}'(w) = v'\}.$$  

(fo2) A random matching $\tilde{\eta}$ first-order stochastically $P_f$—dominates a random matching $\tilde{\eta}'$, denoted by $\tilde{\eta}(f) \succ_P f \tilde{\eta}'(f)$, if for all $P_f^* \in \text{resp}(P_f)$ and all $S \in 2^W$,\footnote{Observe that this definition requires that $\tilde{\eta}$ first-order stochastically dominates $\tilde{\eta}'$ according to all responsive extensions of $P_f$. Note that this requirement is meaningful since the clearinghouse observes firms’ rankings over individual workers only and not which responsive extension they use to compare sets of workers.}

$$\sum_{\{S' \in 2^W | S \sim R_f S\}} \Pr\{\tilde{\eta}(f) = S'\} \geq \sum_{\{S' \in 2^W | S \sim R_f S\}} \Pr\{\tilde{\eta}'(f) = S'\}. \quad (1)$$
All mechanisms used in centralized matching markets are ordinal. In other words the only information available for a clearinghouse are the agents’ ordinal preferences over potential partners. In such an environment a strategy profile is an ordinal Bayesian Nash equilibrium whenever, for any agent’s true ordinal preference, submitting the rank list specified by his strategy maximizes his expected utility for every vNM-utility representation of his true preference. This requires that an agent’s strategy only depends on the ordinal ranking induced by his vNM-utility function (if any). Moreover, ordinal strategies are well-defined if an agent only observes his ordinal ranking and may have (still) little information about his utilities of his potential partners. Below we define the notions of ordinal Bayesian Nash equilibrium

**Definition 3 (Ordinal Bayesian Nash Equilibrium)** Let \( \hat{P} \) be a common belief. Then a strategy profile \( s \) is an ordinal Bayesian Nash equilibrium (OBNE) in the mechanism \( \varphi \) under incomplete information \( \hat{P} \) if and only if for all \( v \in V \) and all \( P_v \in \mathcal{P}_v \) such that \( \Pr\{\hat{P}_v = P_v\} > 0 \),

\[
\varphi[s_v(P_v), s_{-v}(\hat{P}_{-v}|P_v)](v) \succeq_{P_v} \varphi[P_v', s_{-v}(\hat{P}_{-v}|P_v)](v) \quad \text{for all} \ P_v' \in \mathcal{P}_v. \tag{2}
\]

Similarly, a strategy profile \( s \) is a monotonic ordinal Bayesian Nash equilibrium (monotonic OBNE) in the mechanism \( \varphi \) under incomplete information \( \hat{P} \) if and only if condition (2) holds for all \( w \in W \) but, for all \( f \in F \), (2) is replaced by

\[
\varphi[s_f(P_f), s_{-f}(\hat{P}_{-f}|P_f)](f) \succ_{P_f}^{\text{mon}} \varphi[P_f', s_{-f}(\hat{P}_{-f}|P_f)](f) \quad \text{for all} \ P_f' \in \mathcal{P}_f, \tag{2}
\]

where \( \succ_{P_f}^{\text{mon}} \) means that condition (1) in (fo2) holds only for all \( P_f^* \in \text{mres}_P(P_f) \).

Theorem 1 in Ehlers and Massó (2015) characterize OBNE in any stable mechanism under incomplete information \( \hat{P} \) by means of the NE in the stable mechanism under complete information.

**Theorem 1 (Ehlers and Massó, 2015, Theorem 1)** Let \( \hat{P} \) be a common belief, \( s \) be a strategy profile, and \( \varphi \) be a stable mechanism. Then, \( s \) is an OBNE in the stable mechanism \( \varphi \) under incomplete information \( \hat{P} \) if and only if for any profile \( P \)
in the support of \( \tilde{P} \), \( s(P) \) is a Nash equilibrium under complete information \( P \) in the direct preference revelation game induced by \( \varphi \).

### 3.2 Truth-telling with Monotonicity

We will focus on truth-telling. Ehlers and Massó (2007, Corollary 5) shows that truth-telling is an OBNE in a stable mechanism under incomplete information \( \tilde{P} \) if and only if the support of \( \tilde{P} \) is contained in the set of profiles with a singleton core. This is an immediate consequence of Theorem 1.

Here, we first show that for truth-telling to be a monotonic OBNE in a stable mechanism it is still necessary that the common belief puts positive probability only on profiles with singleton core. Therefore, even in monotonic matching markets the core is unique for any realized (or observed) profile if truth-telling is an equilibrium. Note, however, that here we cannot use Theorem 1 because we only allow for monotonic responsive extensions.

**Theorem 2** Let \( \tilde{P} \) be a common belief and assume that truth-telling is a monotonic OBNE in a stable mechanism under incomplete information \( \tilde{P} \). Then, the support of \( \tilde{P} \) is contained in the set of all profiles with a singleton core.

**Proof.** Let \( \varphi \) be a stable mechanism. To obtain a contradiction, suppose that there is some \( P \in \mathcal{P} \) such that both \( \Pr\{\tilde{P} = P\} > 0 \) and \( |C(P)| \geq 2 \). By stability of \( \varphi \), \( \varphi[P] \in C(P) \). Let \( \mu \in C(P) \setminus \{\varphi[P]\} \). If there is some \( w \in W \) such that \( \mu(w)P_w[\varphi[P](w) \), then similarly to Ehlers and Massó (2007) we can show that truth-telling cannot be a monotonic OBNE. Thus, \( \varphi[P](w)R_w\mu(w) \) for all \( w \in W \). since \( DA_W[P] \) is the workers-optimal stable matching,

\[
\varphi[P] = DA_W[P]. \tag{3}
\]

Obviously, (3) needs to hold for all profiles belonging to the support of \( \tilde{P} \).

Since \( \mu \neq \varphi[P] \) and \( \varphi[P] = DA_W[P] \), there exists some \( f \in F \) such that \( \mu(f) \neq \varphi[P](f) \). Then, by Roth and Sotomayor (1990, Theorems 5.12 and 5.13), \( |\mu(f)| = \)
\[ |\varphi[P](f)| = q_f. \] By Roth and Sotomayor (1989, Theorem 4), we have \( wP-fw' \) for all \( w \in \mu(f) \) and all \( w' \in \varphi[P](f) \backslash \mu(f) \). Thus,

\[ \mu(f)P_f^*\varphi[P](f) \]

for all monotonic responsive extensions \( P_f^* \) of \( P_f \).

Let \( \bar{w} \in \mu(f) \) be such that \( wR_f\bar{w} \) for all \( w \in \mu(f) \). Let \( P_f' \in \mathcal{P}_f \) be such that (i) \( A(P_f') = B(\bar{w}, P_f) \) and (ii) \( P_f'|A(P_f') = P_f|A(P_f') \). By construction of \( P_f' \), \( \mu \in C(P_f', P_{-f}) \). Hence, \( \varphi[P_f', P_{-f}](f) = \mu(f) \) or if \( \varphi[P_f', P_{-f}](f) \neq \mu(f) \), then, again by Roth and Sotomayor (1989, Theorem 4), for all \( w \in \varphi[P_f', P_{-f}](f) \) and all \( w' \in \mu(f) \backslash \varphi[P_f', P_{-f}](f) \), \( wP_f'w' \). Thus, by \( P_f'|A(P_f') = P_f|A(P_f') \) and (4),

\[ \varphi[P_f', P_{-f}](f)P_f^*\varphi[P](f) \]

for all monotonic responsive extensions \( P_f^* \) of \( P_f \).

Let \( P_f^* \) be the monotonic responsive extension of \( P_f \) such that for all \( W' \in 2^W \), \( W'P_f^*\mu(f) \) if and only if (i) \( |W'| = q_f \) (note that \( |\mu(f)| = q_f \)) and (ii) there is a one-to-one mapping \( h : W' \rightarrow \mu(f) \) such that \( w'R_fh(w') \) for all \( w' \in W' \) with strict preference holding for some worker in \( W' \). In other words, \( W' \) is strictly preferred to \( \mu(f) \) under \( P_f^* \) only if \( W' \) is unambiguously strictly preferred to \( \mu(f) \) for all monotonic responsive extensions of \( P_f \).

Let \( P_{-f}^* \in \mathcal{P}_{-f} \) be such that both \( \Pr\{\hat{P} = (P_f, P_{-f})\} > 0 \) and \( \varphi[P_f, P_{-f}^*](f) \in B(\mu(f), P_f^*) \). We want to show that \( \varphi[P_f', P_{-f}^*](f) \in B(\mu(f), P_f^*) \). Let \( \hat{\mu} = \varphi[P_f, P_{-f}^*] \) and \( \mu' = \varphi[P_f', P_{-f}^*] \). By the definition of \( P_f^* \) and \( P_f' \), we have that \( \hat{\mu}(f)R_f^*\mu(f) \), \( \hat{\mu}(f) \subseteq A(P_f') \) and \( |\hat{\mu}(f)| = q_f \). Thus, by \( A(P_f') = B(\bar{w}, P_f) \) and \( P_f'|A(P_f') = P_f|A(P_f') \), \( \hat{\mu} \in C(P_f', P_{-f}^*) \). Since \( \mu', \hat{\mu} \in C(P_f', P_{-f}^*) \), Theorem 4 of Roth and Sotomayor (1989) tells us that one of the following two cases needs to hold.

**Case 1:** \( w'R_f\hat{w} \) for all \( w' \in \mu'(f) \) and \( \hat{w} \in \hat{\mu}(f) \backslash \mu'(f) \).

Thus, \( \mu'(f)R_f^*\hat{\mu}(f) \) and since \( \hat{\mu}(f)R_f^*\mu(f) \), \( \mu'(f)R_f^*\mu(f) \). Hence, \( \varphi[P_f', P_{-f}^*](f) \in B(\mu(f), P_f^*) \).
Case 2: $\hat{w}R_f w'$ for all $\hat{w} \in \mu(f)$ and $w' \in \mu'(f) \backslash \hat{\mu}(f)$.

If $\hat{\mu}(f) \neq \mu'(f)$, then

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

But then $\mu'(f) \subseteq A(P_f') = B(\hat{w}, P_f) \subseteq A(P_f)$ and by $P_f'|A(P_f') = P_f|A(P_f)$, we have $\mu' \in C(P_f, P_f')$. Note that $(P_f, P_f')$ is a profile belonging to the support of $\tilde{P}$ and thus, by (3), $\hat{\mu} = \varphi[P_f, P_{-f}'] = DA_W[P_f, P_{-f}']$. Now (6) contradicts the facts that $DA_W$ chooses the stable matching in $C(P_f, P_{-f}')$ which is worst for the firms and $\mu' \in C(P_f, P_{-f})$.

Thus, $\hat{\mu}(f) = \mu'(f)$ and $\mu'(f)R_f^* \hat{\mu}(f)$. Since $\hat{\mu}(f)R_f^* \mu(f)$, $\mu'(f)R_f^* \hat{\mu}(f)$. Hence, $\varphi[P_f', P_{-f}'](f) \in B(\mu(f), P_f')$.

Cases 1 and 2 show that if $P_{-f}' \in \mathcal{P}_{-f}$ is such that $Pr\{\tilde{P} = (P_f, P_{-f}')\} > 0$ and $\varphi[P_f, P_{-f}'](f) \in B(\mu(f), P_f')$, then $\varphi[P_f', P_{-f}'](f) \in B(\mu(f), P_f')$. Since $Pr\{\tilde{P}_{-f'}|P_f = P_{-f}\} > 0$, $\varphi[P](f) \notin B(\mu(f), P_f')$, and $\varphi[P_f', P_{-f}'](f) \in B(\mu(f), P_f')$, it follows that

$$Pr\{\varphi[P_f', \tilde{P}_{-f'}|P_f](f) \in B(\mu(f), P_f')\} > Pr\{\varphi[P_f, \tilde{P}_{-f'}|P_f](f) \in B(\mu(f), P_f')\},$$

which means truth-telling is not a monotonic OBNE in the stable mechanism $\varphi$. ■

It is natural to ask when the link in Theorem 1 breaks in many-to-one matching markets. We will provide an example of a market in which our former main result (stated here as Theorem 1) is not true. More precisely, we will see that Theorem 1 depends on firms having responsive extensions which are not monotonic. In markets where workers are close substitutes, we may restrict the set of responsive extension firms are allowed to have. More precisely, in environments where it is (very) costly to leave some positions unfilled, responsive extensions are necessarily monotonic. For instance, in medical markets for a certain speciality hospitals may not be able to provide full medical service in case a position is left vacant. In such environments it is natural to assume that firms have only monotonic responsive preferences and focus on monotonic OBNE.
Recall that Roth (1985) exhibits an example of a profile with a singleton core where a firm profitably manipulates any stable mechanism and truth-telling is not a NE under complete information. Below we exhibit an example of a many-to-one matching market where (i) truth-telling is a monotonic OBNE for the stable mechanism $DA_W$ (and any profile in the support has a singleton core) and (ii) there is a profile belonging to the support of the common belief at which truth-telling is not a monotonic NE under complete information. Therefore, the link in Theorem 1 is broken in one direction\textsuperscript{18} when all firms have only monotonic responsive extensions and one firm has a capacity of at least two.

**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 $\bar{P}$ with $\Pr\{\bar{P} = P\} = p$ and $\Pr\{\bar{P} = \bar{P}\} = 1 - p$, where $p < 1/2$, and $P$ and $\bar{P}$ are the following profiles:

<table>
<thead>
<tr>
<th>$P_{f_1}$</th>
<th>$P_{f_2}$</th>
<th>$P_{f_3}$</th>
<th>$P_{w_1}$</th>
<th>$P_{w_2}$</th>
<th>$P_{w_3}$</th>
<th>$P_{w_4}$</th>
<th>$\bar{P}_{f_1}$</th>
<th>$\bar{P}_{f_2}$</th>
<th>$\bar{P}_{f_3}$</th>
<th>$\bar{P}_{w_1}$</th>
<th>$\bar{P}_{w_2}$</th>
<th>$\bar{P}_{w_3}$</th>
<th>$\bar{P}_{w_4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_1$</td>
<td>$w_1$</td>
<td>$w_3$</td>
<td>$f_3$</td>
<td>$f_2$</td>
<td>$f_1$</td>
<td>$f_1$</td>
<td>$w_1$</td>
<td>$w_2$</td>
<td>$w_4$</td>
<td>$f_1$</td>
<td>$f_2$</td>
<td>$f_1$</td>
<td>$f_3$</td>
</tr>
<tr>
<td>$w_2$</td>
<td>$w_2$</td>
<td>$w_1$</td>
<td>$f_1$</td>
<td>$f_1$</td>
<td>$f_3$</td>
<td>$f_2$</td>
<td>$w_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$w_3$</td>
<td>$w_3$</td>
<td>$w_2$</td>
<td>$f_2$</td>
<td>$f_3$</td>
<td>$f_2$</td>
<td>$f_3$</td>
<td>$w_3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$w_4$</td>
<td>$w_4$</td>
<td>$w_4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$w_4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note that $P_{f_1} = \bar{P}_{f_1}$. It is straightforward to verify that both profiles have a singleton 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 $\varphi$ be a stable mechanism. Thus, by stability of $\varphi$, $\varphi[P] = \mu$ and $\varphi[\bar{P}] = \bar{\mu}$. First we will show that for the profile $P$ truth-telling is not a monotonic Nash equilibrium.

\textsuperscript{18}It is clear that the other direction is true for any arbitrary game of incomplete information.
under complete information $P$ in the direct preference revelation game induced by $\varphi$. Let $P'_{f_1} \in \mathcal{P}_{f_1}$ be such that $P'_{f_1} : w_1 w_2 w_4 w_3$. 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 $\varphi$, $\varphi[P'_{f_1}, P_{-f_1}] = \mu'$. Obviously, for all responsive (and hence, monotonic) extensions $P^*_f$ of $P_f$ we have $\{w_1, w_4\}P^*_f \{w_3, w_4\}$, which is equivalent to $\varphi[P'_{f_1}, P_{-f_1}](f_1)P^*_f \varphi[P](f_1)$. Therefore, truth-telling is not a monotonic Nash equilibrium in any stable mechanism $\varphi$ under complete information $P$ (and profile $P$ belongs to the support of the common belief $\bar{P}$).

On the other hand we will show that truth-telling is a monotonic OBNE in the stable mechanism $DA_W$ under incomplete information $\bar{P}$. Note that for all $v \in V \setminus \{f_1\}$, 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 that $v$ cannot gain by a deviation.

Next we consider firm $f_1$. All arguments except for the last one apply to any stable mechanism $\varphi$. Observe that $P_{f_1} = \bar{P}_{f_1}$ and the random matching $\varphi[P_{f_1}, \bar{P}_{-f_1}|P_{f_1}]$ 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 $P^*_f$ be a monotonic responsive extension of $P_{f_1}$ and $P''_{f_1} \in \mathcal{P}_{f_1}$. We show that

$$
\varphi[P_{f_1}, \bar{P}_{-f_1}|P_{f_1}](f_1) > P^*_f \varphi[P''_{f_1}, \bar{P}_{-f_1}|P_{f_1}](f_1).
$$

(7)

We distinguish two cases. First, suppose that $|\varphi[P''_{f_1}, P_{-f_1}](f_1)| = 1$ or $|\varphi[P''_{f_1}, \bar{P}_{-f_1}](f_1)| = 1$. Now if (7) does not hold, then by monotonicity of $P^*_f$ 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 must have that $\varphi[P''_{f_1}, P_{-f_1}](f_1)P^*_f \{w_1, w_3\}$ or $\varphi[P''_{f_1}, \bar{P}_{-f_1}](f_1)P^*_f \{w_1, w_3\}$. Obviously, from the definition of $\bar{P}_{-f_1}$, the last is impossible. Thus, $\varphi[P''_{f_1}, P_{-f_1}](f_1)P^*_f \{w_1, w_3\}$ and, by responsiveness of $P^*_f$, we must 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 sta-
ble matching from the firms’ point of view). 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).

Second, suppose that \(|\varphi[P''_{f_1}, P_{-f_1}](f_1)| = 2\) and \(|\varphi[P''_{f_1}, P_{-f_1}](f_1)| = 2\). Then by definition of \(P_{-f_1}\) and \(\varphi[P''_{f_1}, P_{-f_1}](f_1)| = 2\), we must have \(\varphi[P''_{f_1}, P_{-f_1}](f_1) = \{w_1, w_3\}\). Thus, by stability of \(\varphi\), \(\{w_1, w_3\} \subseteq A(P''_{f_1})\).

If 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 (7) holds (because when submitting \(P''_{f_1}\), firm \(f_1\) is matched with probability \(p\) to \(\{w_3\}\) and with probability \(1 - p\) to \(\{w_1, w_3\}\)).

If 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 \(\mu''\) at \((P''_{f_1}, P_{-f_1})\) if \(\mu''(w_4) = f_2\) and the pair \((w_1, f_3)\) would block \(\mu''\) 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, (7) holds for the stable mechanism \(DA_W\).

Note that Example 1 does not contradict Theorem 1. When considering the non-monotonic responsive extension \(P_{f_1}^*\) such that \(\{w_1\}P_{f_1}^*\{w_3, w_4\}\), then firm \(f_1\) gains 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 \(\varphi[\hat{P}_{f_1}, P_{-f_1}](f_1) = \{w_1\}\), which means that truth-telling is not an OBNE in any stable mechanism \(\varphi\) under incomplete information \(\hat{P}\).

---

19If \(DA_F[P''_{f_1}, P_{-f_1}](f_1) \neq \{w_1, w_2\}\), then choose \(P''_{f_1}\) such that \(A(P''_{f_1}) = \{w_1, w_2\}\). Then we obtain \(DA_F[P''_{f_1}, P_{-f_1}](f_1) = \{w_1, w_2\}\).

20Their result says that in a marriage 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).
Furthermore, in Example 1 truth-telling is not a monotonic OBNE in $DA_F$ under incomplete information $\bar{P}$. To see that, consider the preference $P''_{f_1} : w_1w_2w_4w_3\emptyset$ in Example 1. We have

$$DA_F[P''_{f_1}, P_{-f_1}] = \begin{pmatrix} f_1 & f_2 & f_3 \\ w_1, w_4 & w_2 & w_3 \end{pmatrix}$$

and

$$DA_F[P''_{f_1}, \bar{P}_{-f_1}] = \begin{pmatrix} f_1 & f_2 & f_3 \\ w_1, w_3 & w_2 & w_4 \end{pmatrix}.$$

Thus, since

$$\Pr\{DA_F[P''_{f_1}, P_{-f_1}|P_{f_1}](f_1) \in B(\{w_1, w_4\}, P^*_f)\} = 1$$

$$\Pr\{DA_F[P_{f_1}, P_{-f_1}|P_{f_1}](f_1) \in B(\{w_1, w_4\}, P^*_f)\} = 1 - p$$

for all (monotonic) responsive extensions $P^*_f$ of $P_{f_1}$ (i.e. this is an unambiguous deviation for firm $f_1$ because it does not depend on the choice of the responsive extension of $P_f$). This means that truth-telling is not a monotonic OBNE in $DA_F$ under incomplete information $\bar{P}$.

One would think that a naïve way to show that in $DA_F$ truth-telling is a monotonic OBNE (but not for some profile in the support) by enlarging the above example and adding the profile $P''_{w_3\leftrightarrow w_4}$ (where the roles of $w_3$ and $w_4$ are exchanged in $P_{-f_1}$) to the support of $\bar{P}_{-f_1}|P_{f_1}$. By symmetry, then we would have $DA_F[P''_{f_1}, P''_{w_3\leftrightarrow w_4}](f_1) = \{w_3, w_4\}$ and $DA_F[P_{f_1}, P''_{w_3\leftrightarrow w_4}](f_1) = \{w_1, w_3\}$, meaning that under profile $P''_{w_3\leftrightarrow w_4}$ firm $f_1$ strictly prefers reporting $P_{f_1}$ to reporting $P''_{f_1}$. Observe that since $DA_F[P_{f_1}, P''_{w_3\leftrightarrow w_4}]$ is individually rational at $(P_{f_1}, P''_{w_3\leftrightarrow w_4})$ and it matches each worker $w$ to $w$’s best firm, $DA_F[P_{f_1}, P''_{w_3\leftrightarrow w_4}] \in C(P_{f_1}, P''_{w_3\leftrightarrow w_4})$. However, this would then give an immediate contradiction to Theorem 2: $DA_F[P''_{f_1}, P''_{w_3\leftrightarrow w_4}], DA_F[P_{f_1}, P''_{w_3\leftrightarrow w_4}] \in C(P_{f_1}, P''_{w_3\leftrightarrow w_4})$ implies that the core is not a singleton set under profile $(P_{f_1}, P''_{w_3\leftrightarrow w_4})$ which belongs to the support of $\bar{P}$.

### 3.3 The Link for Truth-telling in the $DA_F$

In this subsection we establish that if agents truth-tell in the $DA_F$ stable mechanism, the strong link of Theorem 1 between complete and incomplete information carries over to monotonic matching markets.
Theorem 3 Let $\tilde{P}$ be a common belief. Then, truth-telling is a monotonic OBNE in $DA_F$ under incomplete information $\tilde{P}$ if and only if the support of $\tilde{P}$ is contained in the set of all profiles where truth-telling is a monotonic NE in $DA_F$ under complete information.

3.3.1 Proof of Theorem 3

Let $\tilde{P}$ be a common belief.

($\Leftarrow$) Suppose that for all profiles in the support of $\tilde{P}$, truth-telling is a monotonic NE in $DA_F$ under complete information. Let $P$ be such that $\Pr\{\tilde{P} = P\} > 0$. Then, for all $w \in W$ and all $P'_w \in \mathcal{P}_w$, $DA_F[P](w)R_wDA_F[P'_w, P_{-w}](w)$ and for all $f \in F$ and all $P'_f \in \mathcal{P}_f$, $DA_F[P](f)R_fDA_F[P'_f, P_{-f}](f)$ for all $P'_f \in mres(P_f)$. Hence, for all $w \in W$ and all $P_w \in \mathcal{P}_w$ such that $\Pr\{\tilde{P}_w = P_w\} > 0$, we have that for all $P'_w \in \mathcal{P}_w$,

$$DA_F[P_w, \tilde{P}_{-w}\mid P_w](w) \succ_{P_w} DA_F[P'_w, \tilde{P}_{-w}\mid P_w](w),$$

and for all $f \in F$ and all $P_f \in \mathcal{P}_f$ such that $\Pr\{\tilde{P}_f = P_f\} > 0$, we have that for all $P'_f \in \mathcal{P}_f$,

$$DA_F[P_f, \tilde{P}_{-f}\mid P_f](f) \succ_{P_f} DA_F[P'_f, \tilde{P}_{-f}\mid P_f](f).$$

That is, $P$ is a monotonic OBNE in $DA_F$ under $\tilde{P}$, the desired conclusion.

($\Rightarrow$) Let $P$ be such that $\Pr\{\tilde{P} = P\} > 0$ and assume that $P$ is a monotonic OBNE in $DA_F$ under incomplete information $\tilde{P}$. By Theorem 2, $|C(P)| = 1$. To obtain a contradiction suppose that in addition $P$ is not a monotonic NE in $DA_F$ under complete information $P$. We first show in Lemma 1 below that then some firm with quota of at least two has a profitable deviation such that (i) the firm is only matched to acceptable workers, (ii) the firm is matched to $q_f$ workers and (iii) for some monotonic responsive extension the firm strictly prefers the assigned workers to the the set received under truth-telling.
Lemma 1 Assume $P$ is such that $|C(P)| = 1$. If $P$ is not a monotonic NE in $DA_F$ under complete information $P$, then there exists a firm $f \in F$ such that $q_f \geq 2$ and for some $\hat{P}_f$ we have

(i) $DA_F[\hat{P}_f, P_{-f}](f) \subseteq A(P_f)$,

(ii) $|DA_F[\hat{P}_f, P_{-f}](f)| = q_f = |DA_F[P](f)|$, and

(iii) $DA_F[\hat{P}_f, P_{-f}(f)P_f^*DA_F[P](f)$ for some monotonic responsive extension $P_f^*$ of $P_f$.

Proof of Lemma 1. Let $P \in \mathcal{P}$ be such that $|C(P)| = 1$. Then, using a similar argument than the one used in the sufficiency proof of Theorem 1 in Ehlers and Massó (2007), it can be seen that no worker has a profitable deviation from $P$ in $DA_F$. The same argument applies to any firm with quota one. Thus, if $P$ is not a monotonic NE in $DA_F$ under complete information $P$, then for some $f \in F$ with $q_f \geq 2$ and $\hat{P}_f$, we have

$$DA_F[\hat{P}_f, P_{-f}](f)P_f^*DA_F[P](f)$$

for some $P_f^* \in mresp(P_f)$. Because $P_f^*$ is responsive, we have $DA_F[\hat{P}_f, P_{-f}](f) \cap A(P_f)R_f^*DA_F[\hat{P}_f, P_{-f}](f)$. Thus, by (8) and monotonicity of $P_f^*$,

$$|DA_F[\hat{P}_f, P_{-f}](f) \cap A(P_f)| \geq |DA_F[P](f)|.$$  

But then we may suppose w.l.o.g. that $DA_F[\hat{P}_f, P_{-f}](f) \subseteq A(P_f)$. To see that, assume $DA_F[\hat{P}_f, P_{-f}](f) \nsubseteq A(P_f)$. Then, choose any preference $P_f' \in \mathcal{P}_f$ such that $A(P_f') = DA_F[\hat{P}_f, P_{-f}](f) \cap A(P_f)$. Now consider the matching market where the set of workers $\hat{W} = DA_F[\hat{P}_f, P_{-f}](f) \setminus A(P_f)$ is not present with profile $(P_f', P_{-\hat{W} \cup \{f\}})$. Then, it is easy to see that $DA_F[P_f', P_{-\hat{W} \cup \{f\}}](f) = A(P_f')$. Now letting workers in $\hat{W}$ reenter the market, firms should weakly gain meaning that $DA_F[P_f', P_{-f}](f) = A(P_f')$.  

\[21\] See Theorem 2.25 in Roth and Sotomayor (1990).
only matched to acceptable workers. Thus, \( DA_F[\hat{P}_f, P \_f](f) \subseteq A(P_f) \) holds. But the
inclusion is (i) in the Lemma.

Next we show that \(|DA_F[\hat{P}_f, P \_f](f)| = |DA_F[P](f)|\). To obtain a contradiction,
assume otherwise. By (9) and (i), we have \(|DA_F[\hat{P}_f, P \_f](f)| > |DA_F[P](f)|\). Then
at \( DA_F[P] \) firm \( f \) has some positions unfilled and the ranking of \( P_f \) over \( A(P_f) \)
is irrelevant for the stability of \( DA_F[P] \) under \( P \). Let \( k = |DA_F[\hat{P}_f, P \_f](f)| \). In
particular, \( DA_F[P] \) is stable under \( P \) if firm \( f \)'s quota is reduced from \( q_f \) to \( k \).
W.l.o.g. we may suppose that \( A(\hat{P}_f) = DA_F[\hat{P}_f, P \_f](f) \). Now again in any matching
which is stable under \((\hat{P}_f, P \_f)\) firm \( f \) is matched to \( DA_F[\hat{P}_f, P \_f](f) \). Again firm
\( f \)'s quota may be reduced to \( k \). But now consider \( P'_f \) such that \( A(P'_f) = A(P_f) \) and
\( DA_F[\hat{P}_f, P \_f](f) \) are the first \( k \) most preferred workers under \( P'_f \). But then both
\( DA_F[\hat{P}_f, P \_f] \) and \( DA_F[P] \) must be stable under \((P'_f, P \_f)\), which is a contradiction
because \(|DA_F[\hat{P}_f, P \_f](f)| \neq |DA_F[P](f)|\) and any firm is matched to the same
number of workers at all stable matchings under the profile \((P'_f, P \_f)\).

We have shown \(|DA_F[\hat{P}_f, P \_f](f)| = |DA_F[P](f)|\). Furthermore, if \(|DA_F[P](f)| < q_f \), then, by (i), both matchings \( DA_F[P] \) and \( DA_F[\hat{P}_f, P \_f] \) are stable under \( P \). By
(8), \( DA_F[\hat{P}_f, P \_f](f) \neq DA_F[P](f) \) which means that \(|C(P)| \neq 1 \), a contradiction.
Hence, \(|DA_F[P](f)| = q_f = |DA_F[\hat{P}_f, P \_f](f)| \) (which is (ii)). Now (iii) just follows
from (8). \( \square \)

We now proceed with the proof of Theorem 3. Let \( f \in F \), with \( q_f \geq 2 \), and \( \hat{P}_f \) be
the firm and its preferences identified in Lemma 1, for which (i), (ii) and (iii) hold.
Letting \( \hat{\mu} = DA_F[\hat{P}_f, P \_f] \) and \( \mu = DA_F[P] \), order the workers in \( \hat{\mu}(f) \) and \( \mu(f) \)
according to \( P_f \): let \( \hat{\mu}(f) = \{\hat{w}_1, \ldots, \hat{w}_{q_f}\} \) and \( \mu(f) = \{w_1, \ldots, w_{q_f}\} \). Furthermore,
because of (i)-(iii) and we are using \( DA_F \), we may assume w.l.o.g. that

(a) \( A(\hat{P}_f) = B(w_{q_f}, P_f) \cup B(\hat{w}_{q_f}, P_f) \),

(b) \( \hat{P}_f|\hat{\mu}(f) = P_f|\hat{\mu}(f) \),

(c) \( \hat{P}_f|B(\hat{w}_{q_f}, \hat{P}_f) = P_f|B(\hat{w}_{q_f}, \hat{P}_f) \), and
Case 1: \( \hat{w}_{q_f} P_f w_{q_f} \).

Let \( \hat{w}^* \) consist of the \( q_f \) workers which are \( P_f \)-least preferred in \( B(w_{q_f}, P_f) \). Then choose the monotonic responsive extension \( \tilde{P}^*_f \) of \( P_f \) such that \( W'' \tilde{R}^*_f \hat{w}^* \) iff \( W'' \) contains \( q_f \) workers in \( B(w_{q_f}, P_f) \). Let \( P''_f \) be such that \( A(P''_f) = B(w_{q_f}, P_f) \) and \( P''_f \mid A(P''_f) = \hat{P}_f | A(P''_f) \). Then \( \tilde{w} \in C(P''_f, P_{-f}) \) and by construction, \( \Delta A_f [P''_f, P_{-f}](f) \) contains \( q_f \) workers in \( A(P''_f) \), but \( \Delta A_f [P](f) \) contains at most \( q_f - 1 \) workers in \( A(P''_f) \). Hence,

\[
\Delta A_f [P''_f, P_{-f}](f) \tilde{R}^*_f \hat{w}^* \tilde{P}^*_f \Delta A_f [P_f, P_{-f}](f) = 0. \tag{10}
\]

To obtain a contradiction with the fact that \( P \) is a monotonic OBNE in \( \Delta A_f \) under incomplete information \( \hat{P} \), we consider any subprofile \( P'_{-f} \) such that \( \Pr \{ \hat{P} = (P_f, P'_{-f}) \} > 0 \) and restrict the attention to the upper contour set of \( \hat{P}^*_f \) at \( \hat{W}^* \).

We want to show that if \( \Delta A_f [P_f, P'_{-f}](f) \tilde{R}^*_f \hat{W}^* \), then \( \Delta A_f [P''_f, P'_{-f}](f) \tilde{R}^*_f \hat{W}^* \). By the definitions of \( \hat{W}^* \) and \( \tilde{P}^*_f \), \( \Delta A_f [P_f, P'_{-f}](f) \tilde{R}^*_f \hat{W}^* \) implies that \( \Delta A_f [P_f, P'_{-f}](f) \) contains \( q_f \) workers in \( A(P''_f) \). If \( \Delta A_f [P''_f, P_{-f}](f) \) contains at most \( q_f - 1 \) workers in \( A(P''_f) \), then the ranking does not matter and this matching is also stable (letting \( A(P''_f) = A(P''_f') \) and \( P''_f \mid A(P''_f') = P_f \mid A(P''_f) \) under \( (P''_f', P'_{-f}) \). Furthermore, \( \Delta A_f [P_f, P'_{-f}] \) is stable under \( (P''_f', P'_{-f}) \) and \( \| \Delta A_f [P_f, P'_{-f}] \| = q_f \). But then \( f \) would be matched to different numbers of workers under different matchings belonging to \( C(P''_f', P'_{-f}) \), a contradiction. Thus, \( \Delta A_f [P''_f, P'_{-f}](f) \) contains \( q_f \) workers in \( A(P''_f) \). Now by our choice of \( \hat{W}^* \) and \( P^*_f \), we obtain for any profile \( P'_{-f} \) in the support of \( \hat{P}_{-f} \mid P_f \), \( \Delta A_f [P_f, P'_{-f}](f) \tilde{R}^*_f \hat{W}^* \) implies \( \Delta A_f [P''_f, P'_{-f}](f) \tilde{R}^*_f \hat{W}^* \). Then, together with (10) and \( \Pr \{ \hat{P} = (P_f, P_{-f}) \} > 0 \), this implication means that truth-telling is not an OBNE in \( \Delta A_f \) under \( \tilde{P}^*_f \) of \( P_f \),

\[
\Pr \{ \Delta A_f [P''_f, \hat{P}_{-f} | P_f](f) \in B(\hat{W}^*, \hat{P}_f) \} > \Pr \{ \Delta A_f [P_f, \hat{P}_{-f} | P_f](f) \in B(\hat{W}^*, \hat{P}^*_f) \}.
\]
**Case 2:** \(w_{q_f} R_f \hat{w}_{q_f}\).

Let \(k\) be the first index such that \(\hat{w}_k P_F w_k\). By (iii) in Lemma 1, \(k\) exists. Note that \(k < q_f\) and \(\hat{w}_k P_F w_k R_f w_{q_f}\). Let \(P''_f\) be such that \(A(P''_f) = B(\hat{w}_{q_f}, P_f), B(\hat{w}_k, P_f) \cup \hat{\mu}(f) P''_f A(P''_f) \setminus (B(\hat{w}_k, P_f) \cup \hat{\mu}(f)), \) \(22\) \(P''_f|B(\hat{w}_k, P_f) \cup \hat{\mu}(f) = P_f|B(\hat{w}_k, P_f) \cup \hat{\mu}(f),\) and \(P''_f|A(P''_f) \setminus (B(\hat{w}_k, P_f) \cup \hat{\mu}(f)) = P_f|A(P''_f) \setminus (B(\hat{w}_k, P_f) \cup \hat{\mu}(f))\). We illustrate \(P''_f\) below:

\[
P''_f|B(\hat{w}_k, P_f) \\
P''_f|\hat{\mu}(f) \setminus B(\hat{w}_k, P_f) \\
P''_f|W \setminus (B(\hat{w}_k, P_f) \cup \hat{\mu}(f))
\]

i.e. \(P''_f\) ranks first all elements in \(B(\hat{w}_k, P_f)\) according to \(P_f\), then all elements in \(\hat{\mu}(f) \setminus B(\hat{w}_k, P_f)\) according to \(P_f\), and then the remaining workers according to \(P_f\). Because \(\hat{w}_1, \ldots, \hat{w}_k\) belong to \(B(\hat{w}_k, P_f)\), we have \(\hat{\mu}(f) \setminus B(\hat{w}_k, P_f) = \{\hat{w}_{k+1}, \ldots, \hat{w}_{q_f}\}\) (and this set consists of exactly \(q_f - k\) workers).

If \(f\) is matched to fewer than \(q_f\) workers in any matching belonging to \(C(P''_f, P_{-f})\), then the differences between ranking \(P''_f\) and \(P_f\) do not matter and \(f\) would be matched to fewer than \(q_f\) workers in \(DA_F[P](f)\), the unique matching in \(C(P)\), a contradiction. Thus, \(f\) is matched to \(q_f\) workers in any matching belonging to \(C(P''_f, P_{-f})\).

If \(\hat{\mu} \in C(P''_f, P_{-f})\), then by construction and (8), \(DA_F[P''_f, P_{-f}](f) R_f^* \hat{\mu}(f) P''_f \mu(f)\) implying

\[
DA_F[P''_f, P_{-f}](f) R_f^* \hat{\mu}(f) P''_f \mu(f).
\] \hspace{1cm} (11)

Suppose that \(\hat{\mu} \notin C(P''_f, P_{-f})\). Then, some \((w, \hat{f})\) blocks \(\hat{\mu}\) under \((P''_f, P_{-f})\). By the definition of \(P''_f\), \(\hat{f} = f\) and \(w \in B(\hat{w}_k, P_f)\). We show that \(DA_F[P''_f, P_{-f}](f)\) contains at least \(k\) workers in \(B(\hat{w}_k, P_f)\). We do it through the following two steps:

**Step 1:** Let \(w^1\) be the \(P_f\)-highest ranked worker such that \((w^1, f)\) forms a blocking pair of \(\hat{\mu}\) under \((P''_f, P_{-f})\). Let \(\hat{\mu}^0 = \hat{\mu}\). Then match \(w^1\) to \(f\) and consider \(\hat{\mu}^1\) such that

\[\text{Here we use the convention } SP''_f T \text{ iff } s P''_f t \text{ for all } s \in S \text{ and all } t \in T.\]
that (i) $\hat{\mu}^1(w^1) = f$ and (ii) $\hat{\mu}^1(w) = \hat{\mu}(w)$ for all $w \neq w^1$. In other words, $f$ does not reject any worker and we allow $f$ to have capacity $q_f + 1$. If $\hat{\mu}^1$ does not contain any blocking pair, set $\bar{\mu} = \hat{\mu}^1$ and $\bar{q}_f = q_f + 1$ and then go to Step 2. Otherwise, $\hat{\mu}^1$ contains a blocking pair, say $(w^2, f^2)$. Again, let $w^2$ be the $P_f$-highest ranked worker such that $(w^2, f^2)$ forms a blocking pair. But then it must be that $f^2 = f$ (which would mean that $w^2 \in \hat{\mu}^1(f)$) or $f = \hat{\mu}(w^1)$. Now consider $\hat{\mu}^2$ such that (i) $\hat{\mu}^2(w^2) = f^2$ and (ii) $\hat{\mu}^2(w) = \hat{\mu}^1(w)$ for all $w \neq w^2$. If $w^2 \in \hat{\mu}^1(f)$, then $\hat{\mu}^2$ is stable under $(P''_f, P_{-f})$ and $\hat{\mu}^1(f)$ contains at least $k$ workers in $B(\hat{w}_k, P_f)$. By construction, the same is true for $DA_F[P''_f, P_{-f}](f)$ since for worker $w$, the member of the original blocking pair, $w^1R_fw$ and $w \in B(\hat{w}_k, P_f)$ imply that $w^1 \in B(\hat{w}_k, P_f)$. Otherwise, $w^2 \notin \hat{\mu}^1(f)$. If $f^2 = f$, then we now allow $f$ to have capacity $q_f + 2$, and so on. In other words, we never reject a worker and only vacant positions are filled (where we allow “overbooking” for $f$). At the same time, we reduce $f$’s capacity by one whenever one of the workers assigned to $f$ leaves. Note that this process terminates as the workers’ preference always weakly improves at each iteration (because $f$ rejects no worker). If at some point exactly $q_f$ workers are matched to $f$, then we must have found a stable matching for $(P''_f, P_{-f})$. Let $\bar{\mu}$ denote the resulting matching, and say $|\bar{\mu}(f)| = \bar{q}_f$. Then $\bar{\mu}$ is stable under $(P''_f, P_{-f})$ where $f$ has $\bar{q}_f$ positions (instead of $q_f$). By construction, $\bar{\mu}(f)$ contains at least $k$ workers in $B(\hat{w}_k, P_f)$.

Step 2: Let $\bar{q}_f$ and $\bar{\mu}$ be the outcomes of Step 1. Let $\bar{\mu}(f) = \{\bar{w}_1, \ldots, \bar{w}_{\bar{q}_f}\}$ denote the workers ranked according to $P''_f$. Now consider the matching market where $\bar{W} = \{\bar{w}_{\bar{q}_f+1}, \ldots, \bar{w}_{\bar{q}_f}\}$ is not present. Let $\bar{\mu}'$ be defined by $\bar{\mu}'(f) = \{\bar{w}_1, \ldots, \bar{w}_{\bar{q}_f}\}$ and $\bar{\mu}'(f') = \bar{\mu}(f')$ for all $f' \neq f$. Then $\bar{\mu}' \in C(P''_f, P_{-\bar{W} \cup \{f\}})$. Since $\bar{\mu}'$ contains at least $k$ workers in $B(\hat{w}_k, P_f)$ and by definition of $P''_f$, $DA_F[P''_f, P_{-\bar{W} \cup \{f\}}](f)$ contains at least $k$ workers in $B(\hat{w}_k, P_f)$. Bringing back in $\{\bar{w}_{\bar{q}_f+1}, \ldots, \bar{w}_{\bar{q}_f}\}$, all firms must weakly benefit, i.e. $DA_F[P''_f, P_{-f}](f)$ must contain at least $k$ workers in $B(\hat{w}_k, P_f)$, the desired conclusion.

Now let $\bar{W}^*$ consist of the $k$ workers which are $P_f$-ranked least in $B(\hat{w}_k, P_f)$
and the $q_f - k$ workers $P_f$-ranked least in $B(\hat{w}_{q_f}, P_f)$. Obviously, $\hat{w}_k, \hat{w}_{q_f} \in \hat{W}^*$. Let $\hat{P}_f^*$ be a monotonic responsive extension of $P_f$ such that $W'\hat{P}_f^*\hat{W}^*$ iff $W' \subseteq B(\hat{w}_{q_f}, P_f)$, $|W'| = q_f$ and $W'$ contains at least $k$ workers in $B(\hat{w}_k, P_f)$. Now we have $DA_F[P''_f, P_{-f}](f)\hat{P}_f^*\hat{W}^*$. By the definition of $k$, $\hat{w}_k P_f w_k$. Hence, $DA_F[P_f, P_{-f}](f)$ does not contain at least $k$ workers in $B(\hat{w}_k, P_f)$, which implies that

$$DA_F[P''_f, P_{-f}](f)\hat{P}_f^*\hat{W}^* DA_F[P_f, P_{-f}](f).$$

Summarizing, we have already showed in the case $w_{q_f} R_f \hat{w}_{q_f}$ that there exist $(P_f, P_{-f}) \in \text{supp}(\hat{P})$, $P''_f \in \mathcal{P}_f$, $\hat{W}^* \subseteq W$ and $\hat{P}_f^* \in \text{mres}(P_f)$ such that $DA_F[P''_f, P_{-f}](f) \in B(\hat{W}^*, \hat{P}_f^*)$ and $DA_F[P_f, P_{-f}](f) \notin B(\hat{W}^*, \hat{P}_f^*)$ simultaneously hold.

Now consider $(P_f, P''_f) \in \text{supp}(\hat{P})$. Let $\mu' = DA_F[P_f, P''_{-f}]$ and $\mu'' = DA_F[P''_f, P'_{-f}]$. Suppose that $\mu'(f)\hat{P}_f^*\hat{W}^*$. By construction, $\mu'(f)$ contains at least $k$ workers in $B(\hat{w}_k, P_f)$, $|\mu'(f)| = q_f$ and $\mu'(f) \subseteq B(\hat{w}_{q_f}, P_f)$. We need to show $\mu''(f)\hat{P}_f^*\hat{W}^*$. By construction, $\mu''(f) \subseteq B(\hat{w}_{q_f}, P_f) = A(P''_f)$. Furthermore, if $|\mu''(f)| < q_f$, then the ranking of $P''_f$ does not matter and we would have that $\mu'' \in C(P_f, P''_{-f})$. By Theorem 2, $C(P_f, P''_{-f}) = \{\mu'\}$ which is contradiction because $\mu''(f) < q_f = \mu'(f)$. Thus, $|\mu''(f)| = q_f$.

It remains to be shown that $\mu''(f)$ contains at least $k$ workers in $B(\hat{w}_k, P_f)$. Now if $\mu' \in C(P''_f, P'_{-f})$, then, by definition of $P''_f$, $B(\hat{w}_{q_f}, P''_f) = B(\hat{w}_{q_f}, P_f)$ and the fact that $\mu'(f)$ contains at least $k$ workers in $B(\hat{w}_k, P_f)$, $\mu''(f)$ must contain at least $k$ workers in $B(\hat{w}_k, P_f)$, the desired conclusion. Suppose that $\mu' \notin C(P''_f, P'_{-f})$. Note that we cannot have $\mu'(f) \subseteq \hat{\mu}(f) \cup B(\hat{w}_k, P_f)$ as otherwise $\mu' \in C(P_f, P'_{-f})$, $P''_f |\hat{\mu}(f) \cup B(\hat{w}_k, P_f) = P_f |\hat{\mu}(f) \cup B(\hat{w}_k, P_f)$, and $\hat{\mu}(f) \cup B(\hat{w}_k, P_f)P''_f A(P''_f) \setminus (\hat{\mu}(f) \cup B(\hat{w}_k, P_f))$ would imply $\mu' \in C(P''_f, P_{-f})$, a contradiction.\textsuperscript{23}

To obtain a contradiction, assume $\mu''(f)$ does not contain at least $k$ workers in $B(\hat{w}_k, P_f)$. In particular, this means that $\mu'(f) \neq \mu''(f)$. Then, since $|\hat{\mu}(f) \setminus B(\hat{w}_k, P_f)| = q_f - k$, $\mu''(f) \subseteq B(\hat{w}_k, P_f) \cup \hat{\mu}(f)$ as otherwise $f$ is assigned $k$ workers in $B(\hat{w}_k, P_f)$.

\textsuperscript{23}Here we use again the convention $SP_f^sT$ iff $sP''_f t$ for all $s \in S$ and all $t \in T$. 

30
Let \( w'' \) denote the \( P_f'' \)-least preferred worker in \( \mu''(f) \). By the above we have \( w'' \notin \hat{\mu}(f) \cup B(\hat{w}_k, P_f) \).

**Case 2.1:** \( w'' \) is the \( P_f \)-least preferred worker in \( \mu''(f) \).

Because \( C(P_f, P'_{-f}) = \{ \mu' \} \), we have \( \mu'' \notin C(P_f, P'_{-f}) \). Then some pair \((\hat{w}, f)\) blocks \( \mu'' \) under \((P_f, P'_{-f})\). Because \( w'' \) is the \( P_f \)-least worker in \( \mu''(f) \), we have \( \hat{w}P_ww'' \).

If \( \hat{w} \in B(\hat{w}_k, P_f) \cup \hat{\mu}(f) \), then by construction and \( w'' \notin \hat{\mu}(f) \cup B(\hat{w}_k, P_f) \) we have \( \hat{w}P_w''w'' \) and \((\hat{w}, f)\) blocks \( \mu'' \) under \((P_f'', P'_{-f})\), a contradiction to \( \mu'' \in C(P_f'', P'_{-f}) \).

If \( \hat{w} \notin B(\hat{w}_k, P_f) \cup \hat{\mu}(f) \), then by construction and \( w'' \notin \hat{\mu}(f) \cup B(\hat{w}_k, P_f) \) we have \( \hat{w}P_w''w'' \) and \((\hat{w}, f)\) blocks \( \mu'' \) under \((P_f'', P'_{-f})\), a contradiction to \( \mu'' \in C(P_f'', P'_{-f}) \).

Thus, Case 2.1 cannot occur.

**Case 2.2:** \( w'' \) is not the \( P_f \)-least preferred worker in \( \mu''(f) \).

Let \( \hat{w} \) denote the \( P_f \)-least preferred worker in \( \mu''(f) \). By \( \hat{w} \neq w'' \), we have \( \hat{w} \in B(\hat{w}_k, P_f) \cup \hat{\mu}(f) \). Furthermore, by \( w'' \notin B(\hat{w}_k, P_f) \cup \hat{\mu}(f) \), we cannot have \( \hat{w} \in B(\hat{w}_k, P_f) \). Then by construction we must have \( \hat{w} \in \hat{\mu}(f) \). But now we may just exchange the positions of \( \hat{w} \) and \( w'' \) and consider \( P_f''\hat{w}\hat{w}w'' \). We have \( \mu'' \in C(P_f''\hat{w}\hat{w}w'', P'_{-f}) \). In \( \mu''(f) \) the \( P_f \)-least preferred worker and the \( P_f''\hat{w}\hat{w}w'' \)-least preferred worker coincide and is \( \hat{w} \). But now we are in Case 2.1 and this is a contradiction (where \( \hat{\mu}(f) \) is replaced by \( \hat{\mu}(f)\hat{w}\hat{w}w'' = (\hat{\mu}(f)\{\hat{w}\}) \cup \{w''\}) \).

Thus, \( DA_F[P_f'', P'_{-f}](f) \) contains \( q_f \) workers in \( A(P_f'') \) and at least \( k \) workers in \( B(\hat{w}_k, P_f) \). Now by our choice of \( \hat{W}^* \) and \( \hat{P}_f^* \), we obtain for any profile \( P'_{-f} \) in the support of \( \hat{P}_{-f}[P_f] \), \( DA_F[P_f, P'_{-f}](f)\hat{R}_f^*\hat{W}^* \) implies \( DA_F[P''_f, P'_{-f}](f)\hat{R}_f^*\hat{W}^* \). Since \( DA_F[P''_f, P'_{-f}](f)\hat{R}_f^*\hat{W}^* \hat{P}_f^*DA_F[P_f, P_{-f}](f) \) and \( P \in supp(\hat{P}) \), we have

\[
Pr\{DA_F[P''_f, \hat{P}_{-f}[P_f] \}(f) \in B(\hat{W}^*, \hat{P}_f^*)\} > Pr\{DA_F[P_f, \hat{P}_{-f}[P_f] \}(f) \in B(\hat{W}^*, \hat{P}_f^*)\}.
\]

Hence, truth-telling is not a monotonic OBNE in \( DA_F \) under \( \hat{P} \), a contradiction. □
References


