Implementation of Assortative Matching Under Incomplete Information *

Naomi Utgoff
University of Maryland
Department of Economics
utgoff@econ.umd.edu
October 5, 2014

Abstract

This paper provides a unifying framework for matching markets with incomplete information, when the positive assortative match is the unique efficient stable match. I construct a VCG-like mechanism which implements assortative matching as an ex post Nash equilibrium. It achieves this result using a payment rule that distinguishes between an agent deprived of any match and an agent who merely receives a reduced match surplus. The constructed mechanism recognizes only opportunity costs arising from the former, and not the latter, effect. I also generalize the stronger condition of envy freeness to these incomplete information environments and show that the constructed equilibrium is envy free.

*I am indebted to my advisors, Larry Ausubel, Daniel Vincent, and Erkut Ozbay for their support and guidance. Thanks also to Justin Burkett, Peter Cranton, Daisy Dai, Emel Filiz-Ozbay, John Ham, John Hatfield, Igal Hendel, Ginger Jin, Rob Kulick, Tzu-Yao Lin, Hector Lopez-Carbajal, Kristian Lopez-Vargas, John Shea, Ozlem Tonguc, and John Wallis for helpful discussions and comments. Many thanks to Thayer Morrill for comments on an earlier draft. I am grateful to the University of Maryland Department of Economics for funding this research during Summer 2013.
1 Introduction

Matching markets in which people have incomplete information about the qualities of potential partner(s) occur frequently and there are a variety of practical approaches that (try to) elicit this information prior to match formation to facilitate better or more stable matches. The interview process facilitates job matching; the dating process facilitates marriage. This paper considers matching markets in which all matches are formed through a single central clearinghouse, and in which the positive assortative match is the unique efficient and unique stable match. I model this clearinghouse as a direct revelation mechanism to which each agent reports her privately known partner quality; the mechanism implements the positive assortative match given the agents’ reports. Truthful reporting in this mechanism forms an ex post Nash equilibrium in all such incomplete information matching environments (roommate, marriage, supply chain, both one-to-one and many-to-one versions).

The main contribution of this mechanism is its payment rule. It distinguishes between the opportunity cost an agent imposes on other agents whom her report deprives of membership in a match, and the opportunity cost she imposes when her report merely alters the value of other agents’ respective match surpluses. When each agent’s payment contains the former but not the latter opportunity costs, truthful reporting forms an ex post Nash equilibrium. The simplest example is a roommate market with three agents and a single two-person room. Each person reports her privately known roommate quality to the mechanism. The people with the two higher reported qualities each receive one bed and each of these two roommates receives a match surplus that is increasing and supermodular in her own and her roommate’s qualities. The person who reports the lowest quality receives neither bed, deprives no one else of a bed, and pays 0. Each winning roommate has two effects: first, she deprives the person who submits the lowest report of a bed; second, she affects her roommate’s match surplus. Her payment is the opportunity cost she imposes on the person who did not receive a bed, but does not include her effect on her roommate. This payment rule is the correct generalization of the payment rule in the VCG mechanism (with independent private values) to the interdependent value environment of these matching markets. Under independent private values, a person imposes
opportunity costs on others only when her report deprives others of objects. The payment rule of the VCG-like matching mechanism of this paper recognizes that even when other effects are present due to interdependence, the only opportunity costs that should contribute to an agent’s payment are those due to depriving other agents of objects.

Another way to view the VCG-like mechanism is as a position auction which implements positive assortative matching in this paper’s class of matching markets. Matching theory and auction theory began contemporaneously but independently. Vickrey’s seminal auction paper [14] describes a generalization of the second price auction to multiple identical objects and shows that there is no mechanism that elicits both true supply curves and demand curves. His work lays out the primary concerns addressed in subsequent auction papers: elicitation of buyers’ private valuations, efficiency, and revenue. On the other hand, algorithmic design and match stability are the primary concerns in the (early) matching literature. Gale and Shapley’s seminal matching paper [4] introduces their deferred acceptance algorithm. They show that it always finds a stable match in a marriage market in which preferences are public information, and that a stable match always exists in a marriage market.

Kelso and Crawford [8] were the first authors to link the matching and auction theory literatures. They exhibit a discrete wage adjustment algorithm similar to a clock auction that identifies and implements a stable match between firms and workers in a complete information setting. The intersection of these fields has since emerged as a fruitful area with numerous interesting problems and results. However, Roth [11] demonstrated that when matching theory considers the problem of preference elicitation (akin to the auction-theoretic concern of eliciting valuations) truthful revelation of preferences can be a dominant strategy for only one side of a marriage market. The impossibility arises due to preference heterogeneity across agents.

Current work on matching with incomplete information assumes homogeneous, known preferences and that each agent privately observes her single-dimensional quality as a partner. These markets assume that agents’ payoff functions are such that each agent strictly prefers a higher quality to a lower quality partner. It is well known that in the full information analog of this environment, positive assortative matching is both the unique efficient and unique stable match. Having eliminated preference heterogeneity, there are several interesting results con-
cerning mechanisms which implement positive assortative matching in marriage markets. The two main papers considering these models are by Hoppe, Moldovanu, and Sela [6] and Johnson [7]. Hoppe, Moldovanu, and Sela [6] show that a fully separating signaling equilibrium exists in a marriage market with privately informed agents, but that random matching may be socially superior to a signaling equilibrium due to the signaling costs incurred by agents. Their paper adopts Spence’s [13] convention that signaling costs are wasted; their equilibrium is logically equivalent to an all pay auction. Johnson [7] observes that signaling costs in these environments are often not wasted; many matching markets are run by a self-interested matchmaker serving as an information broker who receives the signaling costs as her payment. Johnson constructs revenue-maximizing auctions that implement truncated assortative matching. Both [6] and [7] implement their solutions as Bayesian Nash equilibria; while they achieve positive assortative matching (or in the case of [7], a truncation thereof) agents experience ex post regret due to the payment structure in Bayesian implementation.

The mechanism of this paper shows that ex post regret is unnecessary in any incomplete information matching market where the match surplus functions are increasing and supermodular in each agent’s one-dimensional partner quality, including the marriage markets of [6] and [7]. I describe a natural generalization of the marriage markets of [6] and [7] to a large class of matching markets in which positive assortative matching is the unique efficient stable match. I then introduce the VCG-like mechanism that unifies these markets in a single framework and prove that this mechanism always implements positive assortative matching as an ex post Nash equilibrium; in particular, this mechanism works in both one-, two-, and many-sided matching markets. This equilibrium is slightly stronger than ex post Nash equilibrium; I show that the idea of locally envy free equilibrium introduced by Edelman, Ostrovsky, and Schwarz [3] generalizes completely to this class of incomplete information environments.

2 Examples

The following examples of roommate and marriage matching give the flavor of the mechanism and intuition for its payment rule. In both examples, the reservation payoff is 0.
2.1 Roommates

There are three agents and one room with two identical beds. Each agent $n$ has privately known roommate quality $x_n \in [0, 1]$. If agents $n$ and $n'$ each receive one of the beds, $n$ receives $s(x_n, x_{n'})$ and $n'$ receives $s(x_{n'}, x_n)$ where $s(\cdot, \cdot)$ is increasing in both arguments and supermodular. These requirements on $s(\cdot, \cdot)$ are sufficient to guarantee that positive assortative matching is the unique efficient stable match. Without loss of generality assume that $x_1 \geq x_2 \geq x_3$. Each agent reports her quality to the mechanism which implements positive assortative matching: the agents who submit the two highest reports receive one bed each. The agent who submits the lowest report receives no bed and the reservation payoff 0. Denote agent $n$’s report $\hat{x}_n$ to distinguish her reported quality from her actual quality. When each agent reports her quality truthfully, agents 1 and 2 each receive a bed in the positive assortative match. The positive assortative match yields the match surpluses shown in the table.

<table>
<thead>
<tr>
<th>Agent</th>
<th>Roommate</th>
<th>Surplus</th>
<th>Payment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>$s(x_1, x_2)$</td>
<td>$s(\hat{x}_3, \hat{x}_2)$</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>$s(x_2, x_1)$</td>
<td>$s(\hat{x}_3, \hat{x}_1)$</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Truthful reporting forms an ex post Nash equilibrium under the payments shown in the table: when any two agents report their qualities truthfully, the third cannot profitably deviate from reporting her own quality truthfully. For example, suppose agents 2 and 3 report their qualities truthfully: if agent 1 reports $\hat{x}_1 \geq x_3$, she receives one of the beds, pays $s(x_3, x_2)$, and receives payoff $s(x_1, x_2) - s(x_3, x_2) \geq 0$ because $s(\cdot, \cdot)$ is increasing in its first argument. If agent 1 reports $\hat{x}_1 < x_3$, she receives no bed and the reservation payoff 0. Since $x_1 \geq x_3$, agent 1 is at least as well off reporting her quality truthfully as she would selecting any other report.

Agent 1’s report has two effects: her report (1) deprives agent 3 of a bed and (2) affects the value of agent 2’s match surplus. The intuition for agent 1’s payment is that she should pay the opportunity cost she imposes on agent 3 by depriving her of a bed, but should not pay for her effect on agent 2, whom she deprived of nothing. Similarly, agent 2’s payment is the opportunity cost she imposes on agent 3, but her effect on agent 1, whom she deprived
of nothing. Agent 3, who deprives neither agent 1 nor 2 of a bed, pays 0. Another way to describe the payment rule is that for each agent $n$, the mechanism identifies the set of other agents with whom $n$ is effectively in competition: in the above example agents 1 and 2 are not ex post in competition for a bed.

### 2.2 Marriage

The marriage market’s structure differs significantly from the roommate market: the list of possible partners for a man (woman) is restricted to women (men), whereas in the roommate market any agent may match with any other. However, despite this difference, all of the observations from the roommate example apply to this marriage example.

Suppose that there are two men and two women. Each man $i$ has privately known quality $m_i \in [0, 1]$ and each woman $j$ has privately known quality $w_j \in [0, 1]$. If man $i$ and woman $j$ are matched, the man receives $s_M(m_i, w_j)$ and the woman receives $s_W(w_j, m_i)$ where both $s_M(\cdot, \cdot)$ and $s_W(\cdot, \cdot)$ are increasing in both arguments and supermodular; these conditions on $s_M(\cdot, \cdot)$ and $s_W(\cdot, \cdot)$ are sufficient to guarantee that positive assortative matching is the unique efficient stable match. Suppose without loss of generality that $m_1 \geq m_2$ and $w_1 \geq w_2$. Each agent reports his (her) type to the mechanism which implements positive assortative matching: the man with the higher report (of the two men) is matched with the woman with the higher report (of the two women), and the man with the lower report is matched with the woman with the lower report. Denote man $i$’s (woman $j$’s) report $\hat{m}_i$ ($\hat{w}_j$) to distinguish his (her) report from his (her) quality. When each agent reports his (her) quality truthfully the positive assortative match pairs man 1 with woman 1 and man 2 with woman 2. The positive assortative match yields the match surpluses shown in the table.

<table>
<thead>
<tr>
<th>Agent</th>
<th>Spouse</th>
<th>Surplus</th>
<th>Payment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$</td>
<td>$W_1$</td>
<td>$s_M(m_1, w_1)$</td>
<td>$s_M(\hat{m}_2, \hat{w}_1) - s_M(\hat{m}_2, \hat{w}_2)$</td>
</tr>
<tr>
<td>$M_2$</td>
<td>$W_2$</td>
<td>$s_M(m_2, w_2)$</td>
<td>0</td>
</tr>
<tr>
<td>$W_1$</td>
<td>$M_1$</td>
<td>$s_W(w_1, m_1)$</td>
<td>$s_W(\hat{w}_2, \hat{m}_1) - s_W(\hat{w}_2, \hat{m}_2)$</td>
</tr>
<tr>
<td>$W_2$</td>
<td>$M_2$</td>
<td>$s_W(w_2, m_2)$</td>
<td>0</td>
</tr>
</tbody>
</table>
Truthful reporting again forms an ex post Nash equilibrium under the payments shown in the table: when any three agents report their qualities truthfully, the remaining agent cannot profitably deviate from reporting truthfully him/herself. The intuition behind the payment rule in the marriage market is identical to the intuition behind the payment rule in the roommate market: an agent’s payment should be the total opportunity cost imposed on other agents whom his (her) report displaced from a marriage. The two-sided structure of the marriage market means that a man (woman) can only displace another man (woman) from a match; a man cannot displace a woman and a woman cannot displace a man. In this example, $M_1$ displaces $M_2$ from marriage with $W_1$. Therefore, his payment is the opportunity cost he imposes on $M_2$ by preventing him from marrying $W_1$. However, $M_1$’s payment does not include the opportunity cost he imposes on either woman because he does not (and cannot) displace either woman from a marriage.

3 Model

There is a set of agents $\{1, \cdots, N\}$, a set of bins $\{1, \cdots, M\}$, and a set of roles $\{1, \cdots, R\}$. Each agent $n$ has privately known type $x_n : \{1, \cdots, R\} \to [0, 1]$ where $x_n(r)$ denotes $n$’s quality if employed in role $r$. Assume that each agent $n$ has one productive role $r_n \in \{1, \cdots, R\}$; for all $r \neq r_n$, $x_n(r) = 0$. There are some matching markets (such as marriage markets) in which an agent’s productive role is public information and others (such as labor markets) in which it may be private. The results do not require that any agent’s productive role be public information, so assume that it is private. Assume also that the $x_n$ are independently distributed. Let $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_N]$, $\mathbf{x}_{-n} = [x_1 \ \cdots \ x_{n-1} \ x_{n+1} \ \cdots \ x_N]$, and $(\mathbf{x}_{-n}, z) = [x_1 \ \cdots \ x_{n-1} \ z \ x_{n+1} \ \cdots \ x_N]$.

Each bin $m$ has a publicly known capacity constraint $c_{mr} \geq 0$ for each role $r$; assume that $c_{ir} \geq c_{jr}$ whenever $i < j$, so that firm $i$ is at least as big as bin $j$ for all $1 \leq i \leq j \leq M$. Each bin also has a minimum of two slots, i.e. $\sum_{r=1}^{R} c_{mr} \geq 2$ for all $m \in \{1, \cdots, M\}$.

Agents are utility-maximizing and each agent seeks a slot in her productive role. An agent who is unmatched or is the sole member of a bin receives the reservation payoff 0. Each matched
agent derives private match surplus from her interactions with other agents in the same bin. For each $r$, there is a common, known surplus function $s_r : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ that is increasing in both arguments and supermodular. Assume that $s_r(0, x) = s_r(x, 0) = 0$ for all $r$ and all $x$; intuitively, an agent whose quality in role $r$ is 0 and who works in role $r$ receives match surplus 0 and contributes 0 to the match surplus of any other agent employed at the same firm.

The conditions on the $s_r(\cdot, \cdot)$ together with the structure of the bins’ capacity constraints imply that positive assortative matching is the unique efficient stable match.\(^1\) In this model, the positive assortative match is the result of implementing the following process in the full information analog of the model environment.

1. Sort agents by productive role; within each role $r$, rank them from highest to lowest by quality.
2. Among agents whose productive role is 1, assign the $c_{11}$ highest quality agents to firm 1, the next $c_{21}$ highest quality agents to firm 2, and so on until there are no more agents whose productive role is 1 or until all role 1 jobs at all $M$ firms are filled.
3. Repeat Step 2 for roles 2, \ldots, $R$.

Let $L \subseteq \{1, \ldots, N\}$ be the set of agents in a bin in their respective productive roles. Agent $n$’s match surplus is $\sum_{i \in L \setminus \{n\}} s_{rn}(x_n(r_n), x_i(r_i))$. Agents’ utilities are quasilinear; when agent $n$ pays $t_n$ for a slot in bin $m$ in role $r$, she receives $\sum_{i \in L \setminus \{n\}} s_{rn}(x_n(r_n), x_i(r_i)) - t_n$.

An assignment is a function $\mu : \{1, \ldots, N\} \rightarrow \{1, \ldots, M\} \times \{1, \ldots, R\} \cup (0, 0)$ which satisfies

1. $\mu(n) = (0, 0)$ if and only if $n$ is unmatched, i.e. $\mu$ assigns $n$ to no slot in any bin
2. $\mu(n) = (m, r)$ if and only if agent $n$ occupies a slot in role $r$ in bin $m$
3. All capacity constraints are satisfied, i.e. $|\mu^{-1}(m, r)| \leq c_{mr}$ for all $m, r$

Let $A$ denote the set of all functions $\mu$ satisfying items 1 - 3 above, and let $\phi : \{1, \ldots, M\} \times \{1, \ldots, R\} \rightarrow \{1, \ldots, M\}$ and $\rho : \{1, \ldots, M\} \times \{1, \ldots, R\} \rightarrow \{1, \ldots, R\}$ denote the canonical

\(^1\)Shapley and Shubik [12] were the first to observe that in a two-sided one-to-one matching market, supermodularity implies that positive assortative matching is the unique efficient stable match.
projections. Let $\overline{\rho} = \rho \circ \mu$. For all agents $n$, denote the set of $n$'s partners by

$$W_{\mu}(n) = \begin{cases} \{i \in \{1, \cdots, N\} \mid \phi \circ \mu(i) = \phi \circ \mu(n)\} \setminus \{n\} & \text{if } \phi \circ \mu(n) \neq 0 \\ \emptyset & \text{if } \phi \circ \mu(n) = 0 \end{cases}$$

Under the assignment $\mu$, agent $n$ receives match surplus

$$\sum_{i \in W_{\mu}(n)} s_{\pi(n)}(x_n(\overline{\mu}(n)), x_i(\overline{\mu}(i)))$$

If $\mu(n) = (0, 0)$, agent $n$’s match surplus is necessarily 0 since $W_{\mu}(n) = \emptyset$. The social surplus under $\mu$ is

$$\sum_{n \in \{1, \cdots, N\}} \left[ \sum_{i \in W_{\mu}(n)} s_{\pi(n)}(x_n(\overline{\mu}(n)), x_i(\overline{\mu}(i))) \right]$$

An unmatched agent or an agent not matched in her productive role receives match surplus 0 and contributes 0 to the social surplus.

## 4 Mechanism Design

A mechanism is a pair of functions $\{q, t\}$ that take as their arguments messages from the $N$ agents, and respectively map those messages into a lottery over assignments and a payment vector. A direct mechanism is one in which each agent’s message space exactly coincides with her type space. By the revelation principle [9], restriction to direct mechanisms proceeds without loss of generality. Let $\hat{x}_n$ denote agent $n$’s report, $\hat{x} = [\hat{x}_1 \: \hat{x}_2 \: \cdots \: \hat{x}_N]$, $\hat{x}_{-n} = [\hat{x}_1 \: \cdots \: \hat{x}_{n-1} \: \hat{x}_{n+1} \: \cdots \: \hat{x}_N]$, and $(\hat{x}_{-n}, z) = [\hat{x}_1 \: \cdots \: \hat{x}_{n-1} \: z \: \hat{x}_{n+1} \: \cdots \: \hat{x}_N]$. The report $\hat{x}_n = 0$ denotes the zero function on $\{1, \cdots, R\}$. Let $\Delta_A$ denote the lottery space over $A$ and let $\Theta = \{x : \{1, \cdots, R\} \to [0, 1] \mid x \text{ is a function}\}$. A direct mechanism in this model is a pair of functions $\{q, t\}$ where $q : \Theta^N \to \Delta_A$ and $t : \Theta^N \to \mathbb{R}^N$. Let $q^\mu(\hat{x})$ be the probability that $\mu$ occurs under the lottery $q(\hat{x})$. The $n$th coordinate $t_n(\hat{x})$ of $t(\hat{x})$ denotes agent $n$’s payment.
A mechanism is feasible if and only if

$$\sum_{\mu \in A} q^\mu(\hat{x}) \leq 1 \text{ for all } \hat{x} \in \Theta^N \quad (1)$$

and

$$q^\mu(\hat{x}) \geq 0 \text{ for all } \hat{x} \in \Theta^N \text{ and for all } \mu \in A \quad (2)$$

Agent $n$’s expected surplus under the lottery $q(\hat{x}_{-n}, \hat{x}_n)$ is

$$H_n(q, \hat{x}_{-n}, \hat{x}_n, x_{-n}, x_n) = \sum_{\mu \in A} \left[ q^\mu(\hat{x}_{-n}, \hat{x}_n) \sum_{i \in W^\mu(n)} s_{i(n)}(x_{n}(\mu(n)), x_{i}(\mu(i))) \right]$$

Agent $n$’s payoff when representing quality $\hat{x}_n$ and the other agents represent $\hat{x}_{-n}$ is

$$U_n(q, t, \hat{x}_{-n}, \hat{x}_n, x_{-n}, x_n) = H_n(q, \hat{x}_{-n}, \hat{x}_n, x_{-n}, x_n) - t_n(\hat{x}_{-n}, \hat{x}_n) \quad (3)$$

A mechanism is ex post individually rational if and only if

$$U_n(q, t, x_{-n}, x_n, x_{-n}, x_n) \geq 0 \text{ for all } n \in \{1, \cdots, N\} \quad (4)$$

In other words, in equilibrium each agent must be at least as well off participating as not.

In order for agents to reveal their respective types truthfully to the mechanism, an incentive compatibility constraint must hold. Ex post incentive compatibility requires that for each agent $n$, when all other agents’ types are revealed, $n$ cannot profitably deviate from truthful reporting. Alternatively, truthful revelation forms an ex post Nash equilibrium. Formally, ex post incentive compatibility requires that for all $x_n, \hat{x}_n \in \Theta$ and all $n \in \{1, \cdots, N\}$

$$U_n(q, t, x_{-n}, x_n, x_{-n}, x_n) \geq U_n(q, t, \hat{x}_{-n}, \hat{x}_n, x_{-n}, x_n) \quad (5)$$
5 VCG-like Mechanism: \( R = 1 \)

This section restricts attention to \( R = 1 \) and suppresses role-specific notation and language; there is one role which by default is the productive role for each agent. This restriction makes it possible to illustrate the payment rule with minimal technical and notational complications.

Agent \( n \)'s type is simply her quality \( x_n \in [0, 1] \). An assignment is a function \( \mu : \{1, \cdots, N\} \rightarrow \{1, \cdots, M\} \cup \{0\} \) such that

1. \( \mu(n) = 0 \) if and only if \( n \) is unmatched, i.e. \( \mu \) assigns \( n \) no slot in any bin
2. \( \mu(n) = m \) if and only if agent \( n \) occupies a slot in bin \( m \)
3. All capacity constraints are satisfied, i.e. \( |\mu^{-1}(m)| \leq c_m \) for all \( m \)

5.1 Existence and Structure of the VCG-like Mechanism

The mechanism of this section is the natural and correct generalization of the well-known VCG mechanism to a one-sided matching environment in which positive assortative matching is the unique efficient stable match. Define

\[
\mu^*_\hat{x} = \arg\max_{\mu \in A} \sum_{n \in \{1, \cdots, N\}} \left[ \sum_{i \in W_\mu(n)} s(\hat{x}_n, \hat{x}_i) \right]
\]

In other words, \( \mu^*_\hat{x} \) is efficient given the report profile \( \hat{x} \). Since the model structure guarantees that the positive assortative match is the unique efficient assignment, \( \mu^*_\hat{x} \) must be the positive assortative match given \( \hat{x} \). The key insight of this extension is that an agent’s payment the opportunity cost she imposes on agent(s) whom she displaces from a bin, rather than every agent on whom she has an effect.

The set of agents displaced from a bin by agent \( n \) is

\[
D_n(\hat{x}) = \{ k \in \{1, \cdots, N\} \setminus \{n\} \mid \mu^*_\hat{x}(k) \neq \mu^*_\hat{x}_{\hat{x}^{-1}(k)} \}
\]
Also define

\[ v_n(\hat{x}) = \sum_{i \in W_{\mu^*}(n)} s(x_n, x_i) \]

\[ \alpha_n(\hat{x}) = \sum_{k \in D_n(\hat{x})} \left[ \sum_{i \in W_{\mu^*}(k)} s(\hat{x}_k, \hat{x}_i) \right] \]

\[ \beta_n(\hat{x}) = \sum_{k \in D_n(\hat{x})} \left[ \sum_{i \in W_{\mu^*}(k)} s(\hat{x}_k, \hat{x}_i) \right] \]

Agent \( n \)'s payment is \( t_n(\hat{x}) = \alpha_n(\hat{x}) - \beta_n(\hat{x}) \).

The VCG-like mechanism operates as follows. Agents simultaneously and independently report their respective types to the mechanism, which selects the socially optimal assignment \( \mu^* \). If there are multiple socially optimal assignments due to identical reports by one or more agents, the mechanism chooses a socially optimal assignment at random. Each agent receives \( v_n(\hat{x}) \) and pays \( t_n(\hat{x}) \).

The VCG-like flavor of this mechanism is now apparent: the mechanism selects a socially optimal assignment given \( \hat{x} \), and charges each agent \( n \) the opportunity cost she imposes on agents in \( D_n(\hat{x}) \). In an independent private values setting, \( D_n(\hat{x}) \) is perforce the set of all agents on whom \( n \) imposes any opportunity cost. In the interdependent value setting of this matching market, the payment rule needs to be more circumspect about which opportunity costs appear in an agent’s payment and which do not in order to satisfy incentive compatibility.

**Theorem 5.1.** *Truthful reporting is an ex post Nash equilibrium in the VCG-like mechanism.*

*Proof.* See Appendix A. \( \square \)

Ex post incentive compatibility is the strongest achievable implementation choice; Williams and Radner [15] showed that interdependent values preclude implementation in dominant strategies.

**Theorem 5.2.** *(Williams and Radner)* Positive assortative matching is never implementable in dominant strategies.
Theorem 5.2 follows from a lemma that generalizes Myerson’s [9] well-known auction implementability result.

**Lemma 5.3.** Truthful reporting in a feasible, individually rational direct mechanism is

1. a dominant strategy if and only if for all \( n \) and \( \bar{x}_n \), \( H_n(q, \bar{x}_n, \bar{x}_n, x_n, x_n) \) is supermodular in \((\bar{x}_n, x_n)\) and

\[
U_n(q, t, \bar{x}_n, x_n, \bar{x}_n, x_n) = U_n(q, t, \bar{x}_n, 0, \bar{x}_n, 0) \tag{6}
\]

\[
+ \int_0^{x_n} \left[ \frac{\partial}{\partial z} H_n(q, \bar{x}_n, \bar{x}_n, \bar{x}_n, z) \right] \bigg|_{\bar{x}_n = z} dz
\]

2. ex post incentive compatible if and only if for all \( n \) and \( x_n \), \( H_n(q, x_n, \bar{x}_n, x_n, x_n) \) is supermodular in \((\bar{x}_n, x_n)\) and

\[
U_n(q, t, x_n, x_n, x_n, x_n) = U_n(q, t, x_n, 0, x_n, 0) \tag{7}
\]

\[
+ \int_0^{x_n} \left[ \frac{\partial}{\partial z} H_n(q, x_n, \bar{x}_n, x_n, z) \right] \bigg|_{\bar{x}_n = z} dz
\]

**Proof.** See Appendix A. \qed

Lemma 5.3 now implies Theorem 5.2.

**Proof.** Suppose that there is a dominant strategy incentive compatible mechanism that implements positive assortative matching. Suppose that \( x_n = 0 \) for \( n = 2, \cdots, N \) and \( \hat{x}_n = 1 - \epsilon \) for \( n = 2, \cdots, N \). Suppose also that \( x_1 = 1 \). If agent 1 reports truthfully \( \hat{x}_1 = 1 \), she wins a slot in bin 1 with \( c_1 - 1 \) randomly assigned bin-mates. She receives \((c_1 - 1)s(1, 0) = 0\). According to the payment rule in Lemma 5.3

\[
t_1(\hat{x}_{-1}, 1) = H_n(q, \hat{x}_{-1}, 1, \hat{x}_{-1}, 1) - \int_0^1 \left[ \frac{\partial}{\partial z} H_n(q, \hat{x}_{-1}, \hat{x}_1, \hat{x}_{-1}, z) \right] \bigg|_{\hat{x}_1 = z} dz
\]

\[
= (c_1 - 1)s(1 - \epsilon, 1 - \epsilon) > 0.
\]

Agent 1’s payoff is \(-(c_1 - 1)s(1 - \epsilon, 1 - \epsilon) < 0\). By reporting \( \hat{x}_1 = 0 \) she would have received non-negative surplus and paid 0. Agents \( 2, \cdots, N \) have types and strategies that make truth-
telling strictly worse for agent 1 than some other option (in this example, reporting 0). This statement is exactly that truthful reporting is not weakly dominant, a contradiction.

In particular, Theorem 5.2 implies that truthful reporting is not, in this environment, a dominant strategy in the naive extension of the VCG mechanism. With interdependence, an agent’s naive VCG payment depends on her own report due to the agent’s externalities she imposes on her partners. Since an agent’s payment depends on her own report, the naive VCG mechanism loses its desirable dominant strategy property. Indeed, in this matching model, truthful reporting does not even form an ex post Nash equilibrium in the naive VCG mechanism.

Truthful reporting in this paper’s mechanism is actually slightly stronger than ex post Nash equilibrium: Edelman, Ostrovsky, and Schwarz’s [3] locally envy free equilibrium extends to this incomplete information environment.

5.2 Envy-Free Equilibria

Definition 5.1. An ex post Nash equilibrium of a mechanism \( \{q, t\} \) is envy free if each agent \( n \) receives identical payoffs under all \( \mu \) that receive positive weight under \( q(x) \).

Envy freeness is a stronger condition than ex post incentive compatibility. Ex post incentive compatibility merely requires that agents be indifferent across tie-breaking in expectation, while envy freeness requires that agents be ex post indifferent across tie-breaking.

Theorem 5.4. Truthful reporting in the VCG-like mechanism is an envy free equilibrium.

Proof. Suppose that two agents, \( n \) and \( n' \) submit identical (truthful) reports. It is sufficient to show that \( v_n(x) - t_n(x) = v_{n'}(x) - t_{n'}(x) \) for all socially optimal \( \mu \). If \( \mu(n) = \mu(n') \) for all socially optimal \( \mu \), then \( D_n(x) = D_{n'}(x) \) and the theorem follows. Now suppose that there is a socially optimal \( \mu \) such that \( \mu(n) \neq \mu(n') \) for some agents \( n, n' \) with identical types.

Suppose that \( x_n = x_{n'}, \mu^*_x(n) = m, \) and \( \mu^*_{x'}(n') = m' \). Without loss of generality, assume
that \( m' = m + 1 \). Then

\[
v_n(x) - t_n(x) = \sum_{i \in W_{\mu_k^*}(n)} s(x_n, x_i) + \sum_{k \in D_n(x)} \left[ \sum_{i \in W_{\mu_k^*}(k)} s(x_k, x_i) - \sum_{i \in W_{\mu_{\mu_{x-n},0}}(k)} s(x_k, x_i) \right]
\]

Since \( \mu(n) = m \) and \( \mu(n') = m + 1 \), \( D_n(x) = \{n'\} \cup D_{n'}(x) \). Therefore

\[
v_n(x) - t_n(x) = \sum_{i \in W_{\mu_k^*}(n)} s(x_n, x_i) + \sum_{k \in D_{n'}(x)} \left[ \sum_{i \in W_{\mu_k^*}(k)} s(x_k, x_i) - \sum_{i \in W_{\mu_{\mu_{x-n},0}}(n')} s(x_k, x_i) \right]
\]

Since \( \mu_{\mu_{x-n},0}(n') = m \) and \( x_n = x_{n'} \), \( \sum_{i \in W_{\mu_k^*}(n)} s(x_n, x_i) = \sum_{i \in W_{\mu_{\mu_{x-n},0}}(n')} s(x_{n'}, x_i) \). Thus

\[
v_n(x) - t_n(x) = \sum_{i \in W_{\mu_k^*}(n')} s(x_{n'}, x_i) + \sum_{k \in D_{n'}(x)} \left[ \sum_{i \in W_{\mu_k^*}(k)} s(x_k, x_i) - \sum_{i \in W_{\mu_{\mu_{x-n},0}}(n')} s(x_k, x_i) \right]
\]

Finally, for all \( k \in D_{n'}(x) \), \( \mu_{\mu_{x-n},0}(k) = \mu_{\mu_{x-n},0}(n') \). Therefore

\[
v_n(x) - t_n(x) = \sum_{i \in W_{\mu_k^*}(n')} s(x_{n'}, x_i) + \sum_{k \in D_{n'}(x)} \left[ \sum_{i \in W_{\mu_k^*}(k)} s(x_k, x_i) - \sum_{i \in W_{\mu_{\mu_{x-n},0}}(n')} s(x_k, x_i) \right]
\]

\[= v_{n'}(x) - t_{n'}(x)\]

\[\square\]

6 Multiple Roles

This section examines the mechanism in the fully general model in which \( R \geq 1 \). With some minor notational modifications and one observation, the results from Section 5 generalize

\[\text{If } m' > m + 1, \text{ then there must be some other agent } n'' \text{ whose quality } x_{n''} = x_n = x_{n'} \text{ and } \mu_{\mu_{x-n}}(n'') = m + 1. \text{ Transitively, it is sufficient to consider } n'' \text{ and } n.\]
completely. Define
\[
\mu^*_x = \arg\max_{\mu \in A} \sum_{n \in \{1, \ldots, N\}} \left[ \sum_{i \in W_{\mu}(n)} s_{\mu(n)}(\tilde{x}_n(\mu(n)), \tilde{x}_i(\mu(i))) \right]
\]

The set of agents displaced from a firm by agent \( n \) is
\[
D_n(\bar{x}) = \{ k \in \{1, \ldots, N\} \setminus \{n\} \mid \mu^*_x(k) \neq \mu^*_{x \setminus n,0}(k) \}
\]

Also define
\[
v_n(\bar{x}) = \sum_{i \in W_{\mu^*_x}(n)} s_{\mu^*_x}(x_n(\mu^*_x(n)), x_i(\mu^*_x(i)))
\]
\[
\alpha_n(\bar{x}) = \sum_{k \in D_n(\bar{x})} \left[ \sum_{i \in W_{\mu^*_x,\setminus n,0}(k)} s_{\mu^*_x}(\tilde{x}_k(\mu^*_x(k)), \tilde{x}_i(\mu^*_x(i))) \right]
\]
\[
\beta_n(\bar{x}) = \sum_{k \in D_n(\bar{x})} \left[ \sum_{i \in W_{\mu^*_x}(k)} s_{\mu^*_x}(\tilde{x}_k(\mu^*_x(k)), \tilde{x}_i(\mu^*_x(i))) \right]
\]

Agent \( n \)'s payment is \( t_n(\bar{x}) = \alpha_n(\bar{x}) - \beta_n(\bar{x}) \). While notationally more complex than the payment defined in Section 5, the intuition is the same: agent \( n \) who wins a slot in bin \( m \) in role \( r \) pays the opportunity cost she imposes on agents whom her report displaces from a bin. An agent’s report can displace another agent from a bin if and only if both agents report the same productive role.

The VCG-like mechanism operates exactly as in Section 5. Agents submit their reports to the mechanism, which selects the socially optimal assignment \( \mu^*_{x^r} \). If there are multiple socially optimal assignments due to identical reports by one or more agents, the mechanism chooses a socially optimal assignment at random. Each agent receives \( v_n(\bar{x}) \) and pays \( t_n(\bar{x}) \).

Fix some \( r \in \{1, \ldots, R\} \). If the types of agents whose productive role is not \( r \) were public information, then the remaining mechanism design problem concerning agents whose productive role is \( r \) is identical to the mechanism design problem in which \( R = 1 \) and in which the types of agents whose productive role is not \( r \) enter the match surplus function \( s_r(\cdot, \cdot) \) as
parameters. The generalization of Section 5’s results relies on the this observation.

**Theorem 6.1.** Truthful reporting is an ex post Nash equilibrium is the VCG-like mechanism \((R \geq 1)\).

*Proof.* Fix an arbitrary \(r \in \{1, \cdots, R\}\) and suppose that all agents whose productive role is not \(r\) report their types truthfully. Then set of all agents whose productive role is \(r\) report to a mechanism identical to the mechanism of Section 5 in which the types of the agents whose productive role is not \(r\) enter as parameters. Theorem 5.1 now implies the result. \(\square\)

Since \(R \geq 1\) includes the case \(R = 1\), \([15]\)’s result that there is no dominant strategy incentive compatible mechanism still holds.

**Theorem 6.2.** Truthful reporting in the VCG-like mechanism is an envy free equilibrium \((R \geq 1)\).

*Proof.* Fix an arbitrary \(r \in \{1, \cdots, R\}\) and suppose that all agents whose productive role is not \(r\) report their types truthfully. The set of all agents whose productive role is \(r\) report to a mechanism identical to the mechanism of Section 5 in which the types of the agents whose productive role is not \(r\) enter as parameters. Within each role, truthful reporting is envy free. Thus when all agents report truthfully, truthful reporting is envy free across all roles. \(\square\)

Last, using a generalization of Lemma 5.3 to \(R\) roles, we show that up to a constant, the VCG-like mechanism is revenue equivalent to all other (Bayesian) incentive compatible mechanisms that implement positive assortative matching.

**Lemma 6.3.** Truthful reporting in a feasible, individually rational direct mechanism is

1. a dominant strategy if and only if for all \(n\) and \(\tilde{x}_n\), \(H_n(q, \tilde{x}_n, \tilde{x}_n, \tilde{x}_n, x_n)\) is supermodular in \((\tilde{x}_n, x_n)\) and

\[
U_n(q, t, \tilde{x}_n, x_n, \tilde{x}_n, x_n) = U_n(q, t, \tilde{x}_n, 0, \tilde{x}_n, 0) + \int_{0}^{x_n} \left[ \frac{\partial}{\partial z} H_n(q, \tilde{x}_n, \tilde{x}_n, \tilde{x}_n, z) \right] \bigg|_{\tilde{x}_n = z} \, dz
\]
2. \textit{ex post incentive compatible} if and only if for all \( n \) and \( x_n \), \( H_n(q, x_n, \hat{x}_n, x_n) \) is supermodular in \((\hat{x}_n, x_n)\) and

\[
U_n(q, t, x_n, x_n, x_n) = U_n(q, t, x_n, 0, x_n, 0) + \int_0^{x_n} \left[ \frac{\partial}{\partial z} H_n(q, x_n, \hat{x}_n, x_n, z) \right]_{\hat{x}_n = z} dz
\]

3. \textit{Bayesian incentive compatible} if and only if for all \( n \) and \( \hat{x}_n \), \( E_{x_n}[H_n(q, \hat{x}_n, \hat{x}_n, \hat{x}_n, x_n)] \) is supermodular in \( \hat{x}_n, x_n \) and

\[
E_{x_n}[U_n(q, t, \hat{x}_n, x_n, \hat{x}_n, x_n)] = E_{x_n}[U_n(q, t, \hat{x}_n, 0, \hat{x}_n, 0)] + \int_0^{x_n} E_{x_n} \left[ \frac{\partial}{\partial z} H_n(q, \hat{x}_n, \hat{x}_n, \hat{x}_n, z) \right]_{\hat{x}_n = z} dz
\]

Proof. Fix an arbitrary \( r \in \{1, \cdots, R\} \) and suppose that all agents whose productive role is not \( r \) report their types truthfully. The set of all agents whose productive role is \( r \) report to a mechanism identical to the mechanism of Section 5 in which the types of the agents whose productive role is not \( r \) enter as parameters. The proof of Lemma 5.3 then applies to each role.

Corollary 6.4. A matching mechanism’s expected revenue is completely determined by \( q \) and the (expected) payoff of to an agent whose quality is 0 in her productive role.

Proof. Recall that Bayesian incentive compatibility implies \textit{ex post} incentive compatibility, so that proving the revenue equivalence of Bayesian incentive compatible mechanisms includes revenue equivalence of \textit{ex post} incentive compatible mechanisms. Equation (10) says that the expected payment of an agent \( n \) depends exclusively on \( q \) and \( E_{x_n}[U_n(q, t, \hat{x}_n, 0, \hat{x}_n, 0)] \). The mechanism’s expected revenue is the sum of expected payments, hence expected revenue depends solely on \( q \) and \( E_{x_n}[U_n(q, t, \hat{x}_n, 0, \hat{x}_n, 0)] \).

6.1 Marriage Markets and Internet Advertising

A marriage market is a matching market in which the matched unit is a household, the two productive roles are man and woman, the \( N \) agents consist of \( K \) men and \( L \) women. The
number of households is not a binding constraint; \( M \geq \min\{K, L\} \). Each household \( m \) has
capacity constraints \( c_{m,\text{man}} = c_{m,\text{woman}} = 1 \). Each man reports his quality and each woman
reports her quality to the mechanism, and the man and woman reporting the \( k \)th highest
qualities in their respective roles are matched.

The complete information version of this marriage market encapsulates the internet ad-
vertising market of [3]. Edelman, Ostrovsky and Schwarz study the generalized second price
(GSP) auction run by internet search engines for sponsored search slots. The search engine
side of the market pays nothing; the qualities of the sponsored search slots (measured as the
probability of a click) are known to all advertisers, with higher-ranked slots having higher
click-through rates. The advertisers bid for slots; the \( k \)th highest bidder wins the \( k \)th best
sponsored search slot and pays the \( k + 1 \)st highest bid. The \( k + 1 \)st highest bid in the GSP
auction exactly coincides with the \( k \)th highest bidder’s payment in the VCG-like mechanism.

6.2 Many-to-One Matching

A labor market is a matching market in which the matched unit is a firm, and each firm
hires agents into one or more productive roles. Classical models consider two roles, manager
and worker. This model and mechanism are capable of handling any finite number of roles.
A university, for example, might hire a one president, several provosts, many deans, full,
associate, and assistant professors and staff. The \( N \) agents consist of all candidates eligible to
fill whatever roles are available.

This example details the mechanism in a matching market between hospitals and doctors.
This market has two productive roles, hospital \( h \) and doctor \( d \). There are two firms. The
capacity constraints are \( c_{1h} = c_{2h} = 1, c_{1d} = 2, \) and \( c_{2d} = 1 \). Hospitals and doctors report their
respective qualities to the mechanism which matches the best hospital with the best doctors
(until \( c_{1d} \) is satisfied or until the supply of doctors is exhausted) and the second best hospital
with the next best doctors. Suppose that there are six agents \{1, 2, 3, 4, 5, 6\} and two firms
\{1, 2\}. Agent 1’s and 2’s productive roles are \( h \), and agent 3’s, 4’s, 5’s, and 6’s productive roles
are \( d \). Without loss of generality, suppose that \( x_1(h) \geq x_2(h) \) and \( x_3(d) \geq x_4(d) \geq x_5(d) \geq
x_6(d) \). The positive assortative match gives the following assignment \( \mu^*_x, D_n(x) \), and \( \mu^*_{x-n,0} \).
The VCG-like mechanism gives the following assignment, surpluses, and payments:

<table>
<thead>
<tr>
<th>Agent</th>
<th>$W_{\mu_k}(n)$</th>
<th>$D_n(x)$</th>
<th>$W_{\mu_{x-0.0}}^*(k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>{3, 4}</td>
<td>{2}</td>
<td>$W_{\mu_{x-1.0}}^*(2) = {3, 4}$</td>
</tr>
<tr>
<td>2</td>
<td>{5}</td>
<td>$\emptyset$</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>3</td>
<td>{1, 4}</td>
<td>{5, 6}</td>
<td>$W_{\mu_{x-3.0}}^*(5) = {1, 4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$W_{\mu_{x-3.0}}^*(6) = {2}$</td>
</tr>
<tr>
<td>4</td>
<td>{1, 3}</td>
<td>{5, 6}</td>
<td>$W_{\mu_{x-4.0}}^*(5) = {1, 3}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$W_{\mu_{x-4.0}}^*(6) = {2}$</td>
</tr>
<tr>
<td>5</td>
<td>{2}</td>
<td>{6}</td>
<td>$W_{\mu_{x-5.0}}^*(6) = {2}$</td>
</tr>
<tr>
<td>6</td>
<td>$\emptyset$</td>
<td>$\emptyset$</td>
<td>$\emptyset$</td>
</tr>
</tbody>
</table>

The table shows the assignment and payments for each agent.
Truthful reporting is an envy free equilibrium: if two agents with the same productive role have the same quality, their payoffs are equal. For example, if \( x_1(h) = x_2(h) \), hospitals 1 and 2 each receive

\[
\sum_{k \in W_{\mu_2}^* (2)} s_H(x_2(h), x_k(\mu_2^*(k))) = \sum_{k \in W_{\mu_2}^* (1)} s_H(x_2(h), x_k(\mu_2^*(k))) - \left[ \sum_{k \in W_{\mu_2}^* (2)} s_H(x_2(h), x_k(\mu_2^*(k))) - \sum_{x \in W_{\mu_2}^* (2)} s_H(x_2(h), x_k(\mu_2^*(k))) \right]
\]

Since \( W_{\mu_2}^* (2) = W_{\mu_2}^* (1) \), these expressions are equal when \( x_1(h) = x_2(h) \). Similarly, if \( x_5(d) = x_6(d) \), doctors 5 and 6 each receive 0. This suggests a natural interpretation of payoffs as wages. If agents in a given role have approximately the same qualities, wages tend to compress. In the limiting case where \( x_1(h) = x_2(h) \) the two hospitals each receive payoff 0. A similar observation obtains for doctors when \( x_3(d) = x_4(d) = x_5(d) = x_6(d) \). If agents’ qualities in some role \( r \) are more homogeneous, there is less to be gained by offering higher wages for higher quality agents.

### 7 Conclusion

This paper shows that all incomplete information matching markets in which positive assortative matching on agent quality is the unique efficient stable match can be unified under a single mechanism design framework. The resulting VCG-like mechanism is new to the literature and always implements positive assortative matching as ex post Nash equilibrium. Indeed, this equilibrium satisfies the slightly stronger condition of envy freeness introduced by [3] and generalizes locally envy freeness to the incomplete information matching environments considered here. The main contribution of this work is the correct generalization of the VCG mechanism to the interdependent values setting of these matching markets: the payment rule distinguishes between the opportunity cost an agent imposes on another by depriving that other agent of an object (a slot in a given bin) and the opportunity cost an agent imposes on another by changing the value of that other agent’s won object (due to interdependence). In
order for truthful reporting to form an ex post Nash equilibrium, each agent’s payment must be based on the former but exclude the latter. One drawback to this mechanism is that in a matching market (unlike a pure auction environment) agents have an incentive to try to match outside the mechanism to avoid making payments. While this problem disappears under the assumption that the mechanism can prevent agents from matching outside, a more elegant approach would be a revenue-neutral extension of the mechanism.

Of particular note is the unification of one-sided and two-sided matching markets. These types of markets typically display very different behaviors; this paper is the only one of which this author is aware that links these two types of markets; further work on such relationships would prove interesting albeit difficult.

Thus far, progress on incomplete information matching problems has been limited to markets in which positive assortative matching is the unique stable efficient match. Interesting avenues for future work include relaxing the restrictions on surplus functions that force us to restrict our attention to positive assortative matching. Another interesting topic is the auction design implications of the VCG-like mechanism.

References


A  Proofs

A.1  Proof of Theorem 5.1

Proof. The proof is similar to the proof that truthful revelation is a weakly dominant strategy in the VCG mechanism.

\[ v_n(x) - t_n(x) = v_n(x) + \beta_n(x) - \alpha_n(x) \]

\[ + \sum_{k \in \{1, \ldots, N\} \setminus [D_n(x) \cup \{n\}]} \left[ \sum_{i \in W_{\mu_k^*}^*(k)} s(x_k, x_i) \right] - \sum_{k \in \{1, \ldots, N\} \setminus [D_n(x) \cup \{n\}]} \left[ \sum_{i \in W_{\mu_k^*}^*(k)} s(x_k, x_i) \right] \]

\[ = \sum_{k \in \{1, \ldots, N\}} \left[ \sum_{i \in W_{\mu_k^*}^*(k)} s(x_k, x_i) \right] - \alpha_n(x) - \sum_{k \in \{1, \ldots, N\} \setminus [D_n(x) \cup \{n\}]} \left[ \sum_{i \in W_{\mu_k^*}^*(k)} s(x_k, x_i) \right] \]

By the definition of \( \mu_k^* \),

\[ v_n(x) - t_n(x) \geq \sum_{k \in \{1, \ldots, N\}} \left[ \sum_{i \in W_{\mu_k^*}^*(k)} s(x_k, x_i) \right] - \alpha_n(x) - \sum_{k \in \{1, \ldots, N\} \setminus [D_n(x) \cup \{n\}]} \left[ \sum_{i \in W_{\mu_k^*}^*(k)} s(x_k, x_i) \right] \]

Rearranging:

\[ v_n(x) - t_n(x) \geq v_n(x_{-n}, z) - t_n(x_{-n}, z) + \alpha_n(x_{-n}, z) - \alpha_n(x) \]

\[ - \sum_{k \in \{1, \ldots, N\} \setminus [D_n(x) \cup \{n\}]} \left[ \sum_{i \in W_{\mu_k^*}^*(k)} s(x_k, x_i) \right] \]

\[ + \sum_{k \in \{1, \ldots, N\} \setminus [D_n(x_{-n}, z) \cup \{n\}]} \left[ \sum_{i \in W_{\mu_k^*_{x_{-n}, z}}^*(k)} s(x_k, x_i) \right] \] (11)
If \( z < x_n \), then \( D_n(x) \supseteq D_n(x_{-n}, z) \), and Equation (11) consequently becomes

\[
v_n(x) - t_n(x) \geq v_n(x_{-n}, z) - t_n(x_{-n}, z) - \sum_{k \in D_n(x) \setminus D_n(x_{-n}, z)} \left[ \sum_{i \in W_{x_{-n}}^*(k)} s(x_k, x_i) \right]
\]

\[
+ \sum_{k \in D_n(x) \setminus D_n(x_{-n}, z)} \left[ \sum_{i \in W_{x_{-n}}^*(k)} s(x_k, x_i) \right]
\]

Agents in \( D_n(x) \setminus D_n(x_{-n}, z) \) are at least as well off when \( n \) reports \( z \) as when \( n \) reports 0, thus

\[
v_n(x) - t_n(x) \geq v_n(x_{-n}, z) - t_n(x_{-n}, z)
\]

Alternatively, if \( z > x_n \), then \( D_n(x) \subseteq D_n(x_{-n}, z) \), and Equation (11) consequently becomes

\[
v_n(x) - t_n(x) \geq v_n(x_{-n}, z) - t_n(x_{-n}, z)
\]

\[- \sum_{k \in D_n(x_{-n}, z) \setminus D_n(x)} \left[ \sum_{i \in W_{x_{-n}}^*(k)} s(x_k, x_i) \right]
\]

\[
+ \sum_{k \in D_n(x_{-n}, z) \setminus D_n(x)} \left[ \sum_{i \in W_{x_{-n}}^*(k)} s(x_k, x_i) \right]
\]

Agents in \( D_n(x_{-n}, z) \setminus D_n(x) \) are at least as well off when \( n \) reports 0 as when \( n \) reports \( x_n \), thus

\[
v_n(x) - t_n(x) \geq v_n(x_{-n}, z) - t_n(x_{-n}, z)
\]

In particular, \( v_n(x) - t_n(x) \geq v_n(x_{-n}, 0) - t_n(x_{-n}, 0) = 0 \), so it is ex post individually rational for agent \( n \) to report truthfully. \( \square \)

### A.2 Proof of Lemma 5.3

**Proof.** This is the proof of part 1 of Lemma 5.3; the proof of part 2 follows analogously.
A mechanism \( \{q,t\} \) is individually rational if for all \( n \in \{1, \cdots, N\} \)
\[
U_n(q, t, \hat{x}_{-n}, x_n, x_{-n}, x_n) \geq 0
\] (12)

A mechanism is dominant strategy incentive compatible if for all \( n \in \{1, \cdots, N\} \) and for all \( x_n, \hat{x}_n \in [0,1] \)
\[
U_n(q, t, \hat{x}_{-n}, x_n, x_{-n}, x_n) \geq U_n(q, t, \hat{x}_{-n}, \hat{x}_n, x_{-n}, x_n)
\] (13)

First, show that if \( \{q, t\} \) is feasible, individually rational, and dominant strategy incentive compatible, it must satisfy the conditions of part 1 of Lemma 5.3. Rewrite the right hand side of Equation (13) as
\[
U_n(q, t, \hat{x}_{-n}, x_n, x_{-n}, x_n) + \sum_{\mu \in A} \left[ q^\mu(\hat{x}_{-n}, \hat{x}_n) \sum_{i \in W_{\mu}(n)} [s(x_n, x_i) - s(\hat{x}_n, x_i)] \right]
\] (14)

Replace (14) into the right hand side of Equation (13). Thus the incentive compatibility constraint is equivalent to
\[
U_n(q, t, \hat{x}_{-n}, x_n, x_{-n}, x_n) \geq U_n(q, t, \hat{x}_{-n}, \hat{x}_n, x_{-n}, x_n)
\] + \sum_{\mu \in A} \left[ q^\mu(\hat{x}_{-n}, \hat{x}_n) \sum_{i \in W_{\mu}(n)} [s(x_n, x_i) - s(\hat{x}_n, x_i)] \right]
\] (15)

Transpose \( x_n \) and \( \hat{x}_n \) in Equation (15):
\[
U_n(q, t, \hat{x}_{-n}, \hat{x}_n, x_{-n}, x_n) \geq U_n(q, t, \hat{x}_{-n}, \hat{x}_n, x_{-n}, \hat{x}_n)
\] + \sum_{\mu \in A} \left[ q^\mu(\hat{x}_{-n}, \hat{x}_n) \sum_{i \in W_{\mu}(n)} [s(\hat{x}_n, x_i) - s(x_n, x_i)] \right]
\] (16)

Combine Equations (15) and (16), and rearrange; we see that incentive compatibility is equiv-
alent to the supermodularity of $H_n(q, \hat{x}_n, x_n, x_n)$ in $\hat{x}_n, x_n$:

$$H_n(q, \hat{x}_n, x_n, x_n, x_n) + H_n(q, \hat{x}_n, x_n, x_n, \hat{x}_n) \geq H_n(q, \hat{x}_n, x_n, x_n, x_n) + H_n(q, \hat{x}_n, x_n, x_n, \hat{x}_n) \quad (17)$$

The payment rule, Equation (6), follows from the standard method of equating the direct and indirect utility functions. Evaluate the incentive compatibility constraint at $x_n = \hat{x}_n$.

The resulting inequality combined with the envelope theorem implies that

$$\frac{d}{dx_n} U_n(q, t, \hat{x}_n, x_n, \hat{x}_n, x_n) = \left[ \frac{\partial}{\partial x_n} U_n(q, t, \hat{x}_n, \hat{x}_n, x_n) \right]_{\hat{x}_n = x_n} \quad (18)$$

Integrating Equation (18)

$$U_n(q, t, \hat{x}_n, x_n, \hat{x}_n, x_n) = U_n(q, t, \hat{x}_n, 0, \hat{x}_n, 0) + \int_0^{x_n} \left[ \frac{\partial}{\partial z} H_n(q, \hat{x}_n, \hat{x}_n, \hat{x}_n, z) \right]_{\hat{x}_n = z} \, dz \quad (19)$$

Since $t_n(\hat{x}_n, x_n)$ is independent of $x_n$, Equation (19) is equivalent to

$$U_n(q, t, \hat{x}_n, x_n, \hat{x}_n, x_n) = U_n(q, t, \hat{x}_n, 0, \hat{x}_n, 0) + \int_0^{x_n} \left[ \frac{\partial}{\partial z} H_n(q, \hat{x}_n, \hat{x}_n, \hat{x}_n, z) \right]_{\hat{x}_n = z} \, dz \quad (20)$$

Now show that if $\{q,t\}$ is feasible, individually rational, and satisfies the conditions of part 1 of Lemma 5.3, it must be dominant strategy incentive compatible. Evaluating the second to last arguments of $U_n$ and $H_n$ in Equation (20) at $\hat{x}_n = x_n$

$$U_n(q, t, \hat{x}_n, x_n, x_n, x_n) = U_n(q, t, \hat{x}_n, 0, x_n, 0) + \int_0^{x_n} \left[ \frac{\partial}{\partial z} H_n(q, \hat{x}_n, \hat{x}_n, x_n, z) \right]_{\hat{x}_n = z} \, dz \quad (21)$$

27
By Equation (21)

\[
U_n(q, t, \hat{x}_{-n}, x_n, x_{-n}, x_n) = U_n(q, t, \hat{x}_{-n}, x'_n, x_{-n}, x_n') \\
+ \int_{x_n}^{x_n'} \left[ \frac{\partial}{\partial z} H_n(q, \hat{x}_{-n}, x_{-n}, z) \right]_{\hat{x}_{-n} = z} \, dz
\]

(22)

By the supermodularity of \( H_n(q, \hat{x}_{-n}, x_{-n}, x_n) \) in \( \hat{x}_{-n}, x_{-n}, x_n \), we conclude that the integrand is increasing in its third argument.\(^3\) Suppose that \( x_n \geq x_n' \). Then

\[
U_n(q, t, \hat{x}_{-n}, x_n, x_{-n}, x_n) \geq U_n(q, t, \hat{x}_{-n}, x'_n, x_{-n}, x_n') \\
+ \int_{x_n}^{x_n'} \left[ \frac{\partial}{\partial z} H_n(q, \hat{x}_{-n}, x'_n, x_{-n}, z) \right]_{\hat{x}_{-n} = z} \, dz
\]

(23)

Integrating

\[
U_n(q, t, \hat{x}_{-n}, x_n, x_{-n}, x_n) \geq U_n(q, t, \hat{x}_{-n}, x'_n, x_{-n}, x_n') \\
+ \sum_{\mu \in A} q^\mu(\hat{x}_{-n}, x'_n) \sum_{i \in W_{\mu}(n)} \left[ s(x_n, x_i) - s(x'_n, x_i) \right]
\]

(24)

Equation (24) is exactly Equation (15) with \( \hat{x}_n \) relabeled as \( x'_n \). Thus part 1 of Lemma 5.3 describes sufficient conditions for a feasible, individually rational mechanism to be dominant strategy incentive compatible.

\(\square\)

---

\(^3\)Since \( H_n(q, \hat{x}_{-n}, \hat{x}_n, x_{-n}, x_n) \) is supermodular in \( \hat{x}_n, x_n \), the cross derivative is non-negative. Therefore, \( \frac{\partial H_n}{\partial x_n} \) must be increasing in \( \hat{x}_n \).