Strategy-Proofness for Hospitals in Matching Markets

Hector Lopez*

Abstract

Strategy-proof implementation is one of the many elements that have contributed to the successful application of matching theory in real life. However, in many-to-one matching markets without transfers (e.g., doctors to hospitals with fixed salaries) there is no stable mechanism which is strategy-proof for hospitals. Furthermore, strategy-proofness and stability cannot be achieved for both hospitals and doctors simultaneously even in one-to-one matching markets. This paper shows that in many-to-one matching markets with transfers it is possible to guarantee stability and strategy-proofness-for-hospitals whenever an opportunity cost condition is satisfied. In addition, it is shown that stability and strategy-proofness are possible for both hospitals and doctors simultaneously. Finally, it is shown that strategy-proofness can be achieved in the interior of the core.

JEL classification: C62; C71; C78; D47.

Keywords: Opportunity cost condition; assignment game; hospital strategy-proofness; VCG auctions; interior strategy-proofness.

1 Introduction

Many matching markets have successfully adopted centralized mechanisms as an alternative to the price system. In these centralized markets, the final allocation is computed using information

*E-mail addresses: hlopezca@umd.edu.

provided by market participants. Typically, the final allocation is chosen to be stable with respect to the provided information. However, it is desirable that stability holds with respect to the actual information. Hence, participants’ incentives to report truthfully are extremely important. In practice, it is often a goal to make participants best strategy to report their actual information, regardless of other agents’ reports. Whenever this is achieved it is said that the mechanism is strategy-proof.

Strategy-proof mechanisms possess several advantages over other mechanisms. First, it is possible to guarantee that efficiency or other properties hold with respect to actual and not only with respect to reported information. This is particularly important for many institutions since it has been shown experimentally\(^2\) that in some non-strategy-proof mechanisms, up to 80% of agents misrepresent their preferences. Furthermore, submitted preferences are often used in welfare assessments,\(^3\) for example by computing the number of participants obtaining their first choice, second choice and so on. Second, no resources are wasted by market participants in order to compute better-than-truthful reports. In an auction, which is a special kind of matching market, bidders spend a lot of time and money devising their strategies and they often hire auction consultants.\(^4\) In school choice, parents spend time and money to obtain better outcomes from non-strategy-proof mechanisms.\(^5\) Third, participants with more information about the market cannot take advantage of less informed participants. This is particularly important in school choice, where equal access to education is often a goal.\(^6\)

In a series of papers, Roth studied several incentive properties of stable mechanisms in two-sided markets without transfers. In these markets there are two kinds of agents: hospitals and doctors. Hospitals want to be matched with a set of doctors and doctors want to be matched with at most one hospital. The market is said to be one-to-one whenever hospitals want to be matched with at most one doctor and many-to-one otherwise. Roth showed that (i) in one-to-one matching markets, it is possible to obtain strategy-proofness for doctors or hospitals\(^,[Roth, 1982]\) but not both simultaneously\(^,[Roth, 1984]\) and (ii) in many-to-one matching markets, strategy-proofness can be guaranteed for doctors, but not for hospitals\(^,[Roth, 1985]\).

\(^2\)See Chen and Sönmez, 2006 for an example
\(^3\)See Featherstone [2011] for a discussion about rankings as a welfare criterion.
\(^4\)Some problems faced by bidders in high-stake auctions are described in Cramton and Schwartz [2000], Milgrom et al. [2009]
\(^6\)See also Pathak and Sönmez [2008, 2013] for a revealing discussion about the issue.
This paper studies many-to-one markets with transfers and shows that stability and strategy-proofness for both hospitals and all agents (doctors and hospitals simultaneously) are possible by characterizing the conditions under which the Vickrey-Clarke-Groves (VCG) mechanism is stable. The VCG mechanism is always strategy-proof.

This paper uses a version of the assignment game proposed by Shapley and Shubik [1971] as generalized by Kelso and Crawford [1982]. In this model, doctors want to be matched to at most one hospital and hospitals can be matched to multiple doctors. In this model, transfers can be made continuously, in discrete quantities, or not at all.

The first contribution of this paper is to show that a VCG mechanism, different from the pivot mechanism, is stable whenever agents’ preferences satisfy an opportunity cost condition. This condition is satisfied whenever every doctor can find a hospital that offers him at least his opportunity cost. In this mechanism, hospital \( i \) receives a payoff \( \pi_i^V = V(A) - V(A \setminus i) \), where \( V \) is the coalitional value function and \( A \) is the set of agents in the market. Leonard [1983] proved the result in one-to-one matching markets and his technique rests completely on the unit demand assumption. Gul and Stacchetti [1999] proved the result for replica economies.

The second contribution is to identify markets where strategy-proofness, together with stability, can be achieved for all agents. The above mechanism cannot be used directly for all agents since, in general, \( \sum_{i \in A} \pi_i^V > V(A) \). However, careful inspection reveals that if agent \( i \) is receiving a payoff of \( \pi_i^V \) he is fully capturing two marginal values. When a new agent enters a matching market, say a doctor, two effects take place. First, a new agent who demands hospitals enters the market; and second, a new object is available for hospitals currently in the market. In order to obtain strategy-proofness we only need agents to be able to capture the marginal value of their information i.e. the marginal value they produce as demanders not as objects. For agent \( i \), this marginal value is captured by \( \pi_i^U = U(S) - U(S \setminus i) \), where \( U(T) \) is the maximum value that can be achieved with all agents present only using the information of agents in \( T \). Thus \( \pi_i^U \) is the marginal contribution of agent \( i \)'s private information and \( \pi_i^U - \pi_i^V \) is the marginal contribution of agent \( i \)'s existence as an object. If no agent’s information is pivotal (the agent himself is pivotal), then the \( U \) mechanism is strategy-proof for all agents and stable whenever the private values of an efficient matching belong to the set of stable

---

7See Vickrey [1961]
The third contribution is to show that strategy-proofness either for hospitals or doctors can be achieved without offering an extremal matching. In general, strategy-proofness is achieved by offering agents their most preferred stable allocation.\(^8\) In continuous transfers models, this payoff is characterized by \(\pi_i^V = V(A) - V(A \setminus i)\). However, as discussed above, strategy-proofness can also be obtained by offering agents the marginal contribution of their information, \(\pi_i^U = U(A) - U(A \setminus i)\). We show that the stability of these payoffs is directly linked to the proportion of surplus generated at each side of the market. In particular, for any vector of stable payoffs \(\{\pi_i\}_A\), there is a division of total surplus that makes the \(U\) mechanism stable and strategy-proof, i.e. \(\pi_i = U(A) - U(A \setminus i)\) for all agents.

Finally, it is shown that strategy-proofness for hospitals can be implemented in discrete transfers. The level to which transfers can be used in a matching markets varies significantly. However, this paper shows that allowing transfers to be negotiated in the matching process not only would improve efficiency (with respect to the reports), but also makes agents more willing to report their true private information. This increases the efficiency (with respect to the actual preferences) and accountability of the market.

\section{Related Literature}

Incentives in matching markets have been studied systematically since Roth’s contributions [1982, 1984, 1985]. A first line of research has been devoted to finding restrictions on preference profiles for mechanisms to achieve strategy-proofness. Demange and Gale [1985] showed that Roth’s conclusions hold in very general one-to-one environments where agents have preferences over each other and all have possibly different valuations over money. Alcalde and Barberà [1994] showed that the Gale-Shapley algorithm is the mechanism that achieves stability and strategy-proofness in the biggest set of preference profiles where both are possible. Sönmez [1997, 1999] introduced two kinds of manipulations observed in real matching markets: capacities misrepresentations and pre-arranged matches. He showed that there is no mechanism capable of avoiding such manipulations in general. Later on, Kojima [2007] and Kesten [2012] showed that some preference domains avoid

\(^8\)See for example, Crawford and Knoer [1981], Kelso and Crawford [1982], Roth and Sotomayor [1992]
those manipulations. This paper continues this line of research and shows that hospital strategy-proofness is possible when an opportunity cost condition holds.

A second line of research has studied incentives in large markets. In particular, Roth and Peranson [1999], Immorlica and Mahdian [2005] and Kojima and Pathak [2009] show that strategy-proofness for almost all agents is possible as the number of agents increase, but the diversity of preferences decreases. Roth and Peranson [1999] showed that this property holds in some physician markets. In small markets, however, manipulability is still possible with current mechanisms. This paper continues the large market spirit by showing that as the number of possible transfers increases, strategy-proofness for all hospitals becomes possible. This fact could have great practical implications since the discreteness of transfers traded in the markets is a market design variable.

This paper is also related to the literature on VCG auctions. It is well known that the VCG payments are lower than the lowest anonymous linear Walrasian equilibrium payments. However, several deviations from anonymous linear prices can achieve VCG payments. Ausubel [2006] uses personalized linear prices. Bikhchandani and Ostroy [2002], de Vries et al. [2007], Mishra and Parkes [2007] use personalized non-linear prices. The opportunity cost condition implies that VCG payoffs can be implemented with anonymous linear prices.

3 Matching Markets

Throughout the paper, capital letters will represent both sets and their cardinality. Similarly, throughout the paper, when an assumption is stated, it is considered as true in all subsequent parts of the paper, including theorems. We denote the union of two sets $Y$ and $X$ by $YX$. The matching market is formed by two kinds of agents, doctors and hospitals. The set of doctors is denoted by $D$ and the set of hospitals by $H$ with typical elements $d$ and $h$, respectively. We denote the set of all agents by $A = HD$. We assume there is a finite number of agents in the market. It will be assumed that hospitals can be matched to several doctors, but doctors can be matched to at most one hospital.

Each doctor $d$ has preferences $u_d : H\{\emptyset\} \rightarrow \mathbb{R}$. Each hospital $h$ has a capacity $c_h$ and preferences $u_h : C_h \rightarrow \mathbb{R}$, where $C_h$ is the set of subsets of $D$ with cardinality at most $c_h$. We assume that there is a common utility metric, that affects agents payoffs linearly, i.e. if agent $a$ is matched to set $T$
and receives a transfer of $t_a$, then his payoff is $\pi_a = u_a(T) + t_a$. The level at which transfers can be made in a particular market varies considerably across applications. In school choice, societies have decided that wealth should not determine who gets the better schools; hence modeled as a no-transfers market. In cadet branching, cadets have the opportunity to serve longer times in order to obtain a branch they like more; hence some discrete transfers are used. Labor markets with flexible transfers would be modeled as continuous transfers matching markets. In order to simplify the distinction between models with different degrees of transfers we introduce the definition of a $q$-market.

**Definition 1.** A $q$-**market** is a market where all transfers made are multiples of $\frac{1}{q} \in \mathbb{R}_+$. The $0$-market represents the no-transfers case and the $\infty$-market represents the continuous case.

A $q$-matching $Mt$ in a $q$-market is a disjoint collection of sets of doctors $M = \{D_i\}_{i \in H+1}$ such that $|D_h| \leq c_h$ and a vector of transfers (multiples of $\frac{1}{q}$) $t \in \mathbb{R}^D$ such that $t_d = 0$ for every $d \in D_{H+1}$. A hospital $h$ is matched to the set of doctors $D_h$ and receives a transfer of $-\sum_{i \in D_h} t_i$. $D_{H+1}$ are unassigned doctors. The set of matchings in a market with $A$ agents is denoted by $M(A)$ and the set of matchings where all transfers are zero is denoted by $M_0(A)$. Given a particular matching $Mt$, the associated payoff of agent $a$ is denoted by $\pi_{Mt}^a$ and the associated private value is $u_{Mt}^a$.

Given a particular matching, every coalition of hospitals and doctors can arrange offers to improve their payoffs from any initial situation; they can try to improve their monetary component with their current partners, they can try to form new partnerships, or reject current ones. If such improving coalitions cannot be formed, then the matching is called stable. Hence stability is the basic notion of equilibrium in matching markets. In addition to its theoretical appeal, stability has been proved fundamental for the correct performance of real matching markets since unstable allocations typically lead to an unraveling of the market.

**Definition 2.** The $q$-matching $Mt$ is **stable** in $q$-market if and only if

- $\pi_{Mt}^a \geq u_{a}(\emptyset)$ for all $a \in A$

- There is no $q$-matching $(Mt)^*$ and $h \in H$ such that $\pi_{a}^{(Mt)^*} \geq \pi_{Mt}^a$ for all $a \in hD_h^*$ with at least one strict inequality.
Notice that the second condition applies only to members of the “blocking coalition” \( hD_h^* \) and hence to find a “blocking matching” \((Mt)^*\) it would be sufficient to assign members of \( hD_h^* \) together and leave everyone else unmatched.

It is well-known that if hospitals’ preferences satisfy the substitutes condition, then the set of stable matchings is not empty.\(^9\) In discrete markets, if agents’ preferences between matchings are strict, then there is a hospital-optimal matching \( Mt_H \) and a doctor-optimal matching \( Mt_D \).\(^{10}\) The associated payoffs for agent \( a \) are denoted by \( \pi_{a}^{Mt_H} \) and \( \pi_{a}^{Mt_D} \), respectively. In the continuous case, a hospital-optimal matching \( Mt_H \) and a doctor-optimal matching \( Mt_D \) always exit and the associated payoffs for agent \( a \) are also denoted by \( \pi_{a}^{Mt_H} \) and \( \pi_{a}^{Mt_D} \). In the continuous case, the optimal matchings might fail to be unique, but their associated payoffs will be. We assume hospitals’ preferences satisfy the substitutes condition. We first define the demand correspondence and value function for hospitals.

**Definition 3.** For every hospital \( h \in H \), \( T_h(t) = \arg\max S \in 2^D u_h(S) - \sum_{i \in S} t_i \) is its demand correspondence and \( v_h(t) = \max_{S \in 2^D} u_h(S) - \sum_{i \in S} t_i \) is its value function.

**Assumption 1.** \( u_h \) satisfies the substitutes condition if and only if for every \( t, t^* \in \mathbb{R}^D \) such that \( t \leq t^* \), for every \( T \in T_h(t) \), there is \( T^* \in T_h(t^*) \) such that \( T \cap S \subset T^* \); where \( S \) is the set of doctors with \( t_s = t^*_s \).

In addition to the substitutes condition we assume that every hospital needs to hire at least one doctor to produce any surplus and every doctor - hospital pair produce more surplus together than the unmatched doctor.

**Assumption 2.** Marginal product. For all \( h \in H \), \( u_h(\emptyset) = 0 \). For all \( h \in H \) and \( d \in D \) and \( D_h \subset D \) such that \( d \notin D_h \) and \( dD_h \in C_h \) we have \( u_h(dD_h) + u_d(h) - u_h(D_h) \geq u_d(\emptyset) \).

The above condition allows for cases with \( u_h(dD_h) < u_h(D_h) \) with \( d \notin D_h \) i.e. a hospital would need to be compensated to hire a doctor. In the context of resident matching, this assumption could feel unnatural, however, in other applications such a school choice, it is the norm. Schools and colleges usually charge students to get admitted.

\(^9\)See Kelso and Crawford [1982]
\(^{10}\)See Hatfield and Milgrom [2005]
4 Strategy-Proofness for Hospitals in the continuous market

We begin this section by providing some standard preliminary definitions. A mechanism $\phi$ is a function that maps preference profiles to $q$-matchings in a $q$-market. The matching at preference profile $\{u_i\}_{i \in A}$ is denoted by $\phi(\{u_i\}_{i \in A}) \in M(A)$. A mechanism $\phi$ is said to be strategy-proof for agent $a$ if there exist no preference profile $u'_a$ and preferences $\{u_i\}_{i \in A \setminus a}$ for all other agents such that $\pi_a(\phi(u'_a, \{u_i\}_{i \in A \setminus a})) > \pi_a(\phi(\{u_i\}_{i \in A}))$. A mechanism $\phi$ is said to be strategy-proof for $B$ (set) if it is strategy-proof for all $b \in B$. That is, no agent in $B$ has an incentive to misreport his preferences under the mechanism. The next proposition, due to Roth [1985], establishes that there is no strategy-proof for Hospitals mechanism in the 0-market. Throughout the paper, “propositions” will be used to state important known results.

Proposition 1. Roth 85. There is no stable and strategy-proof mechanism for Hospitals in the 0-market.

Example 1. Consider a market with three hospitals and four doctors. Hospitals and Doctors have cardinal valuations over each other. The left matrix contains the values each hospital assigns to every doctor and the number of doctors they are willing to hire. The right matrix contains the values for doctors. It is assumed that both groups assign a value of zero to being unassigned.

<table>
<thead>
<tr>
<th>Preferences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hospitals</strong></td>
</tr>
<tr>
<td>$d_1$</td>
</tr>
<tr>
<td>$h_1$</td>
</tr>
<tr>
<td>$h_2$</td>
</tr>
<tr>
<td>$h_3$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Doctors</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_1$</td>
</tr>
<tr>
<td>$h_1$</td>
</tr>
<tr>
<td>$h_2$</td>
</tr>
<tr>
<td>$h_3$</td>
</tr>
</tbody>
</table>

Suppose transfers are prohibited and a stable allocation is to be implemented. In this case, the ordinal representation is sufficient to characterize the set of stable matchings. Under the true preference profile, there is only one stable matching, shown on the left. On the right there is an improving deviation for $h_1$. 
Roth’s proposition applies to 0-markets. In $\infty$-markets, however, there is an efficient and strategy-proof mechanism: VCG. The VCG mechanism is defined below. First we define the coalitional value function.

**Definition 4.** $V(S) = \max_{Mt \in M_0(S)} \sum_{a \in S} \pi_a(Mt)$ is the coalitional value function.

With the coalitional value function at hand, we define the VCG mechanism for a set of agents $B$.

**Definition 5.** The **VCG for agents in $B$** works as follows. It is assumed that the mechanism knows $u_a$ for every $a \in A \setminus B$. Agent $b \in B$ sends $u_b$ to the mechanism and an outcome $(Mt)^* \in \arg\max_{Mt \in M_0(A) a \in A} \sum \pi_a(Mt)$ is implemented. Agent $b$ is charged a payment $p_b = -\sum_{a \in A \setminus b} \pi_a(Mt)^* + W(A \setminus b)$. Where $W$ is a coalitional value function with the reported and known preferences. Let $P = \sum_{b \in B} p_b$ be the total collected payments, let $\{t_a\}_{a \in A \setminus B}$ be a set of real numbers such that $P = \sum_{a \in A \setminus B} t_a$. Then the payoff for agent $a \in A \setminus B$ is given by $\pi_a(Mt)^* + t_a$. Whenever there is a set of transfers $\{t_a\}_{a \in A \setminus B}$ such that the final payoffs are stable, those transfers are implemented.

**Lemma 1.** The VCG mechanism for agents in $B$ is strategy-proof for agents in $B$.

If agents in $B$ play their dominant strategy, then every agent $b \in B$ receives a payoff equal to $\pi_b = V(A) - V(A \setminus b)$ (agent $b$’s VCG payoff). We assume agents always play their dominant strategy.
Notice that VCG payoffs for members of $B$ only depend on the value of the coalitional value function and not on a particular efficient allocation i.e. VCG payoffs and payments are well-defined even when there are multiple optimal allocations. Furthermore, notice that the only condition on payments (and payoffs) for members of $A \setminus B$ is budget balancedness. Whenever there is a set of transfers transfers $\{t_a\}_{a \in A \setminus B}$ such that the final payoffs are stable, we say that the VCG for set $B$ is stable. Example 2 shows that stability for hospitals is possible in some cases while Example 3 shows that it is not always possible.

**Example 2.** Suppose we have the market of example 1, but transfers can be continuously adjusted. Then, VCG for hospitals is strategy-proof and stable.

<table>
<thead>
<tr>
<th>Hospitals</th>
<th>$d_1$</th>
<th>$d_2$</th>
<th>$d_3$</th>
<th>$d_4$</th>
<th>$q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$h_2$</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$h_3$</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Doctors</td>
<td>$d_1$</td>
<td>$d_2$</td>
<td>$d_3$</td>
<td>$d_4$</td>
<td></td>
</tr>
<tr>
<td>$h_1$</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>$h_2$</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$h_3$</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Continuous transfers

- $\pi_{h_1} = 22 - 12 = 10$
- $\pi_{h_2} = 22 - 17 = 5$
- $\pi_{h_3} = 22 - 17 = 5$
- $\pi_{d_1} = 0$
- $\pi_{d_2} = 1$
- $\pi_{d_3} = 1$
- $\pi_{d_4} = 0$

In the previous example, doctors payoffs are obtained after solving a system of equations together with some inequalities. For instance, $h_2$ is matched to $d_2$, together they generate a surplus of 6 and $h_2$’s VCG payoff is 5, then $h_2$’s payoff must be 1. The question is if, in general, the payments collected by the mechanism can be distributed to doctors in a stable way.

In auction theory, Ausubel and Milgrom [2002] have shown that the substitutes condition is sufficient for the VCG payoffs to be stable, i.e. all collected payments by the VCG mechanism can be paid to the auctioneer and the outcome is stable. Unfortunately, in matching markets, the substitutes condition is not sufficient to obtain the same result. Consider the following example.

**Example 3.** There are two identical hospitals and three identical doctors. $u_h(A) = 0$ if $|A| = 0, u_h(A) = 10$ if $|A| = 1, u_h(A) = 18$ if $|A| = 2, u_h(A) = 20$ if $|A| = 3$. $u_d(\emptyset) = u_d(h) = 0$. In this market, both hospitals receive a payoff of $\pi_h = V(A) - V(A/h) = 28 - 20 = 8$ when the VCG mechanism is used. However, in the unique stable allocation both hospitals receive a payoff $\pi_h = 2$ and all doctors receive
a payoff $\pi_d = 8$.

If we want VCG payoffs for Hospitals to agree with their maximum stable payoff, $\pi^{Mt}_h = V(A) - V(A \setminus h)$ for every hospital, we need to be able to construct a payoff equivalent (for all agents different than $h$) allocation in the market with agents $A/h$. In particular, if doctor $d$ is matched to hospital $h$ in the market with $A$ agents, then we need to find an allocation with $A \setminus h$ agents that provides $d$ with $\pi^{Mt}_d$. The natural candidate is one of his blocking coalitions, i.e. one of the coalitions that would be formed if he were offered anything less than $\pi^{Mt}_d$. Unfortunately, it is possible that a hospital $h'$ belongs to a blocking coalition with some doctor $d$ and a different (incompatible) coalition with $d'$. This is illustrated in the following example.

**Example 4.** Consider a market with two hospitals and three doctors. The matrix contains the joint surpluses. It is assumed that both groups assign a value of zero to being unassigned. $t$ and $s$ are a real numbers.

<table>
<thead>
<tr>
<th></th>
<th>$d_1$</th>
<th>$d_2$</th>
<th>$d_3$</th>
<th>$q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>$h_2$</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

**Hospital-optimal stable payoffs**

- $\pi_{h_1} = 4$
- $\pi_{d_1} = 1$
- $\pi_{h_2} = 1$
- $\pi_{d_2} = 2$
- $\pi_{d_3} = 0$

**VCG for Hospitals payoffs**

- $\pi_{h_1} = 5$
- $\pi_{d_1} = 2 + s + t$
- $\pi_{h_2} = 1$
- $\pi_{d_2} = -t$
- $\pi_{d_3} = -s$

- If $\pi_{d_1} < 1$, then $d_1$ and $h_2$ form a blocking coalition.
- If $\pi_{d_2} < 2$, then $d_2$ and $h_2$ form a blocking coalition.
- However, in the market without $h_1$, $h_2$ cannot honor his blocking offers with $d_1$ and $d_2$ simultaneously.

The previous example shows that the substitutes condition is not enough to guarantee the equivalence between the maximum stable payoff and the VCG payoff for hospitals. There are, however, some regularities that the substitutes condition can provide. Proposition 2 states well-known results that are used in our discussion and proof of Theorem 1.

**Proposition 2. In all matching markets,**

- $\pi^{Mt}_d = V(A) - V(A \setminus d)$ for every $d \in D$. Leonard [1983]
\[ \pi_d^{MtH} = V(dA) - V(A) \] for every \( d \in D \). Gul and Stacchetti [1999]. Where \( V(Ad) \) is the value of a market where an identical doctor \( d \) is added.

Proposition 2 characterizes doctors’ payoffs at the hospital-optimal and doctor-optimal matchings. Furthermore, it implies the strategy-proofness and stability of VCG for doctors; as it shows that the VCG payoff for doctors is equal to their doctor-optimal stable payoff. In this section we provide an analogous result for hospitals. Example 4 shows that this cannot be achieved in general, however, we show that if preferences satisfy a joint restriction we call the opportunity cost condition, then the equivalence can be guaranteed.

**Assumption 3.** We say that \( V \), the coalitional value, function satisfies the opportunity cost condition if and only if for all \( d \in D \) and \( h \in H \),

\[ V(A \setminus h) - V(A \setminus hd) \geq V(Ad) - V(A) \]

On the right hand side of the inequality we have \( V(Ad) - V(A) \), this is the opportunity cost of doctor \( d \) in the optimal assignment. Intuitively, if a new copy of doctor \( d \) were added to the market, the copy would go to the second highest value allocation, as the highest value is occupied by the original \( d \).

On the left hand side of the inequality we have \( V(A \setminus h) - V(A \setminus hd) \), this is the marginal value of \( d \) in a market where \( h \) is not present. When the opportunity cost holds, every doctor can find a hospital that offers him at least his opportunity cost.

**Theorem 1.** In the continuous transfers market, \( \pi_h^{MtH} = V(A) - V(A \setminus h) \) for every \( h \in H \).

**Corollary 1.** The VCG for Hospitals is stable and strategy-proof.

Theorem 1 is completely analogous to that of Leonard [1983]. When preferences satisfy the opportunity cost condition it is possible to use VCG for hospitals as a strategy-proof and stable mechanism. The opportunity cost condition is a joint condition on preferences, however, there are some individual preferences that imply it. For instance, linear preferences and unit demand preferences.

**Definition 6.** Hospital preferences are linear if and only if for all \( h \in H \), \( c_h = |H| \) and for all \( D_h \subset D \),

\[ u_h(D_h) = \sum_{d \in D_h} u_h(d) \]

Hospitals preferences are of unit demand if for all \( h \in H \), \( c_h = 1 \).

**Lemma 2.** Unit demand and linear hospital preferences satisfy the opportunity cost condition.
5 Strategy-Proofness for Hospitals And Doctors in the continuous market

We begin this section with another impossibility result due to Roth [1982].

**Proposition 3. Roth 82.** There is no stable and strategy-proof mechanism for hospitals and doctors in the 0-market, even when hospitals have unit demands.

**Example 5.** Consider a market with two hospitals and two doctors. Hospitals and Doctors have cardinal valuations over each other. The left matrix contains the values each hospital assigns to every doctor. All hospital are willing to hire at most one doctor. The right matrix contains the values for doctors. It is assumed that both groups assign a value of zero to being unassigned.

<table>
<thead>
<tr>
<th>Hospitals</th>
<th>Doctors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$d_1$</td>
</tr>
<tr>
<td>$h_1$</td>
<td>2.5</td>
</tr>
<tr>
<td>$h_2$</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$d_1$</th>
<th>$d_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>1.5</td>
<td>4</td>
</tr>
<tr>
<td>$h_2$</td>
<td>2</td>
<td>3.5</td>
</tr>
</tbody>
</table>

Suppose transfers are prohibited and a stable allocation is to be implemented. In this case, the ordinal representation is sufficient to characterize the set of stable matchings. Under the true preference profile, there are two stable matchings, shown on right. Suppose the bottom matching is chosen by the mechanism.

Matching 1

$h_1 : d_1 \quad d_2$  
$d_1 : h_2 \quad h_1$  
$h_2 : d_2 \quad d_1$  
$d_2 : h_1 \quad h_2$

Matching 2

$h_1 : h_2$  
$d_1 : d_2$  
$h_2 : h_1$  
$d_2 : h_2$

Then $h_1$ can manipulate the outcome by manipulating his preferences.
The mechanism provided by Theorem 1 cannot be used directly in this case to obtain strategy-proofness for all agents. In general, \( \sum_{a \in A} V(A) - V(A \setminus a) > V(A) \). Hence, we need to reduce agents payoffs without losing strategy-proofness. Intuitively, agents needs to be able to capture the marginal value generated by their reports. In auction environments, this is precisely achieved by offering every agent \( a \) a payoff equal to \( V(A) - V(A \setminus a) \). In matching environments, this is not the case.

When a new agent enters a matching market, say a doctor \( d \), two effects take place: (1) a new agent who demands hospitals enters the market and (2) a new object is available for hospitals currently in the market. \( V(A) - V(A \setminus d) \) captures both marginal effects. In order to obtain strategy-proofness we only need the first one. Doctor \( d \)'s information's marginal value is captured by \( U(S) - U(S \setminus d) \), where \( U(T) \) is the maximum value that can be achieved with all agents present only using the information of agents in \( T \). The \( U \) function is defined below.

**Definition 7.** \( U(S) = \max_{Mt \in M_0(A)} \sum_{a \in S} \pi^{Mt}_a \) is the optimal matching function

It is possible to define a second VCG mechanism using the \( U \) function.

**Definition 8.** The \( U \)-VCG mechanism for agents in \( B \) works as follows. It is assumed that the mechanism knows \( u_a \) for every \( a \in A \setminus B \). Agent \( b \in B \) sends \( u_b \) to the mechanism and an outcome \( (Mt)^* \in \arg\max_{Mt \in M_0(A)} \sum_{a \in A} \pi^{Mt}_a \) is implemented. Agent \( b \) is charged a payment \( p_b = - \sum_{a \in A \setminus B} \pi^{(Mt)^*}_a + W'(A \setminus b) \).

Where \( W' \) is the optimal matching function with the reported and known preferences. Let \( P = \sum_{b \in B} p_b \) be the total collected payments, let \( \{t_a\}_{a \in A \setminus B} \) be a set of real numbers such that \( P = \sum_{a \in A \setminus B} t_a \). Then the payoff for agent \( a \in A \setminus B \) is given by \( \pi^{(Mt)^*}_a + t_a \). Whenever there is a set of transfers \( \{t_a\}_{a \in A \setminus B} \) such that the final payoffs are stable, those transfers are implemented.

**Lemma 3.** The \( U \)-VCG mechanism for agents in \( B \) is strategy-proof.

If agents in \( B \) play their dominant strategy, then every agent \( b \in B \) receives a payoff equal to \( \pi_b = U(A) - U(A \setminus b) \). We assume agents always play their dominant strategy. The difference between the
VCG and the $U$-VCG mechanisms is the last term in their payoffs: $U(A/b) = \max_{M \in M_0(A)} \sum_{a \in A/b} \pi^M_a$ while $V(A/b) = \max_{M \in M_0(A/b)} \sum_{a \in A/b} \pi^M_a$. If we apply the $U$-VCG to the previous example we achieve stability and strategy-proofness for all agents in the market.

Example 6. Consider the market of example 5 and use the $U$-VCG mechanism.

**Hospitals**

<table>
<thead>
<tr>
<th></th>
<th>$d_1$</th>
<th>$d_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>2.5</td>
<td>1.0</td>
</tr>
<tr>
<td>$h_2$</td>
<td>3.0</td>
<td>4.5</td>
</tr>
</tbody>
</table>

**Doctors**

<table>
<thead>
<tr>
<th></th>
<th>$d_1$</th>
<th>$d_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>1.5</td>
<td>4.0</td>
</tr>
<tr>
<td>$h_2$</td>
<td>4.5</td>
<td>3.5</td>
</tr>
</tbody>
</table>

The payoff for $h_1$ is calculated as follows:

$\pi^U_{h_1} = U(A) - U(A \setminus h_1) = 2.5$

$U(A) = \max_{M \in M_0(A)} \sum_{a \in S} \pi^M_a = V(h_1d_1) + V(h_2d_2) = 4 + 8 = 12$

$U(A \setminus h_1) = \max_{M \in M_0(A \setminus h_1)} \sum_{a \in A \setminus h_1} \pi^M_a = u_{d_1}(h_1) + V(h_2d_2) = 1.5 + 8 = 9.5$

Analogously,

$\pi^U_{d_1} = 12 - 10.5 = 1.5; \pi^U_{h_2} = 12 - 7.5 = 4.5; \pi^U_{d_2} = 12 - 8.5 = 3.5$

It is routine to check that the payoffs in example 6 are stable. It is instructive to compare the $U$-VCG payoffs with the hospital-optimal and doctor-optimal stable matchings. In particular, $\pi^{MT_H}_{h_1} = 4, \pi^{MT_H}_{h_2} = 7, \pi^{MT_H}_{d_1} = 0$ and $\pi^{MT_H}_{d_2} = 1$ for the hospital-optimal stable matching and $\pi^{MT_D}_{h_1} = 0, \pi^{MT_D}_{h_1} = 1, \pi^{MT_D}_{d_1} = 4$ and $\pi^{MT_D}_{d_2} = 7$ for the doctor-optimal stable matching. The $U$-VCG mechanism is achieving strategy-proofness and stability without offering either doctors or hospitals their most preferred stable matching. In the following section we study this property of the $U$-VCG mechanism. In the previous example, the $U$-VCG delivers stable payoffs for all agents. In general, this is not the case. Consider the following example.

Example 7. The $U$-VCG mechanism is not stable for this market, as it removes resources from the market.

**Hospitals**

<table>
<thead>
<tr>
<th></th>
<th>$d_1$</th>
<th>$d_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>4.0</td>
<td>5.0</td>
</tr>
<tr>
<td>$h_2$</td>
<td>4.0</td>
<td>4.0</td>
</tr>
</tbody>
</table>

**Doctors**

<table>
<thead>
<tr>
<th></th>
<th>$d_1$</th>
<th>$d_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>$h_2$</td>
<td>1.0</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Both markets in example 6 and 7 share the same coalitional function, hence they have the same set of stable payoffs. As it can be observed in the examples, the division of surplus between agents plays a fundamental role in the stability of the $U$-VCG mechanism. The next definition formalizes the idea of a division of surplus.
**Definition 9.** Let \( \{u_a\}_A \) be a matching market with hospitals \( H \), doctors \( D \) and \( A = HD \). Consider any other market with the same set of agents \( A \), but different preferences \( \{u'_a\}_A \). We say that \( \{u_a\}_A \) is a **division** of \( \{u'_a\}_A \) if and only if for all \( D_h \subset D \) and \( h \) we have 
\[
 u_h(D_h) + \sum_{d \in D_h} u_d(h) = u'_h(D_h) + \sum_{d \in D_h} u'_d(h)
\]
and for all \( d \in D \) 
\[
 u_d(\emptyset) = u'_d(\emptyset).
\]

Notice that if \( \{u_a\}_A \) is a division of \( \{u'_a\}_A \), then the opposite is also true i.e. divisions form an equivalence relation in the set of preferences profiles. Furthermore, as lemma 4 establishes, divisions have no impact on the set of stable matchings, which only depend on the coalitional value function.

**Lemma 4.** Let \( \Pi \subset \mathbb{R}^A \) be the set of payoffs arising from stable allocations in market \( \{u_a\}_A \). Let \( \{u'_a\}_A \) be a division of \( \{u_a\}_A \) and let \( \Pi' \subset \mathbb{R}^A \) be its set of stable payoffs. Then \( \Pi = \Pi' \).

Whereas the VCG mechanism, and its stability for hospitals, only depends on the coalitional value function, the stability of the \( U \)-VCG mechanism depends on the particular division of surplus in the market. Consider examples 6 and 7. They share the same coalitional value function and set of stable matchings and payoffs. However, when the \( U \)-VCG mechanism is used to elicit preferences, the resulting payoffs are stable only for the division in example 6. The following theorem shows that the payoffs delivered by the \( U \)-VCG mechanism are stable for at least one representative of each equivalence class in the space of preference profiles.

**Theorem 2.** Let \( \pi \in \Pi \), then there is a division such that \( \pi^U_a = \pi_a \), i.e. there is a division of surplus such that the \( U \)-VCG mechanism is stable and strategy-proof for all agents.

Theorem 2 does not say that the \( U \)-VCG mechanism delivers stable payoffs in every market. However, for every market there is a division of surplus that would make the payoffs delivered by the \( U \)-VCG mechanism stable. In other words, for every fixed set of agents \( A \) and coalitional function \( V \) there is market characterized by utility functions on \( \{u_a\}_A \) such that the \( U \)-VCG mechanism would deliver stable payoffs in that market.

Theorem 3 describes the conditions under which utility functions \( \{u_a\}_A \) produce a market for which the \( U \)-VCG mechanism is produces stable payoffs. We first introduce the concept of pivotal information. We say that agent \( i \)'s information is not pivotal if when his information is disregarded, but he is still considered part of the matching market, the optimal allocation does not change.
Definition 10. Let \((Mt)^* \in \arg\max_{Mt \in M_0(A)} \sum_{a \in A} \pi_{Mt}^a\). Agents \(i\)'s information is **not pivotal**, with respect to \((Mt)^*\), whenever \((Mt)^* \in \arg\max_{Mt \in M_0(A) \setminus \{i\}} \sum_{a \in A} \pi_{Mt}^a = \emptyset\).

Theorem 3. Let \((Mt)^* \in \arg\max_{Mt \in M_0(A)} \sum_{a \in A} \pi_{Mt}^a\) and let \(U_a^{(Mt)^*}\) be agent \(a\)'s private value in \((Mt)^*\). Suppose no agent’s information is pivotal with respect to \((Mt)^*\), and \((U_a^{(Mt)^*})_A \in \Pi\), then the \(U\)-VCG mechanism is strategy-proof for all agents and stable.

If any agent’s information is pivotal, then \(U\)-VCG will collect a positive payment from this agent. If the \(U\)-VCG is used to elicit preferences from only one side of the market, then there is the possibility (studied in the next section) of redistributing the payments to the other side to maintain all the surplus in the market. When eliciting preferences from both sides of the market, this possibility disappears. Furthermore, any agent whose information is not pivotal will have a payoff equal to his private surplus at the chosen allocation. Unfortunately, both conditions are independent. The next example shows that there are non-pivotal markets where \((U_a)_A \notin \Pi\) and pivotal markets where \((U_a)_A \in \Pi\).

Example 8. On the left, there is a non-pivotal market where \((U_a)_A \notin \Pi\). On the right, there is a pivotal market where \((U_a)_A \in \Pi\).

<table>
<thead>
<tr>
<th>Hospitals</th>
<th>Doctors</th>
<th>Hospitals</th>
<th>Doctors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(h_1)</td>
<td>3 4</td>
<td>(h_1)</td>
<td>2 4</td>
</tr>
<tr>
<td>(h_2)</td>
<td>0 1</td>
<td>(h_2)</td>
<td>0 4</td>
</tr>
</tbody>
</table>

The \(U\)-VCG payments can be modified in a market where no agent’s information is pivotal but \((U_a)_A \notin \Pi\). Specifically, let \(Mt\) be any stable matching and let \(t_d\) be the minimum salary allowed for doctor \(d\), i.e. regardless of the hiring hospital, \(d\) will charge at least \(t_d\). Of course, this will impact the true preferences in the market, as now, doctor \(d\) will have preferences \(u'_d(h) = u_d(h) + t_d\) and hospital \(h\) will have preferences \(u'_h(D_h) = u_h(D_h) - \sum_{d \in D_h} t_d\). With these new preferences, the market is non-pivotal and \((U'_a)_A \in \Pi\).

6 Non-extremal Strategy-Proofness.

In this section we study the one-sided \(U\)-VCG. As a first motivation, we consider example 6 from the previous section. It can be observed that the \(U\)-VCG mechanism achieves strategy-proofness for all
agents and stability, but does not depend on offering any side their most preferred stable allocation. Example 9 shows that strategy-proofness and stability can be achieved in the interior of the set of stable payoffs.

**Example 9.** Strategy-proofness can be obtained in the interior of the set of stable payoffs.

<table>
<thead>
<tr>
<th>Hospitals</th>
<th>Doctors</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_1$</td>
<td>$d_1$</td>
</tr>
<tr>
<td>2.5</td>
<td>1.5</td>
</tr>
<tr>
<td>$d_2$</td>
<td>$d_2$</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

The $U$-VCG uses the following values:

- $U(A) = V(h_1d_1) + V(h_2d_2) = 4 + 8 = 12$
- $U(A \setminus h_1) = u_{d_1}(h_1) + V(h_2d_2) = 1.5 + 8 = 9.5$
- $U(A \setminus h_2) = V(h_1d_1) + u_{d_2}(h_2) = 4 + 3.5 = 7.5$
- $U(A \setminus d_1) = u_{h_1}(d_1) + V(h_2d_2) = 2.5 + 8 = 10.5$
- $U(A \setminus d_2) = V(h_1d_1) + u_{h_2}(d_2) = 4 + 4 = 8$

Hence,

- $\pi_{h_1}^U = 12 - 9.5 = 2.5$; $\pi_{h_2}^U = 12 - 7.5 = 4.5$
- $\pi_{d_1}^U = 12 - 10.5 = 1.5$; $\pi_{d_2}^U = 12 - 8.5 = 3.5$

We know that the $U$-VCG mechanism is not stable in the market of example 7. However, the following example shows that the $U$-VCG for doctors is stable and strategy-proof for doctors.

**Example 10.** One-sided strategy-proofness can be obtained without offering that side their most preferred stable payoff.

<table>
<thead>
<tr>
<th>Hospitals</th>
<th>Doctors</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_1$</td>
<td>$d_1$</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>$d_2$</td>
<td>$d_2$</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

The $U$-VCG uses the following values:

- $U(A) = V(h_1d_1) + V(h_2d_2) = 4 + 8 = 12$
- $U(A \setminus h_1) = u_{d_1}(h_1) + V(h_2d_2) = 0 + 8 = 8$
- $U(A \setminus h_2) = V(h_1d_1) + u_{d_2}(h_2) = 4 + 4 = 8$
- $U(A \setminus d_1) = u_{h_1}(d_1) + V(h_2d_2) = 4 + 8 = 12$
- $U(A \setminus d_2) = u_{h_1}(d_2) + V(h_2d_1) = 5 + 5 = 10$
Hence the \( U \)-VCG for doctors delivers:

\[
\begin{align*}
\pi_{h_1} &= V(h_1d_1) - 0 = 4; \\
\pi_{h_2} &= V(h_2d_2) - 2 = 6; \\
\pi_{d_1} &= 12 - 12 = 0; \\
\pi_{d_2} &= 12 - 10 = 2
\end{align*}
\]

Hence the \( U \)-VCG for hospitals delivers:

\[
\begin{align*}
\pi_{h_1}^U &= 12 - 8 = 4; \\
\pi_{h_2}^U &= 12 - 8 = 4; \\
\pi_{d_1}^U &= V(h_1d_1) - 4 = 0; \\
\pi_{d_2}^U &= V(h_2d_2) - 4 = 4
\end{align*}
\]

As noted above, the performance of the \( U \)-VCG mechanism depends on the division of surplus between both sides of the market. In this section, a special class of division for which the \( U \)-VCG for hospitals is stable and strategy-proof are studied. It is assumed that preferences are linear. When preferences are linear, surplus can be moved freely from one side to the other. In particular, divisions of the form \( u_h(D_h) = \alpha(u_h(D_h) + \sum_{d \in D_h} u_d(h)) \) for all \( \alpha \in [0, 1] \) are well defined.

**Definition 11.** Let \( \alpha \in [0, 1] \). We say that the hospital and doctors preferences forms an \( \alpha \) division if \( u_h(D_h) = \alpha(u_h(D_h) + \sum_{d \in D_h} u_d(h)) \) for all \( h \) and \( D_h \in D \).

For a fixed alpha, let \( \pi_h^U(\alpha) \) the payoff assigned to hospital \( h \) by the hospital \( U \)-VCG mechanism. The following lemma describes a few properties of \( \pi_h^U \).

**Lemma 5.** Let \( \pi_h^U(\alpha) \) the payoff assigned to hospital \( h \) by the hospital \( U \)-VCG mechanism, then \( \pi_h^U \) is a monotone increasing piecewise linear function satisfying:

\[
\begin{align*}
&\bullet \quad \pi_h^U(0) = 0 \text{ and } \pi_h^U(1) = V(A) - V(A \setminus h) \\
&\bullet \quad \frac{d}{d\alpha} \pi_h^U(\alpha) \geq \alpha(V(A) - V(A \setminus h)) \text{ for all } \alpha \in [0, 1] \\
&\bullet \quad \frac{d}{d\alpha} \pi_h^U(0) = V(hD_h) \text{ and } \frac{d}{d\alpha} \pi_h^U(0) = 0, \text{ where } D_h \text{ is the set of doctors assigned to } h \text{ at the efficient matching.}
\end{align*}
\]

The following theorem establishes several conditions for the one-sided \( U \)-VCG mechanism to deliver stable payoffs in markets characterized by \( \alpha \) divisions.

**Theorem 4.** Suppose preferences are linear and \( \Pi \) has a non-empty interior. Let \( h_d \) be the hospital matched with \( d \) at a fixed efficient matching.

\[
\begin{align*}
&\bullet \quad \text{If the lowest stable payoff for every hospital is zero and } V(h_dd) > V(h'd) \text{ for all } d \text{ and } h' \neq h_d, \\
&\quad \text{then there is } \beta \in (0, 1) \text{ such that for any } \alpha \in [0, \beta] \pi_h^U(\alpha) \text{ for all } h \in H \text{ is stable.}
\end{align*}
\]
• For every doctor there are at least two hospitals \( h, h' \) such that \( V(hd) > 0 \) and \( V(h'd) > 0 \), then there is \( \beta \in (0, 1) \) such that for any \( \alpha \in [\beta, 1] \) \( \pi^U_h(\alpha) \) for all \( h \in H \) is stable.

Suppose that in a particular market all the surplus is generated by hospitals. In this case, hospitals’ preferences could be elicited using the VCG for hospitals mechanism and the hospital-optimal matching would be implemented in a strategy-proof manner. In this case every hospital \( h \) would receive \( \pi^U_h(1) = V(A) - V(A \setminus h) \) as a payoff and doctors would receive their minimal stable payoff. Now suppose that in that market, a small quantity of surplus is shifted from hospitals; for example by offering some wages to doctors. In this new market with wages, the set of stable payoffs has not changed as \( V(hd) \) remains unchanged by the wage offered by \( h \). In this market, using VCG for hospitals would still be possible, however, according to theorem 4 if the \( U \)-VCG for hospitals mechanism is used instead, then an interior solution would be achieved i.e. hospitals would receive less than their maximum stable payoffs and doctors would receive more than their minimum stable payoffs.

7 Hospital Strategy-proofness in the \( q \)-Market

Roth’s theorems show that, in the \( 0 \)-market, it is not possible to achieve strategy-proofness for hospitals or all agents simultaneously. In sharp contrast, in the \( \infty \)-market, strategy-proofness and stability for hospitals is possible. Which model describes better a particular real market depends on the institutional environment. In school choice, transfers are completely prohibited whereas, in residents matching, the monetary component of any transaction is a fundamental part. In markets where transactions are allowed, the level of discreteness of transfers is a market design variable. The following theorem shows that, if strategy-proofness for hospitals is important in a particular market, there is always a discrete level of transactions that achieves it, i.e. transfers only need to be as small as the smallest common factor between the set of possible valuations.

**Theorem 5.** Suppose preferences are integer valued, then strategy-proofness and stability are possible in any \( q \)-market for any \( q \in \mathbb{Z} \) such that \( q \geq 1 \).

When a stable allocation is to be implemented in a market where discrete quantities of money can be exchanged the generalized Gale-Shapley mechanism is the standard solution. For example, in
the formed proposed by Kelso and Crawford [1982]. One important limitation of this mechanism is that it is not strategy-proof for hospitals. Theorem 1 establishes that, when preferences satisfy the opportunity cost condition, and monetary transfers can be continuously adjusted, the VCG for hospitals is stable and strategy-proof. These two properties can be extended to markets with discrete quantities of money by using a lowest common denominator fraction i.e. the smallest tradable quantity of money is sufficiently small to offset the smallest change in valuations. One very simple way of expressing this idea is assuming that preferences are integer-valued. Thus in a market in which preferences are expressed in thousands of dollars, e.g. the value of a one year contract between a hospital and a doctor is in thousands of dollars, a stable allocation can be achieved using the VCG for hospitals mechanism instead of the Gale-Shapley mechanism.

8 Conclusion

In 2002, 16 law firms filed a class action law suit, representing 3 former residents seeking to represent all residents, arguing that the NRMP violated antitrust laws and was a conspiracy to depress resident's wages. The complaint was:

Defendants and others have illegally contracted, combined and conspired among themselves to displace competition in the recruitment, hiring, employment and compensation of resident physicians, and to impose a scheme of restraints which have the purpose and effect of fixing, artificially depressing, standardizing and stabilizing resident physician compensation and other terms of employment.

Defendants' illegal combination and conspiracy has restrained competition in the employment of resident physicians by:

(a) stabilizing wages below competitive levels by exchanging competitively sensitive information regarding resident physician compensation and other terms of employment;

(b) eliminating competition in the recruitment and employment of resident physicians by assigning prospective resident physician employees to positions through the National Resident Matching Program (“NRMP”); and
(c) establishing and complying with anticompetitive accreditation standards and require-
mements through the Accreditation Council for Graduate Medical Education (“ACGME”).

The suit was dismissed on August 12, 2004 in an Opinion, Order & Judgment by Judge Paul L. Friedman based on evidence regarding the structure of wages in other decentralized industries and expert opinions from several economists.

To this date no wage negotiation takes place in the NRMP, however, this paper shows that allowing transfers to be negotiated in the matching process not only would “enhance competition”, it would also make agents more willing to report their true private information. As argued in the introduction, this would increase the efficiency and accountability of the whole program.

9 Appendix

This appendix contains the proofs for the theorems stated in the paper. In order to relate the results of this paper to the literature on VCG auctions we first introduce establish a connection between matching markets and auction markets. The connection establishes that for any stable matching there is an equivalent Walrasian equilibrium, defined as follows.

**Definition 12.** Let \( \{u_a\}_A \) be a matching market with hospitals \( H \) and doctors \( D \). Suppose that for all \( d \in D \) we have \( u_d(h) = 0 \) for all \( h \in H \). A vector \( t \in \mathbb{R}^D \) is a **Walrasian Equilibrium Price** (WEP) if and only if there is \( D_h \in T(t) \) such that \( \bigcup_H D_h = D \) and for all \( h \neq h' \) \( D_h \cap D_{h'} = \emptyset \). \( \{D_h\}_H, t \) is a **Walrasian Equilibrium** (WE).

We first consider a couple of lemmas.

**Lemma. 4** Let \( \Pi \subset \mathbb{R}^A \) be the set of payoffs arising from stable allocations in market \( \{u_a\}_A \). Let \( \{u'_a\}_A \) be a division of \( \{u_a\}_A \) and let \( \Pi' \subset \mathbb{R}^A \) be its set of stable payoffs. Then, \( \Pi = \Pi' \).

**Proof.** By definition, \( \pi \in \Pi \) if and only if \( \sum_C \pi_a \geq V(C) \) for all \( C \subset A \) and \( \sum_A \pi_a = V(A) \). By definition of a division, \( V(C) = V'(C) \) for all \( C \subset A \), hence \( \Pi = \Pi' \). To see this, consider for example \( V(hD_h) = \max_{Mt \in M(h,D_h), a \in hD_h} \sum_{a} u_a(Mt_a) \) and let \( D^* \subset D_h \) such that \( V(hD_h) = V(hD^*) \) i.e. \( D^* \) is the optimal subset
of doctors among $D_h$. Then by definition of a division, for any $D \subset D_h$ we have $u_h(D) + \sum_{d \in D^*} u_d(h) = u'_h(D) + \sum_{d \in D^*} u'_d(h)$, in particular we have $u_h(D^*) + \sum_{d \in D^*} u_d(h) = u'_h(D) + \sum_{d \in D^*} u'_d(h) = V(hD^*) = V'(hD^*) = V(hD_h) = V(D_h)$. □

**Lemma 6.** Let $\Pi \subset \mathbb{R}^A$ be the set of payoffs arising from stable allocations in market $\{u_a\}_A$. Let $\{u'_a\}_A$ be a division of $\{u_a\}_A$ such that $u'_d(h) = 0$ for all $h \in H$. Then $(\pi_H, \pi_d) = \pi \in \Pi$ if and only if $\pi_h = v'_h(\pi_D)$ and $\pi_D$ is a Walrasian Equilibrium Price in the division $\{u'_a\}_A$.

**Proof.** Since $(\pi_H, \pi_D) = \pi \in \Pi$ implies that $\pi_h = V(hD_h) - \sum_{D_h} \pi_d$ if $h$ is matched to $D_h$ and $\pi_h \geq V(hD'_h) - \sum_{D'_h} \pi_d$ for every $D'_h \subset D$, then $\pi_h = v'_h(\pi_D)$ and $\pi_D$ is a Walrasian Equilibrium Price in the division $\{u'_a\}_A$. □

For completion, as the following results are well-known, we include the proofs for Proposition 2.

**Proposition. 2** In all matching markets,

- $\pi_d^{Mt} = V(A) - V(A \setminus d)$ for every $d \in D$. Leonard [1983]
- $\pi_d^{Mt} = V(dA) - V(A)$ for every $d \in D$. Gul and Stacchetti [1999]. Where $V(Ad)$ is the value of a market where an identical doctor $d$ is added.

**Proof.** We show that $\pi_d^{Mt} = V(A) - V(A \setminus d)$ for every $d \in D$. Let $Mt_D$ be the doctors optimal matching. Fix $d$, if $d$ is unassigned in $Mt_D$, then the result follows. If $d$ is assigned to hospital $h$, let $Mt'_D$ be a matching where everything is identical to $Mt_D$ but $t_d$ and $d$ are removed. In this market, $h$ holds whatever doctors he is assigned at the current transfers since $h$ preferences satisfy the substitutes condition. Since the new allocation is stable, it is efficient. Hence $\sum_{A \setminus d} \pi_d^{Mt} = V(A \setminus d)$. If $u_h(D_h \setminus d) - \sum_{D_h \setminus d} \pi_d^{Mt} < u_h(D_h) - \sum_{D_h} \pi_d^{Mt}$, then $d$ can increase his payoff in the original market, contradicting the maximality of $\pi_d^{Mt}$. Thus $\pi_d^{Mt} = V(A) - V(A \setminus d)$. To show that $\pi_d^{Mt} = V(Ad) - V(A)$ for every $d \in D$. Consider a market with agents $A$ plus a copy of doctor $d$ and a new hospital $h'$. Let $u_{h'}(d) < V(Ad) - V(A)$ and zero otherwise. In this new market $h'$ does not get $d$ (or his copy) in any stable allocation and hence $\pi_d^{Mt} \geq V(Ad) - V(A)$ since $(\pi_d^{Mt})$ constitute a Walrasian Equilibrium. Suppose now that $u_{h'}(d) > V(Ad) - V(A)$ and zero otherwise. Now in every stable allocation $h'$ gets $d$ (or his copy) and hence $\pi_d \leq V(Ad) - V(A)$ for any stable allocation in the economy $Ah'd$. Since
any stable allocation in \( Ah' d \) induces a stable allocation in \( A \) we have \( \pi_{d}^{M_{H}} \leq V(Ad) - V(A) \). Hence \( \pi_{d}^{M_{H}} = V(Ad) - V(A) \).

\[ \square \]

**Theorem.** 1 In the continuous transfers market, \( \pi_{h}^{M_{H}} = V(A) - V(A \setminus h) \) for every \( h \in H \).

**Proof.** We first show that \( \pi_{h}^{M_{H}} \leq V(A) - V(A \setminus h) \) for all \( h \in H \). Without loss of generality, consider the division where all surplus belongs to hospitals. To show that \( \pi_{h}^{M_{H}} \leq V(A) - V(A \setminus h) \) for all \( h \in H \). Let \( h \) be any hospital and let \( \pi_{h}^{M_{H}} \) be its maximum stable payoff, then \( \pi_{h}^{M_{H}} = V(A) - \sum_{A \setminus h} \pi_{a}^{M_{H}} \). Since the hospital preferred stable allocation belongs to the core, we have \( \sum_{A \setminus h} \pi_{a}^{M_{H}} \geq V(A/h) \). Hence the result. For the converse, without loss of generality, consider the division where all surplus belongs to hospitals. For notational simplicity let \( s^{*} \) be the set of doctors for hospital \( s \) at the doctor optimal stable allocation when hospital \( h \) is not present and let \( s_{s} \) be the set of doctors for hospital \( s \) at the hospital optimal stable allocation when hospital \( h \) is present. Let \( \pi^{*} \) and \( \pi_{s} \) be the corresponding prices. By the optimality of \( s^{*} \) we have \( v_{s}(\pi^{*}) \geq u_{s}(s_{s}) - \sum_{s_{d}} \pi_{s}^{*} \), this implies that \( v_{s}(\pi^{*}) \geq u_{s}(s_{s}) - \sum_{s_{s}} \pi_{s}^{*} + \sum_{s_{d}} \pi_{s}^{*} = v_{s}(\pi_{s}) + \sum_{s_{s}} (\pi_{s}^{*} - \pi_{s}^{*}) \). Let \( \pi^{*} \) be the corresponding doctor optimal stable allocation when hospital \( h \) is present. Let \( \pi^{*} \) and \( \pi_{s} \) be the corresponding prices. By the optimality of \( s^{*} \) we have \( v_{s}(\pi^{*}) \geq u_{s}(s_{s}) - \sum_{s_{d}} \pi_{s}^{*} \), this implies that \( v_{s}(\pi^{*}) \geq u_{s}(s_{s}) - \sum_{s_{s}} \pi_{s}^{*} + \sum_{s_{s}} (\pi_{s}^{*} - \pi_{s}^{*}) \) i.e \( 0 \geq v_{s}(\pi_{s}) - v_{s}(\pi^{*}) + \sum_{s_{s}} (\pi_{s}^{*} - \pi_{s}^{*}) \) for all \( s \). We also have that \( V(A) = \sum_{h} v_{s}(\pi_{s}) + \sum_{D} \pi_{s}^{*} \) and \( V(A/h) = \sum_{H/h} v_{s}(\pi^{*}) + \sum_{D} \pi_{s}^{*} \). Subtracting we have \( V(A) - V(A/h) = \sum_{h} v_{s}(\pi_{s}) + \sum_{D} \pi_{s}^{*} - \sum_{h} v_{s}(\pi^{*}) - \sum_{D} \pi_{s}^{*} \), reorganizing terms we have \( V(A) - V(A/h) = v_{h}(\pi_{s}) + \sum_{H/h} [v_{s}(\pi_{s}) - v_{s}(\pi^{*}) + \sum_{D} (\pi_{s}^{*} - \pi_{s}^{*})] + \sum_{h} (\pi_{s}^{*} - \pi_{s}^{*}) \). Since the second and third components are non-positive we have \( V(A) - V(A/h) \leq v_{h}(\pi_{s}) \) (the third component is non-positive by the opportunity cost condition). \( \square \)

**Lemma.** 1 and 3. Both VCG and \( U \)-VCG for agents in \( B \) are strategy-proof for agents in \( B \).

**Proof.** Suppose agent \( b \in B \) has preferences \( u_{b} \) and the reported and known preferences for other agents are \( \{u_{a}\}_{A \setminus b} \). Then the VCG payoff for \( b \) when sending \( u_{b} \) is \( u_{b}(M^{*}) + \sum_{a \in A \setminus b} u_{a}(M^{*}) - W(A \setminus b) \) and \( u_{b}(M^{**}) + \sum_{a \in A \setminus b} u_{a}(M^{**}) - W(A \setminus b) \) when sending \( u^{*} \), where \( (M)^{*} \in \arg \max_{Mt \in M_{0}(A) a \in A \setminus b} \sum_{a \in A \setminus b} u_{a}(M) + u_{b}(M) \) and \( (M)^{**} \in \arg \max_{Mt \in M_{0}(A) a \in A \setminus b} \sum_{a \in A \setminus b} u_{a}(M) + u_{b}(M) \). Since \( u_{b}(M^{*}) + \sum_{a \in A \setminus b} u_{a}(M^{*}) \geq u_{b}(M^{**}) + \sum_{a \in A \setminus b} u_{a}(M^{**}) \) we have that \( u_{b}(M^{*}) + \sum_{a \in A \setminus b} u_{a}(M^{*}) - W(A \setminus b) \geq u_{b}(M^{**}) + \sum_{a \in A \setminus b} u_{a}(M^{**}) - W(A \setminus b) \). Hence, VCG is strategy-proof for \( b \). The proof for \( U \)-VCG is analogous. \( \square \)

**Corollary.** 1 The VCG for Hospitals is stable and strategy-proof.
The VCG for hospitals is always strategy-proof and and delivers payoffs equal to \( V(A) - V(A \setminus h) \) for every \( h \in H \). According to the previous theorem these payoffs are identical to the hospital-optimal stable payoffs. Thus, VCG is stable. \(\square\)

**Corollary 2.** In auction markets, VCG payments coincide with the value of the assigned goods at the lowest Walrasian Equilibrium.

**Lemma.** Unit demand and linear hospital preferences satisfy the opportunity cost condition.

**Proof.** Suppose that \( \pi_{h}^{MtH} = V(A) - V(A \setminus h) \) for all \( h \in H \) and fix a hospital \( h^* \), then there is a stable matching in the market without \( h^* \) in which all agents receive \( \pi_{a}^{MtH} \). By construction for any \( C \subset A \setminus h^* \), \( \sum C \pi_{a}^{MtH} \geq V(C) \) and \( V(A \setminus h^*) = \sum \pi_{a}^{MtH} \). In the market without \( h^* \) doctor \( d \) has an optimal stable payoff of \( V(A \setminus h^*) - V(A \setminus h^*d) \) (by Leonard’s theorem) and \( V(A \setminus h^*) - V(A \setminus h^*d) \geq \pi_{d}^{MtH} = V(Ad) - V(A) \), where the inequality comes from the optimality of the doctor-optimal stable payoff and the equality from Gul and Stacchetti’s theorem. Leonard’s theorem shows that \( \pi_{h}^{MtH} = V(A) - V(A \setminus h) \) for all \( h \in H \) in the unit demand case. For the linear case we show directly that \( \pi_{h}^{MtH} = V(A) - V(A \setminus h) \) for all \( h \in H \). Fix a hospital \( h^* \) and let \( B \) the set of doctors matched with \( h^* \). Let every doctor in \( B \) who has a payoff equal to his outside option be unmatched. For every other \( d \in B \), there is a hospital \( h' \), set of doctors \( A \) and \( D \subset B \) such that \( V(h'AD) = \sum \pi_{a}^{MtH} \). Let \( C \) be the set of doctors assigned to \( h' \). We show that \( V(h'CB) = \sum \pi_{a}^{MtH} \). \( V(h'CB) = \sum \pi_{a}^{MtH} + \sum (\sum \pi_{a}^{MtH} + (\sum \pi_{a}^{MtH} + \pi_{h'}^{MtH} - V(h'A))) \), the term in brackets is non-negative by the stability of \( MtH \) and hence \( V(h'CB) \geq \sum \pi_{a}^{MtH} + \sum \pi_{a}^{MtH} \) which together with the stability inequality \( V(h'CB) \leq \sum \pi_{a}^{MtH} + \sum \pi_{a}^{MtH} \) imply the result. \(\square\)

**Theorem.** 2 Let \( \pi \in \Pi \), then there is a division such that \( \pi_{a}^{U} = \pi_{a} \), i.e. there is a division of surplus such that the U-VCG mechanism is stable and strategy-proof for all agents.

**Proof.** Let \( x_{a} \) be the match of agent \( a \) in an efficient allocation. Let \( u_{a}(x_{a}) = \pi_{a} \). For any \( h \) and set of doctors \( D_{h} \neq x_{h} \) let \( u_{h}(D_{h}) = V(D_{h}h) - \sum \pi_{d} \). For any \( d \) and \( h \neq x_{d} \) let \( u_{d}(h) = 0 \). With this division, the allocation is stable whenever any agent reports a zero value for every match, hence efficient. \(\square\)

**Theorem.** 3 Let \( U_{a} \) be agent \( a \)'s private value in an efficient assignment. Suppose no agent’s information is pivotal. i.e. removing one agent’s information does not change the efficient assignment, and \( (U_{a})_{A} \in \Pi \), then the U-VCG mechanism is strategy-proof for all agents and stable.
\textbf{Theorem. 4} Suppose preferences are linear and \( \Pi \) has a non-empty interior. Let \( h_d \) be the hospital matched with \( d \) at the efficient matching.

- If the lowest stable payoff for every hospital is zero and \( V(h_d d) > V(h'd) \) for all \( d \) and \( h' \neq h_d \), then there is \( \beta \in (0, 1) \) such that for any \( \alpha \in [0, \beta] \) \( \pi_h^U(\alpha) \) for all \( h \in H \) is stable.

- If for every doctor there are at least two hospitals \( h, h' \) such that \( V(hd) > 0 \) and \( V(h'd) > 0 \), then there is \( \beta \in (0, 1) \) such that for any \( \alpha \in [\beta, 1] \) \( \pi_h^U(\alpha) \) for all \( h \in H \) is stable.

\textbf{Proof.} Since the lowest stable payoff for every hospital is zero, \( \pi^U(0) \) is stable. Since \( V(h_d d) > V(h'd) \) for all \( d \) and \( h' \neq h_d \), it is possible to increase \( h \) payoff by \( \epsilon > 0 \) and have a set of stable payoffs, i.e. it is possible to reduce the payoff of all doctors matched with hospital \( h \) without them being able to form.
a blocking coalition. Hence \( \epsilon(V(h_1D_{h1}), \cdots, V(h_H D_{hH})) \) is a stable payoff for hospitals for \( \epsilon \) sufficiently small. Let \( \beta = \sup \{ \epsilon > 0 | \epsilon(V(h_1D_{h1}), \cdots, V(h_H D_{hH})) \text{ is a stable payoff} \} \). Since \( \Pi \) is a convex closed set we have the result.

Suppose that for every doctor there are at least two hospitals \( h, h' \) such that \( V(hd) > 0 \) and \( V(h'd) > 0 \). Suppose there is a hospital \( h \) such that \( \frac{d}{d\alpha} \pi^U_h(\alpha) > 0 \) for every \( \alpha \in (\epsilon, 1) \) for every \( \epsilon > 0 \). This implies that for all \( \alpha \in (\epsilon, 1) \), \( \pi^U_h(\alpha) = V(hD^*_h) > 0 \) i.e. hospital \( h \) is assigned doctors in \( D^*_h \), even when they produce a surplus \( V(hD^*_h)(1 - \alpha) \); however, there for every doctor in \( D^*_h \) there is another hospital \( h' \) such that \( V(h'd) > V(h'd)(1 - \alpha) \). Thus for every \( h \) there is \( \epsilon_h \) for which \( \pi^U_h(\alpha) = 0 \) for all \( \alpha \in [\epsilon_h, 1] \). Let \( \beta = \max \{ \epsilon_h \} \).

**Theorem. 5** Suppose preferences are integer valued, then strategy-proofness and stability are possible in any \( q \)-market for any \( q \in \mathbb{Z} \) such that \( q \geq 1 \).

**Proof.** Suppose the VCG mechanism is used. If preferences are integer valued, then agent \( b \) payment is equal to \( p_b = u_b(M_b) - W'(A) + W'(A \setminus b) \in \mathbb{Z} \). If in addition, \( q \geq 1 \) all possible payments are implementable as a matching. Hence, strategy-proofness and stability are possible in the \( \infty \)-market.

**References**


