Principal-Agent VCG Contracts

Ron Lavi∗       Elisheva Shamash†

May 7, 2019

Abstract

We study a game of complete information with multiple principals and multiple common agents. Each agent takes an action that can affect the payoffs of all principals. Prat and Rustichini (Econometrica, 2003) who introduce this model assume first price contracts: each principal offers monetary transfers to each agent conditional on the action taken by the agent. We define a notion of VCG contracts which are a restricted natural class of contractible contracts and study its effect on the existence of efficient pure subgame perfect equilibrium outcomes. We identify a “unitary balancedness” condition that is necessary and sufficient for the existence of a pure subgame perfect equilibrium (SPE) with VCG contracts. As a consequence, we show that the class of instances of this game that admit an efficient SPE with VCG contracts strictly contains the class of instances of this game that admit an efficient SPE with first price contracts. Although VCG contracts broaden the existence of pure subgame perfect equilibria, we show that the worst case welfare loss in any SPE outcome with VCG contracts is not worse than the respective worst case loss with first price contracts.

1 Introduction

Most of the earlier principal-agent settings consider various types of “first-price” contracts where the payment of a principal to an agent can be fully determined given the agent’s action and/or the resulting outcome. For example, the seminal paper of Bernheim and Whinston (1986) considers a complete information common agency model where multiple principals offer first price contracts to a single common agency. Prat and Rustichini (2003) generalize this model to the case of multiple principals and multiple agents, demonstrating that efficient equilibrium outcomes may not necessarily exist with multiple agents.

In fact, it is now well-known that simple first-price contracts limit the existence of desirable equilibrium outcomes and more elaborate contract types have been proposed. For example, Peters and Szentes (2012) and Szentes (2015) define and study contractible contracts – contracts that can depend on the contracts proposed by other principals. The main conclusion is that these types of contracts can significantly expand the set of implementable outcomes (in particular the efficient ones). On the other hand, these papers assume a broad space of contracts that allows for various rather strong collusive contractual statements. There could be settings in which such contracts

∗Technion – Israel Institute of Technology. Email: ronlavi@ie.technion.ac.il.
†Technion – Israel Institute of Technology. Email: elisheva@campus.technion.ac.il.
may not be acceptable or legal.\footnote{In addition, the non-simplicity and the complicated nature of these contracts could be a disadvantage, see for example the discussion in Peters (2014).} In addition, the non-simplicity and the complicated nature of these contracts could be a disadvantage, see for example the discussion in Peters (2014).

This paper analyzes a specific class of contractible contracts that on the one hand seems simple and legally acceptable and on the other hand extends the existence of efficient equilibrium outcomes. We study this in the context of the complete-information setting of “Games Played Through Agents” (GPTA) Prat and Rustichini (2003). There are $M$ principals and $N$ agents, each agent $n$ chooses an action $s_n \in S_n$ and each principal $m$ obtains a value $G^m: S_1 \times \cdots \times S_N \to \mathbb{R}_{\geq 0}$ from the tuple of agent actions. The extensive-form game has two steps: First, principals simultaneously offer contracts to agents that describe the transfer that the agent will receive from the principal. Second, based on these contracts, agents choose actions. Utilities are then realized: a principal’s utility is her value from the tuple of chosen actions minus the sum of transfers that she makes; an agent’s utility is the sum of transfers she receives. Prat and Rustichini (2003) study first-price contracts that are a mapping from the agent’s action to a payment for that action. We refer to this game as the 1st Price GPTA (FP-GPTA).

A special case of this setting is the extensively studied model of multiple simultaneous single-item auctions (e.g., Bikhchandani (1999); Syrgkanis and Tardos (2013); Feldman, Fu, Gravin, and Lucier (2013); Roughgarden, Syrgkanis, and Tardos (2017)). Each buyer has a value for subsets of items and submits separate bids in the different auctions. With first-price single-item auctions this becomes a special case of FP-GPTA: the sellers are the agents, the agent’s actions are the choice of which buyer will receive the item, the buyers are the principals, and the valuation function of buyer $m$ is $G^m(\cdot)$. The bid that buyer $m$ submits specifies the price that will be paid to seller $n$ if her action is to assign the item to $m$. This is exactly the first-price contract of Prat and Rustichini (2003). Second-price auctions cannot be represented in this model since the contract of a buyer in this case depends on the other buyers’ contracts. Namely, principal $m$ submits her willingness to pay but her actual payment is the second highest bid which is stated in another principal’s contract. Thus, to capture second-price auctions one needs to submit contracts that depend on other contracts.

Our definition of VCG contracts follows this intuition. Principal $m$ submits a willingness to pay $p_{m,n}(s_n)$ for every action $s_n$ of every agent $n$. To determine the transfer of principal $m$ to agent $n$, the agent is viewed as a social designer choosing an alternative from $S_n$. The principals are the players of the VCG framework and principal $m$’s value for $s_n \in S_n$ is considered to be $p_{m,n}(s_n)$. The transfer principal $m$ pays agent $n$ is determined according to Clarke’s payment rule. We term the resulting game a VCG-GPTA.

Although this definition is conceptually motivated by the VCG mechanism, there are three significant differences between the two settings. First, agent $n$ is not a social designer but a strategic player. Second, principals pay to multiple agents rather than to a single social designer. Third, a principal’s value is determined by a combination of all agents’ actions and not only by a specific action of a single agent. In particular, the principals often do not have a dominant strategy and a multiplicity of subgame perfect equilibria exist in our setting.\footnote{For example, article 101 of the Treaty on the Functioning of the European Union prohibits anti-competitive behaviour such as price-fixing and production limitations. In 2004 the EU ordered Microsoft to pay about 800 million USD for violating its competition law, and in 2018 the EU fined Google for about 5 billion USD for anti-competitive behaviour.}

As in Prat and Rustichini (2003), our main focus is on the resulting equilibrium welfare (i.e., the
sum of payoffs of all agents and principals) and in particular on the existence of efficient equilibria. While there exist both FP-GPTA as well as VCG-GPTA that do not admit an efficient pure subgame perfect equilibrium, the main message of this paper is that VCG contracts improve the existence of efficient pure subgame perfect equilibria relative to first-price contracts. Specifically, we prove that for every GPTA setting, every pure subgame perfect equilibrium (SPE) in the corresponding FP-GPTA is also a SPE in the corresponding VCG-GPTA. Moreover, for any \( M, N, S_1, \ldots, S_N \) (the cardinality of all these sets must be at least 2), there exist valuations \( G^1(\cdot), \ldots, G^M(\cdot) \) for which the corresponding VCG-GPTA admits an efficient SPE while the corresponding FP-GPTA does not. In fact, the measure of such valuations is positive.\(^3\)

Since our VCG contracts expand the set of subgame perfect equilibria relative to first-price contracts, one might wonder whether additional low welfare SPEs are added. We study this issue using a Price of Anarchy (PoA) analysis (see, e.g., Roughgarden (2005)). The PoA of either a FP-GPTA or a VCG-GPTA is the ratio between the lowest welfare that is obtained in a SPE and the maximal welfare that can be obtained in that game (not necessarily in equilibrium). We first show that the PoA of any FP-GPTA is at least \( \frac{1}{M+1} \) and that for any \( M \) there exists an FP-GPTA with \( M \) principals whose PoA is \( \frac{1}{M} \).\(^4\) We then show that these exact PoA bounds hold for VCG-GPTA as well.

We therefore conclude that the worst-case welfare loss in a VCG-GPTA is not worse than that in a FP-GPTA while the best-case welfare guarantee strictly improves due to the improved existence of efficient equilibria.

Our analysis of pure sub-game perfect equilibria in the VCG-GPTA relies on a necessary and sufficient condition for their existence. We compare our condition, “unitary balancedness”, to a similar balancedness condition given in Prat and Rustichini (2003) (which, as explained there, is conceptually reminiscent of the cooperative concept of balancedness Scarf (1967)). The satisfiability of both conditions can be checked by (two different) linear programs. The comparison between the two conditions makes it clear why some VCG-GPTA admit efficient equilibria while the corresponding FP-GPTA do not.

This characterization approach is conceptually different from the approach taken by most of the literature on multiple simultaneous single-item auctions. The necessary and sufficient conditions for the existence of a pure equilibrium identified both here and in Prat and Rustichini (2003) examine tuples of valuations (i.e., the combination of all principal valuations) and characterize the properties that are equivalent to the existence of a pure equilibrium. In contrast, the auction literature characterizes valuation classes (e.g., gross substitutes) for each single buyer that guarantee existence of a pure equilibrium. For example, if all valuations are gross substitutes then an equilibrium is guaranteed to exist in both first-price and second-price multiple simultaneous auctions Bikhchandani (1999); Fu, Kleinberg, and Lavi (2012). The connections between these two approaches and their implications is an interesting topic for future investigation, especially since the GPTA setting is much more general (for example, it allows for auctions with externalities).

The remainder of the paper is organized as follows. Section 2 reviews the GPTA setting and some relevant results from Prat and Rustichini (2003). Section 3 defines the notion of VCG contracts along with some of their fundamental properties. Section 4 presents a sufficient and necessary

\(^3\)The measure analysis is performed only with respect to weakly truthful outcomes in the FP-GPTA since Prat and Rustichini (2003) only characterize these types of equilibria, see section 5.

\(^4\)Prat and Rustichini (2003) do not discuss equilibrium selection issues and the possible welfare loss in inefficient equilibria, focusing mainly on efficient equilibria. We feel that this issue deserves attention.
condition for the existence of a pure SPE in VCG-GPTA using a notion of unitary balancedness. In section 5 we compare the set of principal valuations that admit an efficient SPE in the corresponding VCG-GPTA to the set of principal valuations that admit a weakly truthful SPE in the corresponding FP-GPTA. We show that the latter is strictly contained in the former, and that the difference between the two sets has positive measure. Section 6 analyzes the price of anarchy of the two games, and section 7 concludes.

2 Games Played Through Agents (GPTA)

Prat and Rustichini (2003) introduced the following complete-information principal-agent setting \( \mathcal{G} = (M, N, S, G) \). There is a set \( M \) of principals and a set \( N \) of agents (abusing notation, \( M \) and \( N \) are also sometimes used to denote the respective cardinalities of these sets). Each agent \( n \in N \) has a finite set of actions \( S_n \), and let \( S = \prod_{n \in N} S_n \). A tuple of actions \( s = (s_1, s_2, \ldots, s_N) \in S \) is termed an outcome. Each principal \( m \in M \) obtains different gains from different outcomes, represented by the function

\[
G^m : S_1 \times S_2 \times \ldots \times S_N \to \mathbb{R}_{\geq 0}.
\]

Tables 1 and 2 show examples with two agents and two principals. Agent’s actions correspond to the rows and columns of the tables, while principal gains are given in the table cells (note that this is not a standard normal-form game as agent utilities are not given).

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>4 , 4</td>
<td>5 , 0</td>
</tr>
<tr>
<td>B</td>
<td>0 , 5</td>
<td>1 , 1</td>
</tr>
</tbody>
</table>

Table 1: “Prisoner’s Dilemma”

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>c , c+1</td>
<td>c+1 , c</td>
</tr>
<tr>
<td>B</td>
<td>c+1 , c</td>
<td>c , c+1</td>
</tr>
</tbody>
</table>

Table 2: “Matching Pennies”

The usual principal-agent extensive-form game is composed of two steps. First, principals offer contracts to agents. Second, given these contracts, agents choose actions. The contract that principal \( m \in M \) offers to agent \( n \in N \) specifies a vector of “principal bids” \( pb^m_n \in \mathbb{R}_{\geq 0}^{S_n} \) where \( pb^m_n(s_n) \) is the maximal willingness of principal \( m \) to pay to agent \( n \) if she will choose action \( s_n \).

Let \( PB^m_n \subseteq \mathbb{R}_{\geq 0}^{S_n} \), let \( PB^m = \prod_{n \in N} PB^m_n \) and let \( PB = \prod_{m \in M} PB^m \).

The actual payment principal \( m \) will pay agent \( n \) is an endogenously-determined payment function \( t_{m,n} : S_n \times \mathbb{R}_{\geq 0}^{\lvert M \rvert \cdot \lvert S_n \rvert} \to \mathbb{R} \), where \( t_{m,n}(s_n, pb^1_n, \ldots, pb^M_n) \) is the payment that agent \( n \) receives from principal \( m \) when her action was \( s_n \) and the principal bids were \( (pb^1_n, \ldots, pb^M_n) \). The game therefore proceeds as follows:

1. Each principal \( m \in M \) chooses a vector of bids \( pb^m = (pb^m_1, \ldots, pb^m_N) \), and let \( pb = (pb^1, \ldots, pb^N) \in PB \).

2. Given these principal bids, each agent \( n \in N \) chooses an action \( s_n \in S_n \).

3. Principal \( m \)’s utility is:

\[
u^m(s, pb) = G^m(s) - \sum_{n \in N} t_{m,n}(s_n, pb^1_n, \ldots, pb^M_n).
\]

Prat and Rustichini (2003) assume more generally that the range of \( G \) includes all the reals, while we assume, for simplicity, that it includes only non-negative reals.
Agent $n$’s utility is $u_n(s, pb) = \sum_{m \in M} t_{m,n}(s_n, pb^1_n, ..., pb^M_n)$.

Prat and Rustichini (2003) use the “First-Price” (FP) payment function $t_{m,n}^{FP}(s_n, pb) = pb_n^m(s_n)$. We refer to this game as the First-Price GPTA (FP-GPTA). Clearly, agent strategies $\hat{s}(\cdot)$ that are part of a SPE of the FP-GPTA must satisfy the following condition:

- **Agent Maximization (AM):** $\forall n, \forall pb : \hat{s}_n(pb) \in \arg\max_{s_n \in S_n} \sum_{m \in M} pb_n^m(s_n)$.

Throughout (for brevity and slightly abusing notation), when we write that $(\hat{p}_n, \hat{s})$ is a SPE we mean that the agent strategies in this SPE (as in all SPEs) satisfy AM and that $\hat{s}$ is the tuple of actions that the agents choose according to AM when the principles play $pb$.

**Lemma 1.** [Prat and Rustichini (2003)] Given a FP-GPTA, $(\hat{p}_n, \hat{s})$ is a SPE if and only if the following three conditions are satisfied:

1. **Agent Maximization (AM):** As was defined above.
2. **Incentive Compatibility (IC):**
   \[
   \forall m \in M, s \in S : G^m(\hat{s}) + \sum_{j \in M \setminus m} \sum_{n \in N} \hat{p}_j^n(s_n) \geq G^m(s) + \sum_{j \in M \setminus m} \sum_{n \in N} \hat{p}_j^n(s_n).
   \]
3. **Cost Minimization (CM):** $\forall m \in M, n \in N : \sum_{m \in M} \hat{p}_n^m(\hat{s}_n) = Max_{s_n \in S_n} \sum_{j \in M \setminus m} \hat{p}_j^n(s_n)$.

For example, in the FP-GPTA based on table 1, one SPE is: $\hat{p}_1^1(T) = 2, \hat{p}_2^1(B) = 3, \hat{p}_1^2(B) = \hat{p}_2^2(B) = 2, \hat{p}_2^2(R) = 0$, with agent strategies $\hat{s} = (T, L)$. Note that AM, IC and CM are satisfied. In contrast, in the FP-GPTA based on table 2 there is no SPE for any choice of $c \in \mathbb{R}$.

In this paper we generalize the setting of Prat and Rustichini (2003) by allowing the actual payment of principal $m$ to agent $n$ to be determined not only by her own bids but also by the bids of the other principals to that agent, according to an exogenously-determined payment function $t_{m,n} : S_n \times \mathbb{R}_0^{[M] | S_n} \rightarrow \mathbb{R}$.

### 3 VCG Contracts

We consider a different payment rule and refer to the resulting game as VCG-GPTA:

\[
\begin{align*}
\hat{t}_{m,n}^{VCG}(s_n, pb) = Max_{x \in S_n} \left\{ \sum_{m' \in M \setminus m} \hat{p}_{n}^{m'}(x) \right\} - Max_{x \in S_n} \left\{ \sum_{m' \in M} \hat{p}_{n}^{m'}(x) \right\} + p_{n}^{m}(s_n) \tag{3}
\end{align*}
\]

Note that $\hat{t}_{m,n}^{VCG}(s_n, pb) \leq p_{n}^{m}(s_n)$ for any $s_n, pb$, since the first term is never smaller than the second term. To see the conceptual connection to VCG, consider a setting in which agent $n$ is a social designer, the set of alternatives to her choice is $S_n$, the principals are the players of the VCG framework, and alternative $s_n \in S_n$ is considered to have a private value of $p_{n}^{m}(s_n)$ to principal $m$. If the agent executes Clarke’s mechanism (Clarke (1971)) in this setting, and

---

*Prat and Rustichini (2003) show this for $c = 0$ and more generally this can be verified for any $c$ in a similar way.*
Proof. AM is necessary and sufficient for action ˆ}{GPTA}.

Theorem 1. Given a VCG-GPTA,

\[ \text{Lemma 2. For any } pb \text{ and } n: \]

\[ \argmax_{x \in S_n} \left\{ \sum_{m \in M} t_{m,n}^VCG(x, pb) \right\} = \argmax_{x \in S_n} \left\{ \sum_{m \in M} pb_{m,n}^m(x) \right\}. \]

Therefore AM holds in any SPE of the VCG-GPTA game.

Proof. Let \( C_{m,n} = \max_{y \in S_n} \left\{ \sum_{m' \in M \setminus m} pb_{n,m'}(y) \right\} - \max_{x \in S_n} \left\{ \sum_{m' \in M} pb_{n,m'}(z) \right\}. \) Then \( \forall n \in N: \)

\[ \argmax_{x \in S_n} \left\{ \sum_{m \in M} t_{m,n}^VCG(x, pb) \right\} = \argmax_{x \in S_n} \left\{ \sum_{m \in M} C_{m,n} + pb_{n,m}(x) \right\}. \]

In addition, \( \argmax_{x \in S_n} \left\{ \sum_{m \in M} C_{m,n} \right\} \)

Lemma 3. For any \( pb, m, n, \) and for any chosen \( s_n \in \argmax_{x \in S_n} \left\{ \sum_{m \in M} pb_{n,m}(x) \right\} : \)

\[ t_{m,n}^VCG(s_n, pb) = \max_{x \in S_n} \left\{ \sum_{m' \in M \setminus m} pb_{n,m'}(x) \right\} - \left\{ \sum_{m' \in M \setminus m} pb_{n,m'}(s_n) \right\}. \]

As a result, two useful properties of \( t_{m,n}^VCG \) are:

(a) for any \( pb, m, n, s_n \), we have \( t_{m,n}^VCG(s_n, pb) \geq 0. \)

(b) \( \forall y_n \in S_n, \forall pb, \) \( pb \in PB : t_{m,n}^VCG(y_n, (pb^m, pb^{-m})) = t_{m,n}^VCG(y_n, (\hat{p}_b^m, pb^{-m})) \)

Proof. According to lemma 2, and equation 3:

\[ t_{m,n}^VCG(s_n, pb) = \max_{x \in S_n} \left\{ \sum_{m' \in M \setminus m} pb_{n,m'}(x) \right\} - \left\{ \sum_{m' \in M} pb_{n,m'}(s_n) \right\} + pb_{n,m}(s_n). \]

Therefore, \( t_{m,n}^VCG(s_n, pb) = \max_{x \in S_n} \left\{ \sum_{m' \in M \setminus m} pb_{n,m'}(x) \right\} - \left\{ \sum_{m' \in M \setminus m} pb_{n,m'}(s_n) \right\}. \)

Note that since there are several agents and the utility of each principal depends on the tuple of actions of all agents, the comparison to VCG is only a non-formal analogy. In the special case of \( N=1 \) (there is only one agent), lemmas 2 and 3 together with the VCG logic imply that in the unique SPE, principals announce \( pb_{n,m}(s_n) = G_{m,n}(s_n) \) for any \( s_n \in S_n. \)

Theorem 1. Given a VCG-GPTA, \( (\hat{p}_b, \hat{s}) \) is a SPE if and only if AM and IC are satisfied.

Proof. AM is necessary and sufficient for action \( \hat{s} \) to be agent \( n \)’s best response to \( \hat{p}_b. \)

In order to show that IC is sufficient, recall that IC states that \( \forall m \in M \) and \( \forall y \in S: \)

\[ G_{m}(\hat{s}) + \sum_{\{m' \in M \setminus m\}} \sum_{n \in N} \hat{p}_{n,m'}(\hat{s}_n) \geq G_{m}(y) + \sum_{\{m' \in M \setminus m\}} \sum_{n \in N} \hat{p}_{n,m'}(\hat{y}_n) \quad (4) \]

By subtracting \( \sum_{n \in N} \max_{x \in S_n} \sum_{m' \in M \setminus m} \hat{p}_{n,m'}(x) \) from both sides of inequality (4), due to lemma 3, we attain \( G_{m}(\hat{s}) - \sum_{n \in N} t_{m,n}^VCG(\hat{s}_n, \hat{p}_b) \geq G_{m}(y) - \sum_{n \in N} t_{m,n}^VCG(y_n, \hat{p}_b). \) Therefore, principal \( m \) cannot increase its utility by declaring a different principal bid \( pb^m \) that results in the agents
choosing an action vector \( y \in S \). This is because in such a case, principal \( m \)'s payment will be
\[
\sum_{n \in N} t_{m,n}^{\text{VCG}}(y_n, (\hat{p}_b^m, \hat{p}_m)) = \sum_{n \in N} t_{m,n}^{\text{VCG}}(y_n, (p_b^m, \check{p}_m))
\]
due to lemma 3.

In order to show that IC is necessary, let us assume negatively that IC does not hold. That is

\[
\exists \hat{s} \in S, \exists m \in M : G^m(\hat{s}) + \sum_{m' \in M \setminus m} \sum_{n \in N} p_b^{m'}(\hat{s}_n) < G^m(\hat{s}) + \sum_{m' \in M \setminus m} \sum_{n \in N} p_b^{m'}(\hat{s}_n)
\]

By announcing \( \forall n \in N, \forall s_n \neq \hat{s}_n : p_b^{n}(s_n) = 0 \) and \( \forall s_n \in \hat{s} : p_b^{n}(s_n) = \sum_{m' \in M \setminus m} p_b^{m'}(\hat{s}_n) - \sum_{m' \in M \setminus m} p_b^{m'}(\hat{s}_n) \), principal \( m \) can cause agents \( n \in N \) to deviate to \( \hat{s} \), and thus increase her utility by at least:

\[
G^m(\hat{s}) - G^m(s) - \sum_{m' \in M \setminus m} \sum_{n \in N} p_b^{m'}(s_n) - \sum_{m' \in M \setminus m} p_b^{m'}(\hat{s}_n),
\]

which is positive, due to equation (5). Therefore \((\hat{s}, \hat{p})\) is not a SPE, which is a contradiction.

Notice that a SPE for VCG-GPTA always exists, because \( \forall m \in M, \forall \hat{s} \in \text{argmax}_x \in S G^m(x) \), principal \( m \) can bid \( \forall n \in N : p_b^{n}(s_n) \) a sufficiently large number and \( \forall s_n \neq \hat{s}_n : p_b^{n}(s_n) = 0 \), while all other principal’s bid zero, on all agent actions. This is the same problem which arises in the 2nd price auction setting. To rule out these equilibria, which are not so interesting, Fu, Kleinberg, Lavi, and Smorodinsky (2017) use the refinement of Nash equilibria with no overbidding.

**Definition 1.** No over-Bidding (NoB): A SPE \((\hat{s}, \hat{p})\) satisfies NoB if

\[
\forall m \in M : \sum_{n \in N} p_b^{m}(\hat{s}) \leq G^m(\hat{s}).
\]

In table 1, the VCG-GPTA has the same SPE with NoB \((\hat{s}, \hat{p})\) as the FP-GPTA with \( \forall m \in M, \forall n \in N : t_{m,n}^{\text{FP}}(\hat{s}, \hat{p}) = t_{m,n}^{\text{VCG}}(\hat{s}, \hat{p}) \). On the other hand, in contrast to the FP-GPTA which does not have a SPE in table 2 where \( c = 1 \), the VCG-GPTA does have a SPE satisfying NoB. For example, \((\hat{s}, \hat{p})\) where \( \hat{s} = TL, \hat{p}_1(T) = \hat{p}_2(L) = 1 \), with all other principal bids equal to zero.

**Theorem 2.** Every pair \((\hat{s}, \hat{p})\) that satisfies AM, IC and CM, also satisfies NoB. As a result, every SPE \((\hat{s}, \hat{p})\) in a FP-GPTA is also a SPE with NoB in a VCG-GPTA.

**Proof.** Prat and Rustichini introduced the following corollary:

**Observation 1.** [Based on Prat and Rustichini (2003) - corollary 1] Given a SPE \((\hat{s}, \hat{p})\) in a FP-GPTA, \( \forall n \in N: \exists \hat{s}_n \in S_n, \hat{s}_n \neq s_n \) s.t. \( \sum_{m \in M} p_b^{n}(s_n) = \sum_{m \in M} p_b^{n}(\hat{s}_n) \). Moreover, \( \forall m \in M, \forall n \in N, \exists a_n \in S_n \) s.t. \( \sum_{m \in M} p_b^{n}(s_n) = \sum_{m \in M} p_b^{n}(a_n) \) and \( \hat{p}_n(a_n) = 0 \).

Consider a SPE \((\hat{s}, \hat{p})\) in a FP-GPTA. Assume negatively, that NoB is not satisfied, that is: \( \exists m \in M : G^m(\hat{s}) < \sum_{n \in N} p_b^{n}(s_n) \). Because \( \hat{s} \) is a SPE in the FP-GPTA, IC is satisfied, so:

\[
\forall a \in S : G^m(\hat{s}) + \sum_{n \in N} \sum_{m \in M} p_b^{m}(s_n) - \sum_{n \in N} p_b^{m}(s_n) \geq G^m(a) + \sum_{n \in N} \sum_{m \in M} p_b^{m}(a_n) - \sum_{n \in N} p_b^{m}(a_n)
\]

Due to observation 1, \( G^m(\hat{s}) - \sum_{n \in N} p_b^{m}(s_n) \geq G^m(a) \). Due to our negative assumption, we have: \( 0 > G^m(\hat{s}) - \sum_{n \in N} p_b^{m}(s_n) \). Due to equation (1), we have: \( G^m(a) \geq 0 \). A contradiction. □

---

7 See also Christodoulou, Kovács, and Schapira (2008).
8 In fact there exists a SPE in the VCG-GPTA for all \( c \geq 1 \) (see theorem 3).
4 Equilibrium Existence via Balancedness

In this section we give a characterization of SPEs in the VCG-GPTA via a notion of unitary balancedness that we will soon define. Prat and Rustichini (2003) defined a similar notion of balancedness and it is useful to start with their notion.

Definition 2. [Prat and Rustichini (2003) - Definition 7] In a FP-GPTA, the vector $\omega$ with dimensions $M \times S$ is a vector of Balanced Weights (BW) with respect to $\hat{s}$ if all its elements are non-negative and $\forall m \in M, \forall n \in N, \forall a_n \in S_n \backslash \hat{s}_n : \sum_{s: s_n = a_n} \omega^m(s) = \sum_{s: s_n = a_n} \omega^1(s)$.

Prat and Rustichini explain that “The matrix $\omega$ assigns a weight to every agent and every deviation... A deviation $s$ involves some agent switching from $\hat{s}_n$ to $s_n$,.. BW requires that for every agent $n$ the sum of weights on a deviation involving a switch on the part of $n$ is constant across principals.” They additionally remark that “The definition of BW is reminiscent of that used in cooperative game theory (e.g. Scarf (1967)) but of course it is different because of the distinction in our game between principals and agents.” Our characterization relies on a slightly different definition of balancedness:

Definition 3. In a VCG-GPTA, the vector $\omega = (\omega^m(s))_{m \in M, s \in S}$ is a collection of Unitary Balanced Weights (UBW) with respect to $\hat{s}$, if $\forall m \in M, \forall s \in S \setminus \hat{s}, \omega^m(s) \geq 0$, and $\forall m \in M, \forall n \in N, \sum_{s: s_n \neq \hat{s}_n} \omega^m(s) = 1$.

<table>
<thead>
<tr>
<th>$S_1 \setminus S_2$</th>
<th>$s_{21}$</th>
<th>$s_{22}$</th>
<th>$s_{23}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{11}$</td>
<td>$\hat{s}$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$s_{12}$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3: A VCG-GPTA with UBW

To illustrate this definition consider table 3, where $S_1 = \{s_{11}, s_{12}\}$ and $S_2 = \{s_{21}, s_{22}, s_{23}\}$. Assume $\hat{s} = (s_{11}, s_{22})$. The sum of the weights in the grey cells in table 3 is equal to one for each $m \in M$, and includes all cells in which agent $n = 2$ is not playing in accordance to the SPE. The weights $\{\omega^m(s)\}$ should satisfy the requirement in definition 3 simultaneously for all $n \in N$.

Definition 4. [Prat and Rustichini (2003) - Definition 4] A FP-GPTA is balanced with respect to $\hat{s}$ if, for every collection of balanced weights $\omega \in BW$, $\sum_{m \in M} \sum_{s \in S} \omega^m(s) [G^m(\hat{s}) - G^m(s)] \geq 0$

Definition 5. A VCG-GPTA is said to be Unitary-Balanced (UB) with respect to $\hat{s}$ if

$$\forall \omega \in UBW : \sum_{m \in M} \left[ \sum_{s \in S} \omega^m(s) [G^m(\hat{s}) - G^m(s)] + \sum_{\{m' \in M \setminus M\}} G^{m'}(\hat{s}) \right] \geq 0$$

Prat and Rustichini (2003) define a notion of a “weakly truthful equilibrium” and show that a FP-GPTA is balanced w.r.t. $\hat{s}$ if and only if there exists such an equilibrium. Specifically, $pb^m$ is weakly truthful relative to $\hat{s}$ if for every $s \in S$, $G^m(s) - \sum_{n \in N} pb^m_n(s) \leq G^m(\hat{s}) - \sum_{n \in N} pb^m_n(\hat{s})$. This notion of a weakly truthful equilibrium generalizes the notion of a truthful equilibrium of Bernheim and Whinston (1986).
Theorem 3. A VCG-GPTA has a SPE that satisfies NoB with outcome $\hat{s} \in S$ if and only if it is UB with respect to $\hat{s}$.

Our result fully characterizes existence of SPE while Prat and Rustichini (2003) only characterize the case where weakly truthful strategies form a SPE (and in fact explicitly mention the full characterization of sub-game equilibria as an open problem).

Proof. Consider the following matrices and vectors (using the indices $m, j \in [1, |M|], i, n \in [1, |N|], s \in [1, |S|])$:

Matrix $A$ is of dimensions $(|S| \cdot |M|) \times (|N| \cdot |M|)$, such that:

\[
A_{(sm,nj)} = \begin{cases} 1 & \text{if } j \neq m, s_n \neq \hat{s}_n, \\ 0 & \text{otherwise} \end{cases} \quad (6)
\]

We demonstrate Matrix $A$, where $S = \{s^1, s^2, ..., s^{|S|}\}$:

| $S \cdot M \setminus N \cdot M$ | $n_1, j_1$ | $m_1, j_2$ | ... | $m_1, j_{|M|}$ | $n_2, j_1$ | ... |
|-----------------|-----------|-----------|----|-------------|-----------|----|
| $s^1, m_1$      | 0         | 1         | ...| 1           | 0         | .  |
| $s^1, m_2$      | 1         | 0         | ...| 1           | 1         | ...|
| $s^i$           | :         | :         | ...| :           | :         | :  |
| $s^1, m_{|M|}$  | 1         | 1         | ...| 0           | 1         | ...|
| $s^2, m_1$      | 0         | 0         | 0  | 0           | 0         | 0  |
|                 | :         | :         | ...| :           | :         | :  |

Table 4: The matrix $A$, where $\hat{s} = s^2$, $\forall i \in \{1, 2\} : \hat{s}_i \neq s^1_i$.

Matrix $B$ is of dimensions $|M| \times (|N| \cdot |M|)$, such that:

\[
B_{(m,nj)} = \begin{cases} 1 & \text{if } j = m, \\ 0 & \text{otherwise} \end{cases} \quad (7)
\]

Vector $a$ is of length $|S| \cdot |M|$, such that:

\[
a_{sm} = G^m(s) - G^m(\hat{s}) \quad (8)
\]

Vector $b$ is of length $|M|$ such that:

\[
b_m = G^m(\hat{s}) \quad (9)
\]

Consider the following two lemmas:

Lemma 4. There exists a vector $x \in \mathbb{R}^{N \cdot M}$ such that:

\[
A \cdot x \geq a \quad (10)
\]

\[
B \cdot x = b \quad (11)
\]

\[
x \geq 0 \quad (12)
\]

iff there exists a SPE $(\hat{s}, \hat{pb})$ in the VCG-GPTA that satisfies NoB.

---

11 Formally speaking, if we denote the $A_{(sm,nj)}$ element as $A_{(k,h)}$, then: $s = (k \text{ div } |M|) + 1$, and $m = (k \text{ mod } |M|) + 1$. $n = (h \text{ div } |M|) + 1$, and $j = (h \text{ mod } |M|) + 1$.

12 Formally speaking, if we denote the $B_{(m,nj)}$ element as $B_{(m,h)}$, then $n = (h \text{ div } |M|) + 1$, and $j = (h \text{ mod } |M|) + 1$.

13 Formally speaking, if we denote the $a_{sm}$ element as $a_k$, then, $s = (k \text{ div } |M|) + 1$, and $m = (k \text{ mod } |M|) + 1$. 

9
Proof. Initially assume that there exists \( x \in \mathbb{R}^{N \times M} \) that satisfies conditions (10), (11) and (12). We show that then there exists a SPE that satisfies NoB. In particular, construct the following \( pb \):

\[
\hat{pb}^m_n(s_n) = \begin{cases} 
\frac{x_{\{M\} \cdot (n-1)+m}}{a_{mn}} & \text{if } s_n = \tilde{s}_n \\
0 & \text{otherwise}
\end{cases} \tag{13}
\]

We shall show that these \( pb \)'s satisfy IC, NoB and AM, which in turn will prove, due to theorem 1, that they maintain a SPE in the VCG-GPTA that satisfies NoB.

Let us prove IC. For every principal \( \forall m \in M \) and for every outcome \( \forall s \in S \), we have due to condition (10) when multiplying the row in matrix \( A \) that corresponds to \( (s,m) \) by the vector \( x \), \( \sum_{m' \in M \setminus m} \sum_{n: s_n \neq \tilde{s}_n} x_{M \cdot (n-1)+m'} \geq a_{ms} = G^m(s) - G^m(\tilde{s}) \). Due to equation (13):

\[
\sum_{m' \in M \setminus m} \sum_{n: s_n \neq \tilde{s}_n} \hat{pb}^m_n(\tilde{s}_n) \geq G^m(s) - G^m(\tilde{s}) + \sum_{m' \in M \setminus m} \sum_{n: s_n = \tilde{s}_n} \hat{pb}^m_n(\tilde{s}_n).
\]

Therefore, by adding \( \sum_{m' \in M \setminus m} \sum_{n: s_n = \tilde{s}_n} \hat{pb}^m_n(\tilde{s}_n) \) to both sides of the equation, we attain:

\[
\sum_{m' \in M \setminus m} \sum_{n: s_n \neq \tilde{s}_n} \hat{pb}^m_n(\tilde{s}_n) + \sum_{m' \in M \setminus m} \sum_{n: s_n = \tilde{s}_n} \hat{pb}^m_n(\tilde{s}_n) \geq G^m(s) - G^m(\tilde{s}) + \sum_{m' \in M \setminus m} \sum_{n: s_n \neq \tilde{s}_n} \hat{pb}^m_n(\tilde{s}_n) + \sum_{m' \in M \setminus m} \sum_{n: s_n = \tilde{s}_n} \hat{pb}^m_n(\tilde{s}_n)
\]

and so, we have IC.

Let us prove NoB. Due to condition (11), \( \forall m \in M : \sum_{n \in N} x_{M \cdot (n-1)+m} = G^m(\tilde{s}) \). Due to equation (13), we have \( \forall m \in M : \sum_{n \in N} \hat{pb}^m_n(\tilde{s}) = \sum_{n \in N} x_{M \cdot (n-1)+m} = G^m(\tilde{s}) \), hence, we have NoB.

Due to condition (12) and equation (13) we see that \( \forall n \in N : s_n \neq \tilde{s}_n \) the principal bids are equal to zero, while all principal bids for \( \tilde{s} \) are non-negative, therefore we have AM.

Now we will show the other direction. That is, if a SPE \((\tilde{s}, \hat{pb})\) exists then there exists a vector \( x \in \mathbb{R}^{N \times M} \) that satisfies conditions (10), (11) and (12).

Consider a SPE with NoB \((\tilde{s}, \hat{pb})\). Due to theorem 1, we have AM and IC. We will construct another SPE with NoB \((\tilde{s}, \hat{pb})\), such that \( \tilde{s} = \tilde{s} \), and,

\[
\forall m \in M, \forall n \in N : \hat{pb}^m_n(s_n) = \begin{cases} 
\hat{pb}^m_n(s_n) + \frac{G^m(s) - \sum_{n \in N} \hat{pb}^m_n(s_n)}{\tilde{s}_n} & \text{if } s_n = \tilde{s}_n \\
0 & \text{otherwise}
\end{cases}
\]

This pair \((\tilde{s}, \hat{pb})\) still satisfies AM, IC and NoB (because IC, AM and NoB inequalities are maintained). And so, by denoting a vector \( \tilde{x} \in \mathbb{R}^{N \times M} : \tilde{x}_{M \cdot (n-1)+m} = \hat{pb}^m_n(\tilde{s}) \), we can see that \( \tilde{x} \) satisfies conditions (10) (because of IC), condition (11) (because of NoB) and condition (12) (because of AM).

Lemma 5. There exists a vector \( x \in \mathbb{R}^{N \times M} \) which satisfies conditions (10), (11) and (12), iff the VCG-GPTA is UB.

Proof. Initially assume there does not exist a vector \( x \in \mathbb{R}^{N \times M} \) that satisfies conditions (10), (11) and (12). We shall use the following Farkas lemma to prove that the game is not UB.

Lemma. The Farkas Lemma [Based on the version in Mangasarian (1969)]: A solution \( x \in \mathbb{R}^{N \times M} \) for conditions (10), (11) and (12), exists if and only if there is no solution in the variables
\((\omega^m(s)_{m \in M, s \in S}, (v_m)_{m \in M})\) (where \(\omega\) is a vector of length of \(|S| \cdot |M|\) and \(v\) is a vector of length \(|M|\))\(^{14}\) for the following system (14)-(16):

\[
\omega \cdot A + v \cdot B = \vec{0}; \quad (14)
\]

\[
\omega^m(s) \geq 0 \quad \forall m \in M, \forall s \in S; \quad (15)
\]

and

\[
\omega \cdot a + v \cdot b > 0. \quad (16)
\]

The system (14) may be rewritten as:

\[
\sum_{\{m' \in M \setminus j\}} \sum_{\{s: s_n \neq \hat{s}_n\}} \omega^m(s) + v_j = 0 \quad \forall j \in M, \forall n \in N; \quad (17)
\]

And therefore,

\[
\forall m \in M, \forall n \in N, \quad \sum_{\{m' \in M \setminus m\}} \sum_{\{s: s_n \neq \hat{s}_n\}} \omega^m(s) = -v_m \quad (18)
\]

Condition (16) can be rewritten (using equation (18)) as:

\[
\forall n \in N: \sum_{m \in M} \sum_{s \in S} \omega^m(s) [G^m(s) - G^m(\hat{s})] - \sum_{m \in M} \sum_{m' \in M \setminus m} \sum_{\{s: s_n \neq \hat{s}_n\}} \omega^m(s_n) \cdot G^m(\hat{s}) > 0 \quad (19)
\]

Due to definition 3, let us substitute: \(\forall n \in N: \sum_{s: s_n \neq \hat{s}_n} \omega^m(s) = 1\). Therefore, inequality (19) can be re-written as

\[
\sum_{m \in M} \sum_{s \in S} \omega^m(s) [G^m(s) - G^m(\hat{s})] - \sum_{m \in M} \sum_{m' \in M \setminus m} G^m(\hat{s}) > 0.
\]

That is, \(\sum_{m \in M} \left[\sum_{s \in S} \omega^m(s) [G^m(\hat{s}) - G^m(s)] + \sum_{m' \in M \setminus m} G^m(\hat{s})\right] < 0\), which is a contradiction to UB.

The proof in the opposite direction is similar. \(\square\)

Theorem 3 is a direct consequence of lemma 4 and lemma 5. \(\square\)

Let us observe tables 1 and 2. Table 1 has the same SPE with NoB in the VCG-GPTA, as in the FP-GPTA. On the other hand, we see that table 2 has a SPE with NoB in the VCG-GPTA, but not in the FP-GPTA. This can be seen by the UB condition, which is similar to the balanced condition for the FP-GPTA (see definition 4). Whereas, UB has an additional positive factor \(\sum_{m \in M} \sum_{m' \in M \setminus m} G^m(\hat{s})\), which enables certain games to attain a SPE with NoB in the VCG-GPTA, while in the FP-GPTA, a SPE (or WT SPE) does not necessarily exist\(^{15}\).

\(^{14}\)Formally speaking, if we denote the \(\omega(s_m)\) as element as \(\omega(k)\), then: \(s = (k \text{ div } |M|) + 1\), and \(m = (k \text{ mod } |M|) + 1\).

\(^{15}\)See Prat and Rustichini (2003) - theorem 2.
5 VCG Contracts Improve Existence of Efficient Pure Equilibria

In this section, we show that VCG contracts expand the set of efficient equilibria relative to the set of “weakly truthful” (WT) equilibria – recall the discussion prior to theorem 3 and additionally note that all WT equilibria are efficient (as shown by Prat and Rustichini (2003)). We restrict attention to games in which $|M| \geq 2$, $|N| \geq 2$, and $\forall n \in N : |S_n| \geq 2$ and define:\[16\]

**Definition 6.**

$WT_{FP}(M, N, S) = \{G \in \mathbb{R}^{M \times S} | \text{The FP-GPTA that corresponds to } (M, N, S, G) \text{ has a WT SPE} \}$,

$SPE_{e_{VCG}}(M, N, S) = \{G \in \mathbb{R}^{M \times S} | \text{The VCG-GPTA that corresponds to } (M, N, S, G) \text{ has an efficient SPE that satisfies NoB} \}$

**Theorem 4.** The set of VCG-GPTA with an efficient SPE that satisfies NoB strictly contains the set of FP-GPTA with a WT SPE. Moreover, $SPE_{e_{VCG}}(M, N, S) \setminus WT_{FP}(M, N, S)$ has positive measure in $\mathbb{R}^{M \times S}$.

**Proof.** We say that a game is strictly UB with respect to $\hat{s}$ if the inequality in definition 5 is strict. We show that an extension (detailed below) of the game $\tilde{gr}$ in table 5 is strictly UB with respect to the efficient $\hat{s} \in S$, but is not balanced with respect to $\hat{s}$. As a result, this extended game belongs to $SPE_{e_{VCG}} \setminus WT_{FP}$.

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>5 , 0 , 0 , ... , 0</td>
<td>0 , 4 , 0 , ... , 0</td>
</tr>
<tr>
<td>B</td>
<td>0 , 4 , 0 , ... , 0</td>
<td>3 , 3 , 0 , ... , 0</td>
</tr>
</tbody>
</table>

Table 5: $\tilde{gr} \in \mathbb{R}^{M \times S}$

For general $M, N, S$ we extend this $\tilde{gr}$ as follows. First, designate a specific action $\tilde{s}_n \in S_n$ for every agent $n \geq 3$. Second, let $G_m(s)$ (for any $m \in M$, $s \in S$) be:

$G^1(T, L, \tilde{s}_3, ..., \tilde{s}_N) = 5$

$G^1(B, R, \tilde{s}_3, ..., \tilde{s}_N) = 3$

$G^2(T, R, \tilde{s}_3, ..., \tilde{s}_N) = 4$

$G^2(B, L, \tilde{s}_3, ..., \tilde{s}_N) = 4$

$G^2(B, R, \tilde{s}_3, ..., \tilde{s}_N) = 3$

Otherwise: $G^m(s) = 0$

We show that the game $\tilde{gr}$ is strictly UB with respect to $\hat{s} = BR$. Note that we may ignore all cells in $\tilde{gr}$ apart from $\{TL, TR, BL, BR\}$ and even in these cells we may ignore all principals besides 1 and 2. Therefore the UB condition can be written as:

$$\forall \omega \in UBW : \sum_{m \in M} \left[ \sum_{s \in S \setminus \hat{s}} \omega^m(s)[\tilde{gr}^m(\hat{s}) - \tilde{gr}^m(s)] + \sum_{m' \in M \setminus \{m\}} \tilde{gr}^{m'}(\hat{s}) \right] =$$

\[16\] Obviously, we can ignore agents with $|S_n| = 1$, and the case of $M = 1$ is straightforward. The case of $|N| = 1$ (common agency) was handled by Bernheim and Whinston (1986). They showed that an efficient SPE always exists in the FP-GPTA, and therefore, due to theorem 2, in the VCG-GPTA.
\[ \omega^1(TL)[3-5] + \omega^1(TR)[3-0] + \omega^1(BL)[3-0] + 3 + \omega^2(TL)[3-0] + \omega^2(TR)[3-4] + \omega^2(BL)[3-4] + 3 \]  

Due to definition 3 we have \( 0 \leq \omega^m(s) \leq 1 \) for every \( m \) and \( s \) and therefore the value of this equation is at least 2. Hence, \( \hat{gr} \) is strictly UB with respect to \( \hat{s} = BR \). Furthermore, it is clear that all games in the “close neighbourhood” of \( \hat{gr} \) are UB with respect to \( BR \in S \) since if we perturbate the values of \( \hat{gr} \) by a small number the UB requirement will still be positive. Specifically, one can verify that all \( \tilde{g} \in \mathbb{R}^{M \times S} \) such that

\[ \forall m \in M, s \in S, \tilde{g}^m(s) + 0.1 \geq \hat{g}^m(s) \geq \tilde{g}^m(s) - 0.1 \]

satisfy UB.

In order to show that \( \tilde{gr} \) is not balanced with respect to \( \hat{s} = BR \), we show that there exists an \( \omega \in BW \) which violates definition 4, specifically \( \forall m \geq 3 : \omega^m(TL) = \omega^2(BL) = \omega^2(TR) = \omega^1(TL) = 1 \), and otherwise \( \omega(\cdot) = 0 \). We show that for these weights,

\[ \sum_{m \in M} \sum_{s \in S} \omega^m(s)(G^m(\hat{s}) - G^m(s)) = -4 < 0 \]

since

\[ \omega^1(TL)[3-5] + \omega^1(TR)[3-0] + \omega^1(BL)[3-0] + \omega^2(TL)[3-0] + \omega^2(TR)[3-4] + \omega^2(BL)[3-4] = 22 \]

\[ 1 \cdot [3-5] + 0 \cdot [3-0] + 0 \cdot [3-0] + 0 \cdot [3-0] + 1 \cdot [3-4] + 1 \cdot [3-4] = -4. \]

Therefore, \( \hat{gr} \) is not balanced with respect to \( BR \) (note that \( BR \) is the unique efficient outcome and therefore there cannot be any WT SPE with any other outcome \( s \in S \)). Therefore \( \tilde{gr} \notin WT^{FP}(M, N, S) \). Furthermore, once again, one can verify that for these weights, all games in the “close neighbourhood” of \( \hat{gr} \) are not balanced. In fact, this holds for the same neighbourhood \( \tilde{g} \in \mathbb{R}^{M \times S} \) such that \( \forall m \in M, s \in S, \tilde{g}^m(s) + 0.1 \geq \hat{g}^m(s) \geq \tilde{g}^m(s) - 0.1 \) that was previously defined in equation (21).

We have shown a subset of games which is UB with respect to \( BR \in S \) but is not balanced with respect to \( BR \). Since this subset has positive measure, the theorem follows.

**Remark.** When \( |S_n| = 2 \) for all \( n \in N \) there is no loss of generality in considering weakly truthful equilibria instead of efficient SPE in FP-GPTA since Prat and Rustichini (2003) showed that all SPE in such a FP-GPTA must be weakly truthful. In contrast, when considering games with \( |S_n| > 2 \), Prat and Rustichini (2003) provided an example in which there exist efficient equilibria in the FP-GPTA which are not WT, and left the characterization of such equilibria as an open question. For example, the game in table 6 has an efficient SPE \((\hat{s}, \hat{pb})\) in the FP-GPTA which is not WT: \( \hat{pb}^3_2(R) = \hat{pb}^3_2(B) = 3, \hat{pb}^1_1(MV) = \hat{pb}^2_1(MV) = \hat{pb}^1_1(MH) = \hat{pb}^2_1(MH) = 1.5 \), and otherwise \( \hat{pb} = 0 \), with \( \hat{s} = (MV, MH) \).

This example also shows that the above proof to theorem 4 is not correct if we change the weakly truthful requirement in definition 6 to a general efficiency requirement.

### 6 A Price of Anarchy Comparison

Since VCG-GPTA significantly expands the set of subgame perfect equilibria of FP-GPTA, one might suspect that extremely inefficient additional equilibria are added. In this section we show
that this is not the case and that in fact the worst-case inefficiency level does not increase. We do this via the formal framework of Price of Anarchy (PoA) (see, e.g., Roughgarden (2005)).

**Definition 7.** Given a principal agent setting $G$ let $POAVCG(G)$ be as follows. Let $s* \in \text{argmax}_{s \in S} G(s)$ be an efficient outcome of $G$. Then,

$$POAVCG(G) = \inf_{\hat{p}b \in PB} \frac{G(s(\hat{p}b))}{G(s*)}.$$ \(s.t. (s(\hat{p}b), \hat{p}b) \text{ is a SPE with NoB in the VCG-GPTA of } G\)

Let $G(M)$ be the class of all principal-agent settings $G$ with $M$ principals and define its Price of Anarchy in a VCG-GPTA, as follows:

$$POAVCG(M) = \inf_{G \in G(M)} POAVCG(G).$$

The Price of Anarchy in an FP-GPTA, $POA_{FP}(M)$, is defined in a similar manner.

**Theorem 5.** For any $M \geq 2$, $\frac{1}{M} \geq POA_{FP}(M) \geq \frac{1}{M+1}$ and $\frac{1}{M} \geq POAVCG(M) \geq \frac{1}{M+1}$.

**Proof.** The proof for the following lemma is based on the smoothness technique of Roughgarden (2015).

**Lemma 6.** $POA_{FP}(M) \geq \frac{1}{M+1}$ and $POAVCG(M) \geq \frac{1}{M+1}$.

**Proof.** We shall begin by proving the case for VCG-GPTA. Let $(s(\hat{p}b), \hat{p}b)$ be a SPE, and let $s*$ be an efficient outcome. Consider the following principal bids:

$$\forall n \in N : bp_{n}^{*m}(s_n) = \begin{cases} \text{Max}_{x \in S_n} \sum_{m' \in M \setminus m} \hat{p}_{m'}^n(x) - \sum_{m' \in M \setminus m} \hat{p}_{m'}^n (s_n^*) & \text{if } s_n = s_n^* \\ 0 & \text{Otherwise} \end{cases} \quad (23)$$

Due to AM, if the principal bids are $(pb^{*m}, \hat{p}b^{-m})$ then all agents $n \in N$ will play $s_n^*$ and the outcome will be $s^*$.

Principal $m$’s payoff function can be bounded above and below as follows, where the first inequality below follows from equation (2) and since $t_{m,n}(\cdot) \geq 0$ (see lemma 3), and the second inequality below follows from the definition of SPE:

$$G^m(s(\hat{p}b)) \geq u^m[\hat{p}b] \geq u^m[(pb^{*m}, \hat{p}b^{-m})] = G^m(s(pb^{*m}, \hat{p}b^{-m})) - \sum_{n \in N} t_{m,n}^{VCG}(s_n(pb^{*m}, \hat{p}b^{-m}), (pb^{*m}, \hat{p}b^{-m})) = G^m(s^*) - \sum_{n \in N} t_{m,n}^{VCG}(s^*, (pb^{*m}, \hat{p}b^{-m})) \quad (24)$$
Due to the definition of $t(\cdot)$ (see equation (3)):

$$
G^m(s^*) - \sum_{n \in N} \max_{x \in S_n} \left\{ \sum_{m' \in M \setminus m} \hat{p}^m_n(x) \right\} + \sum_{n \in N} \left\{ \sum_{m' \in M \setminus m} \hat{p}^m_n(s^*) \right\} \geq 0 \quad \forall m \in M
$$

and for even

$$
G^m(s^*) - \sum_{n \in N} \left\{ \max_{x \in S_n} \sum_{m' \in M \setminus m} \hat{p}^m_n(x) \right\} \geq G^m(s^*) - \sum_{n \in N} \left\{ \sum_{m' \in M} \hat{p}^m_n(s_n(\hat{p})) \right\}
$$

The following equality follows from AM:

$$
\sum \text{to equation (25), which is based on the principal bids in equation (23)}.
$$

The proof for the FP-GPTA is similar to the proof of the VCG-GPTA, save the shift from equation (24) to equation (25), which is based on the principal bids in equation (23).

**Lemma 7.** \( \frac{1}{M} \geq \text{PoA}_{FP}(M) \) and \( \frac{1}{M} \geq \text{PoA}_{VCG}(M) \).

**Proof.** We need to show one principal-agent setting \( G \) with \( M \) principals such that \( \frac{1}{M} \geq \text{PoA}_{FP}(G) \) and \( \frac{1}{M} \geq \text{PoA}_{VCG}(G) \). This is given in table 7.\(^{17}\) Consider the SPE with outcome \( s(\hat{p}) = TL \) in the FP-GPTA and in the VCG-GPTA: where for odd \( m_{\text{odd}} \in M \): \( \hat{p}^{m_{\text{odd}}}_1(T) = \hat{p}^{m_{\text{odd}}}_2(MV) = 1 \), and for even \( m_{\text{even}} \in M \): \( \hat{p}^{m_{\text{even}}}_1(MH) = \hat{p}^{m_{\text{even}}}_2(L) = 1 \), otherwise, \( \hat{p}^{(i)}_1(\cdot) = 0 \). Notice the efficient outcome is \( s^* = BR \), so \( \text{PoA}_{VCG}(M) = \text{PoA}_{FP}(M) = \frac{G(TL)}{G(BR)} = \frac{1}{M} \).

\( \square \)

\( ^{17} \)The example is for an even \( M \). The case that \( M \) is odd is a simple adaptation: in cell (MH, L), instead of “2” in the even spaces have \( G^{m_{\text{even}}}(MH, L) = \frac{M}{(M \text{ div } 2) + 1} \), in cell (T, MV), instead of “2” in the odd spaces have \( G^{m_{\text{odd}}}(T, MV) = \frac{M}{(M \text{ div } 2) + 1} \), in cell (T, L), instead of “1” in the even spaces have \( G^{m_{\text{even}}}(TL) = \frac{M}{(M \text{ div } 2) + 1} \), and in cell (T, L), instead of “1” in the odd spaces have \( G^{m_{\text{odd}}}(TL) = \frac{M}{(M \text{ div } 2) + 1} \).
For intuition regarding theorem 5, consider table 7 with $N = 2$ and $\forall n \in N : |S_n| = 3$. The upper left 4 outcomes can be described as a variation of the “prisoner’s dilemma”; it is in all players best interest to try to deviate from the outcome $\langle TL \rangle$ to the outcomes $\langle T, MV \rangle$ or $\langle MH, L \rangle$. This joint aspiration to deviate prevents the possibility of reaching BR, both in the VCG-GPTA and in the FP-GPTA.

7 Conclusions

In this paper, we extend the principal-agent setting $\mathcal{G}$, with a first price payment rule denoted FP-GPTA (see Prat and Rustichini (2003)) to a second price payment rule denoted VCG-GPTA, where agent’s payments are based on an extension of the VCG mechanism.

We characterize subgame perfect equilibria in the VCG-GPTA, and prove that a VCG-GPTA has a SPE if and only if it satisfies AM and IC. In addition, we present a necessary and sufficient condition for existence of a SPE with NoB in the VCG-GPTA, based on UB weights. We believe there is a connection between UB weights and fractionally sub-additive valuations described in Fu et al. (2017). We leave this study open to future research.

Furthermore, we show that given the set of principal-agent settings, the set of induced VCG-GPTA with an efficient SPE contains the set of induced FP-GPTA with an efficient SPE. Moreover, the symmetric difference between the sets of induced VCG-GPTA with an efficient SPE and the set of induced FP-GPTA games with a WT SPE has positive measure. The precise difference of measure between these two sets is still unknown, and we leave this open for future work.

We proceed to show that although the set of settings $\mathcal{G}$ with a SPE in the corresponding VCG-GPTA is significantly larger than the set of settings $\mathcal{G}$ with a SPE in the corresponding FP-GPTA, the PoA in VCG-GPTA and in the FP-GPTA are of the same order of magnitude.

Two additional topics are left open for future work:

1. Analyzing the SPE in cases where principals or agents may choose the payment rule (FP-GPTA, VCG-GPTA or a different payment rule).

2. Analyzing the VCG-GPTA model, with agent preferences over actions.

References


