Second Order Secret Love*

Xinyang Wang†

November 5, 2019

Abstract

Sometimes, when choosing among strategies that maximize their own payoff, agents choose the strategy that is best for their friends. To study this phenomenon, we study games with lexicographic externalities. The novel ingredient is a set of players preference lists, which represent the order in which players care about the others. The collection of preference lists maps a base game to a game with a lexicographic externality, in which payoff functions are vector-valued and agents compare outcomes according to the lexicographic order. We prove that, for any given preference lists, if the base game has discrete outcomes and upper semi-continuous payoff functions, a Nash equilibrium always exists. In addition, we discuss the efficiency of equilibria in a model with public bads and the epsilon-variations of our formalization.

1 Introduction

In this paper, we study two questions. The first question is an efficiency question: *"Will the world become better if everyone cares about someone else for even just a little bit?"* However, in order to answer this question, we need to first ask an existence question: *"Is there a world in which everyone cares about someone else for just a little bit?"*

*I am grateful to Nick Baigent, Truman Bewley, Mira Frick, John Geanakoplos, Ludovic A. Julien, Xiangliang Li, Jacqueline Morgan, and participants at conferences and seminars for inspiring discussions. I am responsible for all remaining mistakes and typos. First version: March 2018.

†Department of Economics, Yale University, email: xinyang.wang@yale.edu.

A more rigorous way to formulate the second question is whether or not there is an economic model in which every agent cares about someone else for a little bit and an equilibrium of this model exists. In this paper, we define the meaning of "caring about the others for a little bit" and answer these two questions according to our definition.

In the literature, I found two major methods to model how agents care about each other. The first method is qualitative: when agents make decisions, they consider the contamination effects of their behavior to the others. For instance, a person, who knows the negative impact of his own action is negligible in a very large environment, does not litter on a public beach because he thinks this public beach will be destroyed if many others do the same thing. This idea dated back to Immanuel Kant. Roemer [2017] formalized Kant's idea and provided a game model in which all players assume all others will act similarly after they choose their own actions.

The second method is quantitative: agents have additively interdependent payoff functions. For instance, if agent 1 cares about his own payoff u_1 and agent 2's payoff u_2, his happiness level is determined by $u_1 + \varepsilon u_2$, where the number $\varepsilon > 0$ is usually regarded as the degree of care. Naturally, to model that agents care about the others for just a little bit, we take the ε goes to zero. We refer to De Marco and Morgan [2008a,b] for further details. While this additive model is usually easy to implement, there are two problems. Firstly, it is hard to interpret the coefficient ε formally: we note preferences are invariant under positive affine transformations on the payoff functions. However, if we use $\tilde{u_2} = 10u_2$ to denote the payoff of agent 2, the happiness level of agent 1 will become $u_1 + 0.1\varepsilon\tilde{u_2}$. That is, the coefficient ε is not invariant under positive affine transformations on payoff functions. Secondly, when ε goes to 0, the limit preference of a sequence of additively interdependent utility functions is usually regarded as a lexicographic preference. Unfortunately, this claim is not universally true. For instance, when ε goes to 0, the limit of a sequence of equilibria in models with additive utility functions may not be an equilibrium in a model with lexicographic preference.

In this paper, we provide a third method to model how agents care about each other. In our model, agents care about the others in a predetermined lexicographic order. For example, agent 1 cares about himself in the first order and agent 2 in the second order. Therefore, agent 1 will first compute the set of actions maximizing his own payoff. Then, within this set, he chooses an action maximizing agent 2's payoff.

Formally, we propose a game with lexicographic externalities. We call such game lexicographic game. In a lexicographic game, each agent is endowed with a preference list, which is an ordered set of names, and he cares about the others in an order specified by his preference list. A lexicographic game is generated by a base game, which is a classical game with a set of agents who are described by their action spaces and utility functions over the action profiles, and a set of preference lists. Naturally, in a lexicographic game, agents have vector-valued happiness levels, and they compare their happiness levels according to the lexicographic order. We note that the vector-valued happiness levels appear naturally from the lexicographic externalities we study in this paper. For more detailed discussions on the vector-valued utility representations, we refer to Ok [2002].

This formalization of lexicographic games resolves the two problems we mentioned in games with additively interdependent utility functions. Firstly, it is easy to see that the preference represented by vector-valued happiness levels is invariant under positive affine transformations on utility functions. Secondly, we avoid the problem that the limit equilibrium is not a equilibrium of the limit game by studying the limit game directly. In addition, lexicographic game provides a less severe departure from the pure selfishness case, as there is no sacrifice between oneself and others' well-beings.

An explanation of the title "second order secret love" follows. The term "second order" corresponds to the fact that agents compare their happiness levels according to the lexicographic order. The term "love" corresponds to the fact that agents maximize, rather than minimize, the others' payoffs. The term "secret" corresponds to the fact that agents maximize the others payoffs, rather than their happiness levels. For instance, we consider a game with three agents: a father, a son and a girl the son likes. In our setting, if the father's preference list only consists of himself and his son, the father will not take the girl's payoff into account when he acts, even though he can make his son happier by making the girl his son likes happier. Thus, agents' preference lists can be regarded as their private information.

In a lexicographic game, the central solution concept we study is the lexicographic equilibrium, which is defined to be the Nash equilibrium of a game with lexicographic externalities. To compare, we use the term selfish equilibrium to denote the Nash equilibrium of the base game. The following observations on the relationship between lexicographic equilibrium and selfish equilibrium are straightforward. Firstly, if all agents' preference lists consist of only their own names, lexicographic equilibrium coincides with selfish equilibrium. Moreover, if all agents care about themselves in the first order and care about someone else in the second order, lexicographic equilibirum is a refinement of selfish equilibrium. In contrast, if some agent does not care about himself in the first order, lexicographic equilibrium may not be a refinement of selfish equilibrium. It is worth nothing that lexicographic equilibrium is conceptually related to Berge's equilibrium, in which every agent aims to maximize payoffs of all other agents. We refer to Berge [1957], Zhukovskii and Chikrii [1994], Abalo and Kostreva [2004] for further details about Berge's equilibrium.

Having proposed a lexicographic game in which agents care about each other for just a little bit, we wish to answer the existence question we proposed at the beginning of the introduction. Unfortunately, lexicographic equilibrium may fail to exist even when the base game is a finite game. We give a counter example in Example 3. Nevertheless, the counter example suggests that the nonexistence might be due to the fact that the set of best responses is too small. In order words, it is very rare that agents are indifferent between some actions. Consequently, agents do not have a lot of chances to choose an action to maximize their friend's payoff.

In order to ensure the existence of a lexicographic equilibrium, we propose two ways to enlarge the set of best responses. Both ways are inspired by the observation that, when agents are less sensitive (e.g. they cannot distinguish the difference between spending 1.999999 dollars and 2 dollars), their best response sets are larger. The first way is to assume that agents have discrete outcomes in the base game. Since many objects in the world are not infinitely divisible, the model with discrete outcome actually provides an exact formulation for some real world problems. The second way is to assume that agents are indifferent around their optimizers. For instance, an agent who can at most obtain 100 dollars is satisfied and starts to care about his friend as long as he has obtained 99 dollars.

On the technique level, the first way to enlarge the set of best response leads us to a game with discontinuous payoffs. Starting from Dasgupta and Maskin [1986a,b], there are many works on games with discontinuous payoff functions, including Reny [1999], Bagh and Jofre [2006], Barelli and Soza [2009], Bich [2009], Bich and Laraki [2012], Carmona [2009, 2011a,b, 2012], McLennan et al. [2011], Barelli and Meneghel [2013], Prokopovych [2011, 2013]. In this paper, we followed the path led by Reny [1999], McLennan et al. [2011] and Barelli and Meneghel [2013]. In particular, by extending their idea of a continuously secured game to a game with vector-valued payoff functions, we prove that any lexicographic game with discrete outcomes, upper semi-continuous and own strategy quasi-concave payoff functions has a lexicographic equilibrium.

The second way to enlarge the set of best response leads us to the ε-variations of the lexicographic equilibrium. By proving the ε-best response correspondences are upper hemi-continuous, we prove that for any finite base game and any preference lists, the generated lexicographic game has a lexicographic equilibrium.

Lastly, we answer the efficiency question we proposed at the beginning of the introduction. Almost directly from our definition, lexicographic equilibrium will provide a Pareto improvement on the social welfare. In particular, in a game with public bads, we proved that both lexicographic equilibrium and its ε-variations, when ε is small, Pareto dominate the selfish equilibrium.

The rest of this paper is organized as follows. In section 2, we discuss three motivating examples. In section 3, we state the model. In section 4, we study the existence of lexicographic equilibrium and its variations. In section 5, we analyze the social welfare in a game with public bads. In section 6, we conclude.

2 Motivating examples

In this section, we give three examples with two players. Players care about themselves in the first order and each other in the second order. In contrast, we say an action profile is lexicographic equilibrium if no player can deviate to either increase his own payoff or increase the other player's payoff without reducing his own payoff. We say an action profile is a selfish equilibrium if no player can deviate to increase his own payoff.

2.1 Example 1

We start with battle of sexes:

	L	R
T	2,1	0,0
B	0,0	1,2

We note $(\frac{2}{3}T + \frac{1}{3}B, \frac{1}{3}L + \frac{2}{3}R)$ is a mixed strategy selfish (Nash) equilibrium. At equilibrium, both players have equilibrium payoffs $\frac{2}{3}$. However, if player 1 plays B, his own payoff is $\frac{2}{3}$ and but he can increase player 2's payoff to $\frac{4}{3}$. Therefore, this mixed Nash equilibrium is not a lexicographic equilibrium when there is externalities. In contrast, both pure strategy selfish (Nash) equilibria are lexicographic equilibria.

2.2 Example 2

Secondly, a pure strategy selfish equilibrium may not be a lexicographic equilibrium.

	L	R
T	2,2	0,0
B	2,1	1,1

In this game, (B, R) is a selfish equilibrium. However, it is not a lexicographic equilibrium as player 2 can play L to maintain his own payoff and increase the payoff of player 1.

2.3 Example 3

Now, we analyze a modified matching penny game.

	L	R
T	-1,1	1,-1
B	3,-1	-1,1

Since when everyone cares about themselves in the first order, lexicographic equilibria is a subset of selfish equilibria. Thus, to prove there is no lexicographic equilibrium in this game, we only need to show no selfish equilibrium is a lexicographic equilibrium. Firstly, we note there is a unique selfish equilibrium $(\frac{1}{2}T + \frac{1}{2}B, \frac{1}{3}L + \frac{2}{3}R)$. However, fixing player 2's action $\frac{1}{3}L + \frac{2}{3}R$, if player 1 plays $\frac{1}{2}T + \frac{1}{2}B$, $(u_1, u_2) = (\frac{1}{3}, 0)$. If player 1 plays B, $(u_1, u_2) = (\frac{1}{3}, \frac{1}{3}) \succ (\frac{1}{3}, 0)$. Therefore, the unique selfish equilibrium is not an lexicographic equilibrium. Consequently, no lexicographic equilibrium exists in this game.

Next, we prove that the best response correspondence is not upper hemi-continuous in this game. Firstly, we draw player 1's best response correspondence with respect to himself and for player 2.

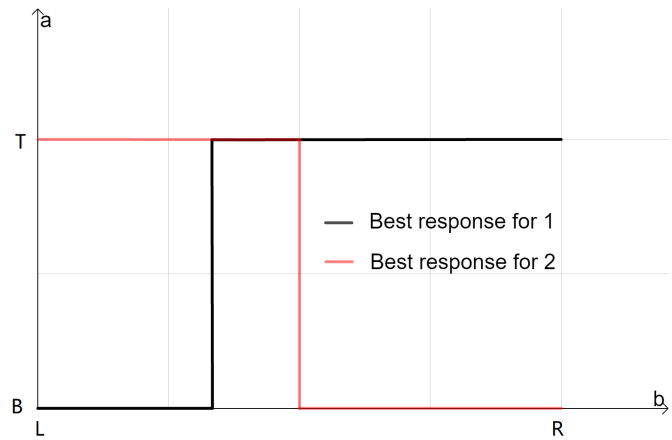

Figure 1: Best response correspondence of player 1 with respect to players 1, 2

The black path denotes player 1's best responses for himself: when player 2 plays L, it is the best for himself to play B; when player 2 plays R, it is the best for himself to play T; and there is one point that he is indifferent. The red path is less standard - it denotes player 1's best responses for player 2: when player 2 plays L, it is the best for player 2 if he plays T; when player 2 plays R, it is the best for player 2 if he plays B; and when player 2 plays $\frac{1}{2}L + \frac{1}{2}R$, player 1's action has no effect on player 2's payoff.

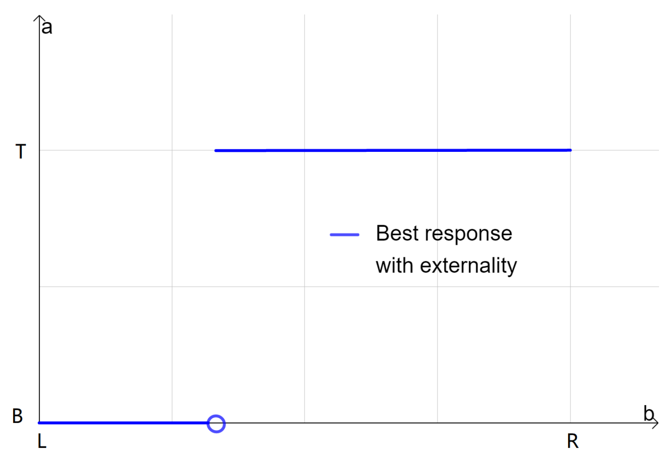

Figure 2: Best response correspondence of player 1

Since player 1 chooses the best responses for player 2 among the best responses for himself. Therefore, the best response correspondence is drawn in blue in Figure 2. And we can see from Figure 2 that player 1's best response correspondence no longer has a closed graph, thus is not upper hemi-continuous.

We observe the non-existence of a lexicographic equilibrium comes from the fact that best response correspondences are not upper hemi-continuous, which is a result of the fact that the set of best responses are not large enough. In this example, player 1's best response with respect to himself is almost everywhere a singleton. Therefore, in order to ensure the existence of a lexicographic equilibrium, we enlarge the set of best responses. We propose two ways. Firstly, we consider utility functions with discrete outcomes. This idea leads us to the field of games with discontinuous payoff functions. Secondly, we consider the ε-variation of the maximization problem.

3 Lexicographic game

We first define a base game, in which there is no externalities. Then, we introduce agents' preference lists which map a base game to a game with lexicographic externalities.

A *base game* is denoted by a tuple $(I, (X_i, u_i)_{i\in I})$. In this tuple, I is a finite set of players. X_i is player i's action set. u_i is player i's *private payoff function*, defined by a real-valued function on the set of action profiles $X = \times_{i\in I} X_i$. For our later convenience, we use the notion X_{-i} to denote the set $\times_{j\neq i} X_j$. A generic element in X is denoted by x and a generic element in X_{-i} is denoted by x_{-i}. We assume, for each $i \in I$, X_i is a non-empty, convex and compact subset of a Hausdorff locally convex topological vector space and u_i is a bounded function on X.

The novel ingredient in this paper is players' preference lists. A preference list is a list of names representing the order a player cares about the others. Formally, a *preference list* is an element in the set $\cup_{n=1}^{|I|} I^n$. Every agent $i \in I$ is endowed with a preference list $\mathcal{L}_i \in \cup_{n=1}^{|I|} I^n$. There are three canonical examples of the preference lists. Firstly, when $\mathcal{L}_i = (i)$, player i is *selfish* as he does not care about the others. Secondly, when $\mathcal{L}_i = (i, i_2)$, player i cares about himself in the first order, player i_2 in the second order and does not care about anyone else. In this case, we say player i_2 is player i's lover. Thirdly, when $\mathcal{L}_i = (i_2, i)$, player i cares about player i_2 in the first order, cares about himself in the second order, and does not care about anyone else.

More generally, if player i has a preference list $\mathcal{L}_i = (i_1, ..., i_n) \in I^n$, then player i has a preference list of size n and he cares about the private payoffs of the others in a lexicographic order specified by $\mathcal{L}_i$. That is, given all others' actions, player i firstly computes the subset of actions which maximizes player i_1's private payoff. Then, among this subset, player i computes the subset of actions which maximizes player i_2's private payoff. This process stops at step n when player i reaches the end of his preference list. Consequently, his best response is defined by the smallest subset of actions obtained at the end of this process. Intuitively, player i cares about player i_1 at the first place, player i_2 at the second place and so on. We say player i cares about player i_k in the k-th order, or player i_k is player i's k-th lover, for $k = 1, 2, ..., n$.

Moreover, as we do not assume the names in a preference list to be disjoint, without loss of generality, in the rest part of the paper, we will regard a preference list as an element in $I^{|I|}$. For instance, $\mathcal{L}_i = (i_1, i_2)$, a preference list of size 2, corresponds to an element $(i_1, i_2, ..., i_2)$ in $I^{|I|}$.

Given a base game $(I, (X_i, u_i)_{i\in I})$ and a set of preference lists $(\mathcal{L}_i)_{i\in I}$, we define a *lexicographic-game* [1] by the tuple $(I, (X_i, u_i, \mathcal{L}_i)_{i\in I})$. Conceptually, any lexicographic game is equivalent to a game $(I, (X_i, U_i)_{i\in I})$ with vector-valued payoff functions. For an agent i with a preference list $\mathcal{L}_i = (i_1, ..., i_{|I|})$, his happiness level U_i is defined by a map from action profiles X to the multi-dimensional ordered set $(\mathbb{R}^{|I|}, \succeq)$ such that

$$U_i(x) = (u_{i_1}(x), ..., u_{i_{|I|}}(x))$$

In order words, player i compares the vector-valued outcomes according to the lexicographic order $\succeq$ on $\mathbb{R}^{|I|}$.[2]

[1] In this paper, we use the prefix "lexicographic-" to highlight all objects with potential externalities of the lexicographic type.

[2] For any $x, y \in \mathbb{R}^{|I|}$, $x \succeq y$ if and only if either $x_1 > y_1$, or there is a $1 \leq n \leq |I|$ such taht $x_k = y_k$ for all $1 \leq k \leq n$ and $x_{n+1} > y_{n+1}$.

4 Lexicographic equilibrium and its variations

In this section, we define our solution concept, lexicographic equilibrium, and its variations. A lexicographic equilibrium is an action profile such that, given all the others' actions, every player plays the best response according his preference. In other word, a lexicographic equilibrium is a Nash equilibrium in the game with vector-valued payoff functions.

4.1 Lexicographic equilibrium

Given a lexicographic game $(I, (X_i, u_i)_{i \in I})$, an action profile $x^* \in X$ is an *lexicographic equilibrium* if, for any $i \in I$ whose preference list is $\mathcal{L}_i = (i_1, ..., i_{|I|})$, we have

$$(u_{i_1}(x^*), ..., u_{i_{|I|}}(x^*)) \succeq (u_{i_1}(x_i, x^*_{-i}), ..., u_{i_{|I|}}(x_i, x^*_{-i})), \forall x_i \in X_i$$

To compare with the no externality case, we define selfish equilibrium as follows. Given a lexicographic game $(I, (X_i, u_i)_{i \in I})$, an action profile $x^* \in X$ is an *selfish equilibrium* if it is a Nash equilibrium of the base game. i.e. for all $i \in I$,

$$u_i(x^*) \geq u_i(x_i, x^*_{-i}) \forall x_i \in X_i$$

It is straightforward to check that, when all players care about themselves in the first order, a lexicographic equilibrium is a refinement of a selfish equilibrium, or equivalently, a Nash equilibrium of the base game. In particular, if all agents only care about themselves, i.e. their profile lists only consist of their own names, a lexicographic equilibrium is a Nash equilibrium of the base game. On the contrary, when some player does not care about themselves in the first order, a lexicographic equilibrium is not necessarily a refinement of a selfish equilibrium.

To state the existence theorems, we give a vector-valued extension of "continuously-secure game" introduced in Barelli and Meneghel [2013].

A lexicographic game $(I, (X_i, u_i, \mathcal{L}_i)_{i \in I})$ is *L-continuous secure at* $x \in X$ if, for every $i \in I$, there is a vector $\alpha^{(i)}_x = (\alpha^{(i)}_{x,1}, ..., \alpha^{(i)}_{x,|I|}) \in \mathbb{R}^{|I|}$, an open neighborhood $V_x \subset X$ of x, and a non-empty valued closed correspondence $\varphi_x : V_x \rightrightarrows X$ such that,

1. for every $i \in I$ and every $y \in V_x$, $\varphi_{x,i}(y) \subset B_i(y, \alpha^{(i)}_{x,1}, ..., \alpha^{(i)}_{x,|I|})$
2. for every $y \in V_x$, there exists an $i \in I$ with $y_i \notin co(B_i(y, \alpha^{(i)}_{x,1}, ..., \alpha^{(i)}_{x,|I|}))$

where $B_i(y, \alpha^{(i)}_{x,1}, ..., \alpha^{(i)}_{x,|I|}) = \{x_i \in X_i : (u_{i_1}(x_i, y_{-i}), ..., u_{i_{|I|}}(x_i, y_{-i})) \succeq (\alpha^{(i)}_{x,1}, ..., \alpha^{(i)}_{x,|I|})\}$.

A lexicographic game $(I, (X_i, u_i, \mathcal{L}_i)_{i\in I})$ is *L-continuous secure* if it is L-continuous secure at every $x \in X$ that is not a lexicographic equilibrium.

The intuitive interpretation is similar to Barelli and Meneghel [2013]: around any non-equilibrium action profile, there is a semi-continuous profitable deviations. The main theorem is as follows:

Theorem 1. *An L-continuous secure lexicographic game has a lexicographic equilibrium.*

Proof. The proof is a modification of the proof in Barelli and Meneghel [2013].

We prove by contradiction and suppose there is no lexicographic equilibrium. By definition, for any $x \in X$, there is a tuple $((\alpha^{(i)}_x)_{i\in I}, V_x, \varphi_x)$ satisfying Condition 1 and Condition 2 of the L-continuous secure, where $\alpha^{(i)}_x \in \mathbb{R}^{|I|}$, V_x is an open neighborhood of x and $\varphi_x : V_x \rightrightarrows X$ is a closed correspondence with non-empty values.

Since X is regular, every $x \in X$ has a closed neighborhood $\tilde{V}_x \subset V_x$. By the compactness of X, X has a finite cover $(\tilde{V}_n)_{n=1,..,N}$. Define $((\alpha^{(i)}_n)_{i\in I}, \varphi_n)_{n=1,...,N}$ to be the corresponding tuple.

Now, for each $i \in I$, we define a function $\beta^{(i)} : X \to \mathbb{R}^{|I|}$ such that

$$\beta^{(i)}(x) = \max\left\{\left(\alpha^{(i)}_{n,1}, ..., \alpha^{(i)}_{n,|I|}\right) : x \in \tilde{V}_n\right\}$$

where the maximum is taken in the lexicographic order.

For any $x \in X$, let $U_x = (\cap_{\{n:x\in\tilde{V}_n\}} V_n) \cap (\cup_{\{n:x\notin\tilde{V}_n\}} \tilde{V}^c_n)$. Clearly, U_x is an open neighborhood of x and, for any $y \in U_x$, $y \notin \cup_{\{n:x\notin\tilde{V}_n\}} \tilde{V}_n$. Consequently, for any $y \in U_x$, $\beta^{(i)}(y) \preceq \beta^{(i)}(x)$ for all $i \in I$. We define a correspondence $\phi_x : U_x \rightrightarrows X$ by the product $\phi_x = \times_{i\in I}\phi_{x,i}$ where

$$\phi_{x,i}(y) = \varphi_{n,i}(y)$$

such that $(\alpha^{(i)}_{n,1}, ..., \alpha^{(i)}_{n,|I|}) = \beta^{(i)}(x)$. By definition of the correspondence φ_x, we have ϕ_x is closed.

Again, for each $x \in X$, there is a closed neighborhood $\tilde{U}_x \subset U_x$ and the cover $\{\tilde{U}_x\}_{x\in X}$ has a finite subcover $(\tilde{U}_\ell)_{\ell=1,2,...,L}$. Let ϕ_ℓ be the corresponding correspondence. Define $\Phi : X \rightrightarrows X$ by

$$\Phi(x) = co \bigcup_{\{\ell\in\{1,2,...,L\}:x\in\tilde{U}_\ell\}} \phi_\ell(x)$$

By definition, Φ is a nonempty, convex and compact-valued correspondence. By Theorem 17.27 in Aliprantis and Border [2007], it is upper hemi-continuous thus closed. Thus, by Kakutani-Fan-Glicksberg theorem (Corollary 17.55 in Aliprantis and Border [2007]), Φ has a fixed point.

On the other hand, for any $x \in X$, define $J = \{\ell \in \{1, 2, ..., L\} : x \in \tilde{U}_\ell\}$ and $J' = \{j \in \{1, 2, ..., N\} : x \in \tilde{V}_j\}$.

Firstly, we claim that $\phi_{\ell,i}(x) \subset B_i(x, \beta^{(i)}(x))$. To prove the claim, we fix $x \in X$. For any $\ell \in J$, by definition, there is some $y \in \tilde{U}_\ell$ such that $\phi_{\ell,i}(x) = \phi_{y,i}(x)$. By the definition of ϕ, there is an $n \in \{1, 2, ..., N\}$ such that, $\phi_{y,i}(x) = \varphi_{n,i}(x)$ where $\beta^{(i)}(y) = (\alpha^{(i)}_{n,1}, ..., \alpha^{(i)}_{n,|I|})$. Since in $x \in \tilde{U}_\ell \subset U_\ell$, $\beta^{(i)}(x) \preceq \beta^{(i)}(y)$. By Condition 1 of the L-continuous secure, we have $\varphi_{n,i}(x) \subset B_i(x, \beta^{(i)}(y)) \subset B_i(x, \beta^{(i)}(x))$. Therefore, we proved the claim.

Taking the convex closure over all $\ell \in J$, we have $\Phi(x) \subset co(B(x, \beta(x)))$ where $\beta = \times_{i \in I} \beta^{(i)}$. As Φ has a fixed point, we have $x \in co(B(x, \beta(x)))$ for some $x \in X$. Recall that $\beta^i(x) = (\alpha^{(i)}_{n,1}, ..., \alpha^{(i)}_{n,|I|})$ for some $n \in \{1, 2, ..., N\}$ such that $x \in \tilde{V}_n \subset V_n$, we have a contradiction of Condition 2 of the L-continuous secure. ■

Applying the above theorem, we can show that there is a lexicographic equilibrium if the base game has discrete outcomes and upper semi-continuous private payoff functions.

Proposition 1. *For a lexicographic game $(I, (X_i, u_i, \mathcal{L}_i)_{i \in I})$, if all action spaces are compact convex subsets of metrizable locally convex topological vector spaces, and all private payoff functions are in the form $u_i : X \to \{0, 1, ..., M\}$ and are upper semi-continuous and quasi-concave in own strategy*[3]*, there is a lexicographic equilibrium.*

We remark that this theorem, even when there is no externalities, does not follow from the classical proof as as the best response correspondence may not be upper hemi-continuous.[4]

[3] i.e. $u_i(x_i, x_{-i})$, as a function of x_i, is quasi-concave for all $x_{-i} \in X_{-i}$.

[4] For example, in a two player case, $X_1 = X_2 = [0, 1]$, $u_1(x_1, x_2)$ is 1 at $(0.5, 0.5)$ and 0 elsewhere. Taking $x_2^n = 0.5 + 0.5/n \to 0.5 = x_2$, $x_1^n = 0 \to 0 = x_1$, we note x_1^n is always a best response to x_2^n, but x_1 is not a best response to x_2.

Proof. We prove the special case when $M = 1$ and $\mathcal{L}_i = (i, i_2)$. The general proof is postponed to the appendix. For any $i \in I$, we define

$$F_i^{(1)} = \{x_{-i} \in X_{-i} : \max_{x_i \in X_i} u_i(x) = 1\}$$

$$F_i^{(2)} = \{x_{-i} \in X_{-i} : \max_{x_i \in X_i} u_{i_2}(x) = 1\}$$

Therefore, when $F_i^{(1)} = \emptyset$, player i's action can not affect his own private payoff. When $F_i^{(2)} = \emptyset$, player i's action can not affect his lover's private payoff.

Firstly, we claim $F_i^{(1)}$ and $F_i^{(2)}$ are closed sets. We prove the $F_i^{(1)}$ is close. The closedness of $F_i^{(2)}$ will follow by symmetry. Take a sequence $x_{-i}^n \to x_{-i}$ and $x_{-i}^n \in F_i^{(1)}$. By the compactness of X and upper semi-continuity of the function u_i, we have a sequence $(x_i^n)_{n \in \mathbb{N}}$ such that $u_i(x_i^n, x_{-i}^n) = 1$. By taking subsequences, $x_i^n \to x_i$. Thus, by upper semi-continuity, $1 = \limsup_n u_i(x_i^n, x_{-i}^n) \leq u_i(x_i, x_{-i}) \leq 1$. Thus, $x_{-i} \in F_i$.

Secondly, we define a set $F_i \subset X_i$ by

$$F_i = \begin{cases} \emptyset & \text{if } F_i^{(1)} = \emptyset, F_i^{(2)} = \emptyset \\ F_i^{(2)} & \text{if } F_i^{(1)} = \emptyset, F_i^{(2)} \neq \emptyset \\ F_i^{(1)} & \text{if } F_i^{(1)} \neq \emptyset, F_i^{(1)} \cap F_i^{(2)} = \emptyset \\ F_i^{(1)} \cap F_i^{(2)} & \text{if } F_i^{(1)} \cap F_i^{(2)} \neq \emptyset \end{cases}$$

and a best reply correspondence $\varphi_i : F_i \rightrightarrows X_i$ such that

$$\varphi_i(x_{-i}) = \big\{x_i \in X_i : (u_i(x_i, x_{-i}), u_{i_2}(x_i, x_{-i})) \succeq (u_i(x_i', x_{-i}), u_{i_2}(x_i', x_{-i})), \forall x_i' \in X_i\big\}$$

We claim φ_i is upper hemi-continuous on F_i. To prove the claim, we take $(x_i^n, x_{-i}^n) \to (x_i, x_{-i})$ where $x_i^n \in \varphi_i(x_{-i})$. There are four cases:

1. $F_i = \emptyset$: $x_i \in \varphi_i(x_{-i})$ trivially.
2. $F_i = F_i^{(2)}$: we have u_i is constant on F_i and $u_{i_2}(x_i^n, x_{-i}^n) = 1$. By the upper semicontinuity of u_{i_2}, $u_{i_2}(x_i, x_{-i}) = 1$. Therefore, $x_i \in \varphi_i(x_{-i})$.
3. $F_i = F_i^{(1)}$: we have u_{i_2} is constant on F_i and $u_i(x_i^n, x_{-i}^n) = 1$. By the upper semicontinuity of u_i, $u_i(x_i, x_{-i}) = 1$. Therefore, $x_i \in \varphi_i(x_{-i})$.
4. $F_i = F_i^{(1)} \cap F_i^{(2)}$: we have $u_i(x_i^n, x_{-i}^n) = u_{i_2}(x_i^n, x_{-i}^n) = 1$. By by upper semicontinuity of u_i and u_{i_2}, $u_i(x_i, x_{-i}) = u_{i_2}(x_i, x_{-i}) = 1$. Therefore, $x_i \in \varphi_i(x_{-i})$.

By Theorem 2.4 in Tan and Wu [2003], there is an upper hemicontinuous, non-empty and closed-valued extension of φ_i to X_{-i}. We abuse the notation and denote this extension by φ_i.

Lastly, we prove there is a lexicographic equilibrium by proving the lexicographic game is L-continuously secure. Suppose $x \in X$ is not a lexicographic equilibrium, and V_x is an open neighborhood of x such that every $y \in V_x$ is not a lexicographic equilibrium. (Otherwise we are done.) For each player $i \in I$, we denote the projection of the space V_x on X_{-i} by $\Pi_{-i}(V_x)$. Define a correspondence $\varphi : V_x \rightrightarrows X$ by $\varphi(x) = (\varphi_i(x_{-i}))_{i\in I}$ and define the pair $(\alpha^{(i)}_{x,1}, \alpha^{(i)}_{x,2})$ as follows:

$$(\alpha^{(i)}_{x,1}, \alpha^{(i)}_{x,2}) = \begin{cases} (0,0) & \text{if } F_i^{(1)} \cap \Pi_{-i}(V_x) = \emptyset, F_i^{(2)} \cap \Pi_{-i}(V_x) = \emptyset \\ (0,1/2) & \text{if } F_i^{(1)} \cap \Pi_{-i}(V_x) = \emptyset, F_i^{(2)} \cap \Pi_{-i}(V_x) \neq \emptyset \\ (1/2,0) & \text{if } F_i^{(1)} \cap \Pi_{-i}(V_x) \neq \emptyset, F_i^{(1)} \cap F_i^{(2)} \cap \Pi_{-i}(V_x) = \emptyset \\ (1/2,1/2) & \text{if } F_i^{(1)} \cap F_i^{(2)} \cap \Pi_{-i}(V_x) \neq \emptyset \end{cases}$$

We verify the two conditions of L-continuously secure. For Condition 1, we have, for any $y \in V_x$ and $x_i \in \varphi_i(y_{-i})$,

$$(u_i(x_i, y_{-i}), u_{i_2}(x_i, y_{-i})) = \begin{cases} (0,0) & \text{if } F_i^{(1)} \cap \Pi_{-i}(V_x) = \emptyset, F_i^{(2)} \cap \Pi_{-i}(V_x) = \emptyset \\ (0,1) & \text{if } F_i^{(1)} \cap \Pi_{-i}(V_x) = \emptyset, F_i^{(2)} \cap \Pi_{-i}(V_x) \neq \emptyset \\ (1,0) & \text{if } F_i^{(1)} \cap \Pi_{-i}(V_x) \neq \emptyset, F_i^{(1)} \cap F_i^{(2)} \cap \Pi_{-i}(V_x) = \emptyset \\ (1,1) & \text{if } F_i^{(1)} \cap F_i^{(2)} \cap \Pi_{-i}(V_x) \neq \emptyset \end{cases}$$

Therefore, $(u_i(x_i, y_{-i}), u_{i_2}(x_i, y_{-i})) \succeq (\alpha^{(i)}_{x,1}, \alpha^{(i)}_{x,2})$. That is, $\varphi_i(y) = \varphi_i(y_{-i}) \subset B_i(y, \alpha^{(i)}_{x,1}, \alpha^{(i)}_{x,2})$ for every $i \in I$ and every $y \in V_x$.

For Condition 2, since the private payoff functions are quasiconcave in own strategies, we have that $B_i(y, \alpha^{(i)}_{x,1}, \alpha^{(i)}_{x,2}) = co(B_i(y, \alpha^{(i)}_{x,1}, \alpha^{(i)}_{x,2}))$ for all $y \in V_x$. Thus, it is sufficient to show, for each $y \in V_x$, there exists a player $i \in I$ such that $y_i \notin B_i(y, \alpha^{(i)}_{x,1}, \alpha^{(i)}_{x,2})$. We prove by contradiction. Suppose there is a point $y \in V_x$ such that $y_i \in B_i(y, \alpha^{(i)}_{x,1}, \alpha^{(i)}_{x,2})$ for all $i \in I$. i.e. $(u_i(y_i, y_{-i}), u_{i_2}(y_i, y_{-i})) \succeq (\alpha^{(i)}_{x,1}, \alpha^{(i)}_{x,2})$. Then, it is easy to check y_i is a best response of agent i. Consequently, y is a lexicographic equilibrium. Contradiction. Thus, Condition 2 holds. By Theorem 1, there is a lexicographic equilibrium. ■

Next, we give an example to illustrate the difference between lexicographic equilibrium and the selfish equilibrium.

Example. *Consider a game with two players* $I = \{1, 2\}$ *who produce and care about each other in the second order. Their action sets are their production levels* $X_1 = X_2 = [0, 10]$. *In addition, their private payoff functions are given by*

$$u_1(x_1, x_2) = 1_{[1,10]}(x_1) - 1_{(1,10]}(x_2) + 1$$ [5]

$$u_2(x_1, x_2) = 1_{[1,10]}(x_2) - 1_{(1,10]}(x_1) + 1$$

That is, both players are satisfied about themselves if and only if they produce no less than 1. Nevertheless, they will be angry if the other player produce more than the least amount to make themselves satisfied.

It is easy to check, both private payoff functions are upper semi-continuous and quasi-concave. Therefore, by Proposition 1, there is a lexicographic equilibrium.

Indeed, it is easy to see that $(x_1 = 1, x_2 = 1)$ *is the unique lexicographic equilibrium and both player obtain 2 units of payoff. In contrast, any* (x_1, x_2), *where* $x_1, x_2 \geq 1$, *is a selfish equilibrium. When* $(x_1, x_2) \neq (1, 1)$, *at least one player just obtains 1 unit of payoff. Consequently, the unique lexicographic equilibrium Pareto dominates every selfish equilibrium that is not a lexicographic equilibrium.*

4.2 ε-lexicographic equilibrium

Next, we define an ε-relaxation of the lexicographic equilibrium. This relaxation will help us to establish the existence result for a wider class of games, and may help the empirical applications of the idea of lexicographic equilibrium.

Formally, given a lexicographic game $(I, (X_i, u_i, \mathcal{L}_i)_{i \in I})$ and any fixed $\varepsilon \in (0, 1)$, an action profile $x^* \in X$ is an $\varepsilon-$*lexicographic equilibrium* if, for any players $i \in I$ whose preference list is $\mathcal{L}_i = (i_1, ..., i_{|I|})$, we have,

$$x_i^* \in \underset{x_i \in B_i^{(|I|-1)}(x_{-i}^*)}{\operatorname{argmax}} u_{i_N}(x_i, x_{-i}^*)$$

where $B_i^{(|I|-1)}$ is the $|I| - 1$-th best response sets defined recursively follows. The 0-th best response set is defined by $B_i^{(0)}(x_{-i}^*) = X_i$. For any $1 \leq k \leq |I| - 1$, the

[5] We add 1 to ensure the payoff functions are non-negative, thus we can apply Proposition 1 directly.

k-th best response set is defined by $B_i^{(k)}(x_{-i}^*) = \{x_i \in B_i^{(k-1)}(x_{-i}^*) : u_{i_k}(x_i, x_{-i}^*) \geq (1-\varepsilon)u_{i_k}(y, x_{-i}^*), \forall y \in B_i^{(k-1)}(x_{-i}^*)\}$.

That is, all players compute their best response recursively. Firstly, they compute the subset of actions giving their 1-st lovers an almost highest private payoff. Secondly, within this subset, they compute the subset of actions giving their 2-nd lovers an almost highest private payoff. This recursive process ends at step $|I|$, when they choose a best action for their I-th lovers among the remaining actions. For instance, when player i has a preference list $\mathcal{L}_i = (i, i_2)$, i.e. he cares about himself in the first order and his lover in the second order, player i chooses an action giving his lover the highest payoff provided they have made themselves almost the best. In contrast, when player i has a preference list $\mathcal{L}_i = (i_2, i)$, i.e. he cares about his lover in the first order and himself in the second order, player i chooses an action giving himself the highest payoff provided they have made their lover almost the best. We note, when ε is close to 1, the second example corresponds to the idea that player i maximizes his own private payoff provided he does not put his opponent in a very bad situation.

Clearly, any lexicographic equilibrium is an ε-lexicographic equilibrium. Now, we state the existence result for a more general class of games.

Theorem 2. *For any finite base game $(I, (X_i, u_i)_{i\in I})$, any preference lists $(\mathcal{L}_i)_{i\in I}$ and any $\varepsilon \in (0,1)$, the lexicographic game $(I, (X_i, u_i, \mathcal{L}_i)_{i\in I})$ has an $\varepsilon-$lexicographic equilibrium.*

Proof. See Appendix B. ■

5 Welfare analysis

In this section, we attempt to provide an answer to the question *"Will the world become better if everyone cares about some other for even just a little bit?"*. We analyze a public bads game $(I, (X_i, u_i, \mathcal{L}_i)_{i\in I})$ defined as follows.

There are finitely many homogeneous players in set I. Each player $i \in I$ chooses his output level (e.g. carbon dioxide emission level) $c_i \in \mathbb{R}_+$. There is a concave return of the individual output. Nevertheless, the total outputs has a convex cost imposed to every player as the output is a public bad. More precisely, the private payoff function is defined by

$$u_i(c_i, c_{-i}) = f(c_i) - g\left(\sum_{i \in I} c_j\right)$$

where f is a differentiable concave function increasing at zero, decreasing at infinity and g is differentiable convex increasing function on the positive axis. We notice, that an increase in one's output level decreases all others' payoffs. For more details about this public bads game, we refer to Roemer [2017] .

In this lexicographic game, every player care about themselves in the first order and care about some other player in the second order. Consequently, $\mathcal{L}_i = (i, i_2)$ for some $i_2 \neq i$.

We have two observations. Firstly, when every player computes the exact maximizer for their own payoff before caring about the others, with some additional assumptions on g, every player is better off when they care about the others. When every player computes an approximate maximizer for their own payoff before caring about the others, with no additional assumption, every player is better off when they care about the others. We present these observations in order.

5.1 Lexicographic equilibrium

We first study the efficiency of lexicographic equilibrium. Note, as players care about some others in the second order, they will produce the least (pollution) level in their argmax sets. i.e. a player i solves:

$$\min\left(\operatorname*{argmax}_{c_i \geq 0} f(c_i) - g(c_i + \sum_{j \neq i} c_j)\right)$$

Nevertheless, if one decreases his output level, the best response of the other players will increase. Thus, it is not immediately clear that the total output $\sum_{j=1}^{N} c_j$ will decrease if everyone produces the least in their argmax sets. Consequently, we work on a special case that the function g is affine. In this special case, the optimal output level does not depends on the total output of the others. Therefore, we have.

Proposition 2. *In a public bad game where g is affine, there is a unique lexicographic equilibrium and it Pareto dominates all selfish equilibrium that is not an lexicographic equilibrium.*

More generally, if g is a function of n variables $c_1, ..., c_N$, the above statement is true if g is additively separable. For example, $g(c_1, ..., c_N) = c_1 + ... + c_N$, $g(c_1, ..., c_N) = c_1^2 + ... + c_N^2$ or $g(c_1, ..., c_N) = c_1 + c_2^2 + ... + c_N^N$.

5.2 ε-lexicographic equilibrium

In this subsection, we study the efficiency of ε-lexicographic equilibrium. We prove that an ε-lexicographic equilibrium Pareto dominates the unique selfish equilibrium. We impose the following assumption on the private payoff functions:

(A) $f \in C^2(\mathbb{R}_+)$ is a strictly concave function satisfies $f'(0) = 0, f'(0) > 0, f'(\infty) < 0$. $g \in C^2(\mathbb{R}_+)$ is a strictly convex function such that $g(0) = g'(0) = 0$. Moreover, f'', g'' are bounded away from zero on $\mathbb{R}_+$.

Proposition 3. *In a public bad model $(I, (X_i, u_i, \mathcal{L}_i)_{i \in I})$ satisfying assumption (A), there is a unique selfish equilibrium, which is Pareto dominated by an ε-lexicographic equilibrium when ε is small enough.*

An intuitive proof is given as follows. When any player only produces an ε-optimal output, he will reduce his production by $\sqrt{\varepsilon}$ as the utility function is very flat around the optimal solution. Consequently, if everyone is willing to give up an ε amount of payoff, the total output will be reduced by a huge number at the level of $\sqrt{\varepsilon}$. That is, the loss from the total output will be reduced by a lot. In conclusion, all players' overall payoff will increase if they care about the others.

To prove this proposition, we study the best response map

$$B(x) = \operatorname*{argmax}_{c \geq 0} f(c) - g(c + x)$$

Firstly, by strict concavity, B is a real-valued function. Furthermore, $B(x)$ is a decreasing function. By implicit function theorem, we know B is differentiable and $B'(x) = \frac{g''(B(x)+x)}{f''(B(x)) - g''(B(x)+x)} < 0$. Actually, since f'' and g'' are bounded away from zero, B' is bounded away from zero.

Firstly, we prove that the unique selfish equilibrium is symmetric. For any selfish equilibrium $(c_i^S)_{i \in I}$ satisfies, we have

$$f'(c_i^S) = g'(\sum_{i \in I} c_j^S), \quad \forall i \in I$$

As f' is monotone, we note c_i^S does not depend on i. Therefore, all selfish equilibrium are symmetric.

Moreover, there is a unique selfish equilibrium. For any symmetric equilibrium $(c_i^S = c^S)_{i\in I}$, c solves the equation $B((|I|-1)c^S) = c^S$. As the left hand side is a strictly decreasing function in c^S, the right hand side is a strictly increasing function in c^S, and $f'(0) > g'(0)$, there is a unique solution.

Next, we prove that the selfish equilibrium is not Pareto efficient. Let the efficient individual production level be c^E. By definition, c^E solves the welfare maximization problem $\max_{c\geq 0} f(c) - g(|I|c)$. Therefore, $f'(c^E) - |I|g'(|I|c^E) = 0$. In contrast, c^S solves the maximization problem $\max_{c\geq 0} f(c) - g(c + (|I|-1)c^S)$. Therefore, $f'(c^S) - g'(|I|c^S) = 0$. Consequently, $c^S > c^E$. That is, the selfish equilibrium is inefficient.

Now we analyze the symmetric ε-lexicographic equilibrium. In this case, people care about the others. Thus, they will produce the least amount $c^{\varepsilon L}$ within the neighborhood of the maximizer. To see whether the decrease in total output can compensate the individual sacrifices, we estimate the deviation $|c^{\varepsilon L} - c^S|$. We note that this deviation must be much larger than ε since the function has very small derivatives around its optimum. Similar observation can be found in Fudenberg et al. [2009]. Formally, we have the following lemma.

Lemma 1. *Suppose $h \in C^2(\mathbb{R}_+)$ is a strictly concave function such that $h(0) = 0$, $h'(0) > 0$, $h'(\infty) < 0$ and h'' is bounded away from zero. Let $x_{\max}$ be the unique maximizer of h. Moreover, let $x^* < x_{\max}$ satisfies $h(x^*) = (1-\varepsilon)h(x_{\max})$. Then, we have*

$$\frac{1}{C}\sqrt{\varepsilon} < x_{\max} - x^* < C\sqrt{\varepsilon}$$

for some constant $C > 0$.

Proof. See Appendix C. ■

Roughly speaking, whenever everyone give up ε of their payoffs, they will output much less and thus the total production becomes about $\sqrt{\varepsilon}$ less. Consequently, everyone is better off when ε is not too large. Formally, the output level $c^{\varepsilon L} > 0$ in a ε-lexicographic equilibrium solves the equation

$$f(c^{\varepsilon L}) - g(|I|c^{\varepsilon L}) = (1-\varepsilon)\max_{c\geq 0}\left(f(c) - g(c + (|I-1|)c^{\varepsilon L})\right)$$

Take $h(c) = f(c) - g(c + (|I-1|)c^{\varepsilon L})$ and apply the lemma, we have

$$B((|I-1|c^{\varepsilon L})) - c^{\varepsilon L} = \Omega(\sqrt{\varepsilon})$$

Recall that

$$B((|I-1|c^{S})) - c^{S} = 0$$

Consider the fixed point graph of the function B, as we have shown that B' is uniformly bounded away from zero around the fixed point c^S, we have

$$0 < c^S - c^{\varepsilon L} \leq C'\sqrt{\varepsilon}$$

for some $C' > 0$. When ε is small enough, we have

$$c^E < c^{\varepsilon L} < c^S$$

Hence, by the strict concavity, we have

$$f(c^E) - g(|I|c^E) > f(c^{\varepsilon L}) - g(|I|c^{\varepsilon L}) > f(c^S) - g(|I|c^S)$$

That is, the ε-lexicographic equilibrium $(c_i = c^{\varepsilon L})_{i \in I}$ Pareto dominates the unique selfish equilibrium.

6 Concluding remarks

In this paper, we provide a new game theoretic model with lexicographic externalities. There are three main differences between lexicographic externalities and the classical additive externalities. Firstly, our model provides a less severe departure from the pure selfishness case in the sense that, unlike the usual formulation, there is no sacrifice between oneself and others' well-beings. Secondly, in our model, agents care about the others only after they maximize their own payoffs. Thirdly, the preferences in our model are invariant under positive affine transformations on payoff functions.

In order to make our model applicable, we provided two existence results. Firstly, by generalizing the idea of "continuously secure game" in Barelli and Meneghel [2013] to vector-valued payoff functions, we proved that a lexicographic equilibrium exists in a lexicographic game with discrete outcomes and upper semi-continuous and own strategy quasi-concave utility functions. Secondly, we proved that a ε-lexicographic equilibrium exists in any lexicographic game generated from a finite base game. Although the ε-variation case is nested in the first case, it may help to apply the idea of lexicographic externalities to experimental investigations.

In addition, the lexicographic externalities improve the social welfare. We proved that lexicographic equilibrium and its variations Pareto dominate the selfish equilibrium in a model with public bads. That is, when agents care about each other for just a little bit, the world will become better in certain cases.

Lastly, we propose two directions of future explorations. Firstly, the current model might be able to be extended to a model with a continuum of players.[6] Secondly, the idea of lexicographic externalities might be able to be embedded in general equilibrium models. In particular, I am curious about whether the idea in this paper can help to understand general equilibrium models with public factors.

References

K. Y. Abalo and M. M. Kostreva. Some existence theorems of nash and berge equilibria. *Applied mathematics letters*, 17(5):569–573, 2004.

C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer Science & Business Media, 2007.

A. Bagh and A. Jofre. Reciprocal upper semicontinuity and better reply secure games: a comment. *Econometrica*, 74(6):1715–1721, 2006.

P. Barelli and I. Meneghel. A note on the equilibrium existence problem in discontinuous games. *Econometrica*, 81(2):813–824, 2013.

P. Barelli and I. Soza. On the existence of nash equilibria in discontinuous and qualitative games. *University of Rochester*, pages 47–116, 2009.

C. Berge. *Théorie générale des jeux à n personnes*, volume 138. Gauthier-Villars Paris, 1957.

[6] One way to extend the current model to a continuum player model is to extend the result in Carmona and Podczeck [2014] to vector-valued payoff functions.

P. Bich. Existence of pure nash equilibria in discontinuous and non quasiconcave games. *International Journal of Game Theory*, 38(3):395–410, 2009.

P. Bich and R. Laraki. A unified approach to equilibrium existence in discontinuous strategic games. 2012.

G. Carmona. An existence result for discontinuous games. *Journal of Economic Theory*, 144(3):1333–1340, 2009.

G. Carmona. Symposium on: Existence of nash equilibria in discontinuous games. *Economic Theory*, 48(1):1–4, 2011a.

G. Carmona. Understanding some recent existence results for discontinuous games. *Economic Theory*, 48(1):31–45, 2011b.

G. Carmona. Reducible equilibrium properties. *Report, University of Cambridge*, 2012.

G. Carmona and K. Podczeck. Existence of nash equilibrium in games with a measure space of players and discontinuous payoff functions. *Journal of Economic Theory*, 152:130–178, 2014.

P. Dasgupta and E. Maskin. The existence of equilibrium in discontinuous economic games, ii: Applications. *The Review of economic studies*, 53(1):27–41, 1986a.

P. Dasgupta and E. Maskin. The existence of equilibrium in discontinuous economic games, i: Theory. *Review of Economic Studies*, 53(1):1–26, 1986b.

G. De Marco and J. Morgan. Friendliness and reciprocity in equilibrium selection. *International Game Theory Review*, 10(01):53–72, 2008a.

G. De Marco and J. Morgan. Slightly altruistic equilibria. *Journal of Optimization Theory and Applications*, 137(2):347–362, 2008b.

D. Fudenberg, D. Levine, and E. Maskin. The folk theorem with imperfect public information. In *A Long-Run Collaboration On Long-Run Games*, pages 231–273. World Scientific, 2009.

A. McLennan, P. K. Monteiro, and R. Tourky. Games with discontinuous payoffs: a strengthening of reny's existence theorem. *Econometrica*, 79(5):1643–1664, 2011.

E. A. Ok. Utility representation of an incomplete preference relation. *Journal of Economic Theory*, 104(2):429–449, 2002.

P. Prokopovych. On equilibrium existence in payoff secure games. *Economic Theory*, 48(1):5–16, 2011.

P. Prokopovych. The single deviation property in games with discontinuous payoffs. *Economic Theory*, 53(2):383–402, 2013.

P. J. Reny. On the existence of pure and mixed strategy nash equilibria in discontinuous games. *Econometrica*, 67(5):1029–1056, 1999.

J. Roemer. *How we do and could cooperate: A Kantian explanation.* Preprint, 2017.

K.-K. Tan and Z. Wu. An extension theorem and duals of gale-mas-colell's and shafer-sonnenschein's theorems. *Set Valued Mappings with Applications in Nonlinear Analysis*, page 449, 2003.

V. I. Zhukovskii and A. A. Chikrii. Linear quadratic differential games. *Naoukova Doumka, Kiev*, 1994.

A Proof of Proposition 2

Proof. We prove the case when the preference order $\mathcal{L}_i$ of length n and no repeat components appears. When there are repeat, we can think the equivalent preference order of length less than n, and the proof is the same.

For $i \in N$, $j \in \{1, 2, ..., N\}$, $m \in \{0, 1, ..., M\}$, we partition X_{-i} into $(M+1)N$ spaces by

$$F_i^{(j,m)} = \{x_{-i} \in X_{-i} : \max_{x_i \in X_i} u_{i_j}(x) = m\}$$

This set denotes the set of others' action profiles such that i_j could have a payoff m. By the same method as in the proof for the special case, we have $F_i^{(j,m)}$ is a closed set.

In addition, we define,

$$F_i^{(j)} = \cup_{m=1}^{M} F_i^{(j,m)}$$

We note whenever $F_i^{(j)} = \emptyset$, i's action can not affect i_j's payoff.

Since a player i will consider his impact on i_j's payoff only if player i is has maximized the payoff vector of i_1 up to i_{j-1} in the lexicographic order, we define the set $F_i \subset X_{-i}$ that highlighting this order by the following:

1. If $F_i^{(1)} = ... = F_i^{(N)} = \emptyset$
$$F_i = \emptyset$$

2. Otherwise,
$$F_i = \bigcap_{(j_s, m_s)_{s=1}^q \in S} \left(F_i^{j_1, m_1} \cap F_i^{j_2, m_2} \cap ... \cap F_i^{j_q, m_q} \right)$$
where $q \in \{1, 2, ..., N\}$, and for any $1 \leq s \leq q$, we have $j_s \in \{1, 2, ..., N\}$, $1 \leq m_s \leq M$, and S is the set of vectors $(j_s, m_s)_{s=1}^q$ satisfying:

   (a) j_1 is the largest j such that $F_i^{(j)} \neq \emptyset$

   (b) m_1 is the largest m such that $F_i^{(j_1, m)} \neq \emptyset$

   (c) For all $1 < s \leq q$, $F_i^{(j_1, m_1)} \cap ... \cap F_i^{(j_{s-1}, m_{s-1})} \cap F_i^{(j_s)} \neq \emptyset$

   (d) For all $1 < s \leq q$, m_s is the largest m such that $F_i^{(j_1, m_1)} \cap ... \cap F_i^{(j_{s-1}, m_{s-1})} \cap F_i^{(j_s, m_s)} \neq \emptyset$

In other word, the set F_i is the minimum one of all set in such form satisfying (a) to (d) or it is the intersection of the longest chain.

Now, we define the best reply correspondence $\varphi_i : F_i \rightrightarrows X_i$ such that,

$$\varphi_i(x_{-i}) = \text{argmax} \left(x_i \in X_i : (u_i(x_i, x_{-i}), ..., u_{i_N}(x_i, x_{-i})) \right)$$

where the max operator is taking over the lexicographic order. By compactness of X_i, φ_i is non-empty valued.

Now we argue φ_i is upper hemi-continuous. Take $(x_i^n, x_{-i}^n) \to (x_i, x_{-i})$ with $x_i^n \in \varphi_i(x_{-i})$: if $F_i = \emptyset$, it is the trivial case. Otherwise,

$$F_i = F_i^{j_1, m_1} \cap F_i^{j_2, m_2} \cap ... \cap F_i^{j_q, m_q}$$

It implies $u_{i_{j_s}}(x_i^n, x_{-i}^n) = m_s$ for all $1 \leq s \leq q$. By upper semi-continuity of the payoff function, $u_{i_{j_s}}(x_i, x_{-i}) = m_s$ for all $1 \leq s \leq q$.

By theorem 2.4 in Tan and Wu [2003] , there is an upper hemicontinuous, non-empty and closed-valued extension of φ_i to X_{-i}. We abuse the notation and denote this extension by φ_i.

Let x not be a lexicographic equilibrium, and let V_x be an open neighborhood of x such that every $y \in V_x$ is not a lexicographic equilibrium. (Otherwise we are done.) Now, for each i, we define $(\alpha_{x,i}^{(1)}, \alpha_{x,i}^{(2)}, ..., \alpha_{x,i}^{(N)})$ by the following. For our convenience, we denote the projection of the space V_x on X_{-i} by $\Pi_{-i}(V_x)$.

If $F_i = F_i^{j_1,m_1} \cap F_i^{j_2,m_2} \cap ... \cap F_i^{j_q,m_q}$, $(\alpha_{x,i}^{(1)}, \alpha_{x,i}^{(2)}, ..., \alpha_{x,i}^{(N)})$ is defined by

$$\alpha_{x,i}^{(j)} = \begin{cases} m_s - 1/2 & \text{if } j = j_s, F_i^{(j_1)} \cap \Pi_{-i}(V_x) = ... = F_i^{(j_{s-1})} \cap \Pi_{-i}(V_x) = \emptyset, F_i^{(j_s)} \cap \Pi_{-i}(V_x) \neq \emptyset, \\ & \text{or, for some } T \subset \{1, 2, ..., s-1\}, \\ & \text{if } t \in T, \cap_{u \notin T, u<t} F_i^{j_u} \cap F_i^{j_t} \cap \Pi_{-i}(V_x) = \emptyset; \text{if } t \notin T, \cap_{u \notin T, u<t} F_i^{j_u} \cap F_i^{j_t} \cap \Pi_{-i}(V_x) \neq \emptyset; \\ & \text{and } \cap_{u \notin T, u<s} F_i^{j_u} \cap F_i^{j_s} \cap \Pi_{-i}(V_x) \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

The second case above is to describe the cases such as $j = j_3$,

$$F_i^{j_1} \cap \Pi_{-i}(V_x) \neq \emptyset, F_i^{j_1} \cap F_i^{j_2} \cap \Pi_{-i}(V_x) = \emptyset, F_i^{j_1} \cap F_i^{j_3} \cap \Pi_{-i}(V_x) \neq \emptyset$$

Now we verify a) and b). For a), we are trying to show $\varphi_i(y) \subset B_i(y, \alpha_{x,i}^{(1)}, ..., \alpha_{x,i}^{(N)})$ for every $i \in N$ and every $y \in V_x$. That is, for any $x_i \in \varphi_i(y_{-i})$,

$$(u_{i_1}(x_i, y_{-i}), ..., u_{i_N}(x_i, y_{-i})) \succeq_i (\alpha_{x,i}^{(1)}, ..., \alpha_{x,i}^{(N)})$$

We first look at non-zero coordinate $\alpha_{x,i}^{j}$ on the right hand side. By definition, it can only be nonzero if $j = j_s$ for some $1 \leq s \leq q$. On the other hand, we know by $x_i \in \varphi_i(y_{-i})$, $u_{i_{j_s}}(x_i, y_{-i}) = m_s > m_s - 1/2$. Thus a) is proved.

Now for b), we note as the private payoff functions are quasiconcave,

$$B_i(y, \alpha_{x,i}^{(1)}, ..., \alpha_{x,i}^{(N)}) = co(B_i(y, \alpha_{x,i}^{(1)}, ..., \alpha_{x,i}^{(N)}))$$

Thus, it is enough to show, for each $y \in V_x$, there exists i with $y_i \notin B_i(y, \alpha_{x,i}^{(1)}, ..., \alpha_{x,i}^{(N)})$. If it is not true, then there is some $y \in V_x$ such that $y_i \in B_i(y, \alpha_{x,i}^{(1)}, ..., \alpha_{x,i}^{(N)})$ for all $i \in N$. That is $(u_{i_1}(y_i, y_{-i}), ..., u_{i_N}(y_i, y_{-i})) \succeq (\alpha_{x,i}^{(1)}, ..., \alpha_{x,i}^{(N)})$,

If $\alpha_{x,i}^{(j)} = 0$, then it means within $\Pi_{-i}(V_x)$, i can not increase i_j's payoff, either because his action has no effect on i_j, $F_i^{(j)} \cap \Pi_{-i}(V_x) = \emptyset$, or because his action has no effect on i_j in the set i maximizes the payoff of players he cares about more. In contrast, if $\alpha_{x,i}^{(j_s)} = m_s - 1/2$ for some $1 \leq s \leq q$, it implies $u_{i_{j_s}} = m_s$, which means i has done the best for j_s within the set he maximizes the payoffs of players he cares about more.

That is, y_i is the best response of i given y. If this is true for all i, y will be a lexicographic equilibrium, contradicts the assumption. Thus, b) holds. Consequently, the result follows from theorem 2.

■

B Proof of Theorem 2

We prove the theorem by applying Kakutani's fixed point theorem. Recall, if a player i has a preference list $(i_1, ..., i_{|I|})$, his best response correspondence is given

$$\operatorname*{argmax}_{x_i \in B_i^{(|I|-1)}(x_{-i})} u_{i_N}(x_i, x_{-i})$$

Applying Berge's theroem, it is sufficient to prove the map $x_i \to B_i^{(|I|-1)}(x_{-i})$ is a continuous correspondence. We prove by induction. Firstly, $x_i \to B_i(0, x_{-i}) = X_i$ is clearly continuous. Suppose $x_i \to B_i^{(k-1)}(x_{-i})$ is continuous for some $1 \leq k \leq |I| - 1$, we prove $x_i \to B_i^{(k)}(x_{-i})$ is continuous.

To prove the upper hemi-continuity, as X_i is compact, it is sufficient to prove that the graph $S = \{(x_{-i}, x_i) : x_i \in B_i^{(k)}(x_{-i})\}$ is closed. Take a sequence $(x_{-i}^n, x_i^n) \in S$ and $(x_{-i}^n, x_i^n) \to (x_{-i}^*, x_i^*)$. By definition, we have $u_{i_k}(x_i^n, x_{-i}^n) \geq (1-\varepsilon) \max_{x_i \in B_i^{(k-1)}(x_{-i}^n)} u_{i_k}(x_i, x_{-i}^n)$. By the continuity of private payoff functions, we have $u_{i_k}(x_i^n, x_{-i}^n) \to u_{i_k}(x_i^*, x_{-i}^*)$. By the induction hypothesis and Berges theorem, we have $\max_{x_i \in B_i^{(k-1)}(x_{-i})} u_{i_k}(x_i, x_{-i})$ is continuous in x_{-i}. Consequently, $u_{i_k}(x_i^*, x_{-i}^*) \geq (1 - \varepsilon) \max_{x_i \in B_i^{(k)}(x_{-i}^*)} u_{i_k}(x_i, x_{-i}^*)$. i.e. $(x_{-i}^*, x_i^*) \in S$.

To prove the lower hemi-continuity, we take $(x_{-i}^*, x_i^*) \in S$ and take $x_{-i}^n \to x_{-i}^*$. That is, $u_{i_k}(x_i^*, x_{-i}^*) > (1 - \varepsilon) \max_{x_i \in B_i^{(k-1)}(x_{-i}^*)} u_{i_k}(x_i, x_{-i}^*)$. By continuity, $u_{i_k}(x_i^*, x_{-i}^n) > (1 - \varepsilon) \max_{x_i \in B_i^{(k-1)}(x_{-i}^n)} u_{i_k}(x_i, x_{-i}^n)$. That is, $(x_i^*, x_{-i}^n) \in S$. Secondly, if $u_{i_k}(x_i^*, x_{-i}^*) = (1 - \varepsilon) \max_{x_i \in B_i^{(k-1)}(x_{-i}^*)} u_{i_k}(x_i, x_{-i}^*)$, we prove by contradiction. If we can not find $(x_i^n, x_{-i}^n) \in S$ and $x_i^n \to x_i^*$, by taking sub sequences, it means there is a neighborhood $U \subset X_i$ containing x_i^* such that $U \cap S(x_{-i}^n) = \emptyset$ for all large n, where $S(x_{-i}) = \{x_i : (x_{-i}, x_i) \in S\}$. However, this is a contradiction: take any $x_i' \in U$ such that

$u_{i_k}(x_i', x_{-i}^*) > u_{i_k}(x_i^*, x_{-i}^*)$. Such x_i' exists as the base game is finite. By continuity, for any large enough n, we have $u_{i_k}(x_i', x_{-i}^n) > u_{i_k}(x_i^*, x_{-i}^*)$. Contradiction.

C Proof of Lemma 1

We note,

$$\int_{x^*}^{x_{\max}} h' dx = h(x_{\max}) - h(x^*) = \varepsilon h(x_{\max})$$

As h is strictly concave, we have h' is decreasing. Thus,

$$h'(x^*)(x_{\max} - x^*) > \varepsilon h(x_{\max})$$

Consequently,

$$x_{\max} - x^* > \frac{\varepsilon h(x_{\max})}{h'(x^*)}$$

However, we note

$$h'(x^*) = h'(x_{\max}) - \int_{x^*}^{x_{\max}} h'' dx = -\int_{x^*}^{x_{\max}} h'' dx$$

Thus,

$$h'(x^*) \leq \left\|h''\right\|_\infty (x_{\max} - x^*)$$

Hence,

$$x_{\max} - x^* > \sqrt{\frac{h(x_{\max})}{\|h''\|_\infty}\varepsilon}$$

Moreover, as $-h'(0) = h'(x_{\max}) - h'(0) = h''(c)x_{\max}$ for some $c \in (0, x_{\max})$. Therefore, $x_{\max} = -\frac{h'(0)}{h''(\xi)} \geq \frac{h'(0)}{\|h''\|_\infty}$. In sum,

$$x_{\max} - x^* > C_1\sqrt{\varepsilon}$$

where $C_1 = (h(\frac{h'(0)}{\|h''\|_\infty})/\|h''\|_\infty)^{1/2}$.

On the other hand, we note, if h'' is bounded away from zero near the maximizer, i.e. $\sup h'' > 0$. We know h' can not grow too slow: for any $t > 0$,

$$h'(x_{\max} - t) = -\int_{x_{\max}-t}^{x_{\max}} h'' dx \geq t \sup h''$$

Therefore,

$$\begin{aligned}\varepsilon h(x_{\max}) &= \int_{x^*}^{x_{\max}} h'(x) dx = -\int_0^{x_{\max}-x^*} h'(x_{\max} - t) dt \\ &\geq -\sup h'' \int_0^{x_{\max}-x^*} t dt = -\frac{\sup h''}{2}(x_{\max} - x^*)^2\end{aligned}$$

That is,

$$x_{\max} - x^* \le \sqrt{\frac{2h(x_{\max})}{-\sup h''}\varepsilon}$$

Recall that, $x_{\max} = -\frac{h'(0)}{h''(\xi)} \le \frac{h'(0)}{-\sup h''}$. We have

$$x_{\max} - x^* < C_2\sqrt{\varepsilon}$$

where $C_2 = (2h(\frac{h'(0)}{-\sup h''})/(-\sup h''))^{1/2}$.