Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

doi:10.3934/xx.xx.xx.xx

pp. **X–XX**

EXTERNALITY EFFECTS IN THE FORMATION OF SOCIETIES

Renato Soeiro

LIAAD-INESC TEC and Department of Mathematics, University of Porto, Portugal
Rua do Campo Alegre, 687, 4169-007, Portugal

Abdelrahim Mousa* and Alberto A. Pinto

*Department of Mathematics, Faculty of Science, Birzeit University, Palestine
LIAAD-INESC TEC and Department of Mathematics, University of Porto, Portugal
Rua do Campo Alegre, 687, 4169-007, Portugal

Abstract. We study a finite decision model where the utility function is an additive combination of a personal valuation component and an interaction component. Individuals are characterized according to these two components (their valuation type and externality type), and also according to their crowding type (how they influence others). We study how positive externalities lead to type symmetries in the set of Nash equilibria, while negative externalities allow the existence of equilibria that are not type-symmetric. In particular, we show that positive externalities lead to equilibria having a unique partition into a minimum number of societies (similar individuals using the same strategy, see [27]); and negative externalities lead to equilibria with multiple societal partitions, some with the maximum number of societies.

1. **Introduction.** A decision is in general a course of action resulting from a process which involves selecting among several possible alternatives. The collective coexistence and constant interaction of individuals necessarily creates a social frame in which decisions are made and a social context to which the decision leads. Regardless of whether these interactions are voluntary or not, they play a significant role in the global patterns of behavior that emerge from the individual decisions. Understanding what underlies a global behavior means understanding not only the interactions among decision-makers, the personal evaluations of alternatives and the interdependence between the two, but also having a grasp on the relation between the characteristics of the decision-makers and the characteristics of the global outcome. In fact, each decision composing this outcome, conveys information about the decision maker, as it reveals a choice, be it either a selection of a product to buy or a public service, be it an economic strategy, a political option, a social behavior, be it a life changing choice or a daily life decision, like choosing a bar to go to friday night. Thus, the study of the global behaviour both presupposes and enhances an understanding of what governs individual decisions. This is particularly relevant if it is assumed that individuals act rationally and the choice is a

2010 *Mathematics Subject Classification.* Primary: 91A06, 91A10, 91A40; Secondary: 91A35, 91A44.

Key words and phrases. Decision model, pure Nash equilibria, social interactions, dyadic interactions, non-market externalities, social congestion, conformity, type-symmetric equilibria, polymatrix games, singleton weighted congestion games.

The first author is supported by FCT grant SFRH/BD/88742/2012 .

best response over the evaluation of the alternatives, in the sense of the existence of a Von Neumann-Morgerstern utility (1944).[1] At the core of a game theoretical approach to the problem is the modelling of interactions between decision-makers. Assuming decisions as a global mutual best response, one may use the concept of Nash equilibrium (1951) to retrieve information not only on a global interacting level, but also on an internal individual level, by analysing the interdependence of these social characteristics and the individual personal evaluation processes. The focus of this work is the relation between the characteristics of individuals and the characteristics of the outcome, where the outcome is seen as the set of pure Nash equilibria of a finite non-cooperative game.

The study of the dependence of global behaviour (as an equilibrium) on the characteristics of individuals, and in particular the study of the dependence of individual decision rules on the strategies of others, has a long tradition. Namely, in the work of Schelling (1971, 1973, 1978) where, for example, different distributions of the level of tolerances of individuals lead to residential segregations with different properties; or in the work of Granovetter (1978), where small differences on the distribution of individual thresholds can lead to completely different collective behaviour; or in Mas-Collel (1984), Pascoa (1993) where an atomless distribution of types leads to symmetric equilibria; or other symmetry properties as in Wooders, Cartwright and Selten (2006) and Wooders and Cartwright (2014), which, like our paper, describe partitions of the set of players into groups that arise in equilibrium.

Our approach is to model the outcome of a decision process as the pure Nash equilibria of a finite (both in players and strategies) non-cooperative, simultaneous move game. The value of a given decision is measured through an utility function that is an additive combination of two components: (i) how much the individual personally values the decision, independently of the strategies of others; (ii) the externalities arising from social interactions with those individuals who make that same decision. This is, of course, a very broad class of utility functions included in many models in the literature. The crucial aspect is the choice of how to model the form of dependence on the strategies of others, i.e. how to model social interactions. Let us highlight three main features of our approach to this choice, and position our work in relation to the different approaches in the literature.

A first key feature is that we consider only dyadic interactions (see for example [4], [11]). Dyadic means that, for any given strategy, the influence/impact of an individual i on an individual j is independent of the decision of others. A class of games that focuses only on this kind of interactions is for example the class of polymatrix games, see [13], [17].[2] Another option would be to introduce (also or instead) a dependence of this influence on the whole strategy profile. In general, excluding such a component usually means excluding some form of non-linear anonymous aggregate dependence on the strategies of others. In fact, many games can be captured by an appropriate dyadic component by using such an excluding assumption, as for example by making the appropriate restriction on singleton weighted congestion games, or on the games presented in [4], [6], [18]. The dyadic component is sometimes refered to as the local component of social interactions and

[1]The issue of rationality is beyond the scope of this work, nevertheless, as we will be looking at the outcome and not at the decision process in itself, and as we will be working in a complete information setting where the parameters are open to interpretation, underlying is in fact a very mild rationality assumption.

[2]A first formal reference appears to be due to E. B. Yanovskaya in 1968.

the latter dependence as the global component.[3] Focusing on dyadic interactions seems like a suitable approach for the case we wish to study.

A second feature is that we assume social interactions to be dichotomic, in the sense of being restricted to whether individuals are using the same strategy or a different one, a type of Independence of Irrelevant Alternatives assumption [19]. In this work it can be better described and motivated in the following manner: given a strategy profile, if an individual i changes her decision, the change only affects those in her new decision, because she will start interacting with them, and those in her old decision, because she will no longer interact with them. Her influence on the rest of the individuals was that she was making a different decision, and that hasn't changed. This is also in the spirit of Independence of Irrelevant Choices as in [14], or no spillovers as in [15]. This assumption is important for some of our results, which would not hold without it.

The third feature is that we will allow for social interactions to give rise to both positive and negative externalities, and then study their effects on equilibria and society formation. The question here could be whether to restrict the dependence to be in some sense 'positive' or complementary, leading to a conformity effect; or 'negative' leading to a congestion effect. On the strand of literature that treats conformity effects (those leading to a common (or symmetric) action which may overcome personal or intrinsic preferences) are for instance the works on behavioral conformity by Wooders, Cartwright and Selten [27], a theory of conformity by Bernheim [3], a model of herd behavior by Banarjee [2], the threshold models of collective action as in Granovetter [12], or even the equilibrium symmetry in supermodular games as in Cooper and John [10]. On the strand of literature focusing on congestion effects is for example the class of congestion games as first proposed by Rosenthal [20], later generalized by Milchtaich [16]; or the works of Quint and Shubik [18], Konishi, Le Breton and Weber [14], to name a few.

Social interactions, regardless of whether they exhibit a conformity or congestion effect, should depend not only on the number of individuals in each choice, but also on the characteristics of those individuals. This is a crucial aspect in the works of Wooders ([24, 25, 26]) and of Conley and Wooders ([7, 8, 9]). Wooders's earlier papers allow preferences to depend on the characteristics of agents (their types), while Conley and Wooders separate two sorts of characteristics: crowding characteristics, which determine the effects of a player on others, and tastes. In our model we will use a type profile that characterizes individuals, or distinguishes, according to three different aspects, or attributes. (Keep in mind though that for us type does not mean Bayesian type, as we will be working on a complete information setting and the type profile is something completely determined a priori.) Following the work of Conley and Wooders, we will start with the use of a crowding space, which distinguishes individuals by their impact on the utility of others. The use of a crowding space has the advantage that allows the characterization of classes of strategies where the relevant information is the number of individuals with the same crowding type in each decision. Observe that there is no restriction here: depending on the choice of the crowding space individuals may be all distinguishable or totally anonymous. We then characterize individuals according to their utility function, i.e. taste type, but we will subdivide the taste type into two components, using the

[3]The use of the terms local and global in this context seem amenable to critique, since one could think of 'global dyadic components' or 'local aggregative components', hence we prefer the terms dyadic and aggregative.

two additive components of the utility function. This allows the characterization of Nash equilibria according to the restrictions imposed on the relation between these two components. Furthermore, dividing the taste type in this way, separates the social part of the model, that captures the social interactions, from the 'personal' part given by the valuation component (sometimes called intrinsic preference, which we wittingly avoid). A key advantage of the separate analysis of the valuation component is that, besides comprising the intrinsic and personal perceived benefit of the decision, it captures exogeneous changes and/or characteristics associated to each decision. Namely, dependending on the decision in question, it may represent prices, taxes, product quality, road quality, marketing, political campaings, bribes, etc...

The present work starts as an extension of the two types dichotomic model by Soeiro et al. [23] to a wider finite setting where there may be any number of types of individuals facing a choice among any number of possible alternatives, but with the focus on pure Nash equilibria. The latter work finds its inspiration from Brida et al. [5], a socio-economic model that analyses how the choice of a service is influenced by the profile of users of that same service; and, on a different line, from Almeida et al. [1], where game theory and the field of social psychology are related through the theories of Planned Behavior or Reasoned Action, proposing the Bayesian-Nash equilibrium as one of the many possible mechanisms behind the transformation of human intentions into behaviors.

In this work, the approach is essentially two-folded. First we start by taking a crowding profile and a strategy profile as given, which we call a social context, and study implications on the utility functions and on society formation for the case when the strategy is a Nash equilibrium. As a natural follow-up, we will add the externality profile, creating what we call a social context extension. The relation between the social context extension and the valuation profile is the basis for our next step in the characterization of equilibria. We will further show that a social context extension where there are only positive externalities always leads to type-symmetric Nash equilibria and societal partitions with a minimum number of societies; on the other hand, negative externalities allow the existence of Nash equilibria that are not type-symmetric and have a high number of societies. Furthermore we provide a procedure to find the personal decision values that turn any admissible strategy profile into a Nash equilibria. Throughout this work, whenever we say Nash equilibrium we shall always mean pure strategy Nash equilibrium, since we are only considering pure strategies.

The work unfolds as follows: in section 2 we set up the model; in section 3 we present the relation of our model to the concept of society in [27]; in section 4 we do the separation of the taste component of the type profile and present a conformity obstruction lemma. The lemma allows us to characterize the conditions in the type profile for a given strategy to be admissible or feasible as a Nash equilibrium; in section 5 we define the Nash domain of a strategy (in terms of utility parameters) and characterize it completely; and finally, in sections 6 and 7 we prove the results (and as such, these are more technical sections).

Notation. Throughout the work we will use in general: boldface for variables that convey information about the whole set of players, called generally profiles; caligraphic letters for spaces of such profiles; capital letters for sets and greek letters for specific parameters of a game. The symbol $\equiv$ is used for definitions.

2. **The model.** The *decision model* we present is based on a finite non-cooperative simultaneous move game. We consider a finite set of individuals (players) $\mathcal{I} \equiv \{1, \ldots, n_{\mathcal{I}}\}$, each having to choose independently and simultaneousy an element from a finite set of alternatives $\mathcal{D} \equiv \{1, \ldots, n_{\mathcal{D}}\}$ (the common strategy set). We describe the decisions of the individuals by a *strategy map* $s : \mathcal{I} \to \mathcal{D}$ associating to each individual $i \in \mathcal{I}$ her decision $s_i \equiv s(i) \in \mathcal{D}$, which in turn defines a (pure) strategy profile $\mathbf{s} = (s_1, \ldots, s_{n_{\mathcal{I}}}) \in \mathcal{S} \equiv \mathcal{D}^{n_{\mathcal{I}}}$. We will use the standard notation $(s_i; \mathbf{s}_{-i})$ to represent strategy profile $\mathbf{s}$, but highlighting the component of individual i and the remaining strategy profile $\mathbf{s}_{-i}$. The utility of a given strategy profile $\mathbf{s}$ for an individual $i \in \mathcal{I}$ is measured through an utility function that is an additive combination of two components: (i) how much the individual personally values decision s_i, independently of the strategies of the others; and (ii) the externalities arising from social interactions with those individuals who made the same decision as her. The latter component thus determines the social impact of the strategy profile $\mathbf{s}_{-i}$ on individual i while making decision s_i.

The value for an individual $i \in \mathcal{I}$ of each alternative $d \in \mathcal{D}$ is given by the *decision value coordinate* $\omega_i^d \in \mathbb{R}$; these coordinates represent how much she likes or dislikes to make a certain decision d. The impact of others on individual's i decision is given by the *social weight coordinates* α_{ij}^d that indicate how much individual i is influenced by an individual j when they are both making decision d. Let us denote the set of individuals who decide d in a strategy profile $\mathbf{s}$ by $s^{-1}(d) \subset \mathcal{I}$. The *utility function* $u : \mathcal{I} \times \mathcal{S} \to \mathbb{R}$ is given by

$$u(i; \mathbf{s}) = \omega_i^{s_i} + \sum_{j \in s^{-1}(s_i) \setminus \{i\}} \alpha_{ij}^{s_i}.$$

The restriction of social interactions to those individuals who make the same decision is in line with some common assumptions in the game theoretic literature, as that of Independence of Irrelevant Choices in [14], or no spillovers in [15]. These are in general assumptions in the spirit of what's most commonly known as a type of Independence of Irrelevant Alternatives assumption (which has long been used, but sometimes differs depending on the context, see for example [19]).[4] Let $\mathcal{U}$ be the space of such utility functions. For a given utility function $u \in \mathcal{U}$, the decision model we presented is a *decision game* $\Gamma \equiv \Gamma(\mathcal{I}, \mathcal{D}, u)$. We sometimes refer to decision games where there are only positive externalities as *conformity games*; and to games where there are only negative externalities as *social congestion games*.

We will study different invariances of a decision game that arise from such a utility function, and then characterize games from different invariance classes. These classes are related to how individuals may be distinguished in the game, be it either because they have different utility functions or because they have different impact on the utility function of others; or both. Following the work of Conley and Wooders ([7, 8, 9]) we start by separating those characteristics of individuals that influence the utility of others, called the crowding type of the individuals. Let C be the set of possible crowding types and let $\mathbf{c} \equiv \mathbf{c}(\Gamma) = (c_1, \ldots, c_{n_{\mathcal{I}}}) \in \mathcal{C} \equiv C^{n_{\mathcal{I}}}$ denote the *crowding profile* of individuals in a decision game Γ. Two individuals $j_1, j_2 \in \mathcal{I}$ have the same crowding type $c_{j_1} = c_{j_2} = c \in C$ if for all $i \in \mathcal{I}$ and $d \in \mathcal{D}$ we have

[4] In a game of complete information, the assumption is in fact one of *dichotomic social influence*, as we stated in the introduction. That is, individuals are influenced by other individuals who make the same decision, and also by those who make a different decision, but just by the fact that they made a different decision, independently of what decision that is. This can be seen by making a variable transformation on the social weights as is done in [23].

$\alpha^d_{ij_1} = \alpha^d_{ij_2} \equiv \alpha^d_{ic}$. The utility function for an individual $i \in \mathcal{I}$ can be rewritten using the crowding space,

$$u_i(s_i; \mathbf{s}_{-i}, \mathbf{c}_{-i}) \equiv u(i; \mathbf{s}) = \omega_i^{s_i} + \sum_{j \in s^{-1}(s_i)\setminus\{i\}} \alpha^{s_i}_{ic_j}.$$

The use of a crowding space in the characterization of a game has the advantage that the utility of an individual i associated with a strategy profile $\mathbf{s}$ is invariant under permutations of strategies of other individuals with the same crowding type. Thus, the crowding space C induces a natural equivalence relation in the strategy space $\mathcal{S}$. This allows the characterization of classes of strategies where the relevant information is the number of individuals with the same crowding type in each decision. We thus define the *crowding-aggregate decision matrix* $\mathbf{L}(\mathbf{s}, \mathbf{c})$ whose coordinates, $l^d_c = l^d_c(\mathbf{s})$, indicate the number of individuals with crowding type $c \in C$ who make decision $d \in \mathcal{D}$ in strategy profile $\mathbf{s}$,

$$\mathbf{L}(\mathbf{s}, \mathbf{c}) \equiv \begin{pmatrix} l^1_1 & \dots & l^1_{n_C} \\ \vdots & \ddots & \vdots \\ l^{n_\mathcal{D}}_1 & \dots & l^{n_\mathcal{D}}_{n_C} \end{pmatrix}.$$

We denote by $\mathcal{L} \equiv \{\mathbf{L}(\mathbf{s}, \mathbf{c}) \in \mathbb{R}^{n_\mathcal{D} \times n_C} : \mathbf{s} \in \mathcal{S}, \mathbf{c} \in \mathcal{C}\}$ the set of all possible crowding-aggregate decision matrices in a given game. Given a matrix $\mathbf{L} \in \mathcal{L}$, there is always a subset of strategy profiles $S \in \mathcal{S}$ such that, for any $\mathbf{s}_1, \mathbf{s}_2 \in S$, we have $\mathbf{L}(\mathbf{s}_1, \mathbf{c}) = \mathbf{L}(\mathbf{s}_2, \mathbf{c}) = \mathbf{L}$. Thus, the set $\mathcal{L}$ characterizes the crowding equivalence relation in the strategy space $\mathcal{S}$ induced by the crowding profile $\mathbf{c}$, and we will refer to the *strategy class* $\mathbf{L} \in \mathcal{L}$ to mean the equivalence class $\{\mathbf{s} \in \mathcal{S} : \mathbf{L}(\mathbf{s}, \mathbf{c}) = \mathbf{L}\}$. The utility function is fully characterized by the following *(reduced) utility matrix* for each individual $i \in \mathcal{I}$,

$$U_i \equiv U(i; \mathcal{D}, C) \equiv \begin{pmatrix} \omega^1_i & \alpha^1_{i1} & \dots & \alpha^1_{in_C} \\ \vdots & \vdots & \ddots & \vdots \\ \omega^{n_\mathcal{D}}_i & \alpha^{n_\mathcal{D}}_{i1} & \dots & \alpha^{n_\mathcal{D}}_{in_C} \end{pmatrix}.$$

The utility matrix defined above defines the taste (or utility) type of an individual, and the utility profile $\mathbf{U} \equiv \mathbf{U}(\mathcal{I}; \mathcal{D}, C) \equiv (U_1, \dots, U_{n_\mathcal{I}}) \in (\mathbb{R}^{n_\mathcal{D} \times (1+n_C)})^{n_\mathcal{I}}$ determines a decision game. The set of Nash equilibria of a decision game will naturally depend on the utility profile. Nevertheless, different utility profiles may lead to the same Nash equilibria. Hence, we will study properties of utility matrices of decision games for which a given strategy class is a Nash equilibrium. For this analysis it will be useful to rewrite the utility function using the strategy classes $\mathcal{L}$. Recall that in this work when we say Nash equilibria we will always mean pure Nash equilibria. As it is natural when dealing with pure Nash equilibria, we will have to make comparisons between pairs of decisions, and this can be done comparing lines in the above matrices, since each line d of those matrices is associated with the utility of the individual i when using strategy $s_i = d$. Hence, it will be useful to introduce a notation for the line vectors associated with each decision. When the choice of an individual $i \in \mathcal{I}$ is $d \in \mathcal{D}$, the social influence that she is subject to, in a given strategy profile $\mathbf{s} \in \mathcal{S}$, may be summarized by two vectors: the *social preferences vector* $\vec{\alpha}_i(d) \in \mathbb{R}^{n_C}$, comprised of the social weights given by individual i to the aggregates of each crowding type in decision d; and the *crowding-aggregate vector* $\vec{l}(d) \in \mathbb{R}^{n_C}$ whose coordinates correspond to the line d of matrix $\mathbf{L}$, and thus indicate the number of individuals with crowding $c \in C$ who make decision d in a given strategy class $\mathbf{L}$,

$$\vec{\alpha}_i(d) \equiv (\alpha_{i1}^d, \dots, \alpha_{in_C}^d), \quad \vec{l}(d) \equiv (l_1^d, \dots, l_{n_C}^d).$$

The utility function can now be rewritten for strategy classes through the above vectors. For an individual $i \in \mathcal{I}$ it is given by

$$u_i(s_i, c_i; \mathbf{L}) \equiv u_i(s_i; \mathbf{s}_{-i}, \mathbf{c}_{-i}) = \omega_i^{s_i} + \vec{\alpha}_i(s_i) \cdot \vec{l}(s_i) - \alpha_{ic_i}^{s_i}$$

where $\cdot$ denotes the usual inner product. Note that determining the utility of an individual using a strategy class $\mathbf{L}$ instead of a specific strategy profile $\mathbf{s}$, forces the need to add some extra information. Namely, each individual needs to know her own crowding type due to the subtraction of coordinate $\alpha_{ic_i}^{s_i}$. This is a consequence of removing individual i from the aggregate $l_{c_i}^{s_i}$ and assigning social weight to $l_{c_i}^{s_i} - 1$ instead. However, this only means that individual i has no social weight on her own utility, rather she has an individual value for that decision, $\omega_i^{s_i}$ (which might nevertheless encompass a social interpretation of decision values). The aforementioned need for the knowledge of an individual's own crowding type, reveals how individuals may retrieve different information from the same aggregated structure of a strategy class.

3. **Societies.** Given a strategy profile $\mathbf{s}$ and a crowding profile $\mathbf{c}$, we define *social context* as the pair $(\mathbf{s}, \mathbf{c})$. In studying social contexts that are based on a Nash equilibrium strategy $\mathbf{s}$, the characterization of the structure of an utility profile $\mathbf{U}$ is naturally limited to studying subsets of individuals that are dintinguishable in that social context, and therefore provide different information. Consider thus a partition $\mathcal{P}(\mathbf{s}, \mathbf{c})$ of the set of individuals $\mathcal{I}$ according to the social context $(\mathbf{s}, \mathbf{c})$, meaning that every pair (d, c) creates a block $P(d, c)$ of the partition whose elements are all the individuals $i \in \mathcal{I}$ with the same crowding type $c_i = c$ and using the same strategy $s_i = d$. That is,

$$P(d, c) \equiv \{i \in \mathcal{I} : (s_i, c_i) = (d, c) \in \mathcal{D} \times C\},$$
$$\mathcal{P}(\mathbf{s}, \mathbf{c}) \equiv \{P(d, c) : (d, c) \in \mathcal{D} \times C\}.$$

This kind of partitions is particularly interesting to relate to the notion of society defined in [27], and in fact inspired by it. A society is an element of a subpartition of a block $P(d, c)$ with an aditional property of convexity as defined properly below. Let us first denote convex hull by $con(\cdot)$ and without ambiguity let us use the same notation for the convex hull formed by the utilities of some individuals $J \subset \mathcal{I}$, thus

$$con(J) \equiv \left\{ \sum_{j \in J} \lambda_j U_j : \lambda_j \in \mathbb{R}_0^+ \text{ and } \sum_{j \in J} \lambda_j = 1 \right\}.$$

A set of individuals $S \in P(d, c)$ is called a *society* if it satisfies the following convexity property: if for $i \in \mathcal{I}$, $c_i = c$ and $U_i \in con(S)$, then $i \in S$ (see [27]). The society is maximal if there is no other society $S' \in P(d, c)$ such that $S \subset S'$. Given a decision game and a block $P(d, c)$ of a social context, let us denote by $SP(d, c) \equiv \{S_1, \dots, S_k\}$ a partition of $P(d, c)$. Let now

$$\mathcal{SP}(\mathbf{s}, \mathbf{c}) \equiv \bigcup_{i \in \mathcal{I}} SP(s_i, c_i).$$

The partition $\mathcal{SP}(\mathbf{s}, \mathbf{c})$ is called a *societal partition* if its blocks $SP(d, c)$ are formed by societies, and it is called a *minimal societal partition* if it is formed by maximal societies.

Definition 3.1 (Global minimum societal partition)**.** A societal partition is a *global minimum* if all its societies coincide with the $P(d,c)$ block, i.e. for all $S \in \mathcal{SP}(\mathbf{s},\mathbf{c})$, $S = P(d,c)$.

We observe that while a partition $\mathcal{P}(\mathbf{s},\mathbf{c})$ is based on a combinatorial concept, societies are based on a topological one. A fundamental question is understanding the minimal societal partition of a Nash equilibrium, and in particular if that partition is a global minimum. We will show that, in the context of our work, when there are only positive externalities between the $P(d,c)$ blocks for a given strategy, the societal partition is a global minimum. In particular, in a conformity game, the minimal societal partition of a Nash equilibrium is always a global minimum, and thus there are at most $n_{\mathcal{D}}n_C$ societies. That is not the case however for games with negative externalities. We will show that social congestion games may not have global minimum societal partitions of its Nash equilibria, and there may be up to $n_{\mathcal{I}}$ maximal societies. For a given block $P(d,c)$, let

$$U(d,c) \equiv \{U_i : i \in P(d,c)\}.$$

We say that two sets $I, J \in \mathcal{I}$ have *positive externalities* in strategy profile $\mathbf{s}$, if for all $i \in I$ and $j \in J$,

$$\alpha_{ic_j}^{s_j} + \alpha_{jc_i}^{s_i} > 0.$$

Theorem 3.2 (Positive externalities)**.** *Let* $(\mathbf{s},\mathbf{c})$ *be a social context and* $\mathbf{s}$ *a Nash equilibrium. If for two distinct decisions* $d, d' \in \mathcal{D}$ *and a crowding type* $c \in C$, $P(d,c)$ *and* $P(d',c)$ *have positive externalities in* $\mathbf{s}$*, then*

$$con(U(d,c)) \cap con(U(d',c)) = \emptyset.$$

Theorem 3.2 relates directly to the notion of societies, and in particular to the concept of global minimum societal partition.

Corollary 1 (Positive externalities)**.** *Let* $(\mathbf{s},\mathbf{c})$ *be a social context and* $\mathbf{s}$ *a Nash equilibrium. For every* $c \in C$ *let* $P(d,c)$ *and* $P(d',c)$ *have positive externalities in* $\mathbf{s}$*, for every* $d, d' \in \mathcal{D}$*, with* $d \neq d'$*. There is a global minimum societal partition.*

In particular, for every Nash equilibrium of a conformity game the minimal societal partition of a Nash equilibrium is a global minimum.

4. **Externalities and valuations.** The second step in our approach is to subdivide each utility matrix into two components, which means we will subdivide the taste type of an individual. The subdivision is done to separate the two additive components of the utility function, namely separating the part that measures the externality effects from the part that measures the individual's independent valuation of the strategy set. We will then categorize individuals according to these two components, so that we can characterize a Nash equilibrium according to the restrictions it imposes on the relation between the two components. These two components are: (i) the column vector of individual decision values $\vec{\omega}_i \equiv \omega_i(\mathcal{D}) \in \mathbb{R}^{n_{\mathcal{D}}}$; and (ii) the submatrix of social weights given to the aggregates of each crowding type, the *social preference (or externality) matrix* $e_i \equiv e_i(\mathcal{D}, C) \in \mathbb{R}^{n_C \times n_{\mathcal{D}}}$;

$$\vec{\omega}_i \equiv \begin{pmatrix} \omega_i^1 \\ \vdots \\ \omega_i^{n_{\mathcal{D}}} \end{pmatrix}, \quad e_i \equiv \begin{pmatrix} \alpha_{i1}^1 & \dots & \alpha_{in_C}^1 \\ \vdots & \ddots & \vdots \\ \alpha_{i1}^{n_{\mathcal{D}}} & \dots & \alpha_{in_C}^{n_{\mathcal{D}}} \end{pmatrix}.$$

We observe that the impact of the decision value vectors $\vec{\omega}_i$ in this relation will not be given by the precise value of their coordinates, but rather by the relative preferences they induce, namely the difference between each pair of coordinates. That is, if a given decision d is a best response for an individual i, then if we changed her vector of decision values by the same amount in each coordinate, d would still be a best response. We will take this into account using a valuation space V with the following property: if two individuals $i, j \in \mathcal{I}$ have the same *valuation type* $v_i = v_j \equiv v \in V$, then their vectors of decision values are in the same relative valuation space. More precisely, the *relative valuation space spanned by* $\vec{\omega}_i$ is

$$W(\vec{\omega}_i) \equiv \{\vec{\omega}_i + k\vec{1} : k \in \mathbb{R}\}.$$

Hence, if the two individuals have the same valuation type v, then $W(\vec{\omega}_i) = W(\vec{\omega}_j)$. However, we do not ask the equivalence class to be maximal, i.e. there might be individuals with diferent valuation types $v_i \neq v_j$ such that the corresponding vectores of decision values $\vec{\omega}_i$ and $\vec{\omega}_j$ satisfy $W(\vec{\omega}_i) = W(\vec{\omega}_j)$. With a slight abuse of notation we will refer to the decision values vector of individuals with the same valuation type v as $\vec{\omega}_v$. The profile of decision value vectors of all individuals is denoted by $\boldsymbol{\omega} \equiv \boldsymbol{\omega}(\mathcal{I}; \mathcal{D}) \equiv (\vec{\omega}_1, \ldots, \vec{\omega}_{n_\mathcal{I}}) \in (\mathbb{R}^{n_\mathcal{D}})^{n_\mathcal{I}}$. The valuation profile is denoted by $\mathbf{v} \equiv \mathbf{v}(\mathcal{I}) \equiv (v_1, \ldots, v_{n_\mathcal{I}}) \in \mathcal{V} \equiv V^{n_\mathcal{I}}$. Note that a profile of decision values might or not be compatible with a valuation profile. The set of all social preference matrices associated with the crowding profile $\mathbf{c}$ of a given game is described by the externality profile $\mathbf{e} \equiv \mathbf{e}(\mathcal{I}; \mathcal{D}, C) \equiv (e_1, \ldots, e_{n_\mathcal{I}}) \in \mathcal{E} \equiv E^{n_\mathcal{I}} \equiv (\mathbb{R}^{n_\mathcal{D} \times n_C})^{n_\mathcal{I}}$. We will use $\alpha^d_{e_i c_j}$ and $\vec{\alpha}_{e_i}(d)$ to refer, respectively, to coordinates $\alpha^d_{ic_j}$ and vector $\vec{\alpha}_i(d)$ of an individual i with externality type e_i.

The categorization of the individuals is thus given by a type map $t \equiv t_\Gamma : \mathcal{I} \to T$ which indicates the type of an individual in the type space $T = C \times E \times V$. The subscript on the type map (which we will omit) is there to reinforce that when we say type we do not mean bayesian type, rather the type map reveals symmetries of the utility profile $\mathbf{U}$, and thus it is something known a priori, since we are considering complete information games. The type map defines a type profile for the game given by the triplet $\mathbf{t} = (\mathbf{c}, \mathbf{e}, \mathbf{v})$ in the space $\mathcal{T} = (C \times E \times V)^{n_\mathcal{I}}$, composed of: (i) a crowding profile $\mathbf{c}$ characterizing individuals according to their *crowding type*; (ii) an externality profile $\mathbf{e}$ characterizing individuals according to their *externality type*; and (iii) a valuation profile $\mathbf{v}$ characterizing individuals according to their *valuation type*. Note that the utility matrix is associated with the pair (e_i, v_i), an individual's taste type. An advantage of separating the taste into two components is that now the pairs (c_i, e_i) are responsible for the 'social' part of the model; they capture the social interactions in the model. We refer to this pair as the *social type* of an individual. Hence, the valuation type component v_i, that represents the way an individual values the possible choices, may be analysed separately. A key advantage of the separate analysis of the valuation component is that it may capture exogeneous changes, as it may represent prices, taxes, product quality, road quality, marketing, political campaings, bribes, etc...

The type profile of a decision game conveys information, or imposes restrictions, on the characteristics of its Nash equilibria. On the subsequent sections we will study the information one can retrieve about the structure of the utility profile of a decision game from studying the restrictions imposed by the type profile. A first natural problem is whether individuals of the same type may use different strategies in a Nash equilibrium, hence, whether all Nash equilibria are type-symmetric. The next lemma is a result on the relation between a social type and its valuations in a Nash equilibrium. Let us define for two individuals i and j the following measure of influence in a strategy profile $\mathbf{s}$,

$$A_{ij}(\mathbf{s}) \equiv \alpha^{s_j}_{e_i c_j} + \alpha^{s_i}_{e_j c_i}.$$

If $A_{ij}(\mathbf{s}) > 0$ we say that the individuals i and j have *positive externalities* in $\mathbf{s}$. Similarly, if $A_{ij}(\mathbf{s}) < 0$ we say that individuals i and j have *negative externalities* in $\mathbf{s}$. Note that if two individuals have positive (resp. negative) externalities in $\mathbf{s}$, then at least one of them would have a positive (resp. negative) externality by changing (unilaterally) her strategy and joining the other in her decision.

Let $dist(\cdot,\cdot)$ be the distance given by the supnorm.

Lemma 4.1 (Conformity obstruction)**.** *Consider a decision game* Γ *and a Nash equilibrium* $\mathbf{s}$*. If* $i, j \in \mathcal{I}$ *and* $s_i \neq s_j$ *then*

$$dist(\vec{\omega}_{v_i}, \vec{\omega}_{v_j}) \geq A_{ij}(\mathbf{s})/2 - n_{\mathcal{I}} dist(e_i, e_j).$$

We call this an obstruction because, in the case of positive externalities, individuals need to be sufficiently different to make different decisions at a Nash equilibrium. This is not the case when negative externalities are in place. The conformity obstruction lemma leads to the following theorem for positive externalities.

Theorem 4.2 (Positive externality)**.** *Let* $\mathbf{s} \in \mathcal{S}$ *be a Nash equilibrium and* $i, j \in \mathcal{I}$ *two individuals of the same social type* $(c_i, e_i) = (c_j, e_j)$ *such that* $s_i \neq s_j$*. If* $A_{ij}(\mathbf{s}) > 0$*, then*

$$dist(W(\vec{\omega}_i), W(\vec{\omega}_j)) \geq A_{ij}(\mathbf{s})/2$$

We say that a strategy profile $\mathbf{s} \in \mathcal{S}$ is *admissible* with respect to a type profile $\mathbf{t} = (\mathbf{c}, \mathbf{e}, \mathbf{v})$ if the following property holds: if $i, j \in \mathcal{I}$ are two individuals of the same social type $(c_i, e_i) = (c_j, e_j)$ with $s_i \neq s_j$ and $A_{ij}(\mathbf{s}) > 0$, then they have different valuation types $v_i \neq v_j$. Equivalently, if $v_i = v_j$ and $A_{ij} > 0$ then $s_i = s_j$.

Corollary 2 (Nash equilibrium admissibility)**.** *A strategy* $\mathbf{s} \in \mathcal{S}$ *to be* $(\mathbf{c}, \mathbf{e}, \mathbf{v})$ *admissible is a necessary condition for* $\mathbf{s}$ *to be a Nash equilibrium.*

Given a type profile $\mathbf{t} = (\mathbf{c}, \mathbf{e}, \mathbf{v})$, we say that a strategy profile $\mathbf{s} \in \mathcal{S}$ is $\mathbf{t}$ *feasible*, if $\mathbf{s}$ satisfies the following two properties: (i) $\mathbf{s}$ is $\mathbf{t}$ admissible; and (ii) if $i, j \in \mathcal{I}$ are two individuals with different social types $(c_i, e_i) \neq (c_j, e_j)$, then $v_i \neq v_j$. (Note that this does not mean i and j have different decision values, but rather that they are allowed to have different ones.)

Theorem 4.3 (Nash equilibrium feasibility)**.** *Given a strategy profile* $\mathbf{s} \in \mathcal{S}$ *and a type profile* $\mathbf{t} \in \mathcal{T}$*, if* $\mathbf{s}$ *is* $\mathbf{t}$ *feasible then there is a profile of decision values* $\boldsymbol{\omega} \in \mathbb{R}^{n_{\mathcal{D}} \times n_{\mathcal{I}}}$ *compatible with the valuation profile* $\mathbf{v} \in \mathcal{V}$*, such that* $\mathbf{s}$ *is a Nash equilibrium.*

We note that given a type profile $\mathbf{t}$ and a strategy profile $\mathbf{s}$, to be $\mathbf{t}$ feasible is not a necessary condition for $\mathbf{s}$ to be a Nash equilibrium.

5. **Nash domains.** The *Nash Domain* $\mathcal{N}(\mathbf{s}, \mathbf{c})$ of a given social context $(\mathbf{s}, \mathbf{c})$ is defined as the set of all utility profiles $\mathbf{U}$ for which $\mathbf{s}$ is a Nash equilibrium. For an individual $i \in \mathcal{I}$ the best response domain $N_i(\mathbf{s}, \mathbf{c})$ of a social context $(\mathbf{s}, \mathbf{c})$ is the set of all utility matrices U_i such that s_i is a best response of individual i to $\mathbf{s}_{-i}$ under the crowding profile $\mathbf{c}$.

Remark 1 (Nash domain cone structure)**.** Let $(\mathbf{s}, \mathbf{c})$ be a social context. We have,

(i) $\mathcal{N}(\mathbf{s}, \mathbf{c}) = N_1(\mathbf{s}, \mathbf{c}) \times \cdots \times N_{n_{\mathcal{I}}}(\mathbf{s}, \mathbf{c})$;
(ii) if $A, B \in N_i(\mathbf{s}, \mathbf{c})$ then $\lambda A + \mu B \in N_i(\mathbf{s}, \mathbf{c})$, for all $\lambda, \mu > 0$;
(iii) if $s_i = s_j$ and $c_i = c_j$ then $N_i(\mathbf{s}, \mathbf{c}) = N_j(\mathbf{s}, \mathbf{c})$.

We note that by condition (ii) Remark 1 the best response domains $N_i(\mathbf{s}, \mathbf{c})$ have a cone structure. Let $s(I)$ (the image by the strategy map s) be the subset of decisions chosen by individuals $I \subset \mathcal{I}$ in the associated strategy profile $\mathbf{s}$. We recall that individuals with the same crowding type retrieve the same information from the aggregated structure of a strategy class $\mathbf{L}$, and if they are using the same strategy, they in fact share a best response (hence (iii)). Therefore, we can rewrite the Nash domain of a social context as follows

$$\mathcal{N}(\mathbf{s}, \mathbf{c}) = \times_{d \in s(\mathcal{I}), c \in C} N(d, c; \mathbf{L}(\mathbf{s}, \mathbf{c}))^{l_c^d}.$$

Given a crowding profile $\mathbf{c}$, a strategy profile $\mathbf{s}$ is a Nash equilibrium if, and only if, for every non-empty block $P(d, c)$ of the partition of the respective social context $(\mathbf{s}, \mathbf{c})$, we have

$$U(d, c) \subset N(d, c; \mathbf{L}).$$

We note that the best response domains $N(d, c; \mathbf{L})$ do not *preserve externalities* in the following sense: given two best response domains $N(d, c; \mathbf{L})$ and $N(d', c; \mathbf{L})$, there are some utilities in $N(d, c; \mathbf{L})$ with 'positive externalities to' some utilities in $N(d', c; \mathbf{L})$, and there are some utilities in $N(d, c; \mathbf{L})$ with 'negative externalies to' some utilities in $N(d, c'; \mathbf{L})$. Since it will be useful to study sets that preserve externalities, we will add an externality profile to the social context, extending it so that we can fiber the best response and utility Nash domains by the externality profile $\mathbf{e}$. Let $(\mathbf{s}, \mathbf{c}, \mathbf{e})$ be the *social context extension* to externality profile $\mathbf{e}$. For an individual $i \in \mathcal{I}$ the *best response valuation domain* $N(s_i, c_i, e_i; \mathbf{L}(\mathbf{s}, \mathbf{c}))$ of a social context extension $(\mathbf{s}, \mathbf{c}, \mathbf{e})$ is the set of all vectors $\vec{\omega}_i$ such that s_i is a best response to $\mathbf{s}_{-i}$, in the profile context $\mathbf{c}, \mathbf{e}$. We observe that if $\vec{\omega}_i \in N(s_i, c_i, e_i; \mathbf{L}(\mathbf{s}, \mathbf{c}))$, then $W(\vec{\omega}_i) \subset N(s_i, c_i, e_i; \mathbf{L}(\mathbf{s}, \mathbf{c}))$. Furthermore, the sets $N(s_i, c_i, e_i; \mathbf{L}(\mathbf{s}, \mathbf{c}))$ are convex, non-empty and preserve externalities. The *Nash valuation domain* of a social context extension $(\mathbf{s}, \mathbf{c}, \mathbf{e})$ is thus given by the cartesian product

$$\mathcal{N}(\mathbf{s}, \mathbf{c}, \mathbf{e}) = \times_{i \in \mathcal{I}} N(s_i, c_i, e_i; \mathbf{L}(\mathbf{s}, \mathbf{c})).$$

Theorem 5.1 (Positive externalities)**.** *Let $\mathbf{s} \in \mathcal{S}$ be a Nash equilibrium. If the individuals $i, j \in \mathcal{I}$ have the same social type $(c_i, e_i) = (c_j, e_j)$ and positive externalities in $\mathbf{s}$, with $s_i \neq s_j$, then*

$$N(s_i, c_i, e_i; \mathbf{L}(\mathbf{s}, \mathbf{c})) \cap N(s_j, c_j, e_j; \mathbf{L}(\mathbf{s}, \mathbf{c})) = \emptyset.$$

Let $\mathcal{I}_t$ be set of individuals with type $t \in T$. For a given type $t = (c, e, v) \in \mathcal{T}$, the type best response valuation domain is

$$N(t; \mathbf{L}(\mathbf{s}, \mathbf{c})) \equiv \bigcap_{i \in \mathcal{I}_t} N(s_i, c, e; \mathbf{L}(\mathbf{s}, \mathbf{c})).$$

In a strategy profile $\mathbf{s}$, individuals of type $t \in T$ are using best responses if $\vec{\omega}_v \in N(t; \mathbf{L}(\mathbf{s}, \mathbf{c}))$. If there are positive externalities within type t, Theorem 5.1 poses a problem for strategies for which $s(\mathcal{I}_t)$ is not a singleton. Thus, it is clear that being admissible with respect to the type profile, is a necessary condition for a strategy profile to be a Nash equilibrium (Corollary 2).

Theorem 5.2 (Nash domain characterization). *If* $\mathbf{s}$ *is* $\mathbf{t}$ *admissible then for every* $t \in T$, $N(t;\mathbf{L}(\mathbf{s},\mathbf{c}))$ *is a (non-empty) convex set that is the closure of an open set and*

$$\mathcal{N}(\mathbf{s},\mathbf{c},\mathbf{e}) = \times_{t\in T} N(t;\mathbf{L}(\mathbf{s},\mathbf{c})) \neq \emptyset.$$

Furthermore, if $\mathbf{s}$ *is* $\mathbf{t}$ *feasible then every* $\boldsymbol{\omega} \in \mathcal{N}(\mathbf{s},\mathbf{c},\mathbf{e}) \neq \emptyset$ *is compatible with* $\mathbf{v}$.

Theorem 4.3 follows from the above theorem. Let $\mathcal{I}_c$ be the set of individuals with a given crowding type $c \in C$. Theorem 5.2 also has an interesting conection to the number of societies in a minimal societal partition of a Nash equilibrium. Take for instance for all the individuals $i \in \mathcal{I}$, $\alpha^d_{ic} = -1$, for every $d \in \mathcal{D}$ and $c \in C$. Hence, all individuals have the same externality type given by the externality matrix with all entries -1. Since for any given $c \in C$, $N(t;\mathbf{L}(\mathbf{s},\mathbf{c}))$ contains an open set, it is possible to choose an utility profile so that we can order the utilities of all the individuals along a line in $N(t;\mathbf{L}(\mathbf{s},\mathbf{c}))$ with the order that we prefer. Each order of the individuals along the line creates a number of societies that only needs to be compatible with the combinatorics imposed by the number of individuals of $\mathcal{I}_c$ that are in each block $P(d,c)$. Thus, taking

$$M_c = \min\{2(n_c - \bar{p}_c) + 1, n_c\},$$

where $n_c = \#\mathcal{I}_c$ and $\bar{p}_c$ is the cardinality of the largest set $P(d,c) \subset \mathcal{I}_c$, we obtain the following corollary.

Corollary 3 (Negative externalities). *Given a social context* $(\mathbf{s},\mathbf{c})$*, for every* $c \in C$ *choose* q_c *such that* $\#s(\mathcal{I}_c) \leq q_c \leq M_c$. *There are utility profiles* $\mathbf{U} \in \mathcal{N}(\mathbf{s},\mathbf{c})$ *such that the minimal societal partition has cardinality* $\sum_{c\in C} q_c$.

As such, for any given social context $(\mathbf{s},\mathbf{c})$, the following minimal societal partitions can arise:

- (global minimum) there are utility profiles $\mathbf{U} \in \mathcal{N}(\mathbf{s},\mathbf{c})$ such that the minimal societal partition is the global minimum societal partition;
- (no global minimum) if for some $c \in C$ there are decisions $d, d' \in \mathcal{D}$, with $d \neq d'$, $\#P(d,c) \geq 1$ and $\#P(d',c) > 1$, then there are utility profiles $\mathbf{U} \in \mathcal{N}(\mathbf{s},\mathbf{c})$ such that there is not a global minimum societal partition;
- (maximality) if $\sum_{c\in C} q_c = n_{\mathcal{I}}$, then there are utility profiles $\mathbf{U} \in \mathcal{N}(\mathbf{s},\mathbf{c})$ such that the cardinality of the minimal societal partition is $n_{\mathcal{I}}$, and thus it is maximal.

6. **Conformity obstruction.** In this section and the next we will prove the results of the previous sections.

Let us define for any two individuals $i, j \in \mathcal{I}$ and $d \in \mathcal{D}$, the following vector,

$$\vec{\varepsilon}_{ij}(d) \equiv \vec{\alpha}_{e_i}(d) - \vec{\alpha}_{e_j}(d).$$

Lemma 6.1. *Consider a decision game* Γ *and a Nash equilibrium* $\mathbf{s}$. *For every* $i, j \in \mathcal{I}$, *if* $s_i \neq s_j$ *then*

$$\omega^{s_i}_{v_i} - \omega^{s_i}_{v_j} + \omega^{s_j}_{v_j} - \omega^{s_j}_{v_i} \geq \alpha^{s_j}_{e_i c_j} + \alpha^{s_i}_{e_j c_i} + \vec{\varepsilon}_{ij}(s_j)\cdot \mathbf{l}^{s_j} - \vec{\varepsilon}_{ij}(s_i)\cdot \mathbf{l}^{s_i}.$$

Proof. Consider a decision game Γ and let $\mathbf{s}$ be a Nash equilibrium of Γ. We have that

$$u_i(s_i;\mathbf{s}_{-i}) \geq u_i(s_j;\mathbf{s}_{-i})$$

and

$$u_j(s_j;\mathbf{s}_{-j}) \geq u_j(s_i;\mathbf{s}_{-j}).$$

Now observe that

$$u_i(s_j;\mathbf{s}_{-i}) = u_j(s_j;\mathbf{s}_{-j}) - \omega_{v_j}^{s_j} + \omega_{v_i}^{s_j} + \vec{\varepsilon}_{ij}(s_j)\cdot \mathbf{l}^{s_j} + \alpha_{e_i c_j}^{s_j}$$

and similarly

$$u_j(s_i;\mathbf{s}_{-j}) = u_i(s_i;\mathbf{s}_{-i}) - \omega_{v_i}^{s_i} + \omega_{v_j}^{s_i} - \vec{\varepsilon}_{ij}(s_i)\cdot \mathbf{l}^{s_i} + \alpha_{e_j c_i}^{s_i}.$$

which concludes the proof. □

Proof (of Lemma 4.1). Lemma 6.1 imples that

$$\omega_{v_i}^{s_i} - \omega_{v_j}^{s_i} + \omega_{v_j}^{s_j} - \omega_{v_i}^{s_j} \geq \alpha_{e_i c_j}^{s_j} + \alpha_{e_j c_i}^{s_i} - 2n_{\mathcal{I}} dist(e_i, e_j)$$

and for d equal to s_i or s_j

$$2|\omega_{v_i}^{d} - \omega_{v_j}^{d}| \geq \alpha_{e_i c_j}^{s_j} + \alpha_{e_j c_i}^{s_i} - 2n_{\mathcal{I}} dist(e_i, e_j).$$

Thus,

$$dist(\vec{\omega}_{v_i}, \vec{\omega}_{v_j}) \geq A_{ij}(\mathbf{s})/2 - n_{\mathcal{I}} dist(e_i, e_j).$$

□

6.1. **Theorem 4.2.**

Proof. Theorem 4.2 can be restated as follows: let $\mathbf{s}$ be a Nash equilibrium and $i, j \in \mathcal{I}$ be two individuals of the same social type $(c_i, e_i) = (c_j, e_j)$, such that $s_i \neq s_j$ and $A_{ij}(\mathbf{s}) > 0$. For all $\hat{\omega}_i \in W(\vec{\omega}_j)$ and $\hat{\omega}_j \in W(\vec{\omega}_i)$ and for all $0 < \varepsilon < A_{ij}(\mathbf{s})/2$, the open balls, in the l_∞ norm, centered at $\hat{\omega}_i$ and $\hat{\omega}_j$, with radius ε and $A_{ij}(\mathbf{s})/2-\varepsilon$ do not intersect,

$$B_{\varepsilon}(\hat{\omega}_i) \cap B_{A_{ij}(\mathbf{s})/2-\varepsilon}(\hat{\omega}_j) = \emptyset.$$

This follows directly from Lemma 4.1. □

6.2. **Theorems 3.2 and 5.1.**

Proof. (of Theorem 3.2) Let $(\mathbf{s}, \mathbf{c})$ be a social context and $\mathbf{s}$ be a Nash equilibrium. Observe that for an individual $i \in P(d, c)$, if her utility is replaced by any utility in $con(U(d, c))$, $s_i = d$ is still a best response. Now note that if for two distinct decisions $d, d' \in \mathcal{D}$ and crowding type $c \in C$, $P(d, c)$ and $P(d', c)$ have positive externalities in $\mathbf{s}$, then for all $U_i \in con(U(d, c))$ and $U_j \in con(U(d', c))$, the individuals i and j would also have positive externalities. Thus, by Lemma 4.1, their utilities must differ at least in decision d and d'. □

Theorem 5.1 follows from Theorem 3.2.

7. **Conformity thresholds.** For the characterization of the Nash valuation domains of a social context extension, let us start by the analysis of individual's best responses. We will then define thresholds for the valuation domains of those best responses in terms of the decision values. Given a decision game Γ and a strategy profile $\mathbf{s}$, the *best response* of individual $i \in \mathcal{I}$ is

$$\mathrm{br}_i(\mathbf{s}_{-i}) \equiv \mathrm{br}(c_i, e_i, v_i; \mathbf{L}(\mathbf{s}, \mathbf{c})) = \arg\max_{d\in\mathcal{D}} \{\omega_{v_i}^{d} + \vec{\alpha}_{e_i}(d)\cdot \vec{l}(d) - \alpha_{e_i c_i}^{d}\}.$$

A strategy profile $\mathbf{s}$ is a *(pure) Nash equilibrium* if, for every $i \in \mathcal{I}$, $s_i = \mathrm{br}_i(\mathbf{s}_{-i})$. In a given social context extension, individuals with a same social type $(c_i, e_i) = (c_j, e_j) = (c, e)$ that make the same decision $s_i = s_j = d$ have the same individual best response valuation domain $N(d, c, e; \mathbf{L}) \equiv N(s_i, c_i, e_i; \mathbf{L}(\mathbf{s}, \mathbf{c}))$. Note that the Nash domains of social contexts are characterized by the best response valuation domains of social context extensions, since

$$N(d, c; \mathbf{L}) = \bigcup_e N(d, c, e; \mathbf{L}).$$

To characterize the best response valuation domains, we are going to define for a given strategy profile $\mathbf{s}$ conformity thresholds $T_{e_i}(s_i \to d; \mathbf{s}_{-i})$, that represent the surplus quantity that individual i has from social preferences, that could create an incentive for her to change from her current decision s_i to decision d. This threshold does not depend on the valuation type of the individual, but rather on the externality context $(\mathbf{s}, \mathbf{c}, \mathbf{e})$. In particular, as refered, it depends on the individual social type and the strategy class to which $\mathbf{s}$ belongs. Let us first define the *auxiliar externality type-threshold* between two decisions $d, d' \in \mathcal{D}$,

$$T_e(d', d; \mathbf{L}) \equiv \vec{\alpha}_e(d) \cdot \vec{l}(d) - \vec{\alpha}_e(d') \cdot \vec{l}(d').$$

Given a strategy profile $\mathbf{s}$, the *conformity thresholds* are given for each individual $i \in \mathcal{I}$ with externality type e_i and for all decisions $d \in \mathcal{D} \setminus \{s_i\}$, by

$$T_{e_i}(s_i \to d; \mathbf{s}_{-i}) \equiv T_{e_i}(s_i, d; \mathbf{L}(\mathbf{s}, \mathbf{c})) + \alpha^{s_i}_{e_i c_i},$$

which will be useful to rewrite using strategy classes,

$$T_{(c_i, e_i)}(s_i \to d; \mathbf{L}(\mathbf{s}, \mathbf{c})) \equiv T_{e_i}(s_i \to d; \mathbf{s}_{-i}).$$

The notation reflects the idea of social incentive towards decision d from strategy s_i. Thus, this is the quantity by which $\omega^{s_i}_{v_i}$ (the value of decision s_i) has to overcome $\omega^d_{v_i}$ (the value of decision d), so that decision s_i is still 'preferable' for an individual with social type (c_i, e_i) in the externality context $(\mathbf{s}, \mathbf{c}, \mathbf{e})$. Observe that when we talk about incentives for player i to change her decision, we might be talking about desincentives, depending upon the sign of the conformity threshold $T_{(c_i, e_i)}(s_i \to d; \mathbf{L}(\mathbf{s}, \mathbf{c}))$. Two opposite extreme cases appear when $\vec{\alpha}_{e_i}(s_i)$ has only positive coordinates and $\vec{\alpha}_{e_i}(d)$ has only negative coordinates, making the threshold negative, thus a desincentive to change; or when the opposite happens, making the threshold positive, thus an incentive to change. Concluding, incentives or desincentives are provoked by the relation between negative and positive coordinates in the social preference matrix.

Lemma 7.1 (Best response valuation domains characterization)**.** *The best response valuation domains $N(d, c, e; \mathbf{L})$ consist of all $\vec{\omega} \in \mathbb{R}^{n_{\mathcal{D}}}$ with the following properties:*

(i) $\omega^d \in \mathbb{R}$*;*
(ii) $\omega^{d'} \in \mathbb{R}$ *satisfying the following threshold inequality*

$$\omega^{d'} \le \omega^d - T_{(c,e)}(d \to d'; \mathbf{L}) \tag{1}$$

for every decision $d' \in \mathcal{D} \setminus \{d\}$.

Hence, $\mathcal{N}(\mathbf{s}, \mathbf{c}, \mathbf{e})$ is non-empty and contains an open set in the space $(\mathbb{R}^{n_{\mathcal{D}}})^{n_{\mathcal{I}}}$.

Proof. The strategy profile $\mathbf{s}$ is a Nash equilibrium if, and only if,

$$u_i(s_i, \mathbf{s}_{-i}) \ge u_i(d, \mathbf{s}_{-i})$$

for every $d \in D$ and $i \in \mathcal{I}$. Let $t_i = (c_i, e_i, v_i) = (c, e, v)$, the utility function can be rewritten explicitly as

$$u(i;\mathbf{s}) = \omega_v^{s_i} + \alpha_{ec}^{s_i}(l_c^{s_i} - 1) + \sum_{c' \neq c}^{n_C} \alpha_{ec'}^{s_i} l_{c'}^{s_i}(\mathbf{s}).$$

Letting $t = t_i$ and $l_t^d = l_t^d(\mathbf{s})$, we get

$$\omega_v^{s_i} - \alpha_{ec}^{s_i} + \sum_{c'=1}^{n_C} \alpha_{ec'}^{s_i} l_{c'}^{s_i} \geq \omega_v^d + \sum_{c'=1}^{n_C} \alpha_{ec'}^{d} l_{c'}^{d} \ .$$

Rearranging the terms, the previous inequality is equivalent to

$$\omega_v^{s_i} \geq \omega_v^d + \alpha_{ec}^{s_i} + \sum_{c'=1}^{n_C} \left(\alpha_{ec'}^{d} l_{c'}^{d} - \alpha_{ec'}^{s_i} l_{c'}^{s_i}\right) \ .$$

Hence, s_i is a best response for i if for every decision d

$$\omega_v^d \leq \omega_v^{s_i} - T_{(c,e)}(s_i \to d; \mathbf{L}).$$

□

7.1. **Theorems 4.3 and 5.2.** Let $\mathbf{t}$ be a type profile and $\mathbf{s}$ be a $\mathbf{t}$ admissible strategy profile, and denote its strategy class by $\mathbf{L} \equiv \mathbf{L}(\mathbf{s}, \mathbf{c})$. For every $t \in T$, let $\mathcal{S}_t \equiv \{s_i \in \mathcal{D} : i \in \mathcal{I}_t\}$ and let us use the subscript t on the parameters to mean the corresponding coordinate of t, for example, if $t = (c, e, v)$, then α_{tt}^d means α_{ec}^d. Since $\mathbf{s}$ is $\mathbf{t}$ admissible, for each type $t \in T$, there is at most one decision $d \in \mathcal{S}_t$ such that $\alpha_{tt}^d > 0$ (if there were two, they would violate the admissibility condition on the valuation map). Let us start by defining, for every type $t \in \mathcal{T}$,

$$d_t^* \equiv \arg\max_{d \in \mathcal{S}_t} \{\alpha_{tt}^d\}.$$

Let $i^* \in \mathcal{I}_t$ be an individual such that $s_{i^*} = d_t^*$, and let

$$\epsilon_t(\mathbf{s}) \equiv \begin{cases} 0 & \text{if } \alpha_{tt}^{d_t^*} \geq 0; \\ \\ -\dfrac{\alpha_{tt}^{d_t^*}}{2} & \text{if } \alpha_{tt}^{d_t^*} < 0. \end{cases}$$

Let $\Omega \equiv \times_{t \in T} \Omega_t$, where for a given type $t \in T$, Ω_t are the open sets of all $\boldsymbol{\omega}_t$ with the following properties:

(i) $\omega_t^{d_t^*} \in \mathbb{R}$;

(ii) if $\mathcal{D} \setminus \mathcal{S}_t \neq \emptyset$ then, for every $d \in \mathcal{D} \setminus \mathcal{S}_t$,

$$\omega_t^d \leq \min_{s_i \in \mathcal{S}_t} \{\omega_t^{s_i} - T_{(c,e)}(s_i \to d; \mathbf{L})\}; \tag{2}$$

(iii) if $\mathcal{S}_t \setminus \{d_t^*\} \neq \emptyset$, then, for every $s_i \in \mathcal{S}_t \setminus \{d_t^*\}$

$$\omega_t^{d_t^*} + T_{(c,e)}(s_i \to d_t^*; \mathbf{L}) + \epsilon_t(\mathbf{s}) \leq \omega_t^{s_i} \leq \omega_t^{d_t^*} - T_{(c,e)}(d_t^* \to s_i; \mathbf{L}) - \epsilon_t(\mathbf{s}). \tag{3}$$

Proof. (of Theorem 5.2) The proof is constructed over one type $t = (c, e, v) \in T$ by showing that $\emptyset \neq \Omega_t \in N(t; \mathbf{L})$, which holds for all $t \in T$, and thus $\Omega \in \mathcal{N}(\mathbf{s}, \mathbf{c}, \mathbf{e})$. As we will be refering always to the same type and to the same strategy, let us, for simplicity of notation, omit the subscript and the strategy class, hence, denote

$$T(d \to d') \equiv T_{(c,e)}(d \to d'; \mathbf{L})$$

Let's start by showing that $\Omega_t \neq \emptyset$. Observe that it is enough to show that equation (3) in the defintion of Ω_t refers to a non-degenerated interval, which translates into

$$-T(d_t^* \to s_i) - T(s_i \to d_t^*) \geq 2\epsilon_t(\mathbf{s}).$$

Hence, as

$$-T(d_t^* \to s_i) - T(s_i \to d_t^*) = -\alpha_{tt}^{s_i} - \alpha_{tt}^{d_t^*},$$

we get

$$-\alpha_{tt}^{s_i} - \alpha_{tt}^{d_t^*} \geq 0 \text{ when } \alpha_{tt}^{d_t^*} \geq 0,$$

and

$$-\alpha_{tt}^{s_i} - \alpha_{tt}^{d_t^*} \geq -\alpha_{tt}^{d_t^*} \text{ when } \alpha_{tt}^{d_t^*} < 0.$$

Now recall that being $\mathbf{t}$ admissible implies that for individuals i and j of the same type using different strategies $A_{ij}(\mathbf{s}) = \alpha_{tt}^{s_i} + \alpha_{tt}^{s_j} \leq 0$. As $\mathbf{s}$ is $\mathbf{t}$ admissible, there is for each type $t \in \mathcal{T}$ at most one decision $d \in \mathcal{S}_t$ such that $\alpha_{tt}^{d} > 0$, and that decision is by definition d_t^*, hence, $\alpha_{tt}^{s_i} \leq 0$.

To see that $\Omega \subset \mathcal{N}(\mathbf{s}, \mathbf{c}, \mathbf{e})$ we will show that the two equations setforth in the definition of the sets Ω_t are sufficient to guarantee that inequalities (1) in lemma 7.1 are satisfied for every individual and every decision. It is straightforward to see from equation (2) that no individual wants to change to decisions $d \notin \mathcal{S}_t$. Let's now check that equation (3) implies that individuals do not want to change between decisions within $\mathcal{S}_t$. Individuals choosing d_t^* do not want to change to other decisions in $\mathcal{S}_t$, since $\epsilon_t(\mathbf{s}) \geq 0$ and

$$\omega_t^{d_t^*} \geq \omega_t^{s_i} + T(d_t^* \to s_i) + \epsilon_t(\mathbf{s}).$$

An individual $i \neq i^* \in \mathcal{I}_t$ doesn't want to change to d_t^*, since $\epsilon_t(\mathbf{s}) \geq 0$ and

$$\omega_t^{s_i} \geq \omega_t^{d_t^*} + T(s_i \to d_t^*) + \epsilon_t(\mathbf{s}).$$

Finally, to see that an individual $i \neq j \in \mathcal{I}_t$ does not want to change to any decision $s_j \neq d_t^*$,

$$\omega_t^{s_i} - \omega_t^{s_j} \geq T(s_i \to d_t^*) + T(d_t^* \to s_j) + 2\epsilon_t(\mathbf{s}),$$

but

$$T(s_i \to d_t^*) + T(d_t^* \to s_j) = T(s_i \to s_j) + \alpha_{tt}^{d_t^*},$$

hence,

$$\omega_t^{s_i} \geq \omega_t^{s_j} + T(s_i \to s_j).$$

□

Theorem 4.3 follows from 5.2.

Acknowledgments. The authors thank the helpful comments of an anonymous referee. The authors would like to thank to LIAAD - INESC TEC and to acknowledge the financial support received by the FCT Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project UID/EEA/50014/2013 and ERDF (European Regional Development Fund) through the COMPETE Program (operational program for competitiveness) and by National Funds through the FCT within Project "Dynamics, optimization and modelling", with reference PTDC/MAT-NAN/6890/2014. Alberto Adrego Pinto also acknowledges the financial support received through project University of Porto and University of São Paulo and the Special Visiting Researcher Program (Bolsa Pesquisador Visitante Especial – PVE) "Dynamics, Games and Applications", with reference 401068/2014-5 (call: MEC/MCTI/CAPES/CNPQ/FAPS), at IMPA, Brazil.

A. S. Mousa thanks the financial support of Birzeit University. Renato Soeiro acknowledges the support of FCT, the Portuguese national funding agency for science and technology, through a Ph.D. grant with reference SFRH/BD/88742/2012.

REFERENCES

[1] L. Almeida, J. Cruz, H. Ferreira and A. A. Pinto, Bayesian-Nash equilibria in theory of planned behaviour, *Journal of Difference Equations and Applications*, **17** (2011), 1085–1093.

[2] A. V. Banerjee, A simple model of herd behavior, *The Quarterly Journal of Economics*, **107** (1992), 797–817.

[3] B. D. Bernheim, A theory of conformity, *Journal of Political Economy*, **102** (1994), 841–877.

[4] M. Le Breton and S. Weber, Games of social interactions with local and global externalities, *Economics Letters*, **111** (2011), 88–90.

[5] J. G. Brida, M. J. Such-devesa, M. Faias and A. Pinto, Strategic Choice in Tourism with Differentiated Crowding Types, *Economics Bulletin*, **30** (2010), 1509–1515.

[6] W. Brock and S. Durlauf, Discrete choice with social interactions, *Review of Economic Studies*, **68** (2011), 235–260.

[7] J. P. Conley and M. Wooders, Taste-homogeneity of optimal jurisdictions in a Tiebout economy with crowding types and endogenous educational investment choices, *Ricerche Economiche*, **50** (1996), 367–387.

[8] J. P. Conley and M. H. Wooders, Equivalence of the core and competitive equilibrium in a tiebout economy with crowding types, *Journal of Urban Economics*, **41** (1997), 421–440.

[9] J. P. Conley and M. H. Wooders, Tiebout economies with diferential inherent types and endogenously chosen crowding characteristics, *Journal of Economic Theory*, **98** (2001), 261–294.

[10] R. Cooper and A. John, Coordinating coordination failures in keynesian models, *The Quarterly Journal of Economics*, **103** (1988), 441–463.

[11] M. S. Granovetter, The strength of weak ties, *American Journal of Sociology*, **78** (1973), 1360–1380.

[12] M. Granovetter, Threshold models of collective action, *The American Journal of Sociology*, **83** (1978), 1420–1443.

[13] J. T. Howson, Equilibria of polymatrix games, *Management Science*, **18** (1972), 312–318.

[14] H. Konishi, M. Le Breton and S. Weber, Equilibria in a model with partial rivalry, *Journal Of Economic Theory*, **72** (1997), 225–237.

[15] H. Konishi, M. Le Breton and S. Weber, Pure strategy nash equilibrium in a group formation game with positive externalities, *Games and Economic Behavior*, **21** (1997), 161–182.

[16] I. Milchtaich, Congestion models with player specific payoff functions, *Games and Economic Behavior*, **13** (1996), 111–124.

[17] L. G. Quintas, A note on polymatrix games, *International Journal of Game Theory*, **18** (1989), 261–272.

[18] T. Quint and S. Shubik, *A Model of Migration*, (1994) Working paper, Cowles Foundation, Yale University.

[19] P. Ray, Independence of irrelevant alternatives, *Econometrica*, **41** (1973), 987–991.

[20] R. W. Rosenthal, A class of games possessing pure-strategy nash equilibria, *International Journal of Game Theory*, **2** (1973), 65–67.

[21] T. C. Schelling, Dynamic models of segregation, *Journal of Mathematical Sociology*, **1** (1971), 143–186.

[22] T. C. Schelling, Hockey helmets, concealed weapons, and daylight savings- a study of binary choices with externalities, *The journal of Conflict Resolution*, **17** (1973), 381–428.

[23] R. Soeiro, A. Mousa, T. R. Oliveira and A. A. Pinto Dynamics of human decisions, *Journal of Dynamics and Games*, **1** (2014), 121–151.

[24] M. H. Wooders, A tiebout theorem, *Mathematical Social Sciences*, **18** (1989), 33–55.

[25] M. H. Wooders, Equivalence of Lindahl equilibria with participation prices and the core, *Economic Theory*, **9** (1997), 115–127.

[26] M. H. Wooders, Multijurisdictional economies, the Tiebout Hypothesis, and sorting, *Proceedings of the National Academy of Sciences*, **96** (1999), 10585–10587.

[27] M. Wooders, E. Cartwright and R. Selten, Behavioral conformity in games with many players, *Games and Economic Behavior*, **57** (2006), 347–360.

[28] M. Wooders and E. Cartwright, Correlated equilibrium, conformity, and stereotyping in social groups, *Journal of Public Economic Theory*, **16** (2014), 743–766.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: desmarque@gmail.com
E-mail address: asaid@birzeit.edu
E-mail address: aapinto@fc.up.pt