THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 1 Capitole
Discipline : Sciences Economiques

Présentée et soutenue par

Yann KERVINIO
Le 22 janvier 2016

Titre :

Assessing the fairness of public policies

Proposal for an approach with an illustration for the location of locally undesirable land uses

---

Ecole doctorale
Toulouse School of Economics

Unité de recherche :
LERNA – TSE

Directeur de Thèse
Stefan AMBEC (Directeur de recherche, INRA)

Rapporteurs

Marc FLEURBAEY (Professeur, Princeton University)

Juan D. MORENO-TERNERO (Professeur, Universidad Pablo de Olavide)

Membres du Jury

Dominique BUREAU (Professeur chargé de cours, Ecole Polytechnique et délégué général, Conseil économique pour le développement durable)

Michel LE BRETON (Professeur, Université Toulouse 1 Capitole)

Thèse réalisée dans le cadre de la formation doctorale des ingénieurs des Ponts, des Eaux et des Forêts

Financée par le Ministère de l'Ecologie, du Développement durable et de l'Energie

Réalisée au Laboratoire d'économie des ressources naturelles
Ecole d'économie de Toulouse
21, allée de Brienne
31000 Toulouse

Sous la direction de Stefan Ambec, directeur de recherche (INRA)
stefan.ambec@toulouse.inra.fr

Cette version en date du 10 avril 2016 inclut des corrections réalisées après la soutenance.

Abstract

This thesis is about the assessment of the fair character of public policies whose distributive impacts are well understood. Because of the diversity of existing judgments on this matter, the relevance of such assessments is a crucial question. This work rests upon the theory of fair allocations and the empirical study of fairness judgments. After presenting each of these approaches and discussing how they could contribute to form practical policy recommendations in three short chapters, three contributions to the literature are presented. All are motivated by the so-called NIMBY ("Not In My Back-Yard") problem, in which a group of communities faces the opportunity of implementing an economically beneficial project that is locally undesirable (e.g. a wastewater treatment plant, a landfill, a wind park, etc).

In the first article, I consider the problem of allocating a single, indivisible project and sharing its benefit among communities with an equal right on it but featuring different provision costs. The differences in these costs may arise from variations in building, operation and maintenance costs for the project but also from differences in the communities' compensation requirements for hosting the project. In this setting, I characterise three allocation rules that correspond to three prominent cooperative solution concepts: the welfare egalitarian solution, the nucleolus and the Shapley value. The principles invoked involve Pareto efficiency, Anonymity, No envy, and axioms of solidarity or reward related to the communities' provision costs. The results clarify how considerations over the nature of the costs could influence fair allocations. The analysis is then extended to settings with asymmetric information and to setting with costs of several kinds. In each extension, I propose and motivate a fair solution. The results of a survey motivated by this analysis are eventually presented.

In the second article, I study individual fairness judgments and preferences for the allocation of an indivisible task and its benefit among two individuals with a different willingness to perform it. My approach relates results from the theory of fair allocations with the empirical observation of fairness judgments and preferences. I first present and motivate four contrasted allocation rules as possible rules of judgment. For some of the participants, a questionnaire was proposed before they knew about their situation. Among the four rules of judgment proposed, the welfare egalitarian allocation rule is the most preferred allocation rule as stated by the participants. Yet, I also observe support for principles that are not compatible with this rule: an important proportion of respondents deem fair to give nothing to someone who would not be willing to perform the task, and another substantial proportion deem fair to split the benefit of the task equally when both participants feature the same compensation requirement. In the experiment, participants had the opportunity to perform a task for pay. However, for any two of them, a single task was available. As required by the Pareto principle, it was allocated to the participant with the lower compensation requirement. In this situation, the stated normative expectations of the task performer are found to be higher, the greater the discrepancy between the compensation requirements. This does not extend to individual distributive preferences as revealed by the offers in a dictator setting. I also find that the task performers who took the questionnaire would deem the equal split fair less often. Overall, few respondents are consistent with any of the four rules proposed.

In the third article, jointly written with Stefan Ambec, we consider the decentralised provision of a global public good with local externalities in a spatially explicit model. Communities decide on the location of a facility that benefits everyone but exhibits costs to the host and its neighbors. They share the costs through transfers. We examine cooperative games associated with this so-called NIMBY problem. We derive and discuss conditions for core solutions to exist. These conditions are driven by the temptation to exclude groups of neighbors at any potential location. We illustrate the results in different spatial settings. These results clarify how property rights can affect cooperation and shed further light on a limitation of the Coase theorem.

Keywords: social choice, justice, fairness, equity, justification, reflective equilibrium, axiomatic analysis, theory of fair allocations, cooperative games, core, empirical social choice, economic experiment, NIMBY, pollution, waste, resource allocation.

Résumé

Cette thèse porte sur l'évaluation du caractère équitable de politiques publiques dont les impacts distributifs sont bien compris. Du fait de la diversité des jugements existants en la matière, la question de la pertinence de telles évaluations est cruciale. Ce travail s'appuie sur la théorie des allocations justes et l'étude empirique des jugements en matière d'équité. Après avoir présenté ce en quoi consistent ces deux approches et en quoi elles peuvent contribuer à la formation de recommandations pratiques en matière de politiques publiques, trois contributions à la littérature sont présentées. Toutes sont motivées par les problèmes, parfois qualifiés de "NIMBY" ("Not In My Back-Yard"), dans lesquels un groupe de communautés a la possibilité d'entreprendre un projet bénéfique sur le plan économique mais qui reste indésirable d'un point de vue local (par exemple, une station d'épuration, un centre d'enfouissement technique, un parc éolien, etc.).

Dans le premier article, je considère le problème de l'allocation d'un unique projet indivisible et des bénéfices associés entre des communautés a priori toutes aussi légitimes pour l'accueillir mais présentant différents coûts de mise en œuvre. Les différences dans ces coûts peuvent résulter de variations dans les coûts de construction, de fonctionnement et de maintenance du projet en leur sein mais aussi de différences dans leurs exigences de compensation pour accueillir le projet. Dans ce cadre, je caractérise trois règles d'allocation qui correspondent à trois concepts majeurs de la théorie des jeux coopératifs : la règle d'allocation égalitaire, le nucléole et la valeur de Shapley. Les principes invoqués comprennent l'efficacité au sens de Pareto, l'anonymité, l'absence d'envie, ainsi que des principes de compensation ou de récompense relatifs aux coûts de mise en œuvre des communautés. Ces résultats clarifient comment des considérations portant sur la nature des différents coûts peuvent amener à recommander des allocations différentes. L'analyse est ensuite étendue à des situations d'asymétrie d'information et à des situations avec des coûts de différente nature. Dans chacune de ces extensions, je propose et motive une solution équitable. Enfin, les résultats d'une enquête motivée par cette analyse sont présentés.

Dans le deuxième article, j'étudie les jugements et les préférences en matière d'équité pour l'allocation d'une tâche indivisible et de ses bénéfices entre deux personnes a priori aussi légitimes l'une que l'autre pour l'entreprendre mais différemment disposées à la réaliser. L'approche proposée met en relation des résultats de la théorie des allocations justes et l'observation empirique des jugements et des préférences en matière d'équité. Je commence par présenter et motiver quatre règles d'allocation contrastées comme des règles de jugement possibles. Pour certains participants, un questionnaire était proposé avant qu'ils ne connaissent leur propre situation. Parmi les quatre règles proposées, les participants adhèrent en majorité au partage égalitaire du bien-être. Cependant, j'observe par ailleurs un support pour des principes qui ne sont pas compatibles avec cette règle : une proportion importante des participants estiment juste de ne rien allouer à quelqu'un qui n'aurait pas souhaité entreprendre la tâche, et une autre proportion substantielle estime juste de partager le bénéfice de la tâche également lorsque les deux participants ont les mêmes exigences de compensation. Dans le contexte de l'expérience, les participants avaient l'opportunité de réaliser la tâche contre rémunération. Cependant, pour chaque paire de participants, une seule tâche était disponible. Comme requis par le principe de Pareto, cette tâche était allouée au participant dont l'exigence de compensation était la plus basse. Dans cette situation, j'observe que les attentes normatives des participants amenés à faire la tâche sont d'autant plus importantes lorsque la différence entre les exigences de compensation des deux participants est élevée. Cette observation ne s'étend pas aux préférences révélées par les offres dans le jeu du dictateur. J'observe aussi que les participants ayant répondu au questionnaire sont moins nombreux à retenir le partage égal. D'une manière générale, peu de participants sont cohérents avec une seule des quatre règles d'allocation proposées.

Dans le troisième article, coécrit avec Stefan Ambec, nous considérons la fourniture décentralisée d'un bien public global présentant des externalités locales dans un modèle spatial explicite. Des communautés décident de la localisation d'une infrastructure qui bénéficie à tous mais présente des coûts pour la communauté d'accueil et ses voisins. Ces coûts peuvent être partagés par des transferts. Nous nous intéressons aux jeux coopératifs induits par ces situations qualifiées de problème "NIMBY". Nous obtenons et discutons des conditions pour que des solutions du cœur du jeu existent. Ces conditions découlent de la tentation d'exclure les groupes de communautés au voisinage de toute localisation potentielle. Nous illustrons ces résultats dans différents contextes spatialisés. Ces résultats permettent de clarifier comment l'allocation des droits de propriété peut affecter la coopération et apportent un éclairage supplémentaire sur une limite du théorème de Coase.

Keywords: choix social, justice, équité, justification, équilibre réfléchi, analyse axiomatique, théorie des allocations justes, jeu coopératif, cœur, choix social empirique, expérience économique, NIMBY, pollution, déchets, allocation des ressources.

Remerciements

C'est avant tout à Stefan Ambec, mon directeur de thèse, que j'adresse mes remerciements pour avoir accepté de me suivre et me conseiller tout au long de ce travail. C'est en grande partie grâce à sa disponibilité, ses conseils et son soutien que j'ai pu mener ce travail à bien.

Plus généralement, je tiens à remercier l'ensemble de l'équipe du Laboratoire d'économie des ressources naturelles et l'ensemble de la communauté de l'école d'économie de Toulouse et du récent institut d'études avancées. L'enthousiasme et le dynamisme de ces équipes contribuent à créer une atmosphère de recherche stimulante dont j'ai grandement bénéficié. Je souhaite en particulier remercier Henrik Andersson, Loïc Batté, Charlotte Cavaillé, Sylvain Chabé-Ferret, Astrid Hopfensitz, Shagun Khare, Thibaut Laurent, Michel Le Breton, Yves Le Yaouanq, César Mantilla, Christoph Rheinberger, François Salanié, Peter Schwardman, Nicolas Treich, Karine Van der Straeten, Boris Van Leuwen et Takuro Yamashita pour des discussions et conseils qui m'ont été précieux dans l'avancement de ce travail.

Parmi ces personnes, il en est à qui je souhaite exprimer une gratitude toute particulière. C'est le cas de Michel Le Breton, pour ses conseils avisés et son suivi tout au long de ce travail. Son enthousiasme communicatif pour les problèmes de partage équitable et les questions de choix social m'a été précieux pour avancer. Je tiens par ailleurs à remercier Astrid Hopfensitz pour son soutien et ses conseils sur le projet d'expérience. En plus d'être pour moi un grand plaisir, les nombreux échanges que nous avons pu avoir ont largement contribué à ma réflexion.

Concernant la réalisation de l'expérience, je tiens aussi à remercier Stéphane Robin pour avoir pris le temps de me fournir de précieux conseils sur le protocole de l'expérience et Stéphane Cézéra pour son appui essentiel dans la mise en œuvre de ce projet.

Du côté de l'école doctorale, il doit être rare de trouver une section de remerciements sans mention d'Aude Schloesing. Cette thèse ne fait pas exception, même si cette mention risque malheureusement de se raréfier à l'avenir. Je lui souhaite une très bonne retraite.

Ces remerciements ne seraient pas complets si j'omettais de remercier les personnes ayant accepté de participer au suivi de ce travail. Je suis reconnaissant aux membres du comité de pilotage, Mathieu Glachant, Michel Le Breton et Alban Thomas pour leur participation en début de projet. Je tiens aussi à remercercier les membres du jury, Dominique Bureau, Marc Fleurbaey, Michel Le Breton et Juan D. Moreno-Ternero, pour avoir accepté de prendre part à l'évaluation de ce travail.

Je tiens par ailleurs à remercier le corps des ingénieurs des ponts, des eaux et des forêts pour avoir rendu ce travail possible. Je remercie les membres du bureau de la commission de formation doctorale des IPEF pour l'animation de la formation doctorale, et en particulier Cyril Kao et Alexandre Pery pour leur suivi tout au long de ce projet.

Je souhaite aussi manifester ma reconnaissance à ceux que j'ai eu le plaisir de fréquenter au quotidien durant ma thèse. A Ananya et Tuba, aux côtés de qui j'ai pris grand plaisir à travailler, à Alissa, pour son organisation du groupe de lecture de philosophie politique, à eux et à tous les autres, merci.

Je tiens enfin à remercier ma famille et mes amis pour leur soutien précieux.

Contents

List of Figures

List of Tables

Introduction

"When the pressure of public opinion seems to force the participants to the obviously fair or reasonable solution, we may exaggerate the pressure or at least misunderstand the way it works on participants unless we give credit to its power to coordinate the participants' expectations."

(Thomas Schelling, 1960, The strategy of conflict)

"Being designed to reconcile by reason, justification proceeds from what all parties to the discussion hold in common. Ideally, to justify a conception of justice to someone is to give him a proof of its principles from premises that we both accept, these principles having in turn consequences that match our considered judgments."

(John Rawls, 1971, A Theory of Justice)

As such, it does not cost anything to say that public decisionmakers should act on behalf of the *common good,* the *public interest*, the *social welfare* or - as will sound familiar to the French citizens - *l'intérêt général.* Nobody would oppose either that *justice*, *equity* and *fairness* are essential requirements of public decisions.[1] Yet, how all these notions operate in trading-off competing interests and conciliating conflicting values in practice is far from consensual. This is precisely here, when defining a fair social objective and deriving its practical implications, that the hard part begins. In this thesis, I seek to propose a method for assessing the relative fairness of different policy options with well-understood distributive impacts. Such an assement may be called a *social judgment*.

In this introduction, the objective is to make explicit what this endeavour requires. I begin by reminding how the classical Arrow theorem establishes that no natural solution exists to the problem of formulating a social judgment out of the diverse individual values that may exist in society. My conclusion therefore is that we cannot avoid being more specific about the nature of the political community concerned by the policies under scrutiny. The two introductory quotes are intended to convey the main intuition for this. As suggested by the first quote by Thomas Schelling, the most natural solution is to prevail in some policy environments. In these instances, a social judgment based on an elaborate justification is most likely to be disparaged by the stakeholder as a counterproductive move away from a settlement. On the contrary, the perspective proposed by Rawls suggests that in other environments, a seemingly natural solution may not resist a justification in favour of an alternative option. Following Sugden (2013), I then propose three contrasted perspectives on the democratic political community. I argue that the relevant methods for motivating a social judgment differ radically in each of these perspectives. I eventually present the outline of this thesis, in which I chose to focus on the methods that are relevant to the deliberative perspective on the democratic political community.

On the need to cope with ethical arguments

The social choice approach is an attempt to provide a compelling rational argument about how a single social objective could be formed out out the diverse, conflicting, and supposedly unalterable individual values that may exist in society in a fair way. It formally reflects on the fair terms of *aggregating* these conflicting (ordinal) views over all the feasible policy options into one consistent ordering over these options, which we may call a *social judgment*. These fair terms essentially consist of four requirements: the universal domain

[1] In this thesis, the terms *justice*, *equity* and *fairness* may next be used indifferently. They denote the reasons, or norms that should guide the adjudication of the conflicting interests or values that exist in society.

condition[2], non-dictatorship[3], the Pareto principle[4], and an independence condition[5]. The Arrow theorem establishes that no aggregation rule meets all these criteria (Arrow, 1951). This means that any attempt to formulate a social judgment over alternative policy options on the basis of the diverse views of the citizens would have to deviate from this framework or to concede on at least one of the former principles.

As an instance, the theory of fair allocations is one such approach that proposed to characterize fair allocations rules in specific contexts out of additional intuitively appealing fairness principles that are compatible with the Pareto principle. It clarifies the logical articulation of these principles along with the necessary trade-offs that exist between them in order to form a well-defined judgment. As emphasised by Fleurbaey et al (2005), this approach most significantly rests upon a relaxation of the independence condition in the Arrowian framework. In this thesis, I will hypothesize that individual values are suceptible of change and I will argue that another significant departure of the theory of fair allocations from the Arrovian framework lies in the different interpretation attached to this particular approach. Indeed, the theory of fair allocations rests upon a representation of individual welfare that is most often interpreted as preferences or tastes, and rarely in terms of individual values. In doing so, the question of conciliating the conflicting value judgments behind the choice of principles is explicitly kept external to the theory. In a way, such an approach returns the question of how to conciliate individual values to the public arena and paves the way for many different perspectives on this problem, maybe recognizing that such a process would have to rely on some form of ethical argument, the validity of which could always be questionned.

Three perspectives on a democratic political community

In order to avoid an intricate and unavoidably fragile meta-ethical exercise, I shall follow Rawls (1993) in viewing this problem as primarily political and presume the existence of a unanimously shared sense of the political community in the policy context at hand. I shall first presume that this political community should be *democratic* in the sense that, not only the welfare, but also the judgments of all the people who have interests at stakes are the relevant judgments to consider in the formation of a social judgment.[6] This is

[2]The universal domain condition states that no restriction should be placed a priori on individual judgments.

[3]Non-dictatorship, requires that no individual shall be able to impose her own view over all pairs of alternatives regardless of other's judgments.

[4]The Pareto principle requires that if everyone prefers one alternative over another one, with some expressing a strict preference, then the social ordering shall strictly favor the former alternative.

[5]This independence condition requires that for any change in the profile of individual judgments that leaves all individual preferences over a given pair of alternatives unchanged, the social ordering over these two alternatives should remain the same.

[6]As one may note, the notion of "legitimate interests" may prove so broad that this mere requirement can be considered as meaningless. I leave undefined the question about what legitimate interests are.

not sufficient for our purpose and I further propose three contrasted conceptions of a democratic political community. They are presented on Figure 0.1. As will be developed afterwards, each of these perspectives requires motivating a social judgment about the fairness of different policy arrangements on drastically different bases.

The three perspectives on the democratic political community proposed differ according to two dimensions, which may be respectively related with a theory of society and a theory of human nature. The first dimension may be called the extent of *the public sphere*. The public sphere denotes the public arenas in which all potentially relevant stakeholders may express their argument publicly. Its extent may vary according to the willingness of the citizens to get involved in political activities or according to the transaction costs involved in public decision making. At some extreme, the citizens may be willing to entrust a single individual or a small group. At another extreme, large panels of stakeholders would directly and systematically get involved in public decision-making. In the first, narrow conception of the public sphere, the fair character of a given policy would result from the conception of the public interest of a given, often elected, individual or community. We may call it the *delegative perspective* on the democratic political community.

The second dimension is the strength of the sense of justice of the people. By the *sense of justice*, I mean both a general capacity of the people to form and revise their own impartial judgments over different policy options and their intrinsic motivation to develop and exercise it. It may also require some degree of tolerance or open-mindedness toward other's values and arguments. At one extreme, still, this sense of justice may remain anecdotal and the citizens primarily motivated by their own self-interest. Only to a lesser extent, the citizens may follow their own moral views or their perceptions of the social norms. We shall call it the *bargaining perspective* on the democratic political community. In this perspective, the relevant bases for social choice are the reasons that are deemed valid by the stakeholders themselves for the allocation of mutual advantage. It remains pessimistic on the possibility of trading-off the conflicting interests in society and presumes the acceptance of the actual situation as a status quo. This makes it a pragmatic and conservative view of politics that corresponds to the contractarian perspective depicted in Sugden (2013).

Finally, individual citizens may both feature a developed sense of justice and have access to a wide public sphere.[7] This creates the conditions for public reasoning and deliberation.

[7] Amartya Sen explicitly favored this perspective to the two others when he said, in his presidential address to the one-hundred seven-th meeting of the American Economic Association:

> "we have to go beyond looking only for the best reflection of given individual preferences, or the most acceptable procedures for choices based on those preferences. We need to depart both from the assumption of given preferences (as in traditional social choice theory) and from the presumption that people are narrowly self-interested homo economicus (as in traditional public choice theory)"
>
> (Amartya Sen, 1995)

In this perspective, the relevant basis for social fairness judgments are the reasons that any citizen would accept, given their own reasonable conception of the good and their understanding of shared values. This perspective, which rests upon the influence of public discussions on individual values and political behavior, may be called the *deliberative perspective* to the democratic community. This corresponds to the liberal social order depicted in Rawls (1993) or Sen (2009).

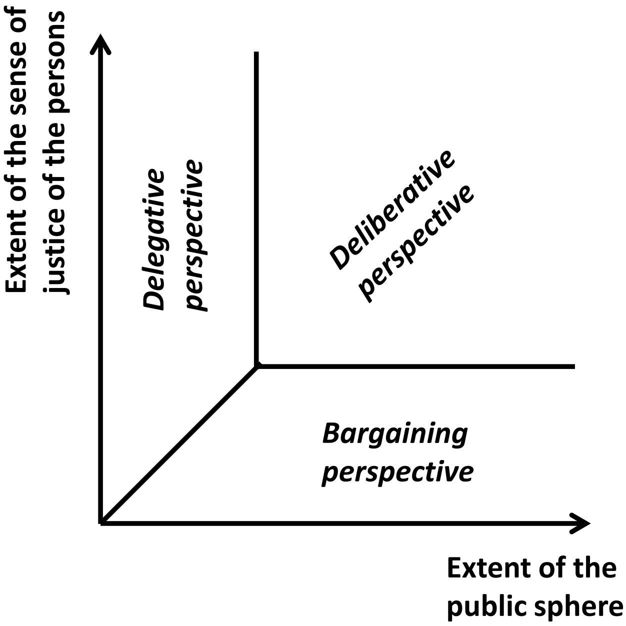


Figure 0.1: Three contrasted perspectives on a democratic political community

Figure 0.1 presents these three contrasted perspectives of a democratic political community. These cases may not be mutually exclusive but rather reflect different sides of a same, complex reality. Still, they are crucial for our discussion as each of these perspectives would ground a social fairness judgment on drastically different bases, as we shall go on to examine.

Implications in terms of analysis

The motivation of a social judgment will differ in the different perspectives adopted according to several dimensions. First, each of these perspectives would recognize different legitimate beholders. On top of this, they would rely on judgments of a different nature. Furthermore, they would impose different constraints on the judgments observed for being worth motivating a social judgment. Last, they will grant a different role to reasoning. The following discussion focuses on these differences. It is summarized in Table 0.1.

In the delegative perspective, social judgments are to result from the moral values of legitimate delegates which may rest upon a conception of the public interest. While some degree of internal consistency may be expected from them, the democratic legitimacy of the delegates is sufficient to establish the validity of a judgment. In this context, the social judgments would stem from the given, supposedly well-defined, conception of the public interest of a given group or person. Normative reasoning may be required or not, depending on the delegates. The role of the analysis is then to assist these delegates in forming their own reasoned judgments over available policy options, and given their own conception of the public interest. Most often, knowledge about individual preferences will constitute a necessary ingredient for the assessment of policies in this perspective. Knowledge about the citizens' values would still matters to the extent that the delegate are held accountable and may be concerned about justifying of their choices.

In the bargaining perspective, the relevant fairness judgments are the judgments of the actual stakeholders. Surprisingly, fairness considerations are not irrelevant though. In this perspective, they are relevant to the extent that they allow to overcome coordination failures and to avoid costly sanctioning behavior in bargaining. In this context, fairness beliefs mainly consist of a notion of salience resulting from shared associations and existing conventions and, to a lower extent, existing social norms and deeply felt individual fairness ideals. A particularity in this perspective is that the actual stakeholders may not find in their own interest to reveal their true perception of a fair bargain so that the actual content of fairness would have to stem from indirect evidence about their perceptions of salience and social norms. Furthermore, it is clear that an individual who would find the most salient possibility in her own interest would seek to avoid any further discussion and reasoning, so this particular perspective precludes ethical reasoning regardless of the extent of the public sphere. The role of the analysis is then to alleviate the coordination failures and mitigate potential wasteful sanctioning behaviors. In this line, the analysis shall provide a convincing evidence that a given option is the one that is the most natural and acceptable candidate for coordinating.

Finally, the deliberative perspective accepts the view of any citizen as a relevant basis under the provision of some degree of competence. First, the citizen may adhere to reasonable values which allow for discussion with other members of society. Second, the citizen should be informed and able to justify their positions. What further makes the validity of a claim for fairness is its certification by an ideal form of public reasoning, notably requiring two features: the first is *logical coherence* as it is a necessary condition for any form of reasoning; the second is *impartiality*, in a sense that will be discussed later on, as it is a necessary requirement to make the content of an argument acceptable for all. This draws the ideal of a judgment that is formed in conditions of impartiality through interpersonal deliberations. The outcome of such hypothetical deliberations determines the requirement of fairness. The eventual social judgment rests upon a plausible *overlapping consensus*, which is a judgment that all citizens may eventually find coherent with their own individual values. As Rawls and Sen admit it, there is no guarantee that such a consensus would always exist. Still, this may constitute the most compelling basis for a common judgment among citizens with different values and interests.[8] The role of the analysis in this context is to ease the deliberation and favor the identification of an overlapping consensus.

The implications of these different perspectives in terms of analysis are summarized in Table 0.1. In the first part of this thesis, I will develop more at length the type of analysis most suited to the deliberative perspective and seek to explore the potential of public reasoning to increase understanding and consensus among the citizens. In this line, Chapter 1 illustrates how different methods for the observation of individual fairness judgments may be more or less relevant in each of the three perspectives proposed, focusing on the deliberative one. Chapter 2 then details what is meant by normative reasoning and how analysis may contribute to propose some schemes of justification.

	Delegative perspective	**Bargaining perspective**	**Deliberative perspective**
Relevant judges	Legitimate delegate	Actual stakeholders	Any reasonable and informed citizen
Condition of validity of observed judgments	Accuracy in reflecting the values of the delegate	Accuracy in reflecting the perception of the actual stakeholders	Conditions of impartiality
Nature of the beliefs at the origin of the social fairness judgment	Moral values of the delegate (which may include a vision of the general interest)	Consciously shared knowledge and modes of inference, social norms	An overlapping consensus
Role for normative reasoning	Dependent on the delegate	Weak	Important

Table 0.1: Implications of different perspectives on a democratic political community on the basis of social judgments

[8] In the case no reasonable consensus emerges out of deliberations, the aggregation of existing views through voting procedures may eventually be required.

Content of this thesis

This thesis is organised into two parts. The first part presents and motivates the approach taken. It consists of three short chapters that are intended to be accessible for a wide audience. It motivates the approach taken and relates it to different branches of the existing research. The second part presents three contributions to the literature. It consists of three research articles that can be read independently from the rest of the thesis.

In the first chapter, I review the main empirical approaches for the observation of individual fairness judgments. These approaches are distinguished according to their relevance in light of the three perspectives proposed previously. I focus more particularly on the deliberative perspective. This drives our focus on contributions that seek to relate the observation of fairness judgments with the theories of fair allocations. Noting the focus on spontaneous intuitive judgments, I then discuss the possibilities of an empirical study of reasoned judgments. An example of a survey motivated by this approach is proposed at the end of Chapter 4, and an experiment is presented in Chapter 5.

In the second chapter, I consider the problem of motivating a fair social objective out of reasoning. I present the theories of fair allocations and discuss their potential role in the formation of reflected judgments. From a precise formulation of principles in a given environment, these theories are helpful to characterize the logical conflicts between norms, to stress the need for prioritisation and to propose ways to conciliate between competing norms. An example of this approach is proposed in Chapter 4. I further argue that the articulation of these theories in a broader framework can help identify new justifications as candidate for an overlapping consensus, which is illustrated in chapter 3.

The third chapter finally illustrates how the overall approach can contribute to the formulation and justification of fair policy proposals in a deliberative perspective. This chapter draws from some of the results presented in the three chapters of the second part. I concentrate on the problem of locating locally undesirable land uses. This context is particularly interesting as it is archetypical of a conflict between general interest and the particular interests of communities. While the qualification NIMBY ("Not In My Back-Yard") is often used as a disparaging qualification of the communities' refusal for hosting the project, we may wonder how a notion of a *general interest* could prevail against *particular interests*.

The second part presents the three articles that compose this thesis.

In the first article, I consider the problem of allocating an indivisible project and sharing its benefit among communities with an equal right on it but different provision costs. The differences in these costs may arise from variations in building, operation and maintenance costs for the project but also from differences in the communities' compensation requirements for hosting the project. In this setting, I characterise three allocation rules that correspond to three prominent cooperative solution concepts: the welfare egalitarian solution, the nucleolus and the Shapley value. The principles invoked involve Pareto efficiency, Anonymity, No envy, and axioms of solidarity or reward related to the communities' provision costs. The results clarify how considerations over the nature of the costs could influence fair allocations. The analysis is then extended to settings with asymmetric information and to setting with costs of several kinds. In each extension, I propose and motivate a fair solution. The results of a survey motivated by this analysis are eventually presented.

In the second article, I study individual fairness judgments and preferences for the allocation of an indivisible task and its benefit among two individuals with a different willingness to perform it. My approach relates results from the theory of fair allocations with the empirical observation of fairness judgments and preferences. I first present and motivate four contrasted allocation rules as possible rules of judgment. For some of the participants, a questionnaire was proposed before they knew about their situation. Among the four rules of judgment proposed, the welfare egalitarian allocation rule is the most preferred allocation rule as stated by the participants. Yet, I also observe support for principles that are not compatible with this rule: an important proportion of respondents deem fair to give nothing to someone who would not be willing to perform the task, and another substantial proportion deem fair to split the benefit of the task equally when both participants feature the same compensation requirement. In the experiment, participants had the opportunity to perform a task for pay. However, for any two of them, a single task was available. As required by the Pareto principle, it was allocated to the participant with the lower compensation requirement. In this situation, the stated normative expectations of the task performer are found to be higher, the greater the discrepancy between the compensation requirements. This does not extend to individual distributive preferences as revealed by the offers in a dictator setting. I also find that the task performers who took the questionnaire would deem the equal split fair less often. Overall, few respondents are consistent with any of the four rules proposed.

In the third article, jointly written with Stefan Ambec, we consider the decentralised provision of a global public good with local externalities in a spatially explicit model. Communities decide on the location of a facility that benefits everyone but exhibits costs to the host and its neighbors. They share the costs through transfers. We examine cooperative games associated with this so-called NIMBY ("Not In My Back-Yard") problem. We derive and discuss conditions for core solutions to exist. These conditions are driven by the temptation to exclude groups of neighbors at any potential location. We illustrate the results in different spatial settings. These results clarify how property rights can affect cooperation and shed further light on a limitation of the Coase theorem.

Part 1 - Presentation of the approach

Chapter 1

Observing individual judgments

The empirical social choice program

"It is the ordering according to values which takes into account all the desires of the individual, including the highly important socializing desires, and which is primarily relevant for the achievement of a social maximum."

(Kenneth J. Arrow, 1951, Social choice and individual values)

Introduction

The common democratic nature of the three background theories proposed in the introduction makes empirical knowledge about the actual judgments that exist in society crucial to the formation of a social fairness judgment. In this line, Yaari and Bar-Hillel (1984) sustained that "the evidence with which the theory [of distributive justice] must be confronted consists of observed ethical judgments or moral intuitions". This initiated a line of research that has been pursued until now and was reviewed by Konow (2003) and, most recently, by Gaertner and Schokkaert (2012). Following these latter, we may call it the *empirical social choice* program.

While initially focused on questionnaire studies, we shall construe this program more generally as the observation of the fairness judgments that exist in society with an intention to test existing theories of justice and motivate a social judgment. Existing attempts in this line actually extend to a large set of methods. Still, all these approaches are likely to differ significantly in their results and interpretations. The objective of this chapter is to contrast the main empirical approaches to the observation of individual fairness judgments regarding rules and allocations and discuss how they may be suited to motivate a social fairness judgment according to the different perspectives on the democratic political community developed earlier.

For each of these approaches, I will discuss its potential to reflect the richness of actual policymaking environments, and to collect observations at a reasonable cost. On top of these practical features, I shall also pay specific attention to the interpretation of the judgments observed. Indeed, even for a same individual, we may expect the judgments observed to differ according to at least four important dimensions.

A first distinction is related to the difference made in the social sciences between *moral* and *social* norms and the related behaviours (Falk and Fischbacher, 2001 ; Bicchieri, 2006 ; Elster, 2009). In this dual perspective, individual judgments could result from two distinct sets of beliefs. Individual *moral norms* is the first one. They manifest as an unconditional preference for some allocation, which may be called a distributive preference. The beliefs underlying such distributive preferences constitute individual *moral judgments*. The second source would result from a common understanding of a set of valid reasons, in a word, from the existence of *social norms*[1], which requires allocating in a given way in a given

[1] According to Bicchieri's (2006) definition, a *social norm* is a behavioural rule R for situations of type S in a population P such that there exists a sufficiently large subset of P such that for any individual i in this subset:

- Contingency hypothesis: i knows that a rule R applies to situations of type S
- Conditional preferences: i prefers to conform to R in situations of type S provided
  - Empirical expectations: i believes that a sufficiently large subset of P will conform to R in situations of type S

context. At the individual level, they will manifest as a preference for some allocation, conditionally on being observed by others, and on the expectation of their disapproving norm-breaching.[2] In this context, the beliefs regarding fairness stem from mutual expectations regarding appropriate allocations. These may be called *normative expectation*s. There is a priori little reason to think that the two sets of individual beliefs would always be identical. On the contrary, the previous discussion suggests that individual judgments may differ depending on whether they relate to one's own moral doctrine, or one's perception of public reasons. I shall therefore make clear whether the nature of the observed fairness judgments are normative expectations or moral judgments.[3]

A second distinction is related to the difference between *intuitive* and *reasoned* judgments. Accounts for observed judgments all rely on some conception akin to the dual process theory of human cognition, which contrasts a rationalist and an intuitionist mode of forming one's own beliefs and decisions. In the first mode, moral judgments primarily result from a rational deliberation out of explicit norms. In the second mode, moral judgments are the direct results of heuristics. The social intuitionist model of Haidt (2001) constitutes a recent exposition and an extension of this view to normative expectations. In this work, the heuristic route is proposed as the main determinant of moral judgments and justifications are seen as the result of an ex-post rationalizing process. On the contrary, it also seems reasonable to expect individual judgments to evolve with reasoning and the consideration of other's perspective and arguments. For now, all what has to be said is that I shall make clear how the judgments observed may have been affected by reasoning. I will present what is precisely meant by reasoning on these matters in the next chapter.

A third distinction relates to the likely considerations that underlie the observed judgments and, more specifically, to their degree of impartiality. Justice concerns are not the sole motives behind one's actions and statements, and surely not the most compelling. Self-interest and some forms of altruism both have been considered to tarnish judgments with partial considerations. Be it by deliberations behind a veil of ignorance as proposed by John Rawls or by using the figure of an impartial spectator as proposed by Adam Smith (1759), a major and common feature of theories of justice is that self-interested considerations - and their instrumental corollaries - shall play no part in it. As these ideal conditions could never be fully realized in the different conditions considered, I shall be clear about the role of partial considerations in accounting for the observations. In partic-

- Normative expectations: i believes that a sufficiently large subset of P expects her to conform to R in situations of type S (and may sanction non-compliance)

[2]Beyond a mere disapproval, strong reciprocity may also be part of the picture. Strong reciprocal preferences feature a willingness to incur a cost for sanctioning norm-breaching. Strong reciprocity belongs to the behavioral traits that are thought to sustain norm-obedience.

[3]It may be tempting to push the claim that *social norms* form the most accurate information for the assessment of public policies. It remains that some social norms may be deemed fundamentally unfair, and one may prefer to argue in such situation by directly appealing to one's moral sentiment.

ular, I shall pay attention to a potential *self-serving bias*, by which one may, consciously or not, find more attractive the criteria that bend toward her own interests.

Finally, a last and more subtle, distinction is about the difference between one's judgment over allocations and one's choice of an allocation in conditions of impartiality. The reasons behind such a difference may not seem clear. Yet, there is some evidence that stated moral judgments may significantly differ from the actual choice of an impartial observer (see e.g. Gold et al, 2015) This possibility was suspected by Miller (1992) who stressed the "danger of what may be called 'Sunday-best' beliefs, that is, the views that people think they ought to hold according to some imbibed theory as opposed to the operational beliefs that would guide them in a practical situation". We may therefore deem that whether a given judgment holds actual consequences is also a relevant feature of the situation on which a judgment is observed.

In this chapter, I start by presenting and discussing the questionnaire studies, which is the seminal and still mainstream approach in the empirical social choice literature. Then, I review the methods that propose to observe the incentivized expression of the fairness judgments of an outside observer. Next, I review the experimental methods that directly observe the judgements and choices of stakeholders. I eventually suggest research directions that may enrich the empirical social choice program.

1.1 Stated judgments of outside observers

In their reflection about how to share the costs of irrigation ditches, Aadland and Kolpin (1998) confronted their choice of equity principles with the view expressed by the ranchers in a direct phone survey. This is definitely the most natural way to start with, and this was also the approach taken by Yaari and Bar-Hillel (1984) in their seminal work. Even now, most of the empirical social choice approach still rests upon questionnaire studies.

In a typical questionnaire, the respondents are proposed a set of *vignettes*. A vignette denotes a description of a hypothetical situation upon which the respondents are invited to express a judgment. The respondent may be invited to express different sorts of judgments, sometimes in the same study. Most often, what is being judged are allocations, described in quantitative terms. The respondents may be asked how fair is a given allocation on a Likert scale (e.g. Kahneman , 1986), to rank different allocations according to their relative fairness, or to propose the fairest distribution in a list. Sometimes, "fair" may be substituted with other formulations such as "just" (e.g. Yaari and Bar-Hillel, 1984) or "acceptable" (e.g. Kahneman, 1986). Sometimes, hypothetical decisions are asked instead of attitudes (e.g. Herrero et al, 2010). Most often, the respondents express their judgments in a particular situations and, only in rare instances, the respondents are invited to express judgments regarding allocation rules (e.g. Herrero et al, 2010) or fairness principles (e.g. Schokkaert et al, 2007).

The seminal work of Yarri and Bar-Hillel (1984) constitutes the archetype of the approach. In their study, the authors surveyed university applicants about how a just allocation should be performed in a hypothetical situation. By varying some details of the situation, they evidence that respondents express significantly different judgments in formally identical situations depending on whether the persons differ according to their needs, their tastes or their beliefs. They further note that none of the prominent allocation rules existing in the axiomatic literature fully accounts for this observation. Many subsequent studies have kept challenging existing axiomatic results through the assessment of spontaneous judgments and further suggested axioms for the development of the theory. These studies mostly adopt a quasi-experimental design in which vignettes are varied randomly *between* the respondents. *Within* designs, that compare the answers of a same respondent to different versions of a vignette are much less frequent. Ideas in this line can be found in studies assessing the individual consistency between one's adhesion to axioms and to allocation rules. For instance, Amiel and Cowell (1999) challenged the existing characterisation of inequality indices on the basis of such an empirical material. The originality of their approach is that they also tried to assess the consistency of answers. They brought evidence that the existing axioms were remote from the actual views expressed by the respondents, which motivated the proposal of axioms more in line with these views in subsequent work (e.g. Ebert, 2009).

Let us now turn to the interpretation of these observations. As judgments are expressed in isolation, they may be interpreted in terms of moral norms. It is also clear that a typical questionnaire study offers little incentive to engage in reasoning. While Yaari and Hillel (1984) initially acknowledged that the evidence with which a theory should be confronted consists in "the ethical judgments made upon reflection by disinterested people" and restricted their attention to axiomatised allocation rules[4], they deemed the characterisation of reasoned judgments as a premature endeavour and focused on the first-thought statements from part of the respondents. Most of the subsequent literature has followed this line. Many subsequent studies have kept a focus on direct answers following the presentation of a vignette. Such an approach would mainly capture heuristics and intuitive judgments. As will be argued in section 1.4, these studies are little informative about the reasoned judgments that would arise following reasoning and argument. Among the handful of studies that addressed this concern, Hurley et al (2011) studied how judgments varied according to a verbal or quantitative framing of the allocation in the vignettes of Yaari and Bar-Hillel's initial study. Depending on the framing, They find significant dif-

[4]More precisely, the authors focused on *non-trivially axiomatisable* allocation rules that is, on allocations rules that were characterised by "axioms which have a force of their own and which can reasonably put forth as fundamental principles".

ferences between the judgments expressed. Their claim is that the verbal formulation is more tantamount to the formulation of principles and the deductive derivation of axioms and allocation rules, while the quantitative description of an allocation is closer to the direct description of a particular allocation. They also report that, when the respondents are presented with both descriptions, their answer conforms more to the quantitative description. In summary, and if the formulation of the question does not explicitly direct the answer otherwise, we shall interpret answers as spontaneous and reasonably impartial moral judgments.

Concerning the general strengths of this method, we first note that the meaning of an observation is rather univocal, provided we presume that, in the absence of any incentives, the respondents would answer sincerely. A difficulty still exists in interpreting the results as the meaning of words such as "just", "fair", "acceptable" or "equitable" is not always clear. Questionnaire studies are also relatively cheap. This opens the possibility for an extensive observation of the judgments held in the general population. Besides, the vignette approach can allow for the description of the many contextual cues that characterise a given policy problem. In addition, the relevance of a given framing may be assessed by asking the respondents whether they felt some information was lacking. An important limit of this approach is that, in the absence of incentives, respondents would not give much considerations to the cases proposed in complex settings. Random answers or heuristics may be checked through the internal consistency of individual answers or the sensitivity of answers to irrelevant cues (such as the position of an answer).

The hypothetical and declarative nature of the judgments observed suggests that they are poor predictors of the actual attitudes of stakeholders, making this method more relevant for the study of justice in a deliberative perspective. Yet, we saw that the relevance of this approach may also be questioned in such a perspective, mainly due to the absence of incentives to reason. This point is further developed in Section 1.4.

1.2 Revealed judgments of outside observers

As was discussed, a strong limitation of questionnaire studies is that they presuppose the sincerity of respondents who, in addition, have little incentives to form thoughtful judgments. One the other hand, it is established that personal stakes significantly bias individual judgments in a self-serving way. Therefore, we may be interested in experimental protocols that incentivize respondents to pay attention to cases in which they have no personal stakes. In this section, I present and discuss two such protocols: quasi-spectator experiments and tacit coordination games.

1.2.1 Quasi-spectator experiments

In the *Theory of Moral Sentiment*, Adam Smith (1759) proposed to infer one's moral conception from the would-be sympathetic feelings of an impartial and well-informed spectator. This thought experiment was, according to the philosopher, the process through which the content of morality could reasonably be inferred, including "the view to the general interest of society" or "the welfare of society" (III-2-3). Initially proposed by Konow (2000), the judgments of individuals involved in the situation of an impartial spectator constitute now a well-accepted empirical material in the empirical social choice approach.

In *quasi-spectator experiments*, a participant, the quasi-spectator, makes a decision that affects the payments of one or several stakeholders in a real situation, most often, in the lab. These experimental conditions are intended to approach the ideal conditions depicted by Adam Smith, which are assumed to be relevant for observing the content of moral rule and judgments (Konow, 2012). These conditions notably involve impartiality, information and sympathy. *Impartiality* is encouraged by fixing the payment of the quasi-spectator in advance and by keeping the spectator and the stakeholders anonymous. Despite a tension between the need for impartiality and the many potential biases that information may bring, a detailed description of the situation along with its many contextual cues is encouraged (Konow, 2009). The provision of *information* intends to ease the plurality of perspective and allows dispensing with a priori judgments regarding the relevant details of a situation. Still, many biases may arise from the identification of the quasi-spectator with some individuals.[5] These are assumed to be drawn away through the statistical nature of the observations. In practice, the provision of information is limited by the third requirement, *sympathy*, which is encouraged by the real consequences of the choice made by the quasi-spectator. While the need for a real experimental setting limits the set of situations that may be contemplated, evidence suggests that stated moral judgments may significantly differ from the actual choice of an impartial observer (see e.g. Gold et al, 2015).

The motivation behind the choice is assumed to be *sympathy*, where this term is to be broadly understood as a mixture of moral motivation, including strong reciprocal feelings toward norm breaching. This later motivational source was evidenced in the third-party punishment situation (Fehr and Fischbacher, 2004), a variant of the quasi-spectator setting. This pleads for interpreting the observed allocations as resulting from the social norms that have been internalized by the participants, including the norms regarding

[5] Actual experiments reveal the sensitivity of the conditions for the impartiality of the judgments expressed by quasi-spectators. By reporting an experiment in which the quasi-spectator and the two stakeholder were publicly allocated a given sum of money beforehand, Aguiar et al (2010) show a tendency to privilege the individual whose endowment was most similar to the spectator's. This shows that the judgments observed in a quasi-spectator setting are not always immune from partial considerations.

praise and punishment for others' actions. In line with Konow (2012), I further argue that sympathy has the epistemic role of incentivizing the spectator to enlarge her judgment to other's experience and perception of a situation, and the motivational advantage of moderating her own moral doctrine and align more closely with the perception of the stakeholders. This seems to suggest that the choices observed are more to be interpreted as normative expectations rather than individual moral beliefs. Still, the anonymity and the lack of social context of the situation may plead for the reverse.

Overall, the judgements observed in quasi-spectator experiments seems suited to the deliberative perspective. In this perspective, a further interest of the quasi-spectator experiment is that, the observed judgments are immune from a self-serving bias. Still, this approach faces challenges. The main challenges to this approach are the diversity of the motivational basis of the choices observed. In spite of its real impact, the allocation of the reasonably moderate sum of money may result from many motives such as a moral motive, altruism, strong reciprocal feelings toward norm breaching, warm glow, a willingness to please the experimentalist, jealousy, and so on. In contrast, tacit coordination games offer a clearer answer to the motivational basis of the choice.

1.2.2 Tacit coordination games

Tacit coordination games were initially introduced in economic discussions by Schelling (1960). While some connection with fairness was discussed, they were not primarily intended to reveal such judgments. Only recently, Camerer and Fehr (2004) and Krupka and Weber (2013) have proposed to use these situation to reveal the existence of social norms.[6] While tacit coordination games have not been much in use for the assessment of fairness judgments at present, I will argue that they can offer informative insights both in the bargaining and deliberative perspectives.

In the version of tacit coordination game that is for interest to us, two individuals are presented with a situation. They have to express a judgment without communicating. If their answers coincide, they get a payment fixed in advance. Answers may be about the appropriateness of an action on a Likert scale (Krupka and Weber, 2013), or the recommendation of a given allocation (Schelling, 1960).

As the sole motive in this situation is to coordinate with the other, we shall assume that tacit coordination games are suited to the observation of consciously shared knowledge,

[6]Other designs have been proposed to elicit social norms. For instance, Bicchieri and Chavez (2010) proposed to elicit the mutual expectations regarding the appropriate behaviours in a given context by directly asking the participants their normative expectations regarding a decision maker's behaviour and to check the consistency of the answer of the decision maker to an incentivized question about her beliefs regarding the other's expectations. They claim that the consistency of these answer signal the presence of a social norm.

among which are social norms. The main limit of coordination games for measuring normative expectations is that many confounding factors may account for the participants' answers, making their interpretation contentious. Indeed, the participants may find in their best interest to coordinate on other salient features of the environment such as visual cues or symmetric allocations. For instance, it may be hard to be convinced that the coordination on an equal split allocation reveals the existence of a norm for this allocation as the symmetry of this allocation may also make it a good candidate for coordinating.

In the bargaining perspective, this criticism may be lessened for two reasons. The first is that salience may in itself constitute a determinant of individual fairness judgments in bargaining environments; As emphasized in Schelling (1960, p.73), "even in those cases in which the only distinguishing characteristic of a bargaining result is its evident fairness, by standard that the participants are known to appreciate, we might argue that the moral force of fairness is greatly reinforced by the power of a fair result to focus attention, if it fills the vacuum that would otherwise exist". A second argument in favour of this method is that, when confronted to competing salient elements, there is some evidence that participants tend to give precedence to normative salience (Isoni et al, 2014).[7] Would the question be directly asked to them, it seems reasonable to assume that social norms are considered as a first basis for coordinating, at least, when the cardinality of the normatively salient allocations does not exceed the cardinality of visually salient allocations. This makes tacit coordination games a promising tool for investigating fairness judgments in the bargaining perspective. Herrero et al (2010) provides an example of the use of this device for the investigation of social norms. In the problem of adjudicating conflicting claims, they show that participants tend to coordinate on the proportional allocation.[8]

The second argument further suggests that, the judgments observed in this situation may also be deemed insightful from a deliberative perspective. In this perspective, a clear interest of this method is that, as payments do not depend on the answer, we are ensured that self-interest does not bias the observed judgments. Furthermore, perspective taking is also encouraged as it is in one's best interest to consider what another person from her group would be most likely to find fair. Thus, this method provides credible incentives for participants to engage in a social form of reasoning. As such this form of reasoning seems to fulfil some of the required conditions of the impartial spectator's view or on deliberations behind a veil of ignorance.

[7] Note that their setting is actually not properly a tacit coordination game as participants are invited to coordinate their choice on an allocation.

[8] In their situation, they incentivize the participants to coordinate on the majoritarian choice. Judgments expressed in this condition may differ from these expressed in two-person tacit coordination games, and depending on the reasons behind these variations, their relative relevance for social choice could be discussed. I do not know any work in this line and leave the comparison of these two variants for further discussion.

In the end, despite not being univocal, this method allows for observations of interest to motivate theories of justice. From a practical point of view, this methods is all the more interesting as it may be implemented in surveys and therefore, allow for a measurement of judgments on reasonably large and representative samples of the population at a reasonable cost. It also allows for an extensive observation of the judgments of a single individual, and the description of detailed contextual cues. All these features make this methods worth including into the realm of empirical social choice.

As an illustration, the climate negotiation arena is one context in which this perspective on fairness may be particularly insightful. In this context, the allocation of greenhouse gas emissions among states remains a debated issue. To my knowledge, most existing studies on this topic have used a survey design to elicit individual judgments. Unsurprisingly enough, they reveal a self-serving bias in the judgments of the negotiators or the citizens of each country (see e.g. Lange, 2010 or Carlsson, 2013). An interest of tacit coordination games is that they may provide a way to counteract this bias and may improve on standard survey results.

1.3 Revealed judgments of stakeholders

As was emphasized, a strong limitation of previous approaches is that observations may both suffer from a hypothetical bias and significantly differ from the actual views of stakeholders. This may rather seem to be an advantage of these in a deliberative perspective. Yet, one may also advocate for the strength of reasons that individual would actually accept to balance against their own self-interest. While not initially designed to the study of fairness judgments, experimental methods have this interest to jointly test for the nature of fairness judgments and their motivating force. In these situations, fairness judgments are observed indirectly through individual choices. This requires discussing how fairness judgments can be disentangled from self-interested motives.

1.3.1 The dictator game

The dictator game is the most common approach to show the existence of distributive preferences. In the simple version of the game, two subjects are paired and one of them, the dictator, decides how to split an exogenous amount of money with another participant, the recipient. The participants then get their payments and leave the experiment. In the standard situation, no communication or direct contact occurs and the experiment is performed using a double blind design, in which the experimenter does not observe the participants' choices and final payoffs. This simple version of the dictator game has been widely performed. In the simple version of this game, the usual observation is that 60%

of the dictators give a positive amount to the recipient, roughly 20% of the endowment (Levitt and List, 2007). Usually, a non-negligible proportion of the respondents is also observed that gives half the endowment.

Interesting variants of this game makes the sum of money to allocate contingent to choices, efforts, or luck. They revealed the sensitivity of fairness judgments and preferences to contextual cues. In a meta-analysis, Engel (2011) shows that distributional concerns are a robust and major source of differences across experiments. Consistently with the observation of questionnaire studies, concerns of needs, desert and social efficiency are all shown to influence the dictators' offers. More precisely, dictators tend to give a lower fraction of the endowment when it resulted from their effort, and the reverse if the endowment results from the recipient's effort, if recipient is needier, or if the offer was inflated (Engel, 2011).

In simple settings, these approaches are said to reveal the extent to which moral norms may account for individual behaviour. In more complex settings, they may also offer interesting information about the individual moral judgments. This requires disentangling self-interested motives from competing fairness motives. Approaches in this line rely on a structural specification of utility, which significantly differs across the different contributions and suggests different representations about the motives at work.

For instance, Andreoni and Miller (2002) propose to generalize the participants' utility functions as a general altruistic utility function that depends both on the dictator and recipient's payments. In this approach the participants are assumed to assess the allocation through the comparison of their own payoff with some social welfare function reflecting their views on fairness, or directly through a social welfare function in which welfare levels are weighted differently between self and others. This approach was shown to fairly well rationalize the data observed in different circumstances (Andreoni and Miller, 2002; Fisman et al, 2007). Still, in a further experiment List (2007) challenged this representation by giving evidence that slight change in the choice set may induce large behavioural changes. He suggests that these changes be due to different underlying norms. This observation cannot be directly accounted for with the altruistic utility function proposed by Andreoni and Miller (2002) unless utility is allowed to vary according to parameters which are independent of the choice set. This emphasizes a limit of such a representation of preferences. Besides, these representations may be limited in conveying the main intuitions behind the motives at work.

An alternative set of approaches has proposed to adopt utility functions that trades off a self-interest motive against the moral worth of the action taken. In particular, some have proposed to model the moral worth of an action by a decreasing function of the distance between this action and the fairness ideals held by the participants (see e.g. Cappelen et al, 2007). Among these models, some allow the fairness ideal to result from an endogenous stage of reasoning (see e.g. Brekke et al, 2003). What distinguishes these approaches from the former are the greater emphasis put on the formation of individual judgments and the non-welfarist representation of moral reasoning adopted. As argued in Tungodden (2004), welfarist reasoning actually contradicts our experience of public and moral deliberations, which relies on verbal statements over rich considerations of desert, needs, freedom, and so on. In a policy perspective, an important interest of this approach is that it explicitly links the observed actions with the underlying judgments. In particular, it opens possibilities for the empirical testing of the theories of fair allocations which will be presented in the next chapter. The relevance of such a representation of preferences and the extent to which it differs from the preferences characterised by Andreoni and Miller (2002) remains to be investigated.

This latter approach was able to provide interesting insights on the formation of actual fairness judgments. Konow (2000) adopts such a representation of utilities that relies on fairness ideals. He further proposes a model a cognitive dissonance to account for the manipulation of one's own belief about fairness in a self-serving way in the dictator setting and provides evidence of a self-serving bias. This suggests that the judgments uncovered by this method may show limited relevance from a deliberative perspective. More recently, Cappelen et al (2007) use a structural model to disentangle the fairness ideals from the role of self-interested motives among subjects facing an allocation problem that requires conciliating a principle of accountability for one's choice with a principle of compensation for circumstantial events. Their model documents the pluralism of fairness ideals. The authors observed behaviours consistent with a fair fraction of individuals favouring each of three different fairness ideals they previously suspected to be considered as such. In a subsequent experiment, they further observed that a preliminary stage of reasoning, consisting in exploring the implications of the three fairness ideals and expressing a judgment about them, did influence both the weight attached to fairness and the proportion of participants referring to each of them (Cappelen et al, 2010). Noticeably enough, they observed a decrease in the proportion of participants referring to a strict egalitarian allocation. Whereas it does not constitutes per se a proof for an equal split heuristics, this is consistent with this idea which has been suggested by many other work (Messik, 1993) and the observations reported in chapter 5.

1.3.2 The ultimatum game and bargaining environments

The ultimatum game is the most common approach to show the existence of reciprocal preferences, which denote a genuine willingness to sanction norm-breaching. Bargaining environments may also provide interesting insights on fairness judgments although observations in these contexts are subject to confounding motives (Roth, 1995). Yet, according to Eisenkopf et al (2013), observed behaviour in bargaining environments provide "the most accurate information about subjectively perceived entitlements". Since the introduction of the simple ultimatum game, bargaining environments have extensively been studied in the experimental literature. They have been used to test for existing theories of distributive justice.

In the standard bargaining situations, two anonymous participants have to agree on how to allocate a given sum of money. A first participant makes an offer about how to allocate a given amount of money, which the other can accept or reject. If the offer is rejected, the total amount to share is reduced and the other participant makes an offer in turn. The game usually proceeds until an offer is accepted or no amount in left to share. Many variants of the bargaining game have been studied, ranging from the standard structured bargaining without communication to unstructured bargaining with communication (e.g. Luhan et al, 2013). The traditional ultimatum game is a particular case of bargaining games which consist in a single stage. In this game, the proponent makes an offer regarding how to share a given endowment, after which the respondent can accept or refuse.

Sophisticated self-interested concerns of reputation building and fairness preferences are both acknowledged to account for behaviour in this context (Roth, 1995). In bargaining environments, fairness preferences are evidenced by a rejection of some offers while only Pareto inferior outcomes could be proposed afterwards. The main interest of the ultimatum game is that a rejection of the offer can directly be interpreted as a manifestation of fairness preferences and reveal existing social norms. Still, the interpretation of the choices in complex settings remains delicate. In their study about the individual beliefs regarding the just division of a division problem with claims, Gächter and Riedl (2006) find that participants declare to find fairest the proportional allocation, both in general and in particular cases. However, in actual bargaining, they settled closest to the constrained equal award allocation rule.

1.3.3 Individual decision-making behind the veil of ignorance

In the Theory of Justice, Rawls (1971) proposed to infer one's political conception of the fair from the likely decision that would be reached after a hypothetical deliberation behind a veil of ignorance. This thought experiment was, according to the philosopher, the process through which the content of social justice could reasonably be inferred. The veil of ignorance was the guarantee of impartiality. It was also taken on board by Harsanyi (1953), who rather framed it as a problem of rational individual choice.

Just as quasi-spectator experiments aim at putting participants in the role of an impartial observer, a branch of experiments have sought to observe judgments in this hypothetical situation. Most experiments reported in the economic literature actually fall closer to Harsanyi's perspective and reduce the deliberation stage to a stage of isolated decision-making (e.g. Krawczyk, 2010) or to direct voting (e.g. Beckman et al, 2002). In a series of questionnaire studies reported in Gaertner and Schokkaert (2012), choices behind the veil of ignorance are framed as an individual choice. They report significant differences with observations in other settings. Still, in this case, instrumental motives and risk attitudes seem to be much more relevant motives than fairness motives in accounting for the observed choices. Absent a direct and consensual link between risk attitudes and fairness, the relevance of these observations for the motivation of a social fairness judgment may be questioned.

A crucial characteristic of Rawls' veil of ignorance thought experiment is the existence of a collective deliberation. This guarantees that the motivation of general rules rests upon an explicit appeal to reasonable and impartial principles. A set of experiments tried to approximate these ideal conditions in a laboratory context. In these conditions, the participants first get involved in a deliberation and choose an allocation rule in a hypothetical situation. The decision after deliberation may ideally be unanimous, but majority decision may also be chosen. In a second stage, they get to know their precise position and the allocation rule is implemented. Frohlich and Oppenheimer (1992) provide an example of such a research program.

In their setting, the strength of the observed choices mainly comes from their motivation as the outcome of an open deliberation process in which self-serving biases and instrumental considerations may reasonably be assumed away. Another strength of this method is that it offers insights regarding the considered judgments on general principles, whereas most previous methods are better suited to elicit considered judgments in particular situations. One weakness is that the laboratory setting may constrain the set of achievable situations and cast doubts on the external validity of the observations.

In conclusion, individual choice behind a veil of ignorance may not have the potential to form convincing arguments about fairness as long as they are framed as an individual decision problem. However, the experimental design as initially proposed by Frohlich and Oppenheimer (1992) has a potential to uncover insights on the fairness judgments held by individuals that seem relevant to the deliberative perspective of the democratic community.

1.3.4 Other experimental protocols

As emphasized in this discussion, the search for fairness may be an effective motive in a large set of situations. Each protocol drives a focus on a different normative source, give different incentives to reason, and rely on different sources of motives for acting fairly. Additional examples include public good games (see e.g. Reuben and Riedl, 2013), the split-the-dollar game, prisoner's dilemma, voting, etc.[9] A common limitation of all these approaches is that they only allow observing fairness judgments indirectly through individual choices. While structural econometric methods have been proposed and successfully implemented to disentangle the fairness ideals from self-interest, the ad-hoc nature of structural models, and the suspicion of a self-serving bias may disqualify the inferred judgments from any normative force in the deliberative perspective. It is indeed a robust finding from the experimental literature that the self-serving bias becomes pervasive as soon as one hold personal stake in the matter. Compelling evidence of this has been given in the context of the dictator game (Konow, 2000 ; Ubeda, 2014) and bargaining environments (Babcock and Loewenstein, 1997).[10] However, this does not make these observations irrelevant to all sorts of social judgments. On the contrary, the connection of the observed judgments with actual behaviours makes them particularly relevant to the bargaining perspective.

1.4 Ways forward for the empirical social choice program

The previous discussion emphasized the diversity of methods that may be used to elicit individual fairness judgments. Given the large literature that is potentially relevant to the empirical social choice program, it focused on the specific contributions that most clearly fall within the scope of economics.[11] Table 1 provides an overview of these methods. Two observations may be made. First, none of these methods systematically investigates the effect of normative reasoning on the judgments. Second, all these methods remain focused on judgments made in isolation. In this section, I discuss how this could limit the relevance of these observation as a basis of social judgments in the deliberative perspective and motivate further developments.

[9]See Camerer and Fehr (2002) for a review of these protocols.

[10]A provision may be given to this claim though. As observed by Bicchieri and Mercier (2010) in the context of a trust game, individuals may bend toward a self-serving option provided it remains publicly justifiable and reasonable. As soon as we are not interested in the observation of fairness ideals, but reasonably fair options, this self-serving bias may appear less of an obstacle.

[11]Its main features may be a focus on individual judgments, a reliance on a formal theoretical background, and the use of monetary incentives. Still, this choice may occult the interdisciplinary nature of this area of research and the complementary insights of other social science and humanities. For instance, Forsé and Parodi (2010) adopt a similar approach and provide a recent sociological perspective on public opinions regarding social justice.

	Beliefs involved	Motives involved	Self-serving bias	Incentive to reason
Individual outside observer				
Questionnaire studies	Fair allocations according to individual moral values	Sincerity	Low	Low
Tacit coordination games	Shared knowledge (including social norms) Salience	Material payoff Sincerity	Low	Medium
Quasi-spectator experiment	Fair allocations according to individual moral values and internalized social norms (Konow, 2012)	Moral satisfaction (Konow, 2009), Strong reciprocity toward norm abidance or breaching (Fehr and Fischbacher, 2004)	Low (Konow, 2012)	Medium (Konow, 2012)
Individual stakeholders				
Dictator game	Fair allocations according to individual moral values	Material payoff Altruism (Andreoni and Miller, 2002) Self image concerns, moral satisfaction & guilt (Levitt & List, 2007) Cognitive dissonance reduction (Konow, 2000)	High (Konow, 2000)	Low (if not repeated) Medium (if repeated or preliminary question-naire)

Ultimatum game	Fair allocations according to social norms Salience (Schelling, 1960)	Proposer : Material payoff (including expected punishment) Responder : Material payoff Strong reciprocity (Bowles et al, 2005)	High	Low
Bargaining	Fair allocations according to social norms (Roth, 1955) Reputation building (Roth, 1995) Salience (Schelling, 1960)	Material payoff (Roth, 1995) Strong reciprocity (Roth, 1995)	High (Babcock and Loewenstein, 1997)	Low
Veil of ignorance without deliberation	Objective or subjective probabilities Beliefs about fairness	Material payoff Risk attitude Fairness motive	Low	Low

Table 1.2: Overview of existing methods for the observation of individual fairness judgments.

Eliciting reasoned judgments

An important reason why the judgments observed through the previous methods would remain unsatisfactory from a deliberative perspective is that none of them is able to give a convincing picture of individual reasoned judgments. There is ample evidence that individual judgments are unstable and subject to change (see e.g. Cappelen et al, 2010) and these changes are at the core of the deliberative ideal. These two reasons emphasize the need for observing reasoned judgments. I will argue in the next chapter that such a study of reasoned judgments shall strive to jointly observe consistent individual judgments over principles and in particular cases. In this endeavour, two types of empirical observations would be particularly interesting to document.

The first are *moral mistakes* (Sunstein, 2005). They consist in jointly held, yet inconsistent judgments. These empirical approaches may actually suggest and confirm moral mistakes by pointing at apparent inconsistencies. The empirical study presented in chapter 5 actually suggests such a mistake when the observation that the equal split is deemed fair, in a situation in which is fails to be satisfy individually rationality, another widely supported principle. What is meant here by "mistake" is that these spontaneous judgments would not resist reasoning and therefore observations of individuals revising their judgments could be expected when made aware of inconsistencies. An empirical account of moral mistakes would therefore have to be related to a contradictory adhesion to principles *and* a change in judgment. In the process of identifying moral mistakes one may arrive at a characterisation of reasoned judgments in a sense that will be made more precise in the next chapter.

Second, judgments still be subject to change while being perfectly consistent. We may call *moral discoveries* (Daniels, 1996, p.348) changes from one consistently justified judgment to another one. These may result from the consideration of additional principles or from a change in the representation of the policy problem. An account of moral discoveries would have to be related to systematic changes in judgments after some well-characterised consideration.

Enriching social interactions

This chapter mostly focused on the observation of individual judgments and choices performed in isolation or in the context of simple and stylized social interactions such as in the tacit coordination game or in the ultimatum game. This may be too rapid an abstraction from the profound influence of institutions and social relation in shaping our own beliefs and preferences (Bowles, 1998). Without going into the fundamental criticism of the possibility of methodological individualism to comprehend what essentially lies in interpersonal relationships, the question whether the judgments observed would differ significantly when performed in isolation or in social settings deserves closer scrutiny.

Of importance to the deliberative perspective, is the potential difference in the nature of reasoning performed in both settings. Actually, there is ample evidence of our limited ability to perform moral reasoning in isolation (see e.g. Haidt, 2001). It may therefore be too optimistic to expect individual judgments to reach a satisfactory degree of consistency and impartiality when performed in isolation, and illusory to expect any elicitation of judgments in reflective equilibrium. That being acknowledged, however, interpersonal argument and scrutiny have been suggested as a a set of conditions under which individuals may be much more willing to achieve consistency (Mercier, 2011 ; Mercier and Landemore, 2012; Vieider, 2011). This suggests that people usually do practice socially-oriented reasoning in order to convince others whereas moral reasoning seems seldom observed in practice. Then, as soon as we are interested in eliciting reasoned judgments, some well-designed social settings may provide relevant conditions.

Proposals in this line may be inspired from experimental deliberative settings such as the one proposed by Frohlich and Oppenheimer (1992). In the spirit of the former approach, deliberative opinions polling is another method of interest. Until now these polls have confined themselves to discuss factual claims and have discarded fairness considerations from the discussions (Fishkin et Luskin, 2005).

Conclusion

The discussion of this first chapter has shown that the perspective taken on the democratic political community bears significant implications regarding the nature of the judgments that are relevant to motivate a social fairness judgment. In bargaining settings in which the role of fairness remains limited to favour coordination over the set of Pareto improving policy options, tacit coordination games may be most suited to point at fair policy proposals. However, experiments have also revealed the existence of settings in which fairness motives may bear some weight in accounting for stakeholders' behaviours. In this perspective, economic experiments are able to reveal some deeply felt fairness views that could trigger sanctioning behaviours or, on the contrary, constitute opportunities to foster cooperative behaviours. While these observations are of great interest in the bargaining perspective, other protocols, such as direct surveys, quasi-spectator experiments or tacit coordination games allow for the observation of individual judgments that are immune from a self-serving bias. These are most interesting in a deliberative perspective. In this perspective, I further argued that existing protocols may benefit from focusing on the characterisation of reasoned judgments, which may have to occur in the context of interpersonal interactions. These results, along with a clear understanding of how reasoning may lead to different reflected judgments, may be particularly helpful in the identification of candidates for an overlapping consensus. This leads to point at the need for clarifying what is meant by reasoning and propose a theory of judgment formation, which is the object of the next chapter.

Chapter 2

Reflecting about fairness

The reflective equilibrium procedure and the axiomatic program

"Of course, we cannot know of how [men's] conception [of justice] vary, or even when they do, until we have a better account of their structure. And this now we lack, even in the case of one man, or homogeneous groups of men."

(John Rawls, 1971, A Theory of Justice)

Introduction

No objective, scientific method of inquiry can tell what *is* fair. Fairness lies in the eyes of the beholder. Absent any anchorage point, one may soon conclude that arguments about fairness are condemned to a permanent drift and leave this analysis to other endeavours. Yet, recognizing that ethical reasoning is fundamentally reflective does not exhaust the possibilities of inquiry. This is one message of John Rawls (1951) when he came to propose the *reflective equilibrium* procedure as a permanent process of mutual adjustment between our adhesion to general principles and our considered judgments in particular situations. This procedure came as an answer to the question whether a reasonable decision procedure, which is sufficiently strong to determine the manner in which competing interests should be adjudicated, could be shown to exist by rational methods of inquiry. As Rawls (1971) puts it, "there is a definite if limited class of facts against which conjectured principles can be checked, namely, our considered judgments in reflective equilibrium." As such our judgments in reflective equilibrium constitute the provisional anchorage point around which normative inquiry can be organized.

This procedure is a widely acknowledged form of ethical inquiry. It also matches our experience of moral inquiry and interpersonal argument. At least to some extent, we must be familiar with a form of deliberation in which one justifies a line of action to oneself or to others from an appeal to intuitively appealing principles that are thought to make sense of our strong intuitions regarding the right thing to do in particular cases. We may also have felt uncomfortable when being revealed inconsistencies between our adhesion on general principles and our intuitions in particular cases. In such a process, we may then find ourselves revising one or the other. In all cases, we seek to get closer to a mutual consistency between both sets of beliefs.

From a practical viewpoint, the reflective equilibrium procedure begins with the definition of a list of desirable properties a rule of judgment shall meet, followed by an assessment of their joint consistency and their practical consequences. The tenability of the rule of judgments identified shall then be assessed in light of their consequences on particular cases, calling for a revision of the underlying principles as soon as untenable implications are identified. A judgment is in reflective equilibrium as soon as it results from a rule of judgment that is retained from such a process. The approach of *equity in economic environments* is a useful guide in the second step of the process. It consists in studying the logical links between general equity principles and rules of judgments in a formal framework. Among this approach, the theory of fair allocations, focuses on the identification of fair allocation rules. This approach has been unified and organized under *the axiomatic program* (Thomson, 2001). While the axiomatic nature of the approach is challenged, its first merit is to operate a clear distinction between the logical inference between the objects and the meaning attached to them through the use of formal representation (Mongin, 2003). Furthermore, by forcing the norms to take a unequivocal form, it avoids the ambiguities attached to the normative statements and enables for a steady progress in ethical reasoning. In this chapter, I start by presenting and motivating the theory of fair allocations. Then I discuss how these theories can contribute to the motivation of individual and social judgments.

2.1 The theory of fair allocations

The theory of fair allocations shares at least two characteristics. The first is about the informational basis of the social judgment. As coined down in Fleurbaey et al (2005), these theories may be interpreted as a relaxation of the independence condition in the Arrovian framework. As soon as we allow the social judgment to depend on more elements than the ordinal preferences of the agents, the choice of a given domain starts to have an ethical dimension, which I discuss. The second dimension is about the form of the judgment. What further distinguishes the theory of fair allocation among the theories of equity in economic environments is that they seek to characterize a binary judgment instead of a full ordering of the alternatives. This contrasts with the evaluation of public policies in practice, which mostly relies on the assessment of quantitative indicators. Still, I will argue that in the deliberative perspective, we may prefer to focus on these binary judgments. I will eventually present the results that are obtained through this approach.

2.1.1 Domains as representations of the situation

The *domain* (or the *economic environment*) is a formal representation of the class of situations that are considered. It can range from the most abstract and general domain, the universal domain[1], to more context-specific domains. These domains consists in a set of micro-economic models in which a set of agents (e.g. individuals, households, communities), their characteristics, including potentially their preferences and beliefs, their property or entitlements, the available resources and existing technologies are explicitly formalised[2].

The choice of a domain defines the level of generality at which reasoning is performed. From an applied perspective, a given policy context is usually very rich and actual judgments would likely rely on many contextual cues that are absent from general theories (see e.g. Yaari and Bar-Hillel, 1984). When it is convincing, we may be willing to consider

[1] The universal domain consists in a set of undefined alternatives, a set of individuals with arbitrary preferences over these alternative.

[2] A discussion of the different domains used to represent the NIMBY problem can be found in Chapter 3.

some forms of interpersonal welfare comparisons. We may also be willing to rely on some individual characteristics that are deemed to elicit reward. As these two examples suggest it, we shall not refrain from considering what is deemed normatively relevant *a priori*. This is why I propose to follow the approach of equity in economic environments and focus on small and context rich domains rather than on general and abstract domains.

Eventually, the choice of a domain will have to result from the arbitrage between several considerations. First, the domain should capture all the *ethically* relevant facts upon which differences in treatments and outcomes may be justified. This is precisely the question of the informational basis of the judgment. Relevant information may notably include the characteristics required to assess individual welfare. Second, the domain should also take into account the *effectively* relevant facts that constrain what allocation can be achieved. This is the problem of determining the set of feasible alternatives. This should notably take into account the likely behaviour that the individuals are expected to follow. Usually, these behaviours will be represented as the maximization of some individual objective that is equated to individual welfare. However, as soon as it is acknowledged that the individuals involved may themselves endorse different view regarding their personal objectives and how society should assess their achievement of their personal objectives, the individual objective may not equate the individual welfare. Finally, the choice of a domain will also result from the necessary simplifications required for conducting a tractable analysis, which have to rely on some judgment regarding the relative importance of each of these facts. This latter constraint should obviously be as little binding as possible.

What comes clear from this discussion is that the choice of a domain is a normative operation in itself. As such, it should therefore be included in the reflective equilibrium process and subject to revision as soon as inconsistencies or untenable judgements are observed. In practice, I propose to start with a list of relevant *generic facts*. These may fall into two categories, according to whether they are deemed *ethically* or *effectively* relevant. The formal description of the situation is then derived to account for these facts. At first, only some of them may be considered.

Note that, depending on the ethically relevant facts identified, a domain may end up being very simple. In the context of a joint production problem, for instance, it may be acknowledged that the only ethically relevant facts are the individual marginal contribution to the total output, or the total cost of a project. Then, the representation of the situation could boil down to the description of a cooperative game with transferable utility, which already involves complicated judgments as the multiple complementarities and substitutabilities between individual actions precludes the unambiguous identification of each person's marginal contribution. For instance, Young proposes to adopt such a perspective for the allocation of the cost of water resource developments, and further presupposes that the outcomes of all possible cooperative arrangements are perfectly understood (Young, 1994). Another example of such a simplification of a domain is provided in Chapter 6.

When many generic facts are identified, richer and more complex descriptions of a situation may gradually be introduced through successive assessments. Note that this progressive and analytic approach is not only motivated by the need to keep the analysis tractable. It also constitutes a well-established process of inquiry in ethics, where reasoning is progressively elaborated from simple and hypothetical situations in which our intuitions are more acute (see e.g. Appiah, 2008). In contexts with asymmetric information, for instance, it is usual to start with a clarification of our judgments under perfect information and, only then, to question the implementability of the identified allocations. An illustration of this progressive approach is provided in Chapter 4. The *relevance* of a domain is achieved as soon as all ethically and effectively relevant facts are taken into account.

2.1.2 Allocation rules as representations of fairness judgments

As was already discussed in the introduction, what is crucial to the social choice approach is the characterisation of a social welfare judgment over the existing alternatives in a given situation. It should be emphasized here that as the reflective equilibrium approach makes clear is that the judgments under scrutiny should be systematic across many situations. This means that they should not be restricted to the single, actual situation, but to a broad range of situation (the domain) in a systematic way. We will speak of *rules of judgment* for such multi-profiles assessments. These are mappings from a domain into a set of single-profile judgments. Our former discussion already dealt with the question of the choice of a domain (their informational content). We now justify why we choose to rely on binary judgments (their form).

In their recent proposal, and consistently with the original Arrovian framework, Fleurbaey and Maniquet (2011) proposed to consider *social ordering functions*. These are mappings from each economy of the domain to the set of orderings of all available alternatives. In contrast, the theory of fair allocations relies on an alternative absolute representation of judgments. This approach focuses on *allocation rules* (or *social choice functions*). These are mappings from each economy of the domain to the set of feasible allocations. This latter theory will be the main focus in this dissertation. We shall try to further motivate this choice.

The main argument that is proposed in favour of it is that social orderings are better suited to accommodate second-best contexts when the set of feasible alternatives is difficult to represent (Fleurbaey and Maniquet, 2011, p. 9).[3] Another argument is that the theory of

[3]Besides, they argue that the approach of fair allocation rule fails to meet the weak Pareto criterion,

fair allocations was claimed to be less informative than the social ordering approach. For instance, *fairness ideals* could easily be derived from an ordering of the former approaches as the maximal elements of the domain.

All this seems to call for a general use of an ordering approach. Yet, this representation of social judgments may even contradict our intuitions regarding the relevant use of fairness considerations in the public sphere. We may have views that some policies could be fairer, but we may also acknowledge the irreducible plurality of individual values. Aware that a wide agreement on every trade-offs that may exist among policy alternatives is an illusory endeavour in public arenas, we may rather be interested in identifying reasonable proposals. These are these tempered judgments that are most relevant in the deliberative perspective, as these are the judgments that enter in the genesis of an overlapping consensus. Of course, binary judgments may be seen as one step toward the design of a full-fledged vision of an ordinal judgment, but the relevance of this view may remain confined to the delegative perspective on the democratic community.

On top of this limitation, the approach of fair allocation rules captures a different nature of a judgment that a purely ordinal approach would miss. Indeed, a purely ordinal approach cannot account for "what we are ready (or not ready) to accept". Yet, social and moral norms are usually such statements, and so are our intuitions regarding fairness criteria. "Everyone should be entitled to the product of her labour"; "No one should starve"; "No inequality should be accepted that is not justified", "No one should prefer someone else's allocation of resources to her own". These are the most likely expressions of what could be required in public. As interpreted here, the theory of fair allocations conveys a different sort of a judgment about whether a given policy is *acceptable*. Where our tolerance ends and outrage arises, a purely ordinal approach does not tell.

2.1.3 Axioms as a representation of equity principles

The last part of the reflective equilibrium procedure is the choice of general principles. The principles are taken as exogenous. They are to be understood as ethical arguments that are relevant to the public debate or the moral convictions that exist in society. They may be motivated in the literature of moral and political philosophy, or by an appeal to intuitions on fairness principles. Some are difficult to oppose. For instance, the two first requirements are logically flow as soon as we accept the relevance of the domain considered, which we proposed to take for granted through this stage of the reflection.

- **Impartiality** is the first requirement. As soon as all normatively relevant information is part of the domain, any different treatment should be justified on the basis of

which require that as soon as an alternative is unanimously strictly preferred, then it should be socially strictly preferred as well. Another argument is that this approach is tantamount the allocation rule approach when domains are rich enough (Fleurbaey and Maniquet, 2011, p. 8-9).

this information. Anonymity formally requires that any permutation of the labels leaves the allocation unchanged.

- **Unanimity** is another consensual requirement. It requires that no alternative should be chosen if another one would be unanimously preferred by the stakeholders. Under the presumption the representation of individual welfare is accepted, this requirement may be assimilated with **Pareto efficiency**, which requires that no allocation is chosen if another one would make all agents at least as well off and some strictly better off. All such allocations are described as *Pareto optima*. In the following, we may systematically restrict our analysis to the set of Pareto optima.

Note that the simplification required in order to conduct a tractable analysis may preclude the direct acceptance of Anonymity or Pareto efficiency. In the approach proposed, these two requirements are temporarily accepted and the assessment of their relevance is reserved for a later stage of reasoning.

On top of these two requirements, additional norms intend to make a finer selection among the Pareto optimal alternatives on the basis of further considerations. These considerations may be called *equity principles* or *distributive norms* as they all focus on the appropriate allocation of goods and bads. Here, I shall briefly describe the main distributive norms studied in the literature and some of the main conflicts identified.[4]

- **No Envy** is the founding principle in the theory of fair allocation rules (Foley, 1967). In its simple formulation, it requires that no agent prefer the allocation of another agent to her own. As such, it may not always be achievable and may conflict with the Pareto criterion and, therefore, appeals to refinements (see e.g. Fleurbaey, 2008).

- **Aggregate efficiency** consists in preferring outcomes that are more efficient in aggregate. It relies on the idea that the virtual possibility of compensation is sufficient to ensure the social desirability of a policy. This principle conduces to criteria such as the cost-benefit analysis, which evaluates policies through the comparison of their net benefit. In settings in which a *numéraire* can be used to achieve any distribution of welfare, it is tantamount with the Pareto principle. Yet, in other settings, it constitutes an independent criterion that enters in sharp conflict with other principles.

- **Solidarity** norms require that all the agents are affected to the same extent, or at least in the same way, by changes in circumstances, such as a decrease in a common resource or a change in the size or the characteristics of the population. These norms may align with other norms in surprising ways. For instance, in the setting considered in Chapter 4, solidarity regarding overall changes in preferences clearly

[4]Thomson (2011) reviews the recent results in the theory of fair allocations.

aligns with an egalitarian requirement. Still, solidarity regarding the size of the population may conduce to require someone not to gain from the increased competition of others and, surprisingly align with allocations that reward the communities with lower costs.

- **Needs and poverty alleviation**, also enter the formation of our judgments. They lead to prefer allocations that grants to all a minimum levels of some basic goods, what distinguishes them from equality principles.
- **Equality** norms lead to prefer outcomes that are more equal. Yet, this notion remains undefined with several respects. First, when perfect equality cannot be obtained, there is no natural and complete ranking of allocations. Second, this theory may conflict with the (ex ante) Pareto criterion in settings with risks. Third, most settings offer multiples possibilities regarding what could be equalized. With this regard, this theory remains undefined.
- **Freedom, rights and autonomy** capture the liberal idea that the allocation rule should not interfere with some individual sphere of autonomy. A first ways to express them is by imposing limits on public intervention. On top of these procedural requirements, lower welfare bounds have been proposed as a formulation of rights. These notions do not go without difficulties. First the question of which welfare bounds best represent existing rights remains undefined. Furthermore, lower welfare bound requirements may emphasize the need to limit the autonomy of individuals in social dilemmas so that arbitrage between the two sort of requirements may be required. This is illustrated in Chapter 6. Also, it was notice early that the protection of a sphere of autonomy is not compatible with the Pareto criterion in general (Sen, 1970).
- **Merit or desert** lead to prefer to reward individuals with a higher contribution. As for equality, the fundamental nature of this category is questionable as desert-based theories need to rely on external values and goals for defining a desert-basis (what makes an individual deserving) and the condition of it validity (e.g. voluntariness) (Lamont, 1994).
- **Consistency** is the requirement that some change shall not affect the one's judgment. For instance, one may acknowledge that for any pair of independent problems such that the welfare of any group of agents can be summed up, the allocation should give to each agent the sum of what it would give to him in each problem. These "technical" requirements are often pretty constraining.

As suggested by the previous presentation, norms are usually too vague to recommend unambiguous practical conclusion. They may state that "people should not be held responsible for circumstances" and leave undefined its practical implications. Some norms may even contradict themselves in the context of a given problem. For instance, a norm that would require treating all people equally may prove too constraining when the problem comes of allocating and indivisible task. The *axioms* are to be understood as a clear and operational expression of these norms to the set of situations described. In the context of the theory of fair allocation rules, they take the following form "An allocation rule satisfies axiom X if and only if Condition A". The condition may bear on the allocation selected in a single situation. For instance, it may be required that all individual welfare shall not reach below or above some level, or that no agent should prefer another's allocation to her own. The condition may also bear on variations across situations. For instance, a robust way to formalize the requirement that "people should not be held responsible for circumstances" is to require that "for any change in circumstances, either all people should benefit, or all should lose". While this may seem to be a weak necessary condition of the former statements, the analysis actually reveals this requirement as a very strong one on the domain we consider in Chapter 4.

We convey the reader to Chapter 4 to get an illustration of how these different notions are formalized into axioms. It should be emphasized here that the formalization of such distributive norms into axioms is not as straightforward as it may seem. In some environments, an axiom may convey very different norms. As an instance, the axiom of Individual Cost Reward that is introduced in chapter 4 may at the same time result from the view that one should be rewarded for socially more beneficial preferences, but also from the view that, everything else equal, transfers should not decrease would a community get poorer. This is developed in subsection 3.2.4 of Chapter 3. On the contrary, a given norms may also be expressed through many axioms, some of which having radically different implications.

This flexibility in the articulation of principles and axioms is especially interesting in the deliberative perspective as it allows to build an independent set of public norms that may be articulated with the multiple individual values that exist in society.[5] An example of this is developed in section 3.2.4. Because it is much more plausible to expect individuals to accept changes in how they express their own values rather than to renege on their values, this gives credibility to the idea that this approach may trigger convergence in

[5]The existence of an independent set of public values is an essential component of most political conceptions of modern democracies. Rawls is only one instance of this and, the following description of the Habermas suggests a convergence on this point:

> "Modern societies are set up so that any agent in any situation can be asked to justify their action and is pre-committed to doing so. In this way reasons provide the invisible lines along which sequences of interactions unfold, and which guide agents away from conflict. As social agents become accustomed to having their actions guided by speech and the mutual recognition of good reasons, so relatively stable patterns of social order begins to form that do not depend directly on credible threat of punishment, on shared religious traditions, or antecedent moral values"
>
> (Finlaysson, 2005 , p.27)

individual adhesions to public norms through deliberation. This in turns reinforces the credibility of the existence of an overlapping consensus.

2.1.4 Characterisations and impossibilities

Once the domain is properly defined and the axioms are stated, the analysis consists in assessing the joint consequence of these conditions. This analysis may yield positive or negative results.

First, it may be the case that, despite the joint desirability of the initial principles, no allocation rule satisfies all the axioms. This establishes an *impossibility result*, which is in itself an indication that trade-offs exist regarding the principles initially imposed. An example of such a situation is to be found in section 4.5.2. This requires revising the initial principles. The theory remains silent regarding the process through which the revision can be performed. Usually some of the principles may be dropped, or logically weakened in order to reach a positive result. By evidencing the source of the impossibility, the proof of the impossibility may in itself provide useful cues for the revision: in some instances, the impossibility will result from a anecdotal counterexample with no moral force so that a marginal adjustment in the principles would suffice; in others, a deeper conflict will be identified between two, equally desirable properties that requires careful consideration.

When an example is found that satisfies all the principles, the analysis may seek to describe exactly the set of allocation rules that satisfies the axioms. This establishes a *characterisation result*. Among these characterisations, the characterisation of a single allocation rule (the *solution*) from different combination of the axioms is particularly informative as it informs us on the limits of what is possible. By making explicit the trade-offs between the axioms, characterisation results draw, progressively, the frontiers of the firm land on which social judgments may be grounded. A characterisation ensures that any alternative allocation rule would fail to meet one of the axioms involved in the characterisation. It can therefore be an exceptionally strong support for an allocation as it tells us that for any different allocation, an example could be found that contradicts these axioms and the underlying principles. When all the axioms involved are intuitively appealing, such an example is likely to constitute a convincing argument in favour of the characterized allocation. Nevertheless, it also signals the proximity of the impossibility and any identification of untenable properties among these characterizing an allocation would require a difficult but necessary reconsideration of the principles endorsed.

2.2 The construction of reasoned judgments

In the context of a policy problem, the reflective equilibrium procedure can be described as follows.

1. Draw a list of the relevant *generic facts* that is, existing constraints on the allocations and differences between individuals that may justify a departure from equality. This allows deriving a *domain*, that is, a formal representation of these situations.
2. Identify a set of desirable principles that an allocation rule should meet in and across all the situations considered. This leads to a set of the *axioms*.
3. In this formal framework, characterize the set of allocation rules that satisfy these properties. As long as no such rule exists, revise the set of principles and axioms.
4. If an allocation rule that satisfies all the axioms is shown to exist, fully characterize the set of such allocation rules.
5. Assess the *tenability* of the rule of judgment on the domain. In case some recommendations conflict with intuitive judgments on some particular case, choose whether to maintain the judgment or to revise the initial set of principles. In this latter case, the conclusion reached is said to be *untenable*.
6. Assess the *relevance* of the formal framework adopted by broadening the assessment to the actual situations in lights of the generic facts identified. This may reveal some missing features of the problem and require revising the description of the domain and/or the principles adopted. As soon as a set of allocation rules pass this latter test, a reflective equilibrium is achieved.

The general framework is summarized on Figure 2.1. It extends the four-stage process proposed by Yaari and Bar-Hillel (1984) to make it applicable to actual judgments in a policymaking perspective. It also clarifies the contribution of the theory of fair allocation in the constitution of *reasoned judgments* (or, in Rawls' terms, *considered judgment in reflective equilibrium*) over alternative policies. This chart shows that many operations are external to the axiomatic approach. In particular, I wish to emphasize that the representation of the problem should be part of the reasoning if we want the method outlined to feed into the policy-making process. Real policy issue may be described in a many alternative ways and a every single axiomatic study on a well-defined domain is likely to raise questions that require to proceeds the analysis further.

In the context of a practical policy problem, it is likely that the mere description of a fairness judgment in reflective equilibrium would fail to convince. Our spontaneous judgments are grounded on informal and often contradictory principles. A judgment in reflective equilibrium may be adopted for lack of a better one, after several reconsiderations of one's intuitions over these principles and their implications in particular cases.

This is through the whole process that we eventually come to a settlement. Therefore, a convincing description shall not only consist in the end points but also in the many paths that lead to them, including the reasons of each successive move. This may seem tedious, but this is the most convincing motivation for a given rule of judgment. An illustration of this is provided in Chapter 3.

Conclusion

In this chapter, I argued in favour of the use of the theories of fair allocations. This approach puts the emphasis on the coherence between our considered judgements in particular situations and our intuitions regarding the regarding the general rules of judgments that generate them. In the context of this operation, the main interest of the axiomatic approach is to bridge the gap between these two sets of judgments. This choice was mainly justified by the possibility of including contextual information and the interest to limit the expression of a judgment to the reasonableness of different policy choices in a deliberative perspective. In the end, the conclusions that may be reached by this approach are very much in the spirit of the following conclusion that Minehart and Nemann (2002) present in the context of the NIMBY problem: “a siting procedure based on the one proposed in [their] paper that in addition is sensitive to issues of “environmental justice” and other moral considerations [...] could well provide an acceptable and satisfactory solution”.

I also stressed that this approach alone is not directly applicable to an actual policy context. Given the many potential controversies regarding the definition of the domain, I argued that this should explicitly be part of the reasoning process. Given the many elements and possibilities involved, the development of a tractable framework for structuring all existing results that are relevant to a given policy issue may be required. Such a scheme would help identify the pending theoretical developments and to formulate assumptions for empirical testing.

The approached proposed clearly aim at the constitution of a narrow reflective equilibrium. In order to be fully convincing, the set of beliefs that should cohere should be further expanded to background theories, in particular about the nature of persons and political institutions. In the conceptual representation of a wide reflective equilibrium as proposed by Daniels (1996, p. 51), a theory of the person and of a theory of society can both lead to different perspectives about the role of morality in society, which in turn impacts our considered judgments about policy options in a wide reflective equilibrium. What is proposed in this thesis amounts to this view provided we endorse one of the three perspectives on the democratic community suggested at the beginning as a background theory. As emphasized earlier, these contrasted background theories differ significantly regarding the level of reasoning involved in a social judgment. We also suggested in

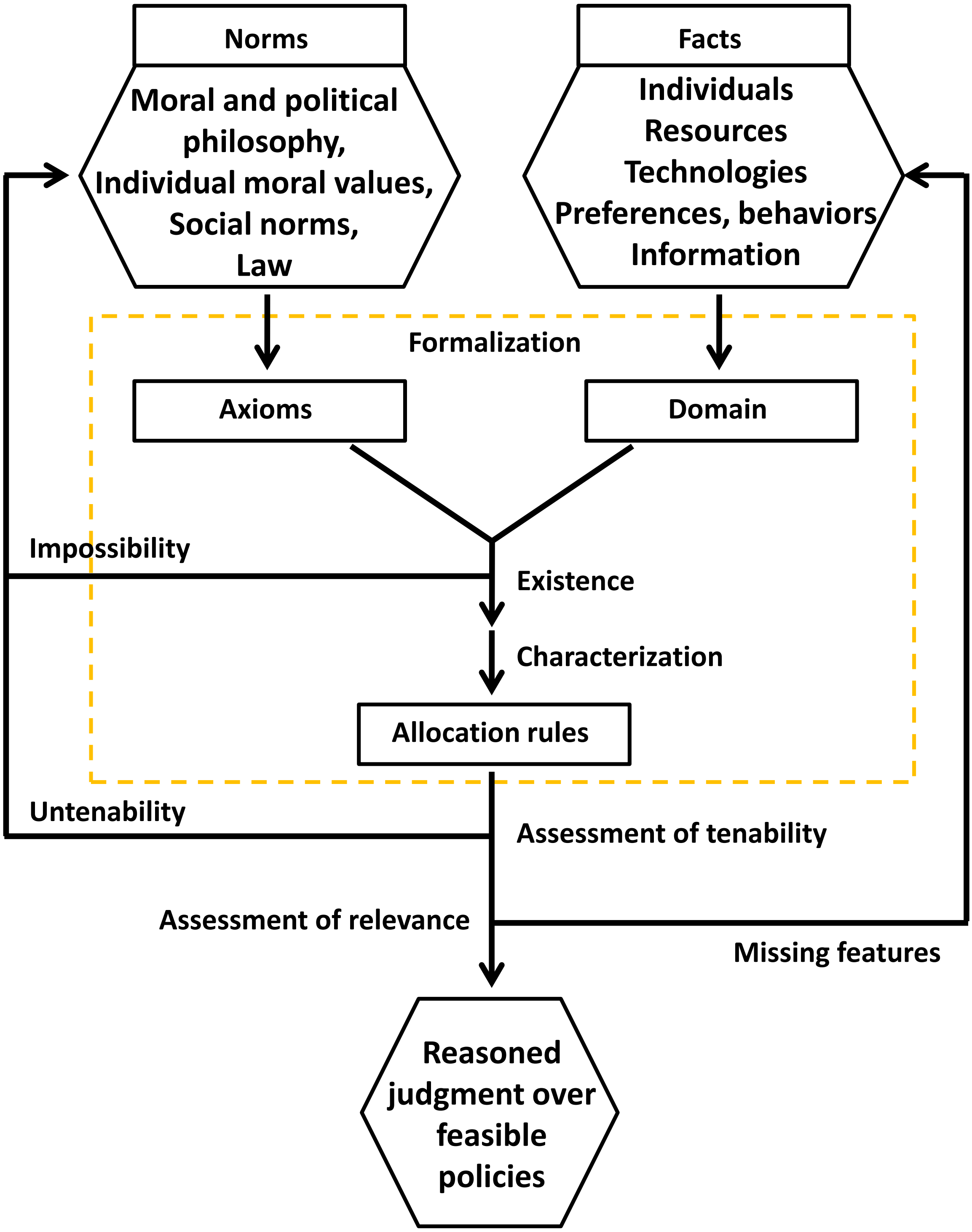


Figure 2.1: Frontiers of the theory of fair allocations in the reflective equilibrium procedure

this chapter that these may rely on a different form of a judgment. This makes the discussions of normative reasoning more or less relevant, depending on the perspective adopted. While at the core of the deliberative perspective, the approach presented here would appear irrelevant in the bargaining perspective.

Chapter 3

Application to the context of locally undesirable land uses

"Justification is a matter of the mutual support of many considerations, of everything fitting together into one coherent view."

(John Rawls, 1971, A Theory of Justice)

Introduction

In June 2008, the French government called for candidates for hosting a site for low-level radioactive waste storage among 3115 technically eligible municipalities. About 40 of them, all located in North Eastern France, showed their interest in the project. After detailed investigations, the Governement announced the selection of two municipalities for complete investigations: Pars-les-Chavanges and Auxon. However, facing strong public opposition in surrounding municipalities and pressure from intermediary political levels, Pars-les-Chavanges withdrew from the process, rapidly followed by the second municipality. The agency in charge of the process stated that "consistently to the approach chosen by the Government and [itself], based on the voluntary participation of municipalities, the municipalities resorted to their right to withdraw from the project"[1]. The site selection process was in a dead end.

Examples abound of oppositions to locally undesirable land use. More surprisingly, opposition is often believed to occur in spite of the presumption that such projects are socially beneficial in the sense of the Hicks-Kaldor criterion (Richman and Boerner, 2006).[2] Such situations are called the “Not In My Backward” (NIMBY) problem. Landfills, incinerators, power plants, windmills or airports are among the multiple potential examples of such locally undesirable facilities.[3] Sometimes, projects such as prisons or refugee camps are presented as potential examples. Generically speaking, NIMBY problems arise when a group of communities[4] can undertake a project that is unanimously recognised as globally beneficial but which remains locally undesirable. These projects often feature a conflict between a small group of people and an institution that upholds a vision of the public interest. They often end up in a gridlock.

In this setting, an assessment of fairness appears relevant to different perspectives of the political community. In the deliberative perspective, the mere issue of balancing the particular loss of the host against the general interest requires the design of a careful and consensual scheme of justification that is valuable in itself. In the bargaining perspective, the location of locally undesirable land uses gives typically raise to conflicts where fairness concerns are prominent. The identification of means to alleviate these conflicts and allow the group to overcome this problem is required for achieving mutual gains. This is suggested by the following analysis. “The key to solving NIMBY, in short, is trust.

[1] the author’s translation.

[2] This criterion consists in comparing the sum of the benefits with the sum of the costs associated with the project as compared to no project. When the former is higher, the project is said to meet such a criterion.

[3] The socially beneficial dimension of such projects is the NIMBY hypothesis; we will not question it. We do not address the question of efficient provision which is the source of an important literature in mechanism design and could also explain the social disapproval of these projects (sometimes referred to as the “Not On Planet Earth” (NOPE) problem).

[4] By community, we mean a group of individuals with a common objective.

Various sources of evidence suggest that individuals can be made receptive to the siting of noxious facilities in their communities *if they can be made to believe that society is committed to treating their interests with respect.* Appropriately structured bottom-up, negotiated-compensation schemes - ones *framed to emphasize* respect for the interests and autonomy of prospective host communities - are one way to reverse deep-seated resentments and thus excite a reciprocal openness to siting decisions. If individuals can't be made to believe that the burden of accepting a noxious facility *is being fairly reciprocated* either in kind or by like sacrifices, the current of resentment that fuels NIMBY will be difficult to reverse, even with financial incentives" (Kahan, 2005, our emphases).

What comes clear from this discussion is that regardless of the perspective adopted, a key question what treating the communities' interest with respect actually requires in this context. The object of this chapter is to conduct an analysis in search for a reasonable proposal that draws both from the methodological discussions of the previous chapters and the contributions presented in the second part of this thesis.

3.1 Representations of the problem

NIMBY problems arise for the location of land uses problems. The basic ingredient for these problems is the existence of at least two communities and a project. We shall further assume that this project is expected to bring local nuisances but that it yields an overall benefit that is deemed greater than its cost. In each case, what has to be decided is

- whether or not to implement the project, and, in the event of its implementation,
- where to locate it, and
- how to design it and whether or not some form of transfers (monetary or not) should be implemented.

As emphasized earlier, reasoning about this problem requires widening the analysis to rules of judgment, which specify a set of fair decisions for *any* potential situation. As was also discussed in the former chapter, the very simple representation we sketched may miss important features of most situations and fail to propose an accurate representation of the feasible alternatives or to capture the normative views most commonly expressed in this context. This is why we first have to consider carefully which features of these situations should be taken into account when making a decision about facility location.

3.1.1 Generic facts

Here follows a proposition of ten generic facts that shall motivate the choice of a domain.

Fact 1 Basic structure: The expected net monetary benefit of the project denotes its monetary benefit net from its construction, operation and maintenance costs. It may vary across communities.

Fact 2 Multiple costs: On top of the monetary costs of the project, the communities may have different compensation requirements for the nuisances that come with the project.

Fact 3 Composite costs: Differences in the compensation requirements of the communities may be due to different objective conditions making some of them more vulnerable to the nuisances associated with the project, but also to their different collective plans, and to their differences in social, demographic and economic characteristics. They may in particular result from unfair existing differences in wealth as poorer communities would tend to have systematically lower compensation requirements.

Fact 4 Non-subsitutabilities: It may be the case that, for some communities, no amount of money could compensate for a deterioration of their living environment.

Fact 5 Private information: Only the citizens of the communities can assess the worth that a degradation of their living environment represents for them. As a result, the compensation requirement of each community is private information.

Fact 6 Individual mobility: Individuals may engage efforts and resources to move across communities.

Fact 7 Externalities: The project may not only be perceived as a nuisance by the host community, but also by the neighboring communities.

Fact 8 Intra-community disagreement: The communities are a collective and may not be able to express a consensual compensation requirement.

Fact 9 Third parties: Third parties, such as future generations, are absent whereas their interest may also be at stake.

Fact 10 Risks and uncertainties: The non-monetary nuisances may range from a known deterioration to an uncertain threat on a community's living environment.

This list aims at capturing the main features of a NIMBY situation. It is to be taken as a fixed point in our reasoning. We shall assume that if we reach a policy proposal that fares well on domains that appropriately reflect these facts, we would have reached a reasonable policy proposal. Still, this list may not be exhaustive and may be discussed and revised in light of identified policies.[5] Taking all these facts into account in the decision clearly constitutes a challenge. The approach proposed consists in introducing progressively these features, sometimes separately.

[5] In the context of a precise NIMBY problem, additional facts may appear relevant such as, for instance, the history of cooperation among the communities. These features could be further introduced within the procedure.

3.1.2 Domains

The following table presents the different domains that are considered in this thesis, as well as in other papers that will be discussed throughout the presentation. All consider a group of communities who can implement an indivisible project and have to decide about its host and transfers. They all capture the basic structure of the problem (Fact 1) and mainly differ according to the way they allow for multiple costs (Fact 2), private information (Fact 5) and externalities (Fact 7). Among the domains considered in the literature, Minehart and Neeman also pay some attention to the nature of the costs (Fact 3) and the potential of intra-community disagreement (Fact 8).

	Verbal description of the domain	Formal description
In this thesis		
Domain 1 (Chapter 4)	A group of n communities can undertake an indivisible project that yields a common benefit B and a private cost c_i **to the host chosen**. This cost is considered **as a whole**. Preferences are quasi-linear. Information is **complete**.	$(B, \boldsymbol{c}) \in \mathbb{R}_+^{n+1}$
Domain 2 (Chapter 4, section 5.1)	A group of n communities can undertake an indivisible project that yields a common benefit B and a private cost c_i **to the host chosen**. This cost is considered **as a whole**. Preferences are quasi-linear. Information is **incomplete on** c_i.	$(B, \boldsymbol{c}) \in \mathbb{R}_+^{n+1}$
Domain 3 (Chapter 4, section 5.2)	A group of n communities can undertake an indivisible project that yields a common benefit B and a private cost c_i **to the host chosen**. This cost **consists in the sum of costs of different nature,** c_i^c **and** c_i^r. Preferences are quasi-linear. Information is **complete**.	$(B, \boldsymbol{c}^c, \boldsymbol{c}^r) \in \mathbb{R}_+^{2n+1}$
Domain 4 (Chapter 6)	A group of n communities can undertake an indivisible project that yields a private benefit to all and a private cost c_{ij} **to the host chosen** i **and each of her neighbors** j. This cost is considered **as a whole**. Preferences are quasi-linear. Information is **complete**.	$(\boldsymbol{b}, \boldsymbol{C}) \in \mathbb{R}_+^n \times \mathcal{M}_n(\mathbb{R}_+)$
In the literature		

Sakai, 2012	Each district of a group of n needs to deal with a private amount of wastes w_i. They can jointly undertake an indivisible project which yields a private cost **to the host only**. This cost **is the sum of a construction cost** c_i **and a disutility** v_i. They are both increasing and concave in the amount of waste processed. Preferences are quasi-linear. Information is **incomplete on** v_i and complete otherwise.	$(\boldsymbol{w}, \boldsymbol{c}, \boldsymbol{v}) \in \mathbb{R}^n \times \mathcal{C}^{2n}$ where $\mathcal{C}$ denotes the set of strictly increasing and weakly concave mapping from $\mathbb{R}_+$ into $\mathbb{R}_+$ that takes value 0 in 0.
Minehart and Neeman, 2002	Each community of a group of n needs to deal with a private amount of wastes w_i. They can jointly undertake an indivisible project which yields a private cost **to the host only**. This cost may be considered **as a whole**, c_i. It is increasing and concave in the amount of waste processed. Preferences are quasi-linear. Information is **incomplete on** c_i and complete otherwise.	$(\boldsymbol{w}, \boldsymbol{c}) \in\in \mathbb{R}^n \times \mathcal{C}^n$
Dehez, 2013	A group of n communities have to undertake an indivisible project that yields a private cost **to the host only**. This cost may be considered **as a whole**, c_i. It may be heterogeneous among communities. Preferences are quasi-linear. Information is **complete**.	$\boldsymbol{c} \in \mathbb{R}^n$

Table 3.2: Domains considered in this thesis and in the literature for the analysis of the NIMBY problem.

One may ask whether these choices of a domain are appropriate or whether they miss important features of the actual situations. We first articulate the results obtained on these different domains. The need for further developments will eventually be discussed.

3.2 Proposal for considered judgments in reflective equilibrium

In the general problem, an allocation consists in the decision to implement the project or not, its location and monetary transfers. We may assume that all relevant information is considered and focus on anonymous allocation rules. This requires us to allow the possibility to select among several locations through a lottery in case two communities have the same cost. We also focus on Pareto optimal allocation rules. They consist in implementing the project whenever its benefits outweigh its cost and at locating it where the cost is lowest. We may next focus on allocation rules that satisfy these two former criteria, and focus on the monetary transfers.

3.2.1 A first line of reasoning

Let us start with Domain 1, the simplest, and consider a situation in which a group of communities may implement a project whose net monetary benefit is the same wherever it is located but which may be perceived differently across location. This perception may be summarized by the communities' willingness to accept the project, or *compensation requirements*, expressed in monetary terms. Let us assume also that the true compensation requirements of the communities are known.

An example of an allocation rule selects a community with a lowest compensation requirement and shares the value of the project equally among the communities. We may call it the strict egalitarian allocation rule. This rule may not seem tenable as the host community may actually find itself worse off after the implementation of the project. As this rule seems to give raise to untenable allocations, we need to get more precise about which additional requirements could be imposed on the problem:

Norm 1 The first comment was related to the fact that the mere existence of the project could be considered as a beneficial circumstance for all. This would a minima require that "all communities should benefit from the existence of the project" (corresponding axiom: Individual rationality).

Norm 2 Perhaps more arguably the previous requirement could be extended to the communities preferences. This could lead to the following requirement "No community should find herself worse off when the all compensation requirements get unambiguously lower". (corresponding axiom: Full solidarity)

Norm 3 No community should envy the others, that is the host should not prefer not to be non-host and get her payment and the non-host should not prefer to be host and get her payment. (corresponding axiom: No envy)

Using the results presented in Chapter 4, it results from the first and last requirements that the allocation rules should give to the host an amount that lies in between her own compensation requirement and the second lowest one (or, in case it is greater than the total value of the project, this value), and share what is left equally among all users, including the host. The second condition further requires that the amount paid to the host's should be his own compensation requirement. This so called welfare egalitarian allocation rule is an attractive solution which is both proposed in Chapter 4 and in Sakai (2012). Yet the tenability of this solution may be questioned in light of the following conclusions.

Observation 1 In some instances, the allocation rule would give all the less to a community as her compensation requirement is low.

Observation 2 In some instances, the allocation rule would give a significant share of the value of the project to a community who would not be willing to implement it on her own.

These two particular conclusions could be accepted or rejected. In this latter case, a revision in the set of principles is required. The formulation of the situation could lead us to formulate the following two additional requirements.

Norm 4 A community should not be paid less when her compensation requirement gets lower. (corresponding axiom: Individual cost reward)

Norm 5 A community that would not benefit from the project in the absence of others should not get any share of its value. (corresponding axiom: No dummy)

The welfare egalitarian allocation rule does not meet these requirements. Actually the first requires, together with No envy and Individual rationality, that the allocation rule should share the second highest compensation requirement among users (or, in case it is greater than the total value of the project, this value). This corresponds to the nucleolus which is discussed in Chapter 4 and in Dehez (2013). This solution no longer satisfies Full solidarity but it satisfies Individual cost reward, so Observation 1 does not hold. For this, it may be deemed attractive by some. Still, Observation 2 still holds for this allocation rule. To show this, consider a three-community case in which two communities feature a compensation requirement that is strictly lower than the net benefit of the project and the third community features a greater compensation requirement than this value.

With the nucleolus, this latter community will get some part of the value of the project. Another potential limitation of this allocation may be that when extended to groups of communities, Observation 1 still holds.

Observation 3 In some instances, would two communities jointly feature lower costs, they may get a lower aggregate payment.

Again, these observations may be deemed tenable, or rejected. In the latter event, the set of principles needs to be revised. As, the latter example reveals it, a fundamental conflict arises between the No envy and the No dummy criteria. Then, imposing the No dummy criterion requires weakening No envy. For instance, it may be restricted between communities that have exactly the same costs. We shall call this latter requirement No envy among equals. Along with Observation 3, this suggests two additional norms.

Norm 6 A group of communities within which each community features a lower compensation requirement should not get a lower aggregate payment. (corresponding axiom: Collective cost reward)

Norm 7 No community should envy another community who features the same cost. (corresponding axiom: No envy among equals)

This weakening of No envy along with the extension of the reward requirement to groups leaves room for many allocation rules. One possibility is to share the whole benefit of the project among the communities with the lowest compensation requirement and give nothing to the others, which may be called the We may call it the strict libertarian allocation rule. The assessment of tenability then raises the following dilemmas.

Observation 4 In some instances, among two communities with arbitrary close compensation requirement, one can get a high payment while the other gets nothing.

Observation 5 In some instances, a community might suffer from the fact as another community with a lower cost than herself lowers its compensation requirement further.

Several routes may be taken depending on whether one or the other observation is considered untenable. Let us propose an additional requirement that seeks to avoid Observation 5.

Norm 8 The reward granted to a community should not make the communities with greater costs worse off (corresponding axiom: Solidarity toward higher-cost communities)

Along with the No envy among equals requirement, Collective cost reward and Solidarity toward higher-cost communities lead to characterize the Shapley value, which is further described in chapter 4 and in Dehez (2013). This allocation rule further satisfies No dummy and Individual rationality. It does not satisfy No envy. Nor does it ensure Full solidarity. But, reaching this stage requires that these axioms have deliberately been deemed less important than the intuitions with which they necessarily conflicted and that motivated the alternative requirements.

The general scheme of the possible combinations of the axioms that results from this discussion is presented on Figure 4.5. One may share some of the intuitions on the norms and dilemmas proposed here or be indifferent between the different allocations. While this scheme is not fixed, it seems reasonable to assume at this stage that at least one of the three allocation rules is deemed tenable.

The next step in the reasoning is then to assess its relevance when applied to a real situation. It is clear that it is not directly applicable to real policy problems. While it captures Facts 2, 4, and partially Fact 1, it fails to reflects the others. In order to get relevant some extensions are required. In particular, Fact 5 leads us to question the implementability of each of these allocation rules. In the literature on fair allocations, this is a natural development of the analysis on which we shall now concentrate.

3.2.2 Accounting for private information

On the specific problem of locally undesirable land uses, the unobservability of the communities' true compensation requirements constitutes a well recognised problem, which led to many proposals (Kunreuther and Kleindorfer, 1986; Minehart and Neeman, 2002; Laurent-Lucchetti and Leroux, 2011; Sakai, 2012). As soon as the costs are not directly observable, the assessment should address the means to implement the allocation rules. These are mechanisms which select an allocation on the basis of the communities' statements of their own compensation requirements. It is usually acknowledged that the communities will not be willing to reveal their true compensation requirement if not in their interest. Given the uncertainty related to the behaviors of others, all sorts of behavior may be expected in the context of a mechanism. Some communities may have different expectation regarding the others' behavior and be willing to take risks, other may be willing to maximize the worse outcome for them, and others may be willing to tell the truth to some extent. We cannot guarantee that the communities always find in their own interest to reveal their true compensation requirement.

I summarize here the discussions which are presented in Chapter 4. Its development is the following. As this information is crucial for having the certainty that we actually achieve these allocations, we conclude that none of the allocation rule that we identified previously can be implemented for sure on the basis of the true compensation requirements. Still, it turns out that the mechanisms that consist is the direct implementation of the allocation rules previously identified all meet a common requirement: they offer the *opportunity* for all communities to get at least some minimal payment that correspond to their identical-cost lower bound. For a given community, this bound corresponds to the level of welfare that would be granted by the welfare egalitarian allocation rule in the hypothetical situation in which all the communities feature the same cost as her. This criterion does not focus on the ex-post allocation but on the perception of the mechanism from the interim stage. If communities do not achieve it at the ex-post stage, it is because they preferred to take risk in the mechanism. This is difficult to prevent but at least, these mechanism prevent that such misreport may compromise the others' opportunities. This criterion leaves many possibilities which could either be discriminated against through a finer representation of the communities' behaviors in the mechanism, or through the requirement of additional procedural criteria. The previous analysis seems particularly adapted to justify these procedural requirements.

3.2.3 Accounting for costs of different kinds

Facts 1 and 2 together emphasize the existence of costs and benefits of different kinds. This can be accommodated in the model as long as we assume that the same norms apply to them. This is actually a strong requirement. An extension is proposed in Chapter 4 that explores the possibility of requiring the two norms identified in the former separately for each part of the cost. This actually leads to an impossibility result. A proposal is made which weakens the reward requirement to situations in which circumstances are uniform. This characterizes a particular allocation rule.

The results end here with a rather complicated scheme. Actually, some generic facts may easily be comprised in this model. Fact 1 brings in the possibility of heterogeneous benefit: these could be incorporated into the model by replacing the costs by a private valuation, as long as this benefit does not vary according to the location of the facility, which we will discuss later on. Fact 4 is accommodated by the absence of any restriction on the costs : costs may be infinite and the results would remain. Fact 6 is accommodated by the fact that the cost may encompass a moving cost. Finally, the model and the solutions proposed are still unable to account for Facts 3 , 7, 8 and 9.

3.2.4 Accounting for differences in wealth

Fact 3 relate to the important claim for environmental justice. The relative deprivation of poorer communities from a good environment is a well established fact which has implications in terms of health and morbidity (Hamilton, 2006). Therefore environmental inequalities tend to exacerbate existing wealth inequalities. A difficulty is that, in spite of this, poorer communities may still express lower compensation requirements for hosting a locally undesirable project, exposing them to a further deterioration of their living environment. This lead some author to plead for considering other locations (see e.g. Hermansson, 2007). A problem is that such allocations would not respect Pareto efficiency. Instead, I propose to require that, the allocation rule should not exacerbate existing inequalities. In this line, I propose the following principle.

Norm 9 Everything else equal, the allocation rule should not give less to a community when it gets poorer.

Along with an assumption of diminishing marginal utility of money, this leads to a condition that is tantamount to the Individual cost reward axioms,[6] and the previous results apply.

3.2.5 Accounting for externalities

Fact 7 raises the question of externalities. The existence of externalities significantly complicates the problem. In chapter 6, we consider a simple setting in which the communities may all bear different cost depending on the location of the project, and show that no allocation rule would pertain to the core of the associated cooperative game. In other words, no allocation rule guarantees that in any situation, all groups of community would get what they could get by themselves. Looking at the reason behind this result, one may propose to apply a polluter pay principle to this specific case and force the communities to internalize the full social cost of their project. Still, this does not bring us back to the simpler, externality-free domain.

To see this, let us consider the possibility to extend the unanimity lower bound requirement in the setting with externalities and perfect information as presented in chapter 6. This requirement is especially interesting as it is met by the 3 allocation rules that we characterized on Domain 1. This requirement was also central in the discussion of private information. Let us consider the simple 2-community case on a domain presented in Chapter 6 (Domain 4). The communities' benefits from the project write b_i, $i \in \{1, 2\}$, and the costs for j of a project located at i writes c_{ij} so that the structure of costs may be represented by the matrix

$$\boldsymbol{C} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

In such a setting, the identical-cost lower bound may be expressed from an analog reasoning as the one presented in Chapter 4. It proceeds as follows. Consider community1

[6]This is an instance of how different principles may motivate a same axiom, which illustrates how individual with radically different values may still find room for consensus without having to concede on their own values.

and an situation in which all communities are identical to her. There exists a single such economy for arbitrary costs.

$$\boldsymbol{C} = \begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{11} \end{pmatrix}$$

In this situation, it seems reasonable to claim that the two individuals should not envy each other. This determines a single distribution of welfare. It turns out that we cannot take this welfare as a feasible welfare guarantee. Consider for instance, the following case:

$$\boldsymbol{C} = \begin{pmatrix} 0 & B \\ 0 & B \end{pmatrix}$$

In this case, community 1 bears no nuisance regardless of the location of the project, and community 2 considers so high a nuisance. Their identical-preference welfare bounds are $\frac{B}{2}$ and 0 respectively. However, the reluctance of community 2 makes it inefficient to implement the project. The maximum total welfare that can be achieved in this setting is 0, so these bounds are not jointly feasible.

This example points to the facts that externalities considerably harden the justification. Still, this fact cannot be assumed away. The first reason is that, as Fact 8 emphasizes it, communities may not exist so that the problem has to be accommodated at the individual level. At this level, assuming externalities away seems excessively restrictive. This lead us to conclude that a reasonable and practical policy proposal shall accommodate Fact 7. In fact, here may be the crux of the opposition against current land developments project. Still, the observation of individual fairness judgments in the externality-free context is required to consolidate our results. Only then could we proceed confidently in extending the scope of this discussion.

3.3 Observed individual judgments

Two studies involving the empirical observations of individual judgments are reported in this thesis. The first study in presented in Section 4.6 of Chapter 4. It consists in an on-line vignette study. In the situation presented, two communities face a common opportunity of implementing a project (a wind park or a waste water treatment plant) that is worth five million euro. They state different compensation requirements for hosting the project. The precise vignette and details regarding the design are presented in the Appendix of Chapter 4. While not representative of the general population, this survey allows us to draw some conclusions. The second study brings the structure of the NIMBY situation studied here in the laboratory. Instead of hosting a project, individuals have to perform an indivisible task and its benefit is allocated. A detailed presentation of the experiment is provided in Chapter 5. Both studies rely on the observation of individual judgments in a situation with a same structure. The main difference is that the on-line survey gathers the expression of judgments in a contextualized framing whereas the experimental setting only shares similarities with the actual NIMBY problems in its structure.

Three conclusions may be reported. Before presenting them, a caution is in order. None of the two studies is representative of the general population. More particularly, the sample surveyed is younger, more educated and more business-oriented than the general population. All the three characteristics can be expected to influence fairness judgments in this setting (see e.g. Keller and Sarin, 1995).

A first observation is that the people seem to be generally supportive for implementing beneficial projects and locating them where the compensation requirement are lowest, at least in our setting in which these compensation requirements are stated and transfers are unconstrained. In the survey, this proportion amounted to an overall 80% in the cases in which the compensation requirements differ significantly. This goes against the idea that “a siting should not be decided depending on who demands the least compensation” (Hermansson, 2007). As soon as communities are allowed to freely state their compensation requirements and that some form of transfers are allowed, this paternalistic position seems difficult to maintain and the answers to the survey seem to go along this line. It should yet be mentioned that Keller and Sarin (1995) reported judgments that go against the Pareto criterion for the allocation of risk and benefits. In their study, respondents tended to prefer an even and proportional distribution of risks and benefits (the objective allocation) to Pareto superior allocations.

A second observation is that a significant proportion of the respondents still refer to an equal split, especially when compensation requirements are equal. We find this observation surprising as it goes against No envy, and more fundamentally Individual rationality. Some argument is favor of this may be found and this is consistent with the observation of many respondents preferring an equal distribution of risks and benefits in Keller and Sarin (1995). However we may note that this observation is most salient in the context-free experimental context. Besides, the experimental results suggest that these judgments are heuristics. They get less frequent after normative reasoning. They may be expected to vanish in the course of deliberation but still constitute important focal points in bargaining.

A last observation concerns the main patterns of the judgments observed in this setting. The overall payments to the host are found to be sensitive to both the lower and the higher cost in the pair. In both empirical settings, I observe that the payment to the host deemed fairest increases which the higher cost, and that it is lowest when both costs are equal. This suggests patterns in the formation of judgments that differ significantly from what the previous discussions proposed. Nevertheless, and beyond the previous observation, we also noted several apparent contradictions in the observed judgments and, at present, we lack a general theory that accounts for a significant proportion of the judgments. Further investigation would be required to overcome these inconsistencies and better characterize judgments in reflective equilibrium. As an example, it would be interesting to test for the adhesion to Individual rationality and observe how individual revise their judgment regarding fair allocation rules when made aware of their breaching this requirement. In the end, this is in this sort of trade-offs that we are interested.

Conclusion

This chapter is intended to illustrate how the approach proposed in this thesis can lead to identify fair policy options and motivate them. As it relies in a large part on reasoning, the approach proposed seems mainly suited to the delegative and the deliberative perspective on the political community and I first highlight the innovative propositions and suggest the ways toward future research. Still, the discussion presented in the introduction suggests that is may also prove relevant to the bargaining perspective.

Consistently with most of the economic literature on the subject, I argued that a solution to the NIMBY problem would have to take the specific preferences of the communities in the decision and that compensation schemes may be a useful device for managing the unfair initial distribution of the benefits of the project. The most innovative outcome of this research is related to its characterizations in the externality-free problem. An interesting output of this analysis is the identification of arguments in favor of some already existing proposals in favor of budget-balanced auctions. In particular, the opportunity for everyone to get, not only some positive level of welfare, but even their own identical-cost welfare lower bound seems to be a strong argument in favor of these auctions schemes.

Of course, many practical details may further be considered before getting to the recommendation, and we may still be far from proposing a solution to the NIMBY problem in all contexts. Among the many questions that remain to be addressed to reach practical recommendations, an important one relates to the treatment of externalities. A second one deals with the addressee of compensation payments as soon as individual may move away.

In the end, we shall note that the discussion presented in this chapter is only one scheme of reasoning. Many other possibilities may actually exist, depending, among other things, on one's initial judgment. For this, and as further generic facts get integrated and more results accumulate, I shall stress the interest of structuring the multiple representations and results in a systematic and practical way. Such a tool remains yet to be invented.

Part 2 - Articles

Chapter 4

Fair allocation rules for sharing the cost of a locally undesirable facility

A discussion of solidarity and reward

Abstract

I consider the problem of allocating an indivisible project and sharing its benefit among communities with an equal right on it but different provision costs. The differences in these costs may arise from variations in building, operation and maintenance costs for the project but also from differences in the communities' compensation requirements for hosting the project. In this setting, I characterise three allocation rules that correspond to three prominent cooperative solution concepts: the welfare egalitarian solution, the nucleolus and the Shapley value. The principles invoked involve Pareto efficiency, Anonymity, No envy, and axioms of solidarity or reward related to the communities' provision costs. The results clarify how considerations over the nature of the cost could influence fair allocations. The analysis is then extended to settings with asymmetric information and to setting with costs of several kinds. In each extension, I propose and motivate a fair solution. The results of a survey motivated by this analysis are eventually presented.

Keywords: NIMBY, fairness, allocation, axiomatic analysis.

JEL codes: D63, Q56.

I wish to thank Stefan Ambec, Henrik Andersson, Michel Le Breton, Christoph Rheinberger, Justin Leroux, François Salanié and Karine Van der Straeten for useful discussions on this project. I am especially grateful to Takuro Yamashita for the discussions that lead to the points presented in section 4.5.1. This paper also benefited from comments by participants at the 12th Meeting of the Society for Social Choice and Welfare in Boston.

Introduction

Examples abound of oppositions to undesirable land use. More surprisingly, opposition is often believed to occur in spite of the presumption that such projects are socially beneficial in the sense of the Hicks-Kaldor criterion (Richman and Boerner, 2006).[1] Such situations correspond to the acronym NIMBY for "Not In My Backward". Landfills, incinerators, power plants, windmills, airports, prisons are among the multiple potential examples of such facilities.[2]

None of the existing approaches for siting locally undesirable land use have proved well suited for a large class of cases. The authoritarian Decide-Announce-Defend approach has usually triggered significant protest and opposition, leading to costly trials and delays. Voluntary negotiations with communities have often ended in a gridlock. Among all accounts for these failures, the perception of the process and its outcomes may play a significant role both in the decision to oppose a project or to refuse a proposal deemed unfair.[3] This calls for a careful justification of the processes involved in this context as well as their distributive outcomes. In order to address these latter, the use of compensation scheme has often be proposed.

In both contexts, the use of compensation schemes remains a much debated issue. Whereas building, operation and maintenance costs are monetary and directly observable, the non-monetary costs of environmental degradations such as odours, noise, landscape degradation or health risks are subjective and cannot be directly observed. Considering this, the economic literature on this question has mainly focused on implementation issues (see e.g. Kunreuther et al, 1987; Kleindorfer and Sertel, 1994; Minehart and Neeman, 2002; Laurent-Lucchetti and Leroux, 2011). Most of these analyses have relaxed the dominant strategy implementability requirements, emphasizing the difficulty to alleviate the problem through the use of compensation schemes. This was reinforced by other work which further emphasized that compensation could conflict with existing norms, and, as a result, crowd out a sense of civic duty, undermine trust, and foster opposition (Frey and Oberholzer-Gee, 1997; Sandel, 2013). In the NIMBY problem, economists are indeed used to observing outright rejection and infinite costs when trying to infer the willingness to

[1] This criterion consists in comparing the sum of the benefits with the sum of the costs associated with the project as compared to no project. When the former is higher, the project is said to meet such a criterion.

[2] The socially beneficial dimension of such projects is the NIMBY hypothesis; I will not question it. I do not address the question of efficient provision which is the source of an important literature in mechanism design and could be at the origin of the social disapproval (sometimes referred to as the "Not On Planet Earth" (NOPE) problem).

[3] The experimental literature documents how normative judgments could lead to wasteful behaviours. See e.g., experimental evidence of strong reciprocity in one-shot ultimatum games (Camerer, 2003), the observation of the refusal of propositions followed by Pareto inferior propositions in bargaining environments (Roth, 1995) or the discussion of how self-serving fairness judgments could lead to bargaining impasse (Babcock and Loewenstein, 1997).

accept of individuals for environmental degradation or health hazards. Some philosophers have also opposed the use of monetary welfare measures and compensations schemes in some conditions.[4] While it is not clear that all these criticisms fully undermine the use of compensation schemes, they highlight the need for a careful justification of transfers on the basis of existing norms. They also emphasizes that we shall not discard the possibility for individuals to refuse to trade-off some dimensions of their welfare against monetary compensations. In this way, answers to these major arguments will be kept explicit.

In this article, we propose a framework that seeks to address this problem. We consider a group of agents (persons, communities) facing an economic opportunity on which they have an equal right a priori. This opportunity has a known economic value. For instance, the production of a jointly undertaken facility or the expected savings associated with a common landfill. This opportunity requires to be undertaken by one of the communities, who will bear its provision cost. It includes the disutility for the local nuisance (local pollution, noise, increased traffic) which will be borne by the host community. Provision costs may vary across communities. Communities may express different willingness to accept the nuisance. In the context of the location of a landfill some communities may even strongly reject their hosting the project. In all these situations, what would be a fair allocation of the object and the benefit associated with it? On the one hand, the equal communities' entitlements on this opportunity plead for an equal sharing of its benefit. However, the diversity of provision costs and considerations over the degree the agents could be deemed responsible for them may also play a role in the answer. We first perform an axiomatic analysis of the problem. An illustrative survey is then motivated and presented.

Related literature

This article relies on the theory of fair allocations. This approach formalizes the intuitions about fairness that may exist in society. It consists in studying the logical links between equity principles (hereafter *axioms*) and allocation rules (or *solutions*) in a formal framework (hereafter a *domain*). The emphasis is on the normative side of the analysis, the trade-offs between principles and the characterisation of the rules of judgments to which they logically conduce.[5] The approach presented here relates to different part of the literature about fair allocations.

First, our approach brings the focus to a set of solutions that have been proposed in the context of the NIMBY problem. Existing approaches related to the NIMBY problem

[4]For interesting philosophical expositions of claims against a perfect substitutability of some dimensions of welfare with money, see Walzer (1983) or, Anderson (1995, section 9.3). Sandel (2013) also contains a recent argument that specifically addresses the NIMBY problem.

[5]See Thomson (2001) for a general presentation of this approach and the *axiomatic program*.

consider a group of communities which benefits from a project. One of them has to host it and to bear its cost. In this setting, an allocation consists in choosing a host and performing monetary transfers. Two recent works have proposed the view that all communities should benefit from a decrease of some of the communities' provision costs. This view is consistent with the idea that provision costs are circumstantial, and therefore, that the communities should be held jointly liable for any change in the distribution of provision costs. It leads them to recommend to compensate the host on the basis of its own provision cost (Laurent-Lucchetti and Leroux, 2010; Sakai, 2012).[6] However, as soon as provision cost result from choices, beliefs or preferences of the communities, one could point at the little recognition of the specific role of the host of the previous approach. By framing the problem in a cooperative framework, Dehez (2013) deviates from these recommendations. He finds that two well known cooperative concepts, the nucleolus and the Shapley value, both recommend to take into account the whole distribution of costs instead of the minimal cost only.[7] Both solution concepts recommend a higher allocation to the host than the previous analyses do. Our analysis will be framed in the same context. It seeks to carry these analyses further by clarifying how each of these allocation rules can result from different considerations over solidarity and reward.

Second, the domain of the analysis is similar to the literature on the fair allocation of a joint production with a convex technology. [8] In this model, a product results from the aggregate input (e.g. work, effort, or investment) of several agents. The total production results from this aggregate input. As soon as the marginal returns are decreasing with the aggregate input, several possibilities arise as for the allocation of the final output. Our setting differs from it as a single input from an agent is required to get a fixed production. It can be seen as a limit case in which the production function is extremely concave. In spite of this difference, the main intuitions behind the analysis performed in this literature are relevant to our setting. As a result, many axioms of this literature will be introduced in our analysis. The main axioms introduced in this literature are *No envy*, *Resource and Population monotonicity*, and some welfare bounds. In our analysis, two notable welfare bounds are introduced and discussed. They are the *Identical-preference lower bound*, which, in our analysis is required by No envy, and the *Stand-alone upper bound* which is required by Population monotonicity. Some results will also appear. This is the case of the conflict between No-Envy and the idea of Population monotonicity (Kim, 2004).

Third, our approach relates to the literature about the fair allocation of the cost of a joint project. As argued in Dehez (2013), our setting is closest to the so-called airport

[6]Note that is both of these works, the communities derive a private and heterogeneous benefit from the project what allows them to focus the reasoning on the benefit side. Here the focus is on the cost side. Therefore, a simpler representation of the benefits will be chosen.

[7]We also depart from Dehez (2013) in that we explicitly introduce the benefit in the analysis. This is why our solution concepts will get a different expression than in his work.

[8]See e.g. Moulin and Roemer (1989) and Moulin (1990, 1992).

problem.[9] This problem considers the allocation of the cost of an infrastructure (e.g. an airport) when it is driven by the maximum capacity required by the agents (e.g. the airline companies). It proposes different allocation rules on the basis the contribution of each individual to the final provision cost. An important difference with the problem presented here lies in that the eventual cost of the project is the minimum instead of the maximum of all the participants' provision costs. In the airport problem, *Individual cost monotonicity* requires that the amount one pays does not decrease with one's cost. It is met by most solutions. In this analysis, a similar axiom, *Individual cost reward,* requires that the amount paid (resp. received) does not increase (resp. decrease) with one's cost. The main difference with this problem is that we consider costs may be *internal.* They are directly borne by the provider and do not reduce the extent of the external benefit. In the airport problem, costs are *external.* They are already part of the benefit to share and the identity of the provider is no longer relevant. Therefore, their analysis is tantamount to the problem of allocating the benefit of the project after having compensated the host for her own provision cost. Our requirement, on the contrary, focuses on the direct transfers including the compensation of the host. As a result, transfers always sum up to the same value and the axioms have different interpretations and consequences. In our analysis, the characterisation of the Shapley value relies on an idea of reward that is absent from this literature. This yields a different characterisation of the Shapley value which, in this literature, is based on axioms of independence from higher costs, additivity or incremental no-subsidy.

Fourth, our approach relates to the problem of fair allocations in economies with a single indivisible good (or bad) along with monetary transfers.[10] This problem deals with the allocation of an object to agents with different preferences. As part of their preferences, the cost is fully acknowledged to be internal to the agents. Several axioms in this analysis are also considered here. First, the axiom of *No envy* requires that agents should not prefer the allocation of others to their own. A weaker version of it, the *Identical-preference lower bound*, restricts this requirement to the hypothetical economy in which all agents have the same preferences. In our analysis, this requirement proves especially interesting as soon as we introduce asymmetric information. We also consider a weaker version: No envy among equals. Another axiom, *Welfare-domination under preference replacement*, was proposed in this context by Thomson (1997). It requires any change in one's preference to make all agents better off. It is a weaker version of the axiom of *Full solidarity* and a stronger requirement of *Solidarity toward higher-cost communities* introduced in this article. The literature established that, within the set of envy-free allocations, this principle leads to characterise the allocation which is least favourable to the provider. This result will also

[9] The literature on this problem is reviewed in Thomson (2007).

[10] The literature on this problem is reviewed in Thomson (2011) and Fragnelli and Gagliardo (2012) a recent relevant contribution to our problem.

appear in this analysis. The main difference with our approach lies in the fact that our axioms of reward are focused on the treatment (the transfers), and not the results (the utilities). To our knowledge, no analog for the axiom of *Individual cost reward* has been proposed in this setting. As a result, the nucleolus and the Shapley value, two allocation rules that will be shown to rest upon such an idea of reward are absent from the analysis of this problem.

Finally, our approach relates to the theories of equality of opportunity and the problem of fair compensation.[11] In this problem, a divisible good (money) can be allocated to agents with differing characteristics. Some of them are deemed to elicit compensation while others not. While the axioms of compensation are similar to the axioms considered here, the axioms of responsibility differ significantly. In their analysis, responsibility mostly relies on an idea of natural reward: the inequalities that naturally arise from the responsibility characteristics of the agents should not be subject to redistribution. In our setting, we are confronted to the particular difficulty that the realization of the agents characteristics do not arise naturally but from the allocation of the resource.[12] In this condition, we propose a different view on responsibility, more akin to an idea of reward. Similarly to the analysis performed in this literature, we explore the possibility of jointly compensating the agents for differences in circumstances and rewarding them for characteristics that are deemed worth rewarding in Section 4.5.2.

Notions of autonomy and reward is often a significant component of the justification of compensation schemes in the NIMBY context.[13] While little emphasis is found in axiomatic literature on the notions of reward and desert, the willingness to punish or reward other's according to their contribution to the group is a robust and widespread observation in the social sciences. In this line, the literature on distributive justice has discussed the notion of desert and emphasized its incomplete character (Lamont, 1994). This notion and related notions such as accountability and responsibility still proved necessary to account for stated fairness judgments in the experimental social choice literature (Konow, 2003; Gaertner and Schokkaert, 2012). To some extent economic experiments also have brought evidence in this line. In an experimental best-shot public good game, close to the problem we consider, Kroll et al (2007) observe that subjects are more reluctant to require the person that features the lowest cost to incur the provision cost as soon as these latter result from past effort. However, they do not allow for transfers so they do not reveal the extent of a fair compensation. In the end, we apparently tend to judge that

[11] This literature is introduced and reviewed in Fleurbaey (2008) and Fleurbaey and Maniquet (2011).

[12] The distribution of the project through a lottery has been proposed in this setting. It would lead to this problem. However, this allocation scheme does not satisfy the Pareto principle.

[13] As an anecdotal illustration, the following justification, reported by Inhaber (1998), was provided by the federal siting task force on Low-Level Radioactive Waste Management in Canada in support for its policy: "Such accusations [...] have been answered by indicating that a bribe is something offered as an illicit payment and that it induces a betrayal of trust. The *'reward'* proposed in this process is well *deserved*" (our emphases).

those whose provision costs are lower would deserve a better treatment in this context. We formalize this as a principle of reward and discuss its implications.

In relation to the survey conducted, this work can be related to the empirical social choice literature. Our approach is closest to Schokkaert et al (2007). The authors consider a problem with the same structure. They briefly discuss allocation rules in this setting and seek to test the empirical relevance of competing axioms through a survey. A difference with this work is that the axioms of reward allow us to go further in the axiomatic analysis and propose characterisations. The eventual survey is also framed differently. We do not observe as strong a support for considerations of responsibility and reward as in their work.

This article is structured as follows. The first section introduces the framework. The second section presents the three allocations rules formerly evoked. The third section presents and motivates a set of axioms. The fourth section presents the characterisation results. These results are extended to a setting with asymmetric information and with composite costs in the fifth section. Finally, the last section presents the results of a survey motivated by the analysis.

4.1 The domain

Let us consider a group of communities. They can cooperate and build a unique facility (e.g. a landfill or a wind farm). The project yields a common monetary benefit. It is rival and excludable. Yet, communities all have an a priori equal right on it. We are interested in the allocation of this benefit when the provision cost for the project depends on the location chosen. We assume that communities are indifferent between locations as long as they do not host the project. Additionally, we assume quasilinear preference over a numéraire and their hosting status, and allow for non-susbstitutability between the numéraire and the hosting status. In this context, we characterise the preference of each community i with a single provision cost c_i, potentially infinite. This cost encompasses both the construction cost at each community and their specific compensation requirements.

Formally, an economy E is defined as a pair $(B, \boldsymbol{c}) \in \mathbb{R}_+ \times \bar{\mathbb{R}}_+^n$ where $\bar{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{+\infty\}$, B denotes the total monetary benefit derived from the project and $\boldsymbol{c} = (c_i)_{i \in N}$ is the vector of individual provision costs. $\mathcal{H} = \arg\min_N\{c_i\}$ is the set of optimal location in E. The set of all such economies is denoted by $\mathcal{E}$. We may distinguish the subset of economies in which the project is beneficial, that is $B \geq \min_N(c_i)$. This subset of economies will be denoted by $\mathcal{E}^*$.

For any coalition $S \subseteq N$, we can define the cost associated with S, $c(S) = \min_S(c_i)$, and the value of the coalition S, $v(S) = \max\{0, B - c(S)\}$. This defines a cooperative game with transferable utility.[14]

Example 4.1. Three similar communities, $i \in \{a, b, c\}$, can jointly undertake a public project. This project is estimated to yield a net monetary benefit of 3, wherever it is located. However, this project is associated to some nuisance to the community that hosts it that cannot be avoided. These effects are respectively evaluated to be worth 1, 2 and 3 for communities a, b and c.[15] This defines the problem $E^0 = (3, (1, 2, 3))$ in $\mathcal{E}^*$. In this problem, $\mathcal{H} = \{a\}$, $B = 3$ and $c(N) = 1$. This economy will be evoked for illustration in the rest of the paper. It is represented on figure 4.1.

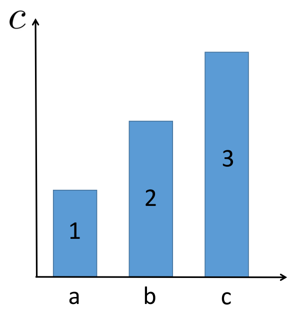


Figure 4.1: A representation of the situation E^0.

4.2 Allocation rules

4.2.1 Definition

In a given economy E, an allocation $\boldsymbol{\phi}$ is a pair of a vector of hosting status $\boldsymbol{h} = (h_i)_{i \in N}$, where h_i takes value 1 for at most one community and 0 for non-hosting communities, and a transfer scheme $\boldsymbol{t} = (t_i)_{i \in N}$. The host community is denoted by h, and $h = 0$ if the project is not implemented.

Definition 4.1. In a given economy $E \in \mathcal{E}$, a **feasible allocation** $\boldsymbol{\phi}$ is a pair $(\boldsymbol{h}, \boldsymbol{t}) \in \{0, 1\}^n \times \mathbb{R}^n$ such that $\sum_N t_i \leq B 1_{\{h>0\}}$.

[14] Note that the marginal contribution of an individual gets lower as a coalition grows: the cooperative game considered is concave.

[15] These "evaluations" for the communities' willingness to accept the nuisance could be thought as being inferred from valuation methods (hedonic pricing, benefit transfers, etc.) or as statements over their compensation requirements. In the latter case, they may not be thought as the true compensation requirement.

For a given allocation $\boldsymbol{\phi}$, $\phi_i = (h_i, t_i)$ denotes the individual allocation of community i. An allocation $\boldsymbol{\phi}$ fully determines the distribution of welfare levels $\boldsymbol{u}(\boldsymbol{\phi})$, where the welfare u_i of the community i under the allocation $\boldsymbol{\phi}$ is defined as follows: $u_i(\phi_i) \equiv u_i(h_i, t_i) = t_i - c_i h_i$. [16]

Definition 4.2. An **allocation rule** Φ is a correspondence that associates with an economy $E \in \mathcal{E}$, a non-empty set $\Phi(E)$ of feasible allocations.

We will restrict our attention to allocations that satisfy the two standard and appealing requirements of Pareto efficiency and Anonymity. Pareto efficiency guarantees that no solution that would make all communities better off exists. In this setting, an allocation rule Φ is Pareto efficient if and only if it always implements efficient projects, always recommends to locate the project where the cost is minimal and transfers to be budget-balanced.

Pareto Efficiency. *An allocation rule Φ is Pareto efficient on $\mathcal{E}$ if and only if for any $E \in \mathcal{E}$ and $\boldsymbol{\phi} \in \Phi(E)$, the three following conditions are met*

- ***Productive efficiency****: $\boldsymbol{\phi}$ recommends to implement the project if and only if it is (strictly) beneficial*
- ***Allocative efficiency****: $\boldsymbol{\phi}$ recommends to locate the project in the set of optimal locations $\mathcal{H}$*
- ***Budget balance****: $\sum_{i \in N} t_i = B$.*

A second appealing requirement is Anonymity. Anonymity guarantees that all differences in treatments are justified. It is a fundamental requirement in normative reasoning and presupposes that all normatively relevant information is included in the description of the problem.

Anonymity. *An allocation rule Φ is anonymous on $\mathcal{E}$ if and only if for any $E \in \mathcal{E}$, $\boldsymbol{\phi} \in \Phi(E)$ and permutation σ of N, $\sigma(\boldsymbol{\phi}) \in \Phi(E^\sigma)$, where $E^\sigma = (B, \sigma(\boldsymbol{c}))$ and $\sigma(\boldsymbol{\phi}) = (\sigma(\boldsymbol{h}), \sigma(\boldsymbol{t}))$.*

4.2.2 Three allocation rules

Three allocations have been proposed in this class of environments. The first allocation rule corresponds to the rule that is characterised in Laurent-Lucchetti and Leroux (2010) and Sakai (2012). The two following rules, the nucleolus and the Shapley value, are derived from the cooperative game perspective adopted in Dehez (2013).

[16] Note that this notion of individual welfare is not intended to capture the actual behaviours of the communities. It is the representation of the welfare that is adopted from an evaluative perspective, and may not be what truly motivates the communities. For instance, multiples arguments have been provided for not including political preferences, altruism or moral satisfaction in the welfare assessments (see e.g. Dworkin (1981), or Diamond (2006)).

4.2.2.1 The welfare egalitarian allocation rule

Welfare egalitarian allocation rule Φ^e. *For any $E \in \mathcal{E} \backslash \mathcal{E}^*$,$\Phi^e(E) = \{(\mathbf{0}, \mathbf{0})\}$, and for any $E \in \mathcal{E}^*$, $\Phi^e(E) =$*

$$\{(\boldsymbol{e_h}, \boldsymbol{t}) | h \in \mathcal{H}, \ \forall i \in N, t_i = \frac{B - c(N)}{n} + h_i c(N)\}$$

It recommends to fully compensate the host to the extent of the cost undergone $c(N)$ and share the remaining benefit $B - c(N)$ equally between all the communities. In the example 4.1, it recommends to locate the project in community a and implement the transfers $(\frac{5}{3}, \frac{2}{3}, \frac{2}{3})$. This yields an equal distribution of individual welfare levels $(\frac{2}{3}, \frac{2}{3}, \frac{2}{3})$.

4.2.2.2 The nucleolus

Consider a coalition S. The deficit of this coalition is the difference between the total welfare obtained from the allocation and the welfare it could get by herself. The nucleolus seeks to make these deficits as equal as possible by maximizing the minimal deficit over all possible coalitions (Schmeidler, 1969). A derivation of the nucleolus on $\mathcal{E}$ is carried out in Appendix 4.B.

Nucleolus Φ^n. *For any $E \in \mathcal{E} \backslash \mathcal{E}^*$,$\Phi^n(E) = \{(\mathbf{0}, \mathbf{0})\}$, and for any $E \in \mathcal{E}^*$, $\Phi^n(E)$ is the set of all pairs $(\boldsymbol{e_h}, \boldsymbol{t})$ such that $h \in \mathcal{H}$ and, for all $i \in N$,*

$$t_i = \frac{\max\{B - c(N \backslash h), 0\}}{n} + h_i \min\{B, c(N \backslash h)\}$$

The nucleolus does not only recommend to compensate the host to the extent of the cost undergone $c(N)$ but also recommends to grant an additional reward to the host to the extent of $c(N \backslash h) - c(N)$ provided it does not exceed the benefit of the project. In case all the communities' provision costs but one exceed the benefit of the project, it recommends to transfer the whole benefit of the project to the community with the lowest cost. In the case where the lowest provision cost is featured by two communities or more, it coincides with the welfare egalitarian allocation rule. In the example 4.1, it recommends to locate the project in community a and implement the transfers $(\frac{7}{3}, \frac{1}{3}, \frac{1}{3})$. This yields the distribution of individual welfare levels $(\frac{4}{3}, \frac{1}{3}, \frac{1}{3})$.

4.2.2.3 The Shapley value

Consider an arbitrary ordering of the communities and, for a given community i, the coalition S formed by all communities preceding it. The marginal contribution of i to S is defined by $v(S \cup \{i\}) - v(S)$. The Shapley value grants to any community its average marginal contribution over all possible orderings (Shapley, 1953). It has been characterised in multiple ways.[17] A derivation of the Shapley value on $\mathcal{E}$ is carried out in Appendix 4.C.

Shapley value $\boldsymbol{\Phi^s}$. *For any $E \in \mathcal{E}\backslash\mathcal{E}^*$,$\Phi^s(E) = \{(\mathbf{0},\mathbf{0})\}$, and for any $E \in \mathcal{E}^*$, $\Phi^s(E)$ is the set of all pairs $(\boldsymbol{e_h}, \boldsymbol{t^s})$ such that $h \in \mathcal{H}$ and transfers $\boldsymbol{t}^s(E)$ can be represented as follows. Let σ be a permutation of N such that $c_{\sigma^{-1}(1)} \leq ... \leq c_{\sigma^{-1}(n)}$ and $\sigma^{-1}(1) = h$. The resulting index will be written in brackets: (i) denotes the initial index of the community with rank i in the ranking induced by σ. Let q be the highest rank such that $c_{(q)} \leq B$. For all $i \in N$, the distribution of welfare levels writes:*

$$u^s_{(i)} = \begin{cases} 0 & \textit{if } i > q \\ \frac{B-c_{(q)}}{q} & \textit{if } i = q \\ \frac{B-c_{(q)}}{q} + \sum_{k=i}^{q-1} \frac{c_{(k+1)}-c_{(k)}}{k} & \textit{if } i < q \end{cases}$$

and the associated transfers:

$$t^s_{(i)} = \begin{cases} 0 & \textit{if } i > q \\ \frac{B-c_{(q)}}{q} & \textit{if } i = q \\ \frac{B-c_{(q)}}{q} + \sum_{k=i}^{q-1} \frac{c_{(k+1)}-c_{(k)}}{k} + h_{(i)}c_{(i)} & \textit{if } i < q \end{cases}$$

Allocations described by the above formula are constructed as follows: first order the communities from the lowest to the highest provision cost. Choose the first community to be the host h and compensate her for the cost c_h undergone. Then, share the remaining benefit $B - c_h$ as follows. For all communities whose costs are higher than B, give nothing. For the community (q) with the highest cost that would allow the project to be implemented, share the resulting benefit $B - c_{(q)}$ among the q communities with a cost lower than $c_{(q)}$. Then, share the difference $c_{(q)} - c_{(q-1)}$ among the $q - 1$ communities with a lower cost than $c_{(q-1)}$, and so on. In the example 4.1, it recommends to locate the project in community a and implement the transfers $(\frac{5}{2}, \frac{1}{2}, 0)$. This yields the distribution of individual welfare levels $(\frac{3}{2}, \frac{1}{2}, 0)$.

4.2.2.4 Comparison of the three solutions

These three allocation rules propose different views about how to share the benefit of the project. Note that for any $E \in \mathcal{E}$ and $h \in \mathcal{H}$, we have for any $(\boldsymbol{\phi^e}, \boldsymbol{\phi^n}, \boldsymbol{\phi^s}) \in \Phi^e(E) \times \Phi^n(E) \times \Phi^s(E)$, $u_h(\boldsymbol{\phi^e}) \leq u_h(\boldsymbol{\phi^n}) \leq u_h(\boldsymbol{\phi^s})$. As noticed in Dehez (2013), the Shapley value could potentially allow for some non-hosts to get more than an equal share

[17] see e.g. Moulin (2003) for a review of these results .

of the benefit, which means that these agents would be paid instead of contributing to the cost of the project. In particular, this means that, contrary to the two other allocation rules, the Shapley value does not pertain to the core of the cooperative game induced by this situation. Intuitively, the nucleolus and the Shapley value both seem to grant more recognition to the communities featuring lower costs. This is precisely this point that the following analysis seeks to clarify.

4.3 Axioms

In this section, we introduce and motivate the axioms that aim at capturing the conflicting notions of solidarity and reward and will constitute the basis of our analysis. We start by presenting the No envy requirement. Then, we present the axioms capturing ideas of reward and solidarity in this setting. Finally, welfare bounds are presented.

4.3.1 No envy

No envy is a cornerstone principle in the theory of fair allocations. Originally introduced by Foley (1967), its implications have been studied in a broad range of environments, and especially in the context of the allocation of indivisible goods. It requires that all communities should prefer their allocation to the allocation of others. This corresponds formally to the following property.

No envy (NE). *For any it $E \in \mathcal{E}$, any $\boldsymbol{\phi} \in \Phi(E)$, and $(i,j) \in N^2$,*

$$u_i(\phi_i) \geq u_i(\phi_j)$$

.

The consequences of this requirement are presented in section 4.4.1. In particular, this precludes any possibility of performing different transfers among the non-hosts. To the extent that this can be considered as too strong a limitation, it will be useful to consider a weaker version of the No envy criterion which applies in settings in which communities features the same provision cost.[18] This leads us to the following weaker version of the No envy criterion.

[18] It is interesting to note that the need to adapt the No envy requirement to accomodate considerations of responsibility and reward is not exclusive to this environment. For instance, Fleurbaey and Maniquet (2011) also propose to adapt this axiom as its application in their setting leads to recommend an equal distribution of the external resource and precludes the satisfaction of their reward requirements.

No envy among equals (NEE). *For any $E \in \mathcal{E}$, any $\boldsymbol{\phi} \in \Phi(E)$, and $(i, j) \in N^2$,*

$$c_i = c_j \Rightarrow u_i(\phi_i) \geq u_i(\phi_j)$$

.

4.3.2 Cost Solidarity or Cost Reward?

This part presents and motivates two sets of axioms which capture, to different degree, how one would like to reward communities for their lower provision costs. What is common to these axioms is that they impose requirements for changes in the cost profile. We start by presenting the reward axioms, which require that communities do not get a lower transfer as their cost decrease. Next, we introduce the solidarity axioms, which require particular changes in the cost profile to affect some communities in the same ways.

4.3.2.1 Reward axioms

Consider the view that each community should be held accountable of its own provision cost. We propose to express the resulting requirement by the following idea of reward: for any decrease in the provision cost of a single community, the transfer received by this community should not decrease. As change in costs may also lead to revise the location choice, we limit this requirement to changes that do not lead to such a revision. The following axiom formalizes this idea:

Individual Cost Reward (ICR). *For any $E = (B, \boldsymbol{c})$ and $E' = (B, \boldsymbol{c'})$ in $\mathcal{E}$, if $\exists j \in N$, $c_j > c'_j$ and, $\forall i \in N \backslash j$, $c_i = c'_i$ then, $\forall (\boldsymbol{\phi}, \boldsymbol{\phi'}) \in \Phi(E) \times \Phi(E')$ such that $\boldsymbol{h} = \boldsymbol{h'}$, we have $t_j \leq t'_j$.*

The former requirement focuses on individual changes. It leaves the possibility that a community gets a lower transfer while her provision cost decreases if some other communities also have decreasing costs. Besides, overlaps between different notions of a community may exist. Therefore, an extension of this requirement to groups can be worth considering. Such an extension requires that for any strict decrease in the individual provision cost of some communities, and provided this change does not lead to revise the optimal site, the aggregate transfer to this group of communities should not decrease. Note that this latter axiom implies the former.

Collective cost reward (CCR). *For any $E = (B, \boldsymbol{c})$ and $E' = (B, \boldsymbol{c'})$ in $\mathcal{E}$, if $\exists S \subset N$, $\forall i \in S$, $c_i > c'_i$ and $\forall i \in N \backslash S$, $c_i = c'_i$ then, $\forall (\boldsymbol{\phi}, \boldsymbol{\phi'}) \in \Phi(E) \times \Phi(E')$ such that $\boldsymbol{h} = \boldsymbol{h'}$, we have*

$$\sum_S t_i \leq \sum_S t'_i$$

.

4.3.2.2 Solidarity axioms

As soon as the provision costs are deemed akin to circumstances, there is no ground to justify a difference in welfare on this basis. The first solidarity principle requires that any decrease in the cost profile should result in a benefit for all. It is equivalent to the axiom of Extended Cost Monotonicity introduced in Laurent-Lucchetti and Leroux (2010).[19]

Full solidarity (FS). *For any* $E = (B, \boldsymbol{c})$ *and* $E' = (B, \boldsymbol{c'})$ *in* $\mathcal{E}$*, if* $\forall i \in N$*,* $c_i \geq c'_i$*, then,* $\forall(\boldsymbol{\phi}, \boldsymbol{\phi'}) \in \Phi(E) \times \Phi(E')$ *and* $i \in N$*,*

$$u_i(\phi_i) \leq u'_i(\phi'_i)$$

where $u'_i(h_i, t_i) = t_i - h_i c'_i$.

The following results will show that it is a strong requirement as it directly requires the transfers to depend solely on the minimal provision cost. Therefore, this axiom leaves no room for the recognition of the specific contribution of communities with lower costs. As argued earlier, we may be willing to reward communities for their low provision costs, which leads us to require that a community should never get a lower transfer as its costs get lower. As we know that this requirement is not compatible from the idea that all communities should benefit from a decrease in the provision costs of some, one possibility is to restrict the solidarity requirement to the communities with initially *greater* provision costs. This is because the reward - and therefore the welfare - of the communities with *lower* provision costs may be justified to get lower when the discrepancy between their provision costs and the others gets reduced. The following requirement captures this idea. It conveys an idea of solidarity while allowing for the recognition of the positive role played by some particular communities.

Solidarity toward higher-cost communities (SHC). *For any* $E = (B, \boldsymbol{c})$ *and* $E' = (B, \boldsymbol{c'})$ *in* $\mathcal{E}$*, if* $\exists S \subset N$*,* $\forall i \in S$*,* $c_i > c'_i$ *and* $\forall i \in N \backslash S$*,* $c_i = c'_i$ *then,* $\forall(\boldsymbol{\phi}, \boldsymbol{\phi'}) \in \Phi(E) \times \Phi(E')$ *and* $i \in N$ *such that* $c_i \geq \max_{j \in N} c_j$*,*

$$u_i(\phi_i) \leq u'_i(\phi'_i)$$

where $u'_i(h_i, t_i) = t_i - h_i c'_i$.

Note that the solidarity axioms rely on a comparison of different utility functions. This makes sense as we compare money-metric utilities.

[19] This axiom is also similar the weaker Monotonicity axiom introduced in Sakai (2012). The main difference is that this latter axiom only requires the welfare levels to be non-decreasing for an *individual* change in the cost that *strictly* decrease the minimal cost. The following results would obtain with such an axiom for continuous allocations.

4.3.3 Welfare bounds

The individual rationality requirement guarantees that no community looses from the project. It is a standard and appealing axiom.

Individual rationality (IR). *For any it $E \in \mathcal{E}$, any $\boldsymbol{\phi} \in \Phi(E)$, and $i \in N$,*

$$u_i(\phi_i) \geq 0$$

Notably, two additional welfare bounds have been proposed in the related literature. These axioms are generally weaker than the axioms we consider and do not lead to characterisations. Yet, they can be intuitively appealing.

The first is a lower bound on welfare. It requires that any community should not be worse off than when it gets an envy-free allocation in the hypothetical economy in which all other communities have the same preferences (here, provision costs). This requirement was first introduced by Moulin (1990) and proposed as a weakening of No envy in the context of the allocation of an indivisible good (Thomson, 2011). Its main justification relies on a two-stage argument that makes clear how this requirement extends the axiom of No envy among equals with a notion of solidarity. For a given community i with provision cost c_i, consider the economy in which all communities feature the same provision cost. No envy among equals requires to equalize all welfare levels to $\max(0; \frac{B-c_i}{n})$ in this setting. Now, consider the change from this distribution of the costs to any arbitrary distribution, holding c_i constant. This resulting environment is unambiguously better. A solidarity principle requires that no community is made worse off from the heterogeneity in the cost profile. This leads to the requirement the Identical-cost Lower bound (Moulin, 1990).

Identical-cost Lower Bound (ICLB). *For any $E \in \mathcal{E}$, any $\boldsymbol{\phi} \in \Phi(E)$, and $i \in N$,*

$$u_i(\phi_i) \geq \max\left(0; \frac{B-c_i}{n}\right)$$

This axiom requires productive efficiency. It further requires that the provision cost of the host is lower than the average provision cost. On the domain we consider, this axiom is weaker than No envy and stronger than No envy among equals. It conveys and idea of solidarity which is weaker than the idea conveyed by the two solidarity axioms formerly introduced. The three allocation rules we characterise in the following satisfy this axiom. In the following, it will become clear that the three allocations formerly introduced rely on these stronger notions of solidarity than the one conveyed by this axiom.

The second is an upper bound on welfare. It requires that no community gets more than what it would have got on its own. This requirement was introduced in the context of this problem to capture the idea of responsibility (Schokkaert et al, 2007). In this situation, the existence of a single project can be seen as a scarcity that is akin to circumstances. If transfers to the non-hosts are intended to compensate the communities from this specific feature of the situation, there is no ground to give more to a community than what it would get on its own. This gives rise to the Stand-alone upper bound (Moulin, 1992).

Stand-Alone Upper Bound (SAUB). *For any $E \in \mathcal{E}$, any $\boldsymbol{\phi} \in \Phi(E)$, and $i \in N$,*

$$u_i(\phi_i) \leq \max(0; B - c_i)$$

This axiom is logically independent from the reward axioms formerly introduced. It is stronger than the No dummy axiom, which requires that a community who would not benefit from the project on its own does not get any benefit. Interestingly, the Shapley value is the only rule to satisfy this requirement among the three we consider (Schokkaert et al, 2007).

4.4 Results

This section starts with the presentation of the implications of No envy. We then sequentially present characterisation results for the three allocation rules presented in Section 4.2. The three characterisations rely on the standard axioms to which differing axioms for solidarity or reward are required. Results are eventually summarized in Subsection 4.4.5.

4.4.1 Implications of No envy

The following lemma characterises the set of individually rational and envy-free allocation rules on $\mathcal{E}$.

Lemma 4.1. *The allocation rule Φ is individually rational and envy-free on $\mathcal{E}$ if and only if, $\forall E \in \mathcal{E}$, $\forall \boldsymbol{\phi} \in \Phi(E)$, $\exists p \in [c(N); \min\{B, c(N \backslash h)\}]$,*

$$\phi \in \{(\boldsymbol{e_h}, \boldsymbol{t}) | h \in \mathcal{H} \textit{ and } \forall i \in N, t_i = \frac{B-p}{n} + h_i p\}$$

As this Lemma shows, No envy considerably reduces the degrees of freedom of the problem. It requires allocative efficiency and reduces the problem to the determination of a single parameter, a premium to the host $p(E)$, in every economy. It further requires this premium to belong to the interval $[c(N); \min\{B, c(N \backslash h)\}]$. The proof of Lemma 4.1 is provided in Appendix 4.D. The intuition behind this result is the following. First, No envy requires that all non-hosts get the same transfer, then an allocation is fully characterised by the premium granted to the host. Preventing envy from the host to the non-hosts requires this premium not be higher than $c(N)$. Furthermore, no-envy from the non-hosts to the host sets a higher bound on this premium: by being host instead, a non-host would incur its own cost but get this premium. Therefore it should not be higher than any of the non-hosts' own costs, which is most constraining for the second lowest cost $c(N\backslash h)$. In case when the second lower cost is so high that it would prevent the project to be implemented, Individual Rationality further requires that the non-hosts do not have to give more than the whole benefit of the project to the host.

Interestingly, the allocation rules Φ^e and Φ^n both satisfy No envy and Individual Rationality on $\mathcal{E}$. They actually correspond to the two extreme compensation values $p = c(N)$ and $p = \min\{B, c(N\backslash h)\}$. Hence, the welfare egalitarian allocation rule (resp. the nucleolus) is the rule that minimizes (resp. maximize) the welfare of the host among all the envy-free and individually rational allocation rules. The next two propositions provide characterisations of these allocations.

4.4.2 Characterisation of the welfare egalitarian allocation rule

The first Proposition provides a characterisation of the welfare egalitarian allocation rule.

Proposition 4.1. *The welfare egalitarian allocation rule Φ^e is the only anonymous and efficient allocation rule which satisfies No envy among equals and Full solidarity on $\mathcal{E}$.*

A formal proof of this characterisation is provided in Appendix 4.E. An illustration on the economy E_0 conveys the main intuition. From the economy E_0, consider the economy E_0' in which all costs are equalized to 1. In this latter economy, Anonymity and No envy among equals together require to choose any location, and equalize the welfare levels across communities. Full Solidarity requires the welfare levels achieved to change uniformly from E_0' to E_0. Because the total welfare to allocate remains unchanged from E_0' to E_0, it requires the distribution of welfares to be the same in both economies. Therefore the allocation in E_0 has to be the welfare egalitarian allocation.

This characterisation is already a well known result. It is much in line with the characterisations performed in Laurent-Lucchetti and Leroux (2010) and Sakai (2012) and emphasizes the crucial role of their monotonicity requirement for the compensation value to be the actual provision cost. Note that Thomson (1997) also suggests an alternative characterisation of this allocation based on No envy and Welfare dominance under preference replacement, where this latter axiom requires the welfare levels to change in the same way only among the communities whose preferences are not changed. This latter characterisation presents the interest of dispensing with interpersonal comparisons of utilities.

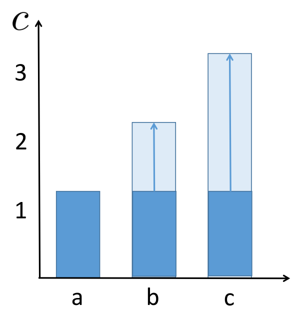


Figure 4.2: Comparisons performed to characterise the welfare egalitarian allocation rule.

As this result suggests it, Full Solidarity precludes the reward and the Stand-alone upper bound requirements introduced previously. In order to leave the possibility of rewarding lower cost communities, we have to weaken this specific requirement.

4.4.3 Characterisation of the nucleolus

Adopting a different perspective on the nature of the costs, the following proposition provides a characterisation of the nucleolus:

Proposition 4.2. *The nucleolus Φ^n is the only anonymous and efficient allocation rule which satisfies No envy, Individual rationality and Individual cost reward on $\mathcal{E}$.*

A complete proof of this characterisation is provided in Appendix 4.F. The intuition of the result is illustrated on the economy E_0 from Example 4.1. From the economy E_0, consider the economy in which the provision cost of community a is increased to 2. No-envy requires that the host be simply compensated to the extent of her own cost in such economy, that is to locate the project in the community a or b and give a premium of 2 to the host. In the allocation where the community a hosts the project, transfers are $(\frac{7}{3}, \frac{1}{3}, \frac{1}{3})$. Now consider the change to E_0. Individual Cost Reward requires that community a does not get a lower transfer than in the previous allocation. No-envy additionally requires the transfer to this community not to exceed that amount. It follows that the transfers have to be exactly $(\frac{7}{3}, \frac{1}{3}, \frac{1}{3})$ in E_0.

Note that the nucleolus does not satisfy the Stand-alone upper bound condition. Actually, Lemma 4.D shows that any envy-free allocation rule would fail to meet this requirement. In order to propose an allocation rule that satisfies this requirement, we must therefore weaken No envy.

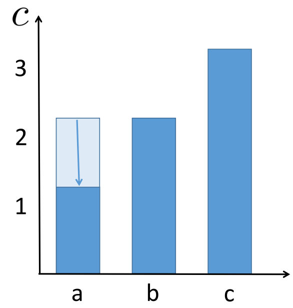


Figure 4.3: Comparisons performed to characterise the nucleolus.

4.4.4 Characterisation of the Shapley value

No envy constrains the solution in two distinct ways. First, it requires equalizing contributions among non-hosts. Second, it puts a higher bound on the extent to which the host could be rewarded. In this context, Solidarity toward higher-cost communities may constitute another argument for placing a higher bound on the reward. Replacing No envy by this axiom leaves room to strengthen Individual cost reward to Collective cost reward. This fully characterises the Shapley value.

Proposition 4.3. *The Shapley value Φ^s is the only anonymous and efficient allocation rule which satisfies No envy among equals, Solidarity toward higher-cost communities, and Collective cost reward on $\mathcal{E}$.*

A complete proof of this characterisation is provided in Appendix 4.G. We give here the intuition of the proof on the example E_0. First, consider the economy where all provision costs are set to 3. In this economy, Anonymity and No envy among equals together require considering all three allocations that consist in locating the project in any of the three communities and grant a premium of 3 to the host. Now consider a decrease in cost from 3 to 2 for the communities a and b. Efficiency, No envy among equals and Anonymity together require to consider the two allocations that consist in locating the project in a or b. Collective cost reward additionally requires that for any of these allocations the total transfer received by the two communities a and b is not lower than the transfer received in the previous economy, that is 1. Besides, Solidarity toward higher-cost communities requires that the community c is not made worse off from such a change. Therefore, c has to get exactly 0 in this economy. Anonymity and No envy among equals then require to equalize the welfares of a and b, that is to give $\frac{1}{2}$ for the non-host and $\frac{5}{2}$ for the host. Finally consider a further decrease of the provision cost of a from 2 to 1. Efficiency requires locating the project in community a, and Collective cost reward requires giving at least $\frac{5}{2}$ to this community. Solidarity toward higher-cost communities requires not giving less than respectively $\frac{1}{2}$ and 0 to the communities b and c. From budget balance, the three communities should get exactly these amounts. This corresponds to the Shapley value.

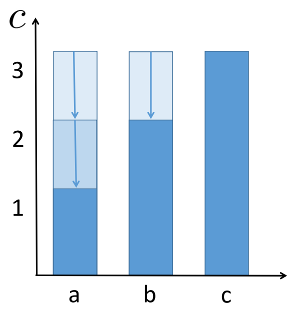


Figure 4.4: Comparisons performed to characterise the Shapley value.

One interesting conclusion that comes out of this result is that the Shapley value actually encompasses an idea of solidarity. An example of an anonymous and efficient allocation rule that jointly satisfies No envy among equals and Collective cost reward is the rule that always share the total surplus equally among the communities in $\mathcal{H}$.

4.4.5 Summary

In this section, we argued that considerations of reward could consistently be evoked and lead to different recommendations in the class of environment considered. All results are summarized in Table 4.1 and Figure 4.5.

	Egalitarian Φ^e	Nucleolus Φ^n	Shapley value Φ^s
Anonymity	Yes$^\bigstar$	Yes$^\blacksquare$	Yes$^\bullet$
Pareto efficiency	Yes$^\bigstar$	Yes$^\blacksquare$	Yes$^\bullet$
Individual rationality (IR)	Yes	Yes$^\blacksquare$	Yes
No envy among equals (NEE)	Yes$^\bigstar$	Yes	Yes$^\bullet$
Collective cost reward (CCR)	No	No	Yes$^\bullet$
Individual cost reward (ICR)	No	Yes$^\blacksquare$	Yes
Full solidarity (FS)	Yes$^\bigstar$	No	No
No envy (NE)	Yes	Yes$^\blacksquare$	No
Solidarity toward higher-cost communities (SHC)	Yes	Yes	Yes$^\bullet$
Identical-cost lower bound (ICLB)	Yes	Yes	Yes
Stand-alone upper bound (SAUB)	No	No	Yes

Table 4.1: Properties of the three allocation rules characterised on $\mathcal{E}$.
Superscripts indicate the characterisations presented in this section.

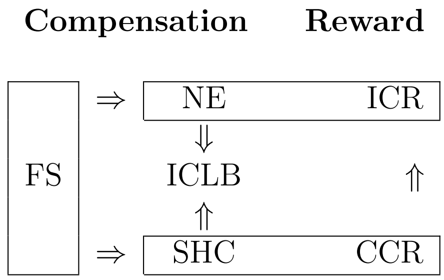


Figure 4.5: Logical relations among axioms for anonymous, individually rational and efficient allocation rules satisfying No envy among equals on $\mathcal{E}$.
Frames indicate the characterisations presented in this section.

4.5 Extensions

On top of the monetary building, maintaining and operating costs, a significant part of the total cost of a facility may result from the communities' disutility from the local nuisance that arise from hosting it. This raises questions about the direct application of the previous results. First, the actual compensation requirements of the communities may not be observable. This is dealt with in subsection 4.5.1. Second, different considerations may hold regarding the responsibility of the communities for different part of the costs. This raises the question of whether the equity principles proposed in the previous section could be applied separately on each part of the costs. This is dealt with in subsection 4.5.2.

4.5.1 Accounting for private information

A significant part of the provision cost of the communities may consist in their subjective disutility for the local nuisance associated with the project. From a practical point of view, such information may not be easily observed. At a more fundamental level, even if such information were actually observed, communities may still be considered as the best judge of their own interest and left free to state their own compensation requirements. This acknowledged, the challenge is to design a procedure in which the communities are incentivized to reveal enough information to argue credibly that a fair outcome is achieved. As we shall soon see, this considerably restricts the set of achievable allocations. Still we may propose arguments for a class of mechanisms that includes the mechanisms consisting in the direct application of the three allocation rules formerly characterised.

4.5.1.1 Mechanisms

We will focus on budget-balanced direct revelation mechanisms.[20] These are mechanisms that select alternatives from the following set :

$$\mathcal{A} = \{(\boldsymbol{e_h}, \boldsymbol{t}) | h \in \{0\} \cup N \text{ and } \sum_N t_i = B1_{\{h>0\}}\}$$

Definition 4.3. A *budget-balanced direct revelation mechanisms* is a function M that associates with any report $\hat{\boldsymbol{c}} \in \mathbb{R}^n_+$ a probability distribution over the set of alternatives $\Delta(\mathcal{A})$.

We denote by $\mathcal{M}$ the set of such mechanisms. Examples may be built from the allocation rules formerly characterised. The mechanism built on the welfare egalitarian allocation rule corresponds to the first price auction that was advocated for by Sakai (2012) in a similar context.[21] The mechanism built on the nucleolus corresponds to the second price auction. It was proposed by Minehart and Neeman (2002).[22] As they do not rely on the actual cost but on the communities stated cost, none of these mechanisms guarantees that the allocation rules are actually implemented through these mechanisms. For instance, in the first-price auction, the optimal community may be tempted to overstate her cost in order to get a higher compensation.

A way to ensure that some allocation is actually achieved is to ensure that truthfully reporting one's compensation requirement is a dominant strategy for all communities. This corresponds to the following requirement, called Strategyproofness.

Strategyproofness. *For any $E \in \mathcal{E}$, $i \in N$, $\hat{c}_{-i} \in \mathbb{R}^{n-1}_+$ and $c_i \in \mathbb{R}_+$,*

$$u_i(M(c_i, \hat{c}_{-i})) \geq u_i(M(\hat{c}_i, \hat{c}_{-i}))$$

It is well-known that no budget-balanced and strategyproof mechanism is ex-post Pareto efficient (Green and Laffont, 1979). This result also holds here and, none of the allocation rules formerly characterised can be implemented in dominant strategies. Ways out of this impossibility have been explored by weakening each of the three requirements. The approach that is adopted here consists in a weakening of ex-post Pareto efficiency and strategyproofness.

[20] We choose to focus on budget balanced mechanisms. This requirement has strong consequences. However, when the unanimity lower bound is required, it may be weakened to a no-deficit requirement and the results would remain.

[21] Sakai's proposal was actually more general as the author argued in favour of the mechanisms that locates the facility at the lowest reported cost and share the this cost proportionally to the communities' individual benefit.

[22] The mechanism proposed was actually slightly more general as the authors argued in favour of the mechanisms that locates the facility at the lowest reported cost and share the second lowest reported cost according to exogenous weights.

4.5.1.2 The identical-cost lower bound

We propose to focus on mechanisms that satisfy a weak axiom met by all the allocation rules formerly characterised: the *identical-preference lower bound*. When imposed at the ex-post stage, it requires a community to weakly prefer the outcome of the mechanism than a monetary transfer of $\max\left(0; \frac{B-c_i}{n}\right)$, whichever rational strategy is chosen. Note that this condition requires Productive efficiency and Individual rationality but not Allocative efficiency at the ex-post stage. Still, it requires the project to be located in a community with a lower provision cost than the average provision cost, which amounts to some degree of allocative efficiency.

Ex-post identical-cost lower bound. *For any $E \in \mathcal{E}$, $i \in N$, and $\hat{c} \in \mathbb{R}_+$,*

$$u_i(M(\hat{c})) \geq \max\left(0, \frac{B - c_i}{n}\right)$$

It is clear that if a mechanism meets this requirement, then it implements the set of allocations satisfying the identical-cost lower bound in dominant strategy. From the revelation principle, we know that there must exist a strategy-proof mechanism that implements this set. Besides this mechanism has to meet productive efficiency and always choose a location where the actual cost is lower than the average cost. In the two agent case, this requires ex-post efficiency. Yet, this condition is too strong to be met by budget-balanced mechanisms.

Proposition 4.4. *No budget-balanced mechanism satisfies the ex-post identical-preference lower bound.*

A way to weaken the previous requirement is to require it at the interim stage. This requires understanding how the communities' form their decision in the context of a mechanism, which requires to consider communities' preferences over a course of actions for which outcomes are uncertain. Many possibilities arise. One is the general Bayesian framework adopted in Borgers and Smith (2014), in which rationality is common knowledge and the communities are expected utility maximizers characterised by their preferences, their risk attitudes and their subjective beliefs. Another possibility is to consider that communities maximize the worst possible outcome of the mechanism. This maximin behavior was proposed and axiomatized by Minehart and Neeman (2002) in the context of the NIMBY problem. From this discussion, we may rather accept that there exist many possibilities for defining a rational decision criterion and for assessing the communities welfare in this setting. Therefore, we may be willing to design mechanisms that are robust in the sense that they satisfy some properties independently from these many possibilities. One such possibility is to require, for all communities, the existence of an opportunity to achieve their identical-cost lower bounds, independently from the others' choices.

Interim identical-cost lower bound. *For any $E \in \mathcal{E}$, $i \in N$, there exists $\hat{c}_i \in \mathbb{R}_+$ such that for any $\hat{c}_{-i} \in \mathbb{R}_+^{n-1}$,*

$$u_i(M(\hat{c}_i, \hat{c}_{-i})) \geq \max\left(0, \frac{B - c_i}{n}\right)$$

We first note that this is achievable.

Proposition 4.5. *There exist budget balanced mechanisms that satisfy the interim identical-preference lower bound.*

The proof of this result is provided in Appendix. It relies on the presentation of the first and the second-price auctions, previously introduced. It turns out that they both satisfy the interim identical-preference lower bound condition as, by reporting truthfully one's type, every community is ensured to get at least her identical-preference lower bound. To put it differently, no community can deprive another from the *opportunity* of achieving this welfare level. To get oa sense of what this axiom requires, consider a community whose true compensation requirement is c_2 and assume that this community is convinced to actually feature the second lowest cost. In a second-price auction, this community would find beneficial to understate her compensation requirement. If it turns out that she stated the lowest cost so that she actually hosts the project. This choice would be inefficient if there exists a community with a lower true compensation requirement than her. Yet, all other communities would actually benefit from this misreport: community 2 will always bear at least the entire inefficiency cost of her misreport. Of course, this community could have benefited from this misreport but in any case, the potentially large efficiency loss resulting from her risk taking would have limited consequences on the other's opportunities.

We are interested in characterising the whole class of mechanisms that satisfy the interim identical-preference lower bound. The next result establishes that there is a sense according to which this property extends exactly to the following class of mechanisms $\mathcal{U}$.

Definition 4.4. A mechanism M belongs to $\mathcal{U}$ if and only if

- It always implements a project that is beneficial according to the claims:

$$\min(\hat{c}_i) \geq B \iff M(\hat{\boldsymbol{c}}) = (0, \mathbf{0})$$

- When a project is implemented:
  - Transfers are budget-balanced : $\sum_N t_i = B$
  - $t_h \geq \frac{n-1}{n}(B - \hat{c}_h)$
  - $\forall i \in N \backslash \{h\}$, $t_i \geq \max\left(0, \frac{B - \hat{c}_i}{n}\right)$

The characterisation writes as follows:

Proposition 4.6. *A budget-balanced mechanism M satisfies the interim identical-preference lower bound if and only if it is payoff equivalent to a mechanism in $\mathcal{U}$ for maximin players.*

The proof of this result is provided in Appendix.

4.5.1.3 Conclusion

It comes with no surprise that the former allocation rules cannot be implemented under asymmetric information. Still, we showed that the mechanisms built upon the three allocation rules formerly characterised would still meet some reasonable distributive requirements. In the end, such mechanisms are not ex-post efficient, but, from an interim perspective, they are not too inefficient either. They do not ensure that any community will achieve some level of welfare at the ex-post stage, but they all offer to the communities the *opportunity* to benefit from the project to the extent of their own identical-cost lower bound.

At this stage, a crucial question still remains open: are all mechanisms in $\mathcal{U}$ Pareto optimal at the interim stage? An answer to this question would have to consider more precisely the likely behaviors of the communities in this context. I conjecture that further improvements may not be achievable without restricting the set of plausible behavior or imposing additional structure to this problem. For instance, the core of the argument proposed by Minehart and Neeman (2002) in favor of the second price auction relies on its greater ex-post efficiency. Still, this result is only established in a standard Bayesian framework. As soon as we allow for more general type spaces, little room seems to be left for Pareto improvement at the interim stage (Borgers et al, 2015, p.194).

This analysis brings additional arguments in favour of mechanisms that were advocated for in the context of this problem (Minehart and Neeman, 2002; Sakai, 2012). It also suggests and motivates the direct implementation of the Shapley value as another potentially desirable mechanism in this setting. An important remark is that the design of a mechanism upon the three allocation rules formerly characterised constitutes a significant departure from distributive to procedural justice. As was noted by Young (1995, chap. 8), a focus on fair *processes* rather than on fair *allocations* may be required in this specific context. It is actually in such a procedural perspective that the reward axioms make most sense as a manifestation of reciprocity.

4.5.2 An extension to additive composite costs

The results presented in Section 5.3 rely on a simple description of the cost and a unique perception of how each communities could be deemed responsible for it. Nevertheless, the communities' provision costs may result from the combination of many different sources. As in Sakai (2012), we may consider that the cost of a project is composed of at least two elements. The first is the monetary cost of constructing, operating and maintaining the project at a given community. This may depend on purely circumstantial characteristics such as location or the geological properties of land. The second source of cost is the community compensation requirement for bearing the nuisance. This may as well depend on many factors among which socio-demographic characteristics may play an important role. It seems unlikely that the same consideration would hold for each of these costs. Therefore, we shall look at the possibilities of combining compensation and reward when costs result from both characteristics. In this extension, we assume that the communities' provision costs can be decomposed additively into a part that elicit compensation, and another that elicit reward. To fix ideas, one can think of the cost of constructing, operating and maintaining the project to be tied to compensation but the communities' compensation requirement to elicit reward. Of course, this decomposition of the cost may be challenged and alternative decompositions could be considered. For instance, the cost of constructing, operating and maintaining the project may also depend on past investments in infrastructures, which are deemed to elicit reward.

When costs are the sum of a circumstantial and a responsibility costs, a natural extension of the former analysis would seek to conciliate Full Cost Solidarity for the circumstantial part of costs, and Collective cost reward and Solidarity toward higher-cost communities for the part of the cost that is deemed to elicit reward. Yet, in many settings, a tension between compensation and reward has been identified (see e.g. Fleurbaey, 2008). This extension proposes a solution to this problem.

4.5.2.1 The domain

In this framework, communities are characterised by two types of costs. One which is deemed circumstantial, denoted by $\boldsymbol{c^c} \in \mathbb{R}^n_+$, and one that elicits reward, denoted by $\boldsymbol{c^r} \in \mathbb{R}^n_+$. The total provision cost in community i therefore writes $c_i = c^c_i + c^r_i$. Let us denote by $\mathcal{E}^+$ the set of such economies. Most concepts defined on the domain $\mathcal{E}$ can be transposed to $\mathcal{E}^+$.

Example 4.2. Three similar communities, $i \in \{a, b, c\}$, can jointly undertake a public project. This project is estimated to yield a total monetary benefit of 6. It is associated to negative effects to the community that hosts it. Some additional monetary costs would be required to build the project in some communities. These additional costs are deemed circumstantial.[23] They are estimated to amount to 1, 0 and 4 in a, b and c respectively.

[23] As an example, a landfill is more expensive to build on sandy soil than clay. The geological properties of the land could be an example of a characteristic akin to circumstances.

Besides, some nuisances cannot be avoided. These nuisances are objectively similar on the three sites. Yet the communities' compensation requirements are evaluated to be 2, 4 and 1 for communities a, b and c respectively.[24] Considering these latter amounts as worth of reward, this defines the problem $E_0^+ = (6, (1, 0, 4), (2, 4, 1))$ in $\mathcal{E}^+$. In this problem, $\mathcal{H} = \{a\}$, $B = 6$ and $c(N) = 3$. This economy will be evoked for illustration in the rest of the section. It is represented on Figure 4.6.

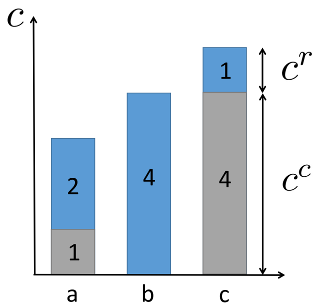


Figure 4.6: A representation of the situation E_0^+.

The crux of the problem is that the two costs and their correlation jointly determine the actual cost of the project. In the example 4.2, Pareto efficiency requires to locate the project in community a, regardless of the nature of her costs. A direct application of the welfare egalitarian solution to this problem would fail to to recognize the specific contribution of the community a. Still, a direct application of the Shapley value in this economy would unduly sanction community c as her featuring the highest cost only results from poor circumstances. As we will see, a direct application of the former axioms on each part of the costs is not possible. All this suggests that an application of the Shapley value would only be consistent with an idea of compensation for different circumstances in a setting in which the communities' circumstantial costs are equalized.

4.5.2.2 An allocation

We focus on allocation rules that allocate the total welfare among communities according to the Shapley value in the hypothetical economy in which all communities feature a same circumstantial cost. This is the sense of the following allocation Φ^+:

Allocation Φ^+. *Let $E = (B, \boldsymbol{c}^c, \boldsymbol{c}^r) \in \mathcal{E}^+$ and define the hypothetical economy $E' = (B, \boldsymbol{c}') \in \mathcal{E}$ such that $\forall i \in N$, $c'_i = c^{ref} + c^r_i$ where $c^{ref} = \min_N c_i - \min_N c^r_i$.*

[24] These "evaluations" for the communities' willingness to accept the nuisance could be thought as being inferred from valuation methods (hedonic pricing, benefit transfers, etc.) or as direct statements over their compensation requirements.

$\Phi^+(E)$ is the set of anonymous and efficient allocation rules that yields the distribution of welfare of the Shapley value $u^{Sh}(E')$.

In the Example 4.2, the allocation rule Φ^+ recommends to locate the project in a. The reference cost is $c^{ref} = 2$ so the distribution of welfare levels is similar to the distribution induced by the Shapley value in the hypothetical economy $(B, (4, 8, 3))$. Therefore, the distribution of welfare levels induced by Φ^+ in E is $(1, 0, 2)$. This is achieved by implementing the transfers $(4, 0, 2)$.

Several comments can be made. First, note that while Φ^+ yields a single allocation of utilities, some degrees of freedom are left regarding how to actually allocate the project is the situations for which there exists several optimal locations with different composition of the costs. In the 2-agent case, the allocation set Φ^+ contains an infinite number of allocation rules which allocates the project to $h \in \mathcal{H}$ and allocates the whole benefit of the project to the host if $c^c_h + c^r_{-h} \geq B$ and, otherwise, allocates $t_h = \frac{B+c^c_h+c^r_{-h}}{2}$ to the host and $t_{-h} = \frac{B-c^c_h-c^r_{-h}}{2}$ to the other community.

4.5.2.3 Axioms

The transposition of the axioms of Efficiency, Anonymity, Individual Rationality and No envy among equals on $\mathcal{E}^+$ is straightforward. The axioms related to compensation and reward are defined only for the part of the cost to which they are deemed to apply. This entails the following definitions.

Circumstance Solidarity. *For any $E = (B, \boldsymbol{c^c}, \boldsymbol{c^r}) \in \mathcal{E}^+$ and $E' = (B, \boldsymbol{c^{c\prime}}, \boldsymbol{c^r})$ in $\mathcal{E}^+$, if $\forall i \in N$, $c^c_i \geq c^{c\prime}_i$ then, $\forall (\boldsymbol{\phi}, \boldsymbol{\phi'}) \in \Phi(E) \times \Phi(E')$ and $i \in N$,*

$$u_i(\phi_i) \leq u'_i(\phi'_i)$$

where $u'_i(h_i, t_i) = t_i - h_i c'_i$.

Collective Reward for Responsibility Cost. *For any $E = (B, \boldsymbol{c^c}, \boldsymbol{c^r})$ and $E' = (B, \boldsymbol{c^c}, \boldsymbol{c^{r\prime}})$ in $\mathcal{E}^+$, if $\exists S \subset N$, $\forall i \in S$, $c^r_i > c^{r\prime}_i$ and $\forall i \in N \backslash S$, $c^r_i = c^{r\prime}_i$ then, $\forall (\boldsymbol{\phi}, \boldsymbol{\phi'}) \in \Phi(E) \times \Phi(E')$ such that $\boldsymbol{h} = \boldsymbol{h'}$, we have*

$$\sum_S t_i \leq \sum_S t'_i$$

.

Solidarity toward higher-responsibility-cost communities. *For any $E = (B, \boldsymbol{c^c}, \boldsymbol{c^r})$ and $E' = (B, \boldsymbol{c^c}, \boldsymbol{c^{r\prime}})$ in $\mathcal{E}^+$, if $\exists S \subset N$, $\forall i \in S$, $c^r_i > c^{r\prime}_i$ and $\forall i \in N \backslash S$, $c^r_i = c^{r\prime}_i$ then,*

$\forall(\boldsymbol{\phi}, \boldsymbol{\phi'}) \in \Phi(E) \times \Phi(E')$ *and* $i \in N$ *such that* $c_i^r \geq \max_{j\in N} c_j^r$,

$$u_i(\phi_i) \leq u_i'(\phi_i')$$

where $u_i'(h_i, t_i) = t_i - h_i c_i'$.

Corollary 4.1 that follows establishes that no anonymous, efficient allocations meet these former three requirements along with No envy among equals. Therefore, we will have to consider weaker requirements. We propose to weaken the requirements on the responsibility cost to situations in which circumstantial costs are uniform. This gives rise to the two following axioms.

Collective Reward for Responsibility Cost in Uniform Circumstances. *For any* $E = (B, \boldsymbol{c^c}, \boldsymbol{c^r})$ *and* $E' = (B, \boldsymbol{c^c}, \boldsymbol{c^{r\prime}})$ *in* $\mathcal{E}^+$, *if* $\forall i \in N$, $c_i^c = c$ *and* $\exists S \subset N$, $\forall i \in S$, $c_i^r > c_i^{r\prime}$ *and* $\forall i \in N \backslash S$, $c_i^r = c_i^{r\prime}$ *then,* $\forall(\boldsymbol{\phi}, \boldsymbol{\phi'}) \in \Phi(E) \times \Phi(E')$ *such that* $\boldsymbol{h} = \boldsymbol{h'}$, *we have*

$$\sum_S t_i \leq \sum_S t_i'$$

Solidarity toward higher-responsibility-cost communities in Uniform Circumstances. *For any* $E = (B, \boldsymbol{c^c}, \boldsymbol{c^r})$ *and* $E' = (B, \boldsymbol{c^c}, \boldsymbol{c^{r\prime}})$ *in* $\mathcal{E}^+$, *if* $\forall i \in N$, $c_i^c = c$ *and* $\exists S \subset N$, $\forall i \in S$, $c_i^r > c_i^{r\prime}$ *and* $\forall i \in N \backslash S$, $c_i^r = c_i^{r\prime}$ *then,* $\forall(\boldsymbol{\phi}, \boldsymbol{\phi'}) \in \Phi(E) \times \Phi(E')$ *and* $i \in N$ *such that* $c_i^r \geq \max_{j\in N} c_j^r$,

$$u_i(\phi_i) \leq u_i'(\phi_i')$$

where $u_i'(h_i, t_i) = t_i - h_i c_i'$.

4.5.2.4 A characterisation

We have the following characterisation:

Proposition 4.7. *An anonymous and efficient allocation rule satisfies Individual rationality, No envy among equals, Circumstance solidarity, Collective reward for responsibility cost in uniform circumstances, and Solidarity toward higher-responsibility-cost communities in uniform circumstances on* $\mathcal{E}^+$ *if and only if it belongs to* Φ^+ .

A proof of this results is provided in Appendix 4.K. We provide an intuition of it on Example 4.2. It relies on the fact that there exists a single reference level for the circumstantial costs that, when equalized across the communities, keeps the total net benefit of the project constant. It is this reference level that defines Φ^+.

A direct Corollary of this result relates to the difficulty to achieve Circumstance Solidarity, Collective Reward for Responsibility Cost, and Solidarity toward higher-responsibility-cost communities on $\mathcal{E}^+$.

Corollary 4.1. *No anonymous and efficient allocation rule satisfies Individual rationality, No envy among equals, Circumstance solidarity, Solidarity toward higher-responsibility-cost communities, and Collective reward for responsibility cost on $\mathcal{E}^+$.*

The proof of this result is provided in Appendix 4.L. We show that the allocation Φ^+ does not satisfies Collective reward for responsibility cost on a specific example.

4.6 An illustrative survey

Any axiomatic analysis remains speculative until it is confronted with the observation of actual judgments and this is what lacks from this analysis. In particular, the axioms of reward were only motivated based on intuition and anecdotal evidence. As a further confrontation, we implement a survey in order to test for the relevance of these axioms and the overall analysis in accounting for actual judgments. This survey was administered on-line among 257 University students in Toulouse between November, 13th and December, 13th, 2014. The main results are presented in this section. Further details are presented in Appendix 4.M.

The vignettes presented to the participants are intended to capture situations in $\mathcal{E}$. In order to keep the problem simple, the situation of the vignette corresponded to the simplest two-community problem.[25] The costs are presented as a direct expression, from the communities, of their willingness to accept the nuisance associated with the project. Two framings are proposed. In one of them, the project is a wind park. In the other, it is a wastewater treatment plant.[26] In the main vignette the project is expected to yield a benefit of five million euros and the respective provision costs of communities a and b are one and three million euros. This defines the situation $E_1 = (5, (1, 3))$.

After being presented with this situations, respondents are also presented with the problems $E_2 = (5, (1, 6))$, $E_3 = (5, (3, 3))$ and $E_4 = (5, (1, 2))$ in order. For each of these problems, they were asked whether they think the project should be implemented, and, when so, in which community, and how its benefit should be shared. The distribution of the answers for each case is presented in Table 4.2, and for the participants who chose to implement a project and locate it in "A" or in "A or B", the average amount given to A for hosting the project is reported on Figure 4.7.[27] We observe the average amount granted to the host is greater in E_2 than in E_4 (t-test for difference of the means, p-value=0.0025). This is consistent with the idea of reward. Nevertheless, another difference is that the

[25] Note that in the 2-agent case, the nucleolus and the Shapley value coincide.

[26] As the answers related to the allocation of the benefit does not significantly differ between the two situations, answers are reported jointly.

[27] When the participants answered "A or B, indifferently" for the location choice, they were asked to state how much A should get, would it be chosen to be the host.

amount granted to the host is significantly lower in E_3 than in E_1 and E_2 (t-test for difference of the means, p-value<0.001 in both cases). This result is surprising as none of the three allocation rules that have been considered requires this: the welfare egalitarian allocation rule requires this amount to increase from E_1 to E_3 whereas the nucleolus and the Shapley value both require it to remain constant. An increase in the allocation from E_4 to E_3 is also required by the two latter rules. Instead, a decrease is observed (t-test for difference of the means, p-value=0.011). In the end, none of the allocation rules introduced in this analysis is able to fully account for the observed answers. There may be at least two explanations for that. One possible explanation is the more frequent and intuitive appeal of the equal split heuristics in the situation E_3. Another one is the adhesion to a principle of ex-ante fairness in the situation in which the host is to be decided through a lottery, which is suggested by the answer “A or B, indifferently”.

%	No project	Project located in “A”	“A or B”	“B”
$E_1 = (1,3)$	9	81	9	1
$E_2 = (1,6)$	11	77	10	2
$E_3 = (3,3)$	17	2	81	0
$E_4 = (1,2)$	8	65	23	4

Table 4.2: Fraction of the repondents choosing each location for the facility in the different problems.

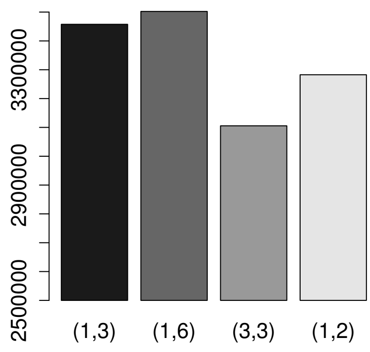


Figure 4.7: Average compensation chosen for the host in four different cases.

Next, respondents were asked to state their adhesion to simple statements. These were intended to reflect the main axioms chosen in the previous analysis.

- **Principle A (Pareto efficiency):** "The project should be located in the community A as its estimation of the damage borne is lowest."
- **Principle B:** "It is a project of general interest. The hosting community should not receive any particular compensation for hosting it."
- **Principle C:** "Only the hosting community bears its nuisance. It should therefore get the whole benefit of it."
- **Principle D (No envy among equals (NEE)):** "Would both community have the same estimation for the damage borne, they should benefit from it to the same extent. Then, the community which hosts the project should be compensated to the extent to its own estimation of the damage and the remaining benefit should be shared equally."
- **Principle E (Full solidarity (FS)):** "Would they have an estimation of the damage low or high, all communities should benefit from the project to the same extent."
- **Principle F (Individual cost reward (ICR))[28]:** "The community whose estimation of the damage is lowest contributes to the interest of all. It should be rewarded for that."

(in %)		Agree	Disagree	No opinion
Principle A	Windpark	86	14	0
	Wastewater plant	75	24	2
Principle B	Windpark	13	86	0
	Wastewater plant	13	86	2
Principle C	Windpark	49	50	1
	Wastewater plant	39	60	2
Principle D	Windpark	67	30	3
	Wastewater plant	71	26	3
Principle E	Windpark	34	62	4
	Wastewater plant	40	55	5
Principle F	Windpark	71	23	6
	Wastewater plant	57	33	10

Table 4.3: Adhesion to the different principles.

We observe that the principles of Pareto efficiency and No envy among equals are widely accepted. A majority of respondents seems to disagree with the principle of full cost solidarity while agreeing with the principle of individual cost reward. However, these differences could be due to the ambiguities in the statement of the principles and the unusual framing of the problem. In order to get more insights about the perception of

[28]Note that in the two-agent case, ICR and CCR are equivalent.

the principles, we investigate how the adhesion to a given principle correlates with the allocation to the host for the respondents who chose A as the host.[29] Regression results are presented in Table 4.4.

Allocation to A (million euros)	Case (1,3)	Case (1,2)	Case (1,6)
Constant	2.97^{***}	2.99^{***}	3.71^{***}
	(0.20)	(0.32)	(0.33)
Adhesion A	0.36^{**}	0.20	−0.30
	(0.18)	(0.26)	(0.27)
Adhesion B	-0.27^{*}	-0.37^{*}	-0.63^{***}
	(0.15)	(0.19)	(0.21)
Adhesion C	0.50^{***}	0.53^{***}	0.51^{***}
	(0.09)	(0.12)	(0.13)
Adhesion NEE	0.02	−0.10	−0.14
	(0.11)	(0.14)	(0.15)
Adhesion FS	−0.14	-0.36^{***}	-0.42^{***}
	(0.10)	(0.13)	(0.14)
Adhesion ICR	0.00	0.11	0.15
	(0.11)	(0.14)	(0.15)
R^2	0.18	0.20	0.21
Adj. R^2	0.16	0.17	0.19
Num. obs.	207	168	198

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 4.4: Relationship between the adhesion to the principles and the allocation to the host for the respondents who choose A as the host.

The main observation that comes out of this table is the expected correlation between the adhesion to FS and a lower allocation to the host. Yet we do not observe that the adhesion to ICR correlates with this allocation. However, the expected correlation are observed when we look at the respondents who completely agreed with the principles as shown on Table 4.5. A puzzling observation is that this correlation is not observed in all these cases and, particularly, not in the case (1,3) which, contrary to the two other cases, was decided after being shown the principles and the different allocation rules. This suggests an effect of the order of the different elements which could be investigated further.

Figures 4.8 and 4.9 show the judgments expressed regarding different allocations in the problem E_1. The allocations proposed all consider the community A as the host community. The difference lies in the allocation of the benefit of the project:

- **Allocation 1**: "The community A should get half the benefit, which is 2.5 million euro."

[29]The total number of respondents is 257, so choosing A as the host in the cases (1,3), (1,2), and (1,6) gathers respectively 81, 65 and 77% of the respondents.

Allocation to A (million euros)	Case (1,3)	Case (1,2)	Case (1,6)
Constant	3.31***	3.04***	3.33***
	(0.07)	(0.09)	(0.09)
Strong adhesion A	0.20**	0.10	0.06
	(0.10)	(0.12)	(0.13)
Strong adhesion B	−0.27	−0.14	−0.46
	(0.33)	(0.35)	(0.44)
Strong adhesion C	0.71***	1.01***	0.93***
	(0.15)	(0.19)	(0.20)
Strong adhesion NEE	−0.01	0.01	−0.15
	(0.11)	(0.13)	(0.14)
Strong adhesion FS	−0.11	−0.43**	−0.59***
	(0.16)	(0.21)	(0.22)
Strong adhesion ICR	0.03	0.29**	0.35**
	(0.11)	(0.14)	(0.15)
R^2	0.15	0.22	0.20
Adj. R^2	0.13	0.19	0.17
Num. obs.	207	168	198

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 4.5: Relationship between a strong adhesion to the principles and the allocation to the host for the respondents who choose A as the host.

- **Allocation 2 (Egalitarian allocation)**: "The community A should get her own valuation of the damage borne, which is 1 million euros, plus half the remaining benefit, which is 2 million euros, thus, in total, 3 million euro."
- **Allocation 3**: "The community A should get the average valuation of the damage borne, which is 2 million euros, plus half the remaining benefit, which is 1.5 million euros, thus, in total, 3.5 million euro."
- **Allocation 4 (nucleolus, Shapley value)**[30]: "The community A should get the other's valuation of the damage borne, that is 3 million euros, plus half the remaining benefit, which is 1 million euros, thus, in total, 4 million euro."
- **Allocation 5**: "The community A should get the whole benefit of the project, which is 5 million euro."

Both figures actually suggest that the welfare egalitarian allocation is the allocation deemed fairest by a majority of the respondents in this situation. However, according to Figure 4.9, fairness ideals are diverse: 45% of them would deem fairest to give a greater amount to the host than the amount necessary to equalize welfare levels. This is also surprising as we observed a much lower adhesion to the idea of full solidarity than to the idea of reward. This suggest that the judgments that are observed are not

[30]Note that in the two-agent case, the nucleolus and the Shapley value are confounded.

fully consistent with the individual adhesion to principle. A consequence of this it that the judgments reported here may be expected to evolve with the awareness between the logical relationship of the different elements.

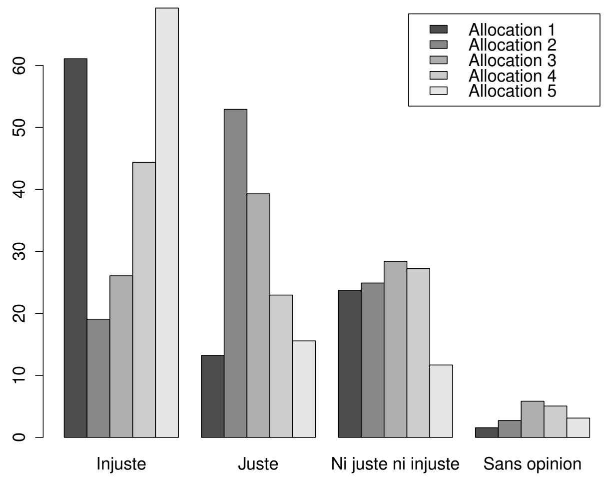


Figure 4.8: Percentage of respondents deeming each of the different allocations proposed (respectively) “unfair”, “fair”, “neither fair, nor unfair” or “without opinion” in the situation E_1.
The allocations 1, 2, 3, 4 and 5 respectively correspond to an allocation of 2.5, 3, 3.5, 4 and 5 millions euros to the first community (the host).

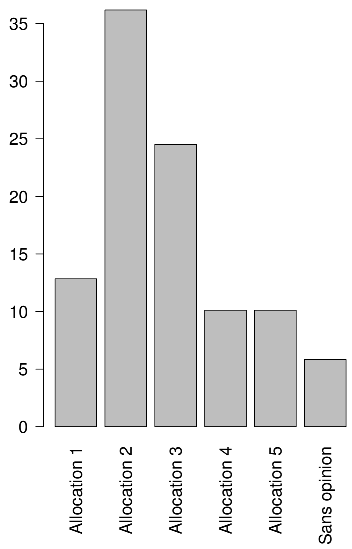


Figure 4.9: Percentage of respondents chosing each allocation as “the fairest” in the situation E_1.

This survey suggests that the welfare egalitarian allocation remains the most reasonable allocation in this setting. Yet, it also suggests that considerations of reward play a role in the formation of individual judgments. In a more thorough analysis, we may be willing to test whether the adhesion the welfare egalitarian allocation resists the potential arguments in favor of other allocations that have been suggested in this analysis. In particular, we may be willing to observe the adhesion to the view that the host should not get a lower transfer would he have a lower compensation requirement as this is a crucial element of this analysis.

Conclusion

In the simple description of NIMBY environments three allocations proposed in the literature are characterised on the basis of No envy, and axioms related to considerations of solidarity and reward. We further discussed the consequences of asymmetric information and brought additional arguments in favor of two classical mechanisms that were proposed in the context of this problem. We further discussed the necessary arbitrage between principles in settings where the provision costs consist in the sum of circumstantial and responsibility costs.

Despite its initial motivation in the context of the NIMBY problem, this analysis may prove relevant in other problems. A more general model would allow each type to encompass community-specific valuations that may be observable or unobservable. Part of these valuations may be deemed to elicit reward while another part may be deemed worth compensating. The resulting framework is presented in Table 4.6. Among the examples that would fit this description, the problem of allocating of property right on the commons, like water, fish, or clean air may be of great interest. Concerns about overexploitation of these resources have led to the establishment of property rights in the form of water quotas, fishing quotas or emissions rights. While guaranteeing the efficient use of these resources, the establishment of property requires the exclusion of former users and has to cope with heated debate and vivid opposition (Raymond, 2003; Hanesson, 2004). This makes fair allocations particularly relevant to this problem. In the case of a fishery, the market value of the catch associated with the fishing quota is an observable common value B. Fisherman may express various willingness to undertake the effort to fish v_i^{int} and may incur various operation and maintenance costs c_i^{ext}. In this example, a community's observable valuation v_i^{ext} is defined as $B - c_i^{ext}$. Both the observable and the unobservable valuation can be decomposed into a part that elicit reward and another that elicit compensation.

While this discussion suggests that the framework presented could be relevant to many problems of current interest, we also have to acknowledge that this framework may require

	Communities valuation	
	Observable	Unobservable
Reward	$v_i^{ext,r}$	$v_i^{int,r}$
Compensation	$v_i^{ext,c}$	$v_i^{int,c}$

Table 4.6: Characteristics of the communities in a general context.

to be extended further to tackle the crux of NIMBY conflicts. A first limitation of this model is that we assumed that the project was yielding a certain and well-defined common benefit. This may not be the case. In some instances such as in the case of prisons or refugee camps, this benefit may not easily be expressed in monetary terms. Our analysis does not directly carry over to these settings. A second - and maybe most crucial - limitation of this approach is that we assumed away externalities in the costs. Yet, these externalities may actually be the crux of the issue. Developments along this line seem to be a necessary step toward the design of justifications applicable to public decision making in NIMBY contexts.

This being said, we shall be confident with the conclusion of this analysis before considering any further extension. For this, a confrontation with the actual judgments held in society is required. In a tentative survey, we find that the welfare egalitarian allocation attracts the support of a majority of the respondents but we also observe a large adhesion for an idea of rewarding the host. This suggests that the observed judgments may be susceptible of change. These potential changes would be particularly interesting to investigate further.

Appendix

4.A Notations

N	Set of communities
$\mathcal{E} = \mathbb{R}_+ \times \mathbb{R}_+^n$	Set of economies
$E \in \mathcal{E}$	Particular economy
B	Benefit of the project
$\boldsymbol{c} = (c_i)_{i \in N}$	Provision costs
$\mathcal{H} = \arg\min_N c_i$	Set of optimal locations
$\boldsymbol{\phi} = (\boldsymbol{h}, \boldsymbol{t})$	Allocation
$\boldsymbol{h} = (h_i)_{i \in N}$	Vector of hosting status
h	Host ($h = 0$ if no project is undertaken)
$\boldsymbol{t} = (t_i)_{i \in N}$	Vector of transfers
$F(E) = \{(\boldsymbol{h}, \boldsymbol{t}) \vert \sum_{i \in N} t_i \geq B 1_{\{h>0\}}\}$	Feasible allocations
$\boldsymbol{\Phi} : \mathcal{E} \to 2^{F(E)}$	Allocation rule
$\boldsymbol{\Phi^e}$	Welfare egalitarian allocation rule
$\boldsymbol{\Phi^n}$	nucleolus
$\boldsymbol{\Phi^s}$	Shapley value
$\mathcal{E}^+ = \mathbb{R}_+ \times \mathbb{R}_+^n \times \mathbb{R}_+^n$	Set of economies with additive costs
$\boldsymbol{e}$	Vector with all components equal to 1
$\boldsymbol{e_i}$	Vector with i's component equal to 1, and others 0
$\boldsymbol{0}$	Null vector

4.B Derivation of the nucleolus on $\mathcal{E}$

Let $E \in \mathcal{E}$ and consider $h \in \mathcal{H}$, $\boldsymbol{t}$ be a transfer scheme and $S \subset N$. The deficit associated with the coalition S is:

$$d(S) = \begin{cases} -\min(B, c(S)) - t(S) & \text{if } h \in S \\ -t(S) & \text{if } h \in S \end{cases}$$

It is straightforward to check that all deficits associated with the equal sharing rule are negative. Hence, all deficits associated with the nucleolus are negative as well. Let us consider $\boldsymbol{t^n}$, the nucleolus. Note that the deficits decrease with the inclusion of any $j \neq 1$. It follows that deficits are minimal for coalitions $N\backslash\{i\}$, $i \in N$.

We have:

$$\begin{cases} d(N\backslash\{i\}) = t_i^n & \text{for all } i \neq h \\ d(N\backslash\{1\}) = t_1^n - \min(B, c(N\backslash\{h\})) \end{cases}$$

When these deficits are equalized, the minimal deficit is maximized. This gives the following system :

$$\begin{cases} t_1 - \min(B, c(N\backslash\{h\})) & = t \\ t_i & = t \quad \text{for all } i \neq h \\ \sum_{i\in N} t_i & = B \end{cases}$$

This system is invertible and its solution yields the nucleolus.

4.C Derivation of the Shapley value on $\mathcal{E}$

First note that, for any coalition $S \subseteq N$, the value of a coalition writes: $v(S) = B - \min\left(B, \min_S(c_i)\right)$. Using the additivity property of the Shapley value and the result presented in Dehez (2013), an expression of the welfare distribution induced by the Shapley value follows:

$$u_{\sigma^{-1}(i)} = \frac{B}{n} - \frac{\min(B, c_{\sigma^{-1}(n)})}{n} + \sum_{k=i}^{n-1} \frac{\min(B, c_{\sigma^{-1}(k+1)}) - \min(B, c_{\sigma^{-1}(k)})}{k}$$

This distribution of welfare is implemented by choosing $h \in \mathcal{H}$ and through the following transfers:

$$t_{\sigma^{-1}(i)} = \frac{B}{n} - \frac{\min(B, c_{\sigma^{-1}(n)})}{n} + \sum_{k=i}^{n-1} \frac{\min(B, c_{\sigma^{-1}(k+1)}) - \min(B, c_{\sigma^{-1}(k)})}{k} + c(N)h_i$$

4.D Proof of Lemma 4.1

Let Φ anonymous, individually rational and envy-free allocation rule, $E \in \mathcal{E}$ and $\boldsymbol{\phi} = (\boldsymbol{e_h}, \boldsymbol{t}) \in \Phi(E)$. First, let's consider $(i, j) \in N\backslash h^2$ two different non-hosts. No envy from i to j implies $t_i \geq t_j$ and No envy from j to i, $t_i \leq t_j$. Hence, $t_i = t_j$: all non-host should get the same transfer. Let's denote by t_h and t_{nh} the respective transfers to the host and the non-hosts. No envy from the host to the non-hosts implies $t_h - c_h \geq t_{nh}$ or, equivalently, $t_h - t_{nh} \geq c_h$. Finally, No envy from the non-hosts to the host implies $t_{nh} \geq t_h - c_j$ for all $j \in N \backslash h$ or, equivalently, $c(N \backslash h) \geq t_h - t_{nh}$. This requires $h \in \mathcal{H}$ and $t_h - t_{nh} \in [c(N); c(N \backslash h)]$. Let us define $p = t_h - t_{nh}$, the premium to the host. The budget balance condition writes $t_h + (n-1)t_{nh} = B$, which can be rewritten $t_h = \frac{B}{n} + \frac{n-1}{n}p$. Therefore, the welfare of the host writes $u_h(\phi_h) = \frac{B}{n} + \frac{n-1}{n}p - c(N) \geq \frac{B-c(N)}{n} \geq 0$, where the first inequality comes from the fact that $p \geq c(N)$ and the second from the efficiency of the project: $B \geq c(N)$. Hence the rationality of the host is required by No envy. The welfare of the non-hosts writes $u_i(\boldsymbol{\phi}) = \frac{B-p}{n}$, which is positive if and only if $p \leq B$. Therefore, $p \in [c(N); \min(B, c(N \backslash h))]$. In summary, we have $\forall i \in N$, $t_i = \frac{B-p}{n} + h_i p$. Conversely, it is straightforward to check that such a solution is envy-free.

4.E Proof of Proposition 4.1

Let Φ be an anonymous and efficient allocation rule satisfying No envy among equals (NEE) and Full Solidarity (FS) on $\mathcal{E}$ and $E = (B, \boldsymbol{c}) \in \mathcal{E}$. If $B < \min\limits_N c_i$, Pareto efficiency requires that no project is implemented. Otherwise, define $E' = (B, \boldsymbol{c'})$ such that for all $i \in N$, $c'_i = \min\limits_N c_i$. We have, from Anonymity and NEE, $\Phi(E') = \{(\boldsymbol{e_h}, \boldsymbol{t}) | h \in N \text{ and } t_i = (h_i - \frac{1}{n}) \min\limits_N c_i\}$. Now consider $\mathcal{H} = \arg\min\limits_N \{c_i\}$, the set of optimal location in E. Efficiency and Anonymity require that $h \in \mathcal{H}$ if and only if there exists $\boldsymbol{\phi} \in \Phi(E)$ such that h is the host in $\boldsymbol{\phi}$. Besides, FS requires that all communities have a higher welfare level in E' than in E. Because the net value of the project is the same in both economies, budget balance further requires the allocation of welfare to be the same as in E'. Therefore, Φ can only be Φ^e. Conversely, Φ^e is anonymous, efficient and satisfies NEE and FS on $\mathcal{E}$.

4.F Proof of Proposition 4.2

4.F.1 Characterisation

Let Φ be an anonymous and efficient allocation rule which satisfies No envy, Individual rationality (IR) and Individual cost reward (ICR) on $\mathcal{E}$, $E \in \mathcal{E}$ and $h \in \mathcal{H}$. If $B < \min\limits_N c_i$, Pareto efficiency requires that no project is implemented. If $B = \min\limits_N c_i$, either no project is implemented, or a project is implemented and all welfare levels set to 0. Otherwise, let $\epsilon > 0$ such that $\epsilon < B - c(N)$ and define $E' = (B, \boldsymbol{c'})$ such that $c'_h = \min(B - \epsilon, c(N \backslash h))$ and $\forall i \in N \backslash h$, $c'_i = c_i$. Consider $\mathcal{H} = \arg\min\limits_N \{c_i\}$, the set of optimal location in E. Productive efficiency first requires that a project is implemented. Together with Anonymity it further requires that $h \in \mathcal{H}$ if and only if there exists $\boldsymbol{\phi} \in \Phi(E)$ such that h is the host in $\boldsymbol{\phi}$. The same applies to the set of optimal locations in E'. As any optimal location in E in also optimal in E', we know that for any $h \in \mathcal{H}$, there exists $(\boldsymbol{\phi}, \boldsymbol{\phi'}) \in \Phi(E) \times \Phi(E')$ such that h is the host in $\boldsymbol{\phi}$ and $\boldsymbol{\phi'}$. By Lemma 4.1, $t'_h \geq \frac{n-1}{n} \min(B - \epsilon, c(N \backslash h))$ in E'. From ICR, we know, $t_h \geq t'_h$. Besides, $t_h \leq \frac{n-1}{n} \min(B, c(N \backslash h))$ by Lemma 4.1. As the former inequality must hold true for any $\epsilon > 0$, we must have $t_h = \frac{n-1}{n} \min(B, c(N \backslash h))$ and, from Lemma 4.1 $t_i = -\frac{1}{n} \min(B, c(N \backslash h))$ for all $i \in N \backslash h$. Then, we must have $\boldsymbol{\phi} \in \Phi^n(E)$. Conversely, the nucleolus meets all previously stated axioms.

4.F.2 Independence of the axioms

- The allocation rule defined by $\forall E \in \mathcal{E}$, $\Phi(E) = \{(\boldsymbol{e_h}, \boldsymbol{t}) | h \in \mathcal{H} \text{ and } \forall i \in N, t_i = (h_i - \frac{1}{n}) c(N \backslash h)\}$ satisfies all axioms but IR on $\mathcal{E}$.
- The welfare egalitarian allocation rule Φ^e satisfies all axioms but ICR on $\mathcal{E}$.
- The Shapley value Φ^s satisfies all axioms but No envy on $\mathcal{E}$.

4.G Proof of Proposition 4.3

4.G.1 Characterisation

First note that it is straightforward to check that Φ^s satisfies No envy among equals (NEE), Solidarity toward higher-cost communities (SHC), and Collective cost reward (CCR) on $\mathcal{E}$.

Let Φ be an anonymous and efficient allocation rule which satisfies NEE, SHC and CCR on $\mathcal{E}$. Let $E = (B, \boldsymbol{c}) \in \mathcal{E}$ and σ be a permutation of N such that $c_{\sigma^{-1}(1)} \leq ... \leq c_{\sigma^{-1}(n)}$. In order to simplify the notations, indexes are redefined according to this new ordering.

If $E \in \mathcal{E} \backslash \mathcal{E}^*$, Pareto efficiency requires that $\Phi(E) = \{(\boldsymbol{0}, \boldsymbol{0})\}$. Otherwise, for any $k \leq n$, define $E^k = (B, \boldsymbol{c^k})$ where $\boldsymbol{c^k}$ is defined by $c_i^k = \min(B, c_k)$ if $i < k$ and $c_i^k = \min(B, c_i)$ otherwise. Let m be the highest index such that $c_m \leq B$ and consider the property H_k: "$\Phi(E^k) = \Phi^s(E^k)$". We show that H_k is true for any $k \leq m$ by decreasing induction. First consider $E^m = (B, B\boldsymbol{e})$. Pareto efficiency and Anonymity together require that all communities are considered as potential hosts. No envy among equals further requires, $\Phi(E^n) = \{(\boldsymbol{e_h}, \boldsymbol{t}) | h \in N \text{ and } \forall i \in N, t_i = (h_i - \frac{1}{n}) \min(B, c_n)\}$. Therefore $\Phi(E^n) = \Phi^s(E^n)$. Assume, now, that H_{k+1} is true, that is $\Phi(E^{k+1}) = \{(\boldsymbol{e_h}, \boldsymbol{t}) | h \in \mathcal{H} \text{ and } \forall i \in N, t_i = \sum_{j=i}^{n-1} \frac{min(B, c_{j+1}^{k+1}) - min(B, c_j^{k+1})}{j} - \frac{min(B, c_n)}{n} + c_{k+1} h_i\}$. If $c_k = c_{k+1}$, the previous reasoning can be iterated and H_k is trivially true. If $c_k < c_{k+1}$, let $\boldsymbol{\phi^k} = (\boldsymbol{h^k}, \boldsymbol{t^k}) \in \Phi(E^k)$. From efficiency and anonymity, we know that $\boldsymbol{\phi^k}$ recommends to locate a project at any $h \leq k$.

For any such h, let $\boldsymbol{\phi^{k+1}} = (\boldsymbol{h^{k+1}}, \boldsymbol{t^{k+1}})$ be the allocation in $\Phi(E^{k+1})$ that recommends the same location. SHC first requires that, for all $i > k$, $t_i^k \geq t_i^{k+1}$. Besides, CCR requires that $\sum_{i \leq k} t_i^k \geq \sum_{i \leq k} t_i^{k+1}$. Assume one of these inequalities is strict, then, summing all these inequalities and using Budget balance, we get a contradiction. Therefore, these inequalities must hold with equality. In particular,

$$\sum_{i \leq k} t_i^k = \sum_{i \leq k} t_i^{k+1} = k \left(\sum_{j=k+1}^{n-1} \frac{\min(B, c_{j+1}) - \min(B, c_j)}{j} - \frac{\min(B, c_n)}{n} \right) + \min(B, c_{k+1})$$

Besides, Anonymity requires that welfare levels are equalized for all $i \leq k$. Given the expression of the total welfare to allocate among these communities, we get for any $i \leq k$

$$u_i(\phi_i^k) = \sum_{j=k}^{n-1} \frac{\min(B, c_{j+1}) - \min(B, c_j)}{j} - \frac{\min(B, c_n)}{n}$$

This yields $\Phi(E^k) = \Phi^s(E^k)$. Then H_k is true for any positive k. In particular, H_1 writes $\Phi(E) = \Phi^s(E)$ which establishes unicity.

4.G.2 Independence of the axioms

- The nucleolus Φ^n meets all axioms but CCR.
- Consider the allocation rule Φ defined by $\forall E \in \mathcal{E} \backslash \mathcal{E}^*, \Phi(E) = \{(\mathbf{0}, \mathbf{0})\}$, and $\forall E \in \mathcal{E}^*$, $\Phi(E)$ is the set of all pairs $(\boldsymbol{e_h}, \boldsymbol{t})$ such that $h \in \mathcal{H}$, $\forall i \in \mathcal{H}$, $t_i = \frac{B - c_h}{|\mathcal{H}|} + h_i c_h$ and $\forall i \in N \backslash \mathcal{H}$, $t_i = 0$. This allocation meets all axioms but SHC.

4.H Proof of Proposition 4.4

Note that the ex-post identical-preference lower bound has to be satisfied on any restricted type space. In particular, in the standard Bayesian, risk neutral type space. On this domain, we know that if a mechanism implements a social choice function, then there exists a Bayesian incentive compatible (BIC) direct mechanism that also implements it (the revelation principle). We show that there does not exist such a mechanism.

Assume there exists a direct revelation mechanism M that satisfies the ex-post identical-preference lower bound and consider the following 2-agent case ($i \in \{1, 2\}$), with a common prior with full support on $[0; B)$ and common knowledge of all the features of the environment. Let $c_2 \in [0; B)$. From ex-post ULB, we know that $M(c_2, c_2)$ chooses a community to host the project with some strictly positive probability and requires a transfer $\frac{B + c_2}{2}$ to this community. Without loss of generality, let us say it is community 1 and consider the economy (c_1, c_2), where $c_1 < c_2$. Ex-post ULB requires some degree of allocative efficiency. More precisely, it requires to always locate the project where the cost is lower than the average cost. Therefore, ex-post ULB requires the project is located in community 1 for sure in $M(c_1, c_2)$. In addition, BIC requires that community 1 should get at least $x_1 = \frac{B+c_2}{2}$ as a transfer in $M(c_1, c_2)$. Budget balance and the ex-post ULB requirement for the other community, further require that it should get exactly $x_2 = \frac{B-c_2}{2}$. It is clear that this allocation rule is not BIC as the community 2 would benefit from under-reporting her cost. Therefore, there exists no mechanism that satisfies the ex-post ULB in this case. This establishes Proposition 4.4.

4.I Proof of Proposition 4.5

We respectively show it in the first and second price auctions.

First, consider the first price auction. If all stated costs are greater than B, no project is implemented. Otherwise, this mechanism selects a host $h \in \arg\min_N \hat{c}_i$ and implements the transfers $x_h = \frac{B+(n-1)\hat{c_h}}{n}$ and $\forall i \in N\backslash\{h\}, x_i = \frac{B-\hat{c_h}}{n}$. Consider what can be expected by the community i with cost $c_i < B$ when truthfully reporting her cost. We consider two cases. First, would all the others' report a greater cost than her, she would have to host the project and would end up with a transfer $\frac{B+(n-1)c_i}{n}$ and incur her own provision cost c_i. Therefore, she would achieve a exactly the level of welfare required by her identical preference lower bound. In the event another community j states a strictly lower cost that her, $\hat{c}_j$, she would get $\frac{B-\hat{c}_j}{n}$. As the community truthfully reports her type, we have $\hat{c}_j < c_i$ so this transfer is greater than $\frac{B-c_i}{n}$. Therefore, by reporting c_i, the agent i is ensured to get a least $\frac{B-c_i}{n}$.

The same reasoning applies to the second price auction. If all stated costs are greater than B, no project is implemented. Otherwise, this mechanism selects a host $h \in \arg\min_N \hat{c}_i$ and implements the transfers $x_h = \min\left(B, \frac{B}{n} + \frac{n-1}{n} \min_{N\backslash\{h\}} \hat{c}_i\right)$ and $\forall i \in N\backslash\{h\}, x_i = \max\left(0, \frac{1}{n}\left(B - \min_{N\backslash\{h\}} \hat{c}_i\right)\right)$. Consider what can be expected by the community i with cost $c_i < B$ when truthfully reporting her cost. We consider two cases. First, would all the others' report a greater cost than her, she would have to host the project and would end up with a transfer $\frac{B+(n-1)\hat{c}_j}{n}$, where $\hat{c}_j \geq c_i$, and incur her own provision cost c_i. Therefore, she would achieve a greater level of welfare that what is required by her identical preference lower bound. If on the contrary, if another community j states a strictly lower cost that her, she would get $\frac{B-\hat{c}_j}{n}$. As the community truthfully reports her type, we are ensured that this transfer is greater than $\frac{B-c_i}{n}$. Therefore, by reporting c_i truthfully, the community i is ensured to get a least $\frac{B-c_i}{n}$.

4.J Proof of Proposition 4.6

Let $i \in N$ be a community with cost c_i and M be a mechanism that satifies the interim identical-preference lower bound. We know that there exists a strategy s that guarantees an ex-post utility of at least $\max(0, \frac{B-c_i}{n})$ to i. We also know that no strategy can guarantee a strictly higher ex-post utility to i for sure. Indeed, assume there exists such a strategy s' and consider the situation in which all communities feature the same cost c_i. In this setting $M(s', ..., s')$ would not be feasible. Therefore, maximin players with cost c_i would always (weakly) prefer strategy s. We label this strategy $\hat{c}_i$ and extend the strategy space by duplicating it. We do the same for all costs $c \in \mathbb{R}_+$. Eventually, we delete all remaining strategies. As we know that these strategies are never strictly prefered by maximin players, the resulting mechanism is payoff equivalent and truthfull for maximin players. Conversely, any mechanism in $\mathcal{U}$ satisfies the interim identical-cost lower bound as, for all mechanisms in $\mathcal{U}$, truthfully reporting one's cost guarantees an ex-post utility level of $\max\left(0, \frac{B-c_i}{n}\right)$ regardless of the others' statements.

4.K Proof of Proposition 4.7

Let Φ be an anonymous and efficient allocation rule which satisfies Individual rationality, No envy among equals, Circumstance solidarity, Collective reward for responsibility costs in uniform circumstances, and Solidarity toward higher responsibility-cost communities in uniform circumstances on $\mathcal{E}^+$. Let $E_0 = (B, \boldsymbol{c^c}, \boldsymbol{c^r}) \in \mathcal{E}^+$. Define $c^{ref} = \min_N c_i - \min_N c_i^r \geq 0$ and $E_1 = (B, c^{ref}\boldsymbol{e}, \boldsymbol{c^r})$. Following the proof of Proposition 4.3, Efficiency, Anonymity, Individual Rationality, No envy among equals, Solidarity toward higher-responsibility-cost communities in uniform circumstances, and Collective Reward for Responsibility Cost in Uniform Circumstances together require that $\forall \phi \in \Phi(E_1)$ and $i \in N$, $u_i(\phi) = u_i^{Sh}((B, c^{ref}\boldsymbol{e} + \boldsymbol{c^r}))$. Define $E_2 = (B, \boldsymbol{c^{c\prime}}, \boldsymbol{c^r})$ such that $\forall i \in N$, $c_i^{c\prime} = c_i^c$ if $c_i^c < c^{ref}$, and $c_i^{c\prime} = c^{ref}$ otherwise. Circumstance solidarity requires that individual welfare levels are all weakly greater in E_2 than in E_1. However, the minimal cost in E_2 has to be the same as in E_1 by construction. As the minimal cost is the same in both economies, individual welfare levels have to be identical. From the same reasoning, the individual welfare levels have to be weakly greater in E_2 than in E_0 and, as the minimal cost is the same in both economies, individual welfare levels have to be identical in all the three economies. Therefore, $\forall \phi \in \Phi(E_0)$ and $i \in N$, $u_i(\phi) = u_i^{Sh}((B, c^{ref}\boldsymbol{e} + \boldsymbol{c^r}))$. This establishes $\Phi \in \boldsymbol{\Phi^+}$.

Conversely, we show that any $\Phi \in \boldsymbol{\Phi^+}$ satisfies all the axioms. Let $\Phi \in \Phi^+$.

We start by showing that Φ satifies CiS. Consider $E_2 = (B, \boldsymbol{c^{c2}}, \boldsymbol{c^r}) \in \mathcal{E}^+$ and $E_3 = (B, \boldsymbol{c^{c3}}, \boldsymbol{c^r}) \in \mathcal{E}^+$. Define $E_2' = (B, c_2^{ref}\boldsymbol{e} + \boldsymbol{c^r})$ and $E_3' = (B, c_3^{ref}\boldsymbol{e} + \boldsymbol{c^r})$, where $c_2^{ref} =$

$\min_N \{c_i^{c2} + c_i^r\} - \min_N \{c_i^r\}$ and $c_3^{ref} = \min_N \{c_i^{c3} + c_i^r\} - \min_N \{c_i^r\}$. Without loss of generality, assume $c_2^{ref} \leq c_3^{ref}$. We show that the welfare levels induced by the Shapley value are non-increasing from E_2' to E_3'. For this, we reorder the communities by increasing c_i^r. The resulting index will be denoted in parenthesis. The welfare levels induced by the Shapley value can be defined iteratively as follows:

$$u_{(n)} = \frac{max(0, B - c^{ref} - c^r_{(n)})}{n}$$

$$\forall i \in \{1, ..., n-1\}, u_{(i)} = u_{(i+1)} + \frac{\min(B - c^{ref}, c^r_{(i+1)}) - \min(B - c^{ref}, c^r_{(i)})}{i}$$

It is clear that $u_{(n)}$ decreases with c^{ref}. Now, assuming $u_{(i+1)}$ decreases with c^{ref}, we show that $u_{(i)}$ decreases with c^{ref}. Define $a = B - c^{ref}$, $b = c^r_{(i+1)}$ and $c = c^r_{(i)}$. We have $b \geq c$. Consider $a' > a$ and define : $\Delta = \min(a', b) - \min(a', c) - \min(a, b) + \min(a, c)$. Six cases are possible:

1. If $a' > a > b \geq c$, $\Delta = 0$
2. If $a' > b \geq a \geq c$, $\Delta = b - a \geq 0$
3. If $a' > b \geq c > a$, $\Delta = b - c \geq 0$
4. If $b \geq a' > a \geq c$, $\Delta = a' - a \geq 0$
5. If $b \geq a' \geq c > a$, $\Delta = a' - c \geq 0$
6. If $b \geq c > a' > a$, $\Delta = 0$

Therefore $u_{(i)}$ decreases with c^{ref} and, by induction, this is true for all $i \in N$. In particular, we have $\forall i \in N$, $u_i^{Sh}(E_2') \geq u_i^{Sh}(E_3')$. As, for all i in N, the welfare levels in E_2 and E_3 are respectively $u_i^{Sh}(E_2')$ and $u_i^{Sh}(E_3')$, CiS is satisfied.

We now show that Φ^+ satisfies Solidarity toward higher-responsibility-cost communities in uniform circumstances, and Collective Reward for Responsibility Cost in Uniform Circumstances. Consider $E_4 = (B, \boldsymbol{c^c}, \boldsymbol{c^r}) \in \mathcal{E}^+$ such that $\forall i \in N$, $c_i^c = c$. We have $\Phi^+(E^4)$ contains a single allocation that implements the same transfer and yields the same distribution of welfare as $\Phi^s(E_4')$, where $E_4' = (B, c\boldsymbol{e} + \boldsymbol{c^r})$. We know that Solidarity toward higher-cost communities and Collective cost reward are met by Φ^s. Then, they are met by Φ^+ as well. This establishes Proposition 4.7.

4.L Proof of Corollary 4.1

Let $E \in \mathcal{E}^+$ with two agents, and such that $c_1 < c_2 \leq B$, $c_1^c < c_2^c$ and $c_1^r > c_2^r$. The set Φ^+ contains a single allocation rule that yields the following allocation of welfare levels :

$$\begin{cases} u_1^s &= \frac{B-c_1}{2} + \frac{c_2^r - c_1^r}{2} \\ u_2^s &= \frac{B-c_1}{2} + \frac{c_1^r - c_2^r}{2} \end{cases}$$

As $c_1^r > c_2^r$, Solidarity toward higher-responsibility-cost communities requires that any decrease in c_2^r does not make community 1 worse off, which is contradicted by the expression of the welfare of community 1.

4.M Presentation of the survey

4.M.1 Implementation

The survey was carried out online among 257 University students in Toulouse. The following subsections respectively present the vignettes, the structure of the survey and its results. The detailed screens are presented at the end of the appendix.

4.M.2 The vignettes

The survey relies on two vignettes that are presented below.

Vignette 1 (windpark)

Two neighboring communities, A and B, consider the implementation of a wind park. A study is undertaken. Two sites, one in A and one in B, could host the project. After a consultation about landscape issues, it is admitted that only one of the two sites could be implemented.

In top of the environnemental benefits brought by the project, that spread beyond the limits of the communities, the benefit of the project to the communities is evaluated worth €5 millions, wherever it is located. This benefit corresponds to the value of the produced electricity during the project lifetime net of the payment to an operator who takes charge of all construction, operation and maintenance costs. If the project is implemented, this benefit will have to be shared between the communities.

The project is devised in order to limit as much as possible the nuisance to its neighbors. However, some remaining nuisance, mainly noise, are unavoidable. They will entirely be borne by the citizens of the community that hosts the project.

All this considered, both communities decide to estimate the damage they would bear. The citizens in the community A unanimously agree to estimate the damage worth €1 million and the citizens in the community B, €3 millions. These amounts will next be called the “estimation by the hosting community of the damage borne”.

From an outside perspective, the two communities are very similar. They have the same number of inhabitants and a comparable wealth level. The study established that the level of noise perceived will be the same on both sites. The difference between the estimation of the damage borne seems to only be explained by the fact that citizens in B are less inclined to accept the noise associated with the project than citizens in A.

Vignette 2 (wastewater treatment plant)
Two neighboring communities, A and B, consider the implementation of a wastewater treatment plant. A study is undertaken. Two sites, one in A and one in B, could host the project.
In top of the environnemental benefits brought by the project, that spread beyond the limits of the communities, the benefit of the project to the communities is evaluated worth €5 millions, wherever it is located. This benefit corresponds to the savings as compared to alternative solutions fo treating the sewage. If the project is implemented, this benefit will have to be shared between the communities.
The project is devised in order to limit as much as possible the nuisance to its neighbors. However, some remaining nuisance, mainly odors and noise, are unavoidable. They will entirely be borne by the citizens of the community that hosts the project.
All this considered, both communities decide to estimate the damage they would bear. The citizens in the community A unanimously agree to estimate the damage worth €1 million and the citizens in the community B, €3 millions. These amounts will next be called the "estimation by the hosting community of the damage borne".
From an outside perspective, the two communities are very similar. They have the same number of inhabitants and a comparable wealth level. The study established that the level of the nuisance perceived will be the same on both sites. The difference between the estimation of the damage borne seems to only be explained by the fact that citizens in B are less inclined to accept the noise associated with the project than citizens in A.

4.M.3 Outline of the survey

The survey consists in the eight following screens:

1. **Welcome screen:** Respondents are presented the object of the survey. They answer their day of birth within the month in order to be allocated into the different versions of the questionnaire.
2. **Presentation of the vignette:** Respondents are presented with the first part of the vignette. During the presentation of the vignette, the respondents are asked the following question:
   - With which of the following statements do you agree most?
     - These are mostly the citizens of concerned communities that are capable to estimate what a degradation of their living environment is worth. The decision should proprietarily rely on such estimations.
     - These are mostly objective measures, comprising for instance an estimation of the decline in residential value associated with the nuisance, which can

constitute a reliable measure of the damage borne. The decision should proprietarily rely on such estimations.

- None of them.

3. **Presentation of the vignette:** Respondents are shown the rest of the vignette. Half of them are asked to make an allocation choice which consists in answering the three following questions:

   a) Should the project be implemented?

   b) Which community should host the project?

   c) How should the benefit of €5 millions be shared?

4. **Diverse situations:** Respondents are asked to choose an allocation for three different situations.

5. **Principles:** Respondents are asked to state their degree of approval for different principles.

6. **Propositions:** Respondents are asked their degree of approval for five different allocation rules.

7. **Final choice:** Respondents are asked to choose an allocation for the situation of the vignette.

8. **Final screen:** Respondents answer some socio-demographic characteristics.

4.N Detailed screens

Bienvenue !

Merci de participer à cette enquête.

Nous sommes intéressés par votre opinion concernant la réalisation d'un projet d'intérêt public susceptible de générer des nuisances locales. Qu'il s'agisse d'éoliennes, d'incinérateurs ou de stations d'épuration, de tels projets soulèvent de nombreuses questions d'équité et font souvent l'objet d'une forte opposition de la part des riverains. Nous nous concentrons sur certains aspects de ces problèmes.

Cette enquête dure entre 15 et 20 minutes. Elle est réalisée à des fins de recherche uniquement. A l'issue de l'enquête, vous pourrez participer à un tirage au sort pour le versement d'un don de 150€ à une organisation de votre choix parmi une dizaine d'organisations. L'enquête est anonyme et la participation au tirage au sort, facultative et indépendante de l'enquête. Les conditions détaillées de ce tirage au sort sont accessibles ici.

Lorsque vous êtes prêt à commencer, vous pouvez répondre à la question suivante.

Vous êtes né: *

- Entre le 1 et le 8 du mois
- Entre le 9 et le 16 du mois
- Entre le 17 et le 24 du mois
- Entre le 25 et le 31 du mois

Page suivante

La situation

Nous vous invitons à considérer la situation suivante.

Deux communautés voisines, A et B, s'intéressent à la réalisation d'un parc éolien. Une étude est réalisée. Deux sites, l'un en A, l'autre en B, permettraient d'accueillir le projet. A la suite d'une concertation au sujet des enjeux paysagers, il est convenu que *seul l'un des deux sites pourra être retenu.*

En plus des bénéfices environnementaux apportés par le projet, qui s'étendent au-delà des frontières des communautés, le **bénéfice du projet pour les communautés** est évalué à **5 millions d'euros**, quel que soit le site retenu. Celui-ci correspond à la valeur de l'électricité produite sur toute la durée de vie du projet nette de la rémunération d'un opérateur qui prend en charge la totalité des coûts de construction, d'opération et de maintenance. Si le projet est réalisé, ce bénéfice sera à partager entre les communautés.

Le projet est conçu afin de limiter autant que possible les nuisances pour les riverains. Cependant, des nuisances, principalement du bruit, restent inévitables. *Elles seront entièrement supportées par les citoyens de la communauté d'accueil du projet.*

Des délibérations publiques sont organisées au cours desquelles les positions suivantes sont invoquées.

Avec laquelle de ces deux positions êtes-vous le plus en accord ? *

- "Ce sont avant tout des mesures objectives, comprenant par exemple une estimation de la baisse des prix immobiliers découlant des nuisances, qui peuvent constituer une évaluation fiable du préjudice subi. La décision devrait en priorité s'appuyer sur de telles évaluations."
- "Ce sont avant tout les citoyens des communautés concernées qui sont à mêmes d'évaluer le préjudice que constitue une dégradation de leur cadre de vie. La décision devrait en priorité s'appuyer sur de telles évaluations."
- Aucune de ces propositions

Page suivante

La situation

La description de la situation se poursuit ci-dessous.

Deux communautés voisines, A et B, s'intéressent à la réalisation d'un parc éolien. Une étude est réalisée. Deux sites, l'un en A, l'autre en B, permettraient d'accueillir le projet. A la suite d'une concertation au sujet des enjeux paysagers, il est convenu que *seul l'un des deux sites pourra être retenu.*

En plus des bénéfices environnementaux apportés par le projet, qui s'étendent au-delà des frontières des communautés, le **bénéfice du projet pour les communautés** est évalué à **5 millions d'euros**, quel que soit le site retenu. Celui-ci correspond à la valeur de l'électricité produite sur toute la durée de vie du projet nette de la rémunération d'un opérateur qui prend en charge la totalité des coûts de construction, d'opération et de maintenance. Si le projet est réalisé, ce bénéfice sera à partager entre les communautés.

Le projet est conçu afin de limiter autant que possible les nuisances pour les riverains. Cependant, des nuisances, principalement du bruit, restent inévitables. *Elles seront entièrement supportées par les citoyens de la communauté d'accueil du projet.*

Compte tenu de cela, les deux communautés décident d'évaluer le préjudice subi. Les citoyens de la communauté A conviennent à l'unanimité de l'évaluer à hauteur de 1 million d'euros, et ceux de la communauté B, à hauteur de 3 millions d'euros. Ces montants constituent ce que l'on appellera par la suite l'**évaluation par la communauté d'accueil du préjudice subi**.

D'un point de vue extérieur, les deux communautés sont très similaires. Elles ont le même nombre d'habitants et un niveau de richesse comparable. L'étude établit que le niveau de bruit perçu sera le même sur chaque site. La différence entre les évaluations du préjudice subi semble ne pouvoir s'expliquer que par le fait que les citoyens en B sont moins disposés à accepter le niveau de bruit associé au projet que les citoyens en A.

Les alternatives sont rappelées dans le tableau suivant :

(en million d'€)	**Projet localisé en A**	**Projet localisé en B**
Evaluation par la communauté d'accueil du préjudice subi	1	3

Au final, vous serez invité à exprimer ce que vous considérez être une décision juste dans cette situation. Les écrans qui suivent visent à encourager la réflexion à ce sujet. A chaque fois, essayez de répondre en toute honnêteté ce que vous pensez être juste d'un point de vue extérieur. Il n'est pas essentiel de rester cohérent d'une partie à l'autre si vous changez d'opinion.

Avant de commencer, nous vous invitons à formuler une première recommandation. Vous pourrez revenir dessus par la suite.

Quelle communauté devrait accueillir le projet ? *

- (•) A
- ( ) B
- ( ) A ou B, indifféremment
- ( ) Aucune, le projet ne devrait pas être réalisé

Etant donné la communauté retenue pour accueillir le projet, comment devrait être partagé le bénéfice de 5 millions d'euros résultant du projet ? *

Déplacez le curseur pour répondre.

Part de A — Part de B

2500000

Page suivante

Cliquez ici pour créer gratuitement votre propre questionnaire en ligne !

Partie 1/4 : Situations diverses

La situation considérée est rappelée ci-dessous. Il n'est pas nécessaire de relire cet encadré.

Deux communautés voisines, A et B, s'intéressent à la réalisation d'un parc éolien. Une étude est réalisée. Deux sites, l'un en A, l'autre en B, permettraient d'accueillir le projet. A la suite d'une concertation au sujet des enjeux paysagers, il est convenu que *seul l'un des deux sites pourra être retenu.*

En plus des bénéfices environnementaux apportés par le projet, qui s'étendent au-delà des frontières des communautés, le **bénéfice du projet pour les communautés** est évalué à **5 millions d'euros**, quel que soit le site retenu. Celui-ci correspond à la valeur

Les alternatives sont présentées dans le tableau suivant :

(en million d'€)	Projet localisé en A	Projet localisé en B
Evaluation par la communauté d'accueil du préjudice subi	1	3

Nous vous invitons à considérer les trois situations similaires suivantes. Dans toutes ces situations, seules les évaluations du préjudice subi diffèrent. Le bénéfice monétaire du projet est toujours évalué à 5 millions d'euros.

A chaque fois, nous vous invitons à proposer la décision qui vous semble la plus juste.

Dans la situation suivante : *

(en million d'€)	Projet localisé en A	Projet localisé en B
Evaluation par la communauté d'accueil du préjudice subi	1	6

Quelle communauté devrait accueillir le projet ?

- A
- B
- A ou B, indifféremment
- Aucune, le projet ne devrait pas être réalisé

Dans la situation suivante : *

(en million d'€)	Projet localisé en A	Projet localisé en B
Evaluation par la communauté d'accueil du préjudice subi	3	3

Quelle communauté devrait accueillir le projet ?

- A
- B
- A ou B, indifféremment
- Aucune, le projet ne devrait pas être réalisé

Dans la situation suivante : *

(en million d'€)	Projet localisé en A	Projet localisé en B
Evaluation par la communauté d'accueil du préjudice subi	1	2

Quelle communauté devrait accueillir le projet ?

- A
- B
- A ou B, indifféremment
- Aucune, le projet ne devrait pas être réalisé

Page suivante

Partie 2/4 : Principes

La situation considérée est rappelée ci-dessous. Il n'est pas nécessaire de relire cet encadré.

Deux communautés voisines, A et B, s'intéressent à la réalisation d'un parc éolien. Une étude est réalisée. Deux sites, l'un en A, l'autre en B, permettraient d'accueillir le projet. A la suite d'une concertation au sujet des enjeux paysagers, il est convenu que *seul l'un des deux sites pourra être retenu.*

En plus des bénéfices environnementaux apportés par le projet, qui s'étendent au-delà des frontières des communautés, le **bénéfice du projet pour les communautés** est évalué à **5 millions d'euros**, quel que soit le site retenu. Celui-ci correspond à la valeur

Les alternatives sont présentées dans le tableau suivant :

(en million d'€)	**Projet localisé en A**	**Projet localisé en B**
Evaluation par la communauté d'accueil du préjudice subi	1	3

Au cours des délibérations, plusieurs principes sont invoqués. Nous vous invitons à vous prononcer à leur sujet.

Concernant la localisation du projet, le principe invoqué est le suivant *

Principe A : "Le projet devrait être réalisée au sein de la communauté A car l'évaluation du préjudice subi y est la plus basse."

Vous êtes

- Totalement d'accord
- Plutôt d'accord
- Plutôt pas d'accord
- Pas d'accord
- Sans opinion

Concernant le partage des bénéfices, cinq principes sont invoqués.

Principe B : "Il s'agit d'un projet d'intérêt général. La communauté d'accueil ne devrait pas recevoir de compensation particulière du fait qu'elle accueille le projet." *

Vous êtes

- Totalement d'accord
- Plutôt d'accord
- Plutôt pas d'accord
- Pas d'accord
- Sans opinion

Principe C : "Seule la communauté d'accueil en subit les nuisances. Le bénéfice du projet devrait donc lui revenir en totalité." *

Vous êtes

- Totalement d'accord
- Plutôt d'accord
- Plutôt pas d'accord
- Pas d'accord
- Sans opinion

Principe D : "Si les deux communautés avaient une évaluation des nuisances identiques, elles devraient bénéficier du projet dans la même mesure. Il faudrait alors compenser la communauté qui accueille le projet à hauteur de sa propre évaluation du préjudice subi et partager le bénéfice restant en parts égales." *

Vous êtes

- Totalement d'accord
- Plutôt d'accord
- Plutôt pas d'accord
- Pas d'accord
- Sans opinion

Principe E : "Qu'elles aient une évaluation du préjudice subi basse ou haute, toutes les communautés devraient bénéficier du projet dans la même mesure." *

Vous êtes

- Totalement d'accord
- Plutôt d'accord
- Plutôt pas d'accord
- Pas d'accord
- Sans opinion

Principe F : "La communauté dont l'évaluation du préjudice subi est la plus basse contribue à l'intérêt de tous. Elle devrait être récompensée pour cela." *

Vous êtes

- Totalement d'accord
- Plutôt d'accord
- Plutôt pas d'accord
- Pas d'accord
- Sans opinion

Pensez-vous que d'autres principes devraient être invoqués ?

- Non
- Oui

Page suivante

Partie 3/4 : Propositions

La situation considérée est rappelée ci-dessous. Il n'est pas nécessaire de relire cet encadré.

Deux communautés voisines, A et B, s'intéressent à la réalisation d'un parc éolien. Une étude est réalisée. Deux sites, l'un en A, l'autre en B, permettraient d'accueillir le projet. A la suite d'une concertation au sujet des enjeux paysagers, il est convenu que *seul l'un des deux sites pourra être retenu.*

En plus des bénéfices environnementaux apportés par le projet, qui s'étendent au-delà des frontières des communautés, le **bénéfice du projet pour les communautés** est évalué à **5 millions d'euros**, quel que soit le site retenu. Celui-ci correspond à la valeur

Les alternatives sont présentées dans le tableau suivant :

(en million d'€)	Projet localisé en A	Projet localisé en B
Evaluation par la communauté d'accueil du préjudice subi	1	3

Au cours des délibérations, il est retenu que la communauté pour laquelle l'évaluation des nuisances est la plus basse, c'est à dire A, accueille le projet. Cinq propositions de partage du bénéfice monétaire sont alors évoquées.

Nous vous invitons à donner votre avis les concernant.

Première proposition: La communauté A devrait recevoir la moitié du bénéfice, c'est-à-dire 2,5 millions d'euros. *

Part de A — Part de B

2500000

- Juste
- Ni juste, ni injuste
- Injuste
- Sans opinion

Deuxième proposition: La communauté A devrait recevoir sa propre évaluation du préjudice subi, c'est-à-dire 1 millions d'euros, plus la moitié du bénéfice restant, c'est-à-dire 2 millions d'euros, soit au total 3 millions d'euros. *

Part de A — Part de B

3000000

- Juste
- Ni juste, ni injuste
- Injuste
- Sans opinion

Troisième proposition: La communauté A devrait recevoir l'évaluation moyenne du préjudice subi, c'est-à-dire 2 millions d'euros, plus la moitié du bénéfice restant, c'est-à-dire 1,5 millions d'euros, soit au total 3,5 millions d'euros. *

Part de A — Part de B

3500000

- Juste
- Ni juste, ni injuste
- Injuste
- Sans opinion

Quatrième proposition: La communauté A devrait recevoir l'évaluation du préjudice subi la plus haute, c'est-à-dire 3 millions d'euros, plus la moitié du bénéfice restant, c'est-à-dire 1 million d'euros, soit au total 4 millions d'euros. *

Part de A — Part de B

4000000

- Juste
- Ni juste, ni injuste
- Injuste
- Sans opinion

Cinquième proposition: La communauté A devrait recevoir la totalité du bénéfice monétaire du projet, c'est-à-dire 5 millions d'euros. *

Part de A — Part de B

5000000

- Juste
- Ni juste, ni injuste
- Injuste
- Sans opinion

Parmi ces cinq propositions, laquelle vous paraît la plus juste ? *

- La première
- La deuxième
- La troisième
- La quatrième
- La cinquième
- Sans opinion

Page suivante

Partie 4/4 : Proposition finale

La situation considérée est rappelée ci-dessous. Il n'est pas nécessaire de relire cet encadré.

Deux communautés voisines, A et B, s'intéressent à la réalisation d'un parc éolien. Une étude est réalisée. Deux sites, l'un en A, l'autre en B, permettraient d'accueillir le projet. A la suite d'une concertation au sujet des enjeux paysagers, il est convenu que *seul l'un des deux sites pourra être retenu.*

En plus des bénéfices environnementaux apportés par le projet, qui s'étendent au-delà des frontières des communautés, le **bénéfice du projet pour les communautés** est évalué à **5 millions d'euros**, quel que soit le site retenu. Celui-ci correspond à la valeur

Les alternatives sont présentées dans le tableau suivant :

(en million d'€)	**Projet localisé en A**	**Projet localisé en B**
Evaluation par la communauté d'accueil du préjudice subi	1	3

Enfin, nous vous invitons à formuler votre propre opinion concernant la décision qui vous semble la plus juste dans la situation considérée.

Quelle communauté devrait accueillir le projet ? *

- A
- B
- A ou B, indifféremment
- Aucune, le projet ne devrait pas être réalisé

Pourquoi avez-vous choisi cette recommandation ?

Avez-vous manqué d'information ?

- Non
- Oui

Page suivante

Informations générales

Nous vous demandons finalement de répondre aux questions suivantes.

Un briquet et un stylo coûtent ensemble 2,50€. Le stylo coûte 2€ de plus que le briquet. Combien coûte le briquet ? *

Un plan de nénuphar double de surface toutes les semaines. Le plan de nénuphar a recouvert la moitié d'un étang en 48 semaines. En combien de semaines le nénuphar aura-t-il recouvert la totalité de l'étang ? *

Nous vous demandons enfin quelques informations vous concernant

Vous êtes *

- un homme
- une femme

Quelle est votre année de naissance ? *

Quel est votre statut d'emploi ?

- Etudiant (y compris doctorat)
- Travailleur
- Sans emploi
- Retraité

Si vous êtes étudiant, quel est votre domaine d'étude principal ?

- Santé et Sport
- Sciences (Physique, Chimie, etc.)
- Ingénieurie
- Art, Littérature et Sciences Humaines
- Droit, Sciences de Gestion et Sciences politiques
- Economie
- Autres

Voulez-vous participer au tirage au sort ? *

La participation est soumise à l'acceptation du règlement, accessible ici.

- Oui
- Non

Si vous avez des commentaires avant de conclure l'enquête, vous pouvez les laisser ici:

Terminer

Chapter 5

Fairness judgments for the allocation of an indivisible task and its benefit

An experimental study

Abstract

I study individual fairness judgments and preferences for the allocation of an indivisible task and its benefit among two individuals with a different willingness to perform it. My approach relates results from the theory of fair allocations with the empirical observation of fairness judgments and preferences. I first present and motivate four contrasted allocation rules as possible rules of judgment. For some of the participants, a questionnaire was proposed before they knew about their situation. Among the four rules of judgment proposed, the welfare egalitarian allocation rule is the most preferred allocation rule as stated by the participants. Yet, I also observe support for principles that are not compatible with this rule: an important proportion of respondents deem fair to give nothing to someone who would not be willing to perform the task, and another substantial proportion deem fair to split the benefit of the task equally when both participants feature the same compensation requirement. In the experiment, participants had the opportunity to perform a task for pay. However, for any two of them, a single task was available. As required by the Pareto principle, it was allocated to the participant with the lower compensation requirement. In this situation, the stated normative expectations of the task performer are found to be higher, the greater the discrepancy between the compensation requirements. This does not extend to individual distributive preferences as revealed by the offers in a dictator setting. I also find that the task performers who took the questionnaire would deem the equal split fair less often. Overall, few respondents are consistent with any of the four rules proposed.

Keywords: experiment, empirical social choice, vignette study, fairness judgments, distributive preference, allocation rule, reasoning, resource allocation

JEL codes: D63

I wish to thank Stefan Ambec, Astrid Hopfensitz, César Mantilla, Stéphane Robin and Boris Van Leuwen for their advice and challenging discussions on this project, Stephane Cezera for his support with the implementation of the experiment, Shagun Khare for her precious help in running the sessions, and Stefan Ambec and Astrid Hopfensitz for funding this experiment.

Introduction

In most real situations, the diversity and conflicting nature of existing norms makes the mere nature of what is a fair distribution subject to doubt and argument. Think, for instance, of a situation in which an indivisible resource held in common is to be allocated among communities, individuals or firms. Among other problems, one could think of the allocation of spectrum, or fishing, pollution or water right allocation. This may also apply to the allocation of an economic opportunity among competing agents. In this situation, it seems reasonable to guarantee that the agent who actually exerts himself to make the resource productive reaps the benefit of it. Still, it also seems reasonable to compensate the others for being deprived from accessing the resource. There is no straightforward solution to this conflict. In the experiment presented here, the participants were confronted with such a situation. They had the opportunity to perform a task that was paid €20. However, for any two of them, a single task was available. The individual with the lower compensation requirement was offered the opportunity to perform it while the other individual was given the option to leave. The payment of the task could be shared among the two participants. In this situation, we are interected in the fairness judgments held by the people. Still, the empirical observation of individual fairness judgments in this situation is confronted by several alternatives and two important challenges.

The first challenge is that there is no straightforward solution to the conflict mentioned. As a result, a confident and thoughtful judgment regarding what a fair distribution is in this situation may require reasoning beyond what individuals are inclined to perform. The judgments and behaviors observed in this situation may rely on diverse heuristics[1] and be susceptible to change through reasoning. By reasoning, we refer to the reflective equilibrium process proposed by John Rawls (1951) and its positive counterpart, the *dual-process theory* of moral reasoning.[2] In these theories, reasoning consists in seeking consistency between one's adhesion to general fairness principles and one's considered judgments in particular cases through a process of mutual adjustment. We first propose and motivate four contrasted allocation rules as conjectured rules of judgments in reflective equlibrium. Allocation rules are systematic rules of judgment regarding how the benefit of the task should be allocated across situations. Their being in reflective equlibrium means that they both derive from plausible intuitive fairness principles and are consistent with the likely most compelling intuitions regarding how to allocate the benefit of the task in particular situations. In a preliminary questionnaire, some participants were invited to express fairness judgments regarding these *rules* and in four contrasted *particular cases.* We are interested in assessing the consistency between their adhesion to allocation rules

[1] A heuristic is a practical method for finding a satisfactory solution in complex settings that eases the cognitive load of making a decision.

[2] see Saunders (2009) for a discussion of this.

and their answers in particular cases. We are also interested in the effect of performing the questionnaire on their judgments and behavior when involved in the situation for real.

Second, the fairness judgments and preferences observed may come from different sources. Experimental evidence reveals that, in some environments, individuals behavior cannot be explained by a mere concern about one's own material payoff, suggesting the existence of a genuine preference for fair allocations. Such behaviors would have to result from some form of *fairness preferences*. The existing literature traditionally distinguishes pure *distributive preferences* and *reciprocal preferences* (Falk and Fischbacher, 2001). The former denotes a fundamental preference for some ideal distribution. The latter denotes a preference to punish or reward someone according to her perceived intentions, where, in some situations, intentions would have to be inferred from an allocation choice. In the traditional ultimatum game, for instance, the proponent makes an offer regarding how to share a given endowment, after which the respondent can accept or refuse. In this situation, intentions are derived from the choice of an allocation, and therefore, also rest upon some judgment about what a fair allocation is. As theories of distributive preferences and, to some extent, theories of reciprocity relate fairness behaviors to fairness ideals, this raises the question of what these ideals are. Following this dual perspective on fairness preference, beliefs about distributive justice could result from two sources. Individual moral values are the first. We will call these beliefs underlying distributive preferences, *moral judgments*. A second basis for fairness preferences can rely on the assumption of a common understanding on a set of acceptable premise in a given group, in a word, on the existence of *distributive norms*. This individual belief regarding others' *normative expectations* may differ substantially from one's own *moral judgments*. In this experiment, we propose to measure and contrast both of them. Moral judgments are measured through a questionnaire answered in isolation and moral preferences are observed in a dictator setting. Normative expectations are measured in two ways. First, we change the instructions to the same questionnaire and ask respondents to coordinate with another participant (Krupka and Weber, 2013). Second, in their situations, each participant is invited to state what choice of allocation from the other she would find *defendable* and we elicit her beliefs regarding the most likely statement of another participant in a different situation (Bicchieri and Chavez, 2010).

Related literature

The experimental study of distributive preferences originally focused on inequality aversion (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) and the traditional ultimatum game has proposed that sanctioning behavior depends on how far the allocation is from the equal split. There is now clear evidence that equal allocations are not always the most favored distributions in situations featuring different levels of global efficiency (Engelmann and Strobl, 2004) or individual heterogeneity regarding choices (Cappelen et al, 2013), effort (Konow, 2000; Cherry et al, 2002; Cappelen et al, 2007; Krawczyk, 2010), or individual characteristics (Cappelen et al, 2010). Following this line of research, some recent experiments have studied situations of normative conflict. For instance, Cappelen et al (2007) designed a situation in which two individuals with different productivities can make an investment decision. The resulting payment is pooled and has to be allocated. In this situation, a conflict arises between holding the participants accountable for their investment choice, while compensating them for their differences on productivity. The observed behaviors were consistent with a fair fraction of individuals favoring each of three different fairness ideals. In a subsequent experiment, Cappelen et al (2013) designed a situation in which two individuals separately choose their exposure to risk. Risk is realized and the resulting payment is pooled. In this situation, a conflict arises between holding the participants accountable for their choice, and limiting the ex-post level of inequalities. The observed behaviors were consistent with a moderate level of inequality aversion and a desire to hold the individual accountable for their choice. Similarly, our setting features a conflict between guaranteeing that the participant who actually performs the tasks reaps the benefit of it, while compensating the other for being deprived from the opportunity of performing it. No experiment exists to our knowledge that focuses on how individuals deal with this trade-off. Similarly to the previous studies, we find that respondents express a large diversity of views, ranging from an equal division of the monetary benefit to the allocation of the whole benefit to the task performer.

Several studies further suggest that the direct measurement of fairness judgments and preferences would fail to capture how these are sensitive to reasoning. Following Cappelen et al (2007), Cappelen et al (2010) observed that a stage of reasoning, consisting of observing the consequence and expressing a judgment regarding each of the three fairness ideals, did influence the proportion of participants referring to each of them. Following this approach, we also design a stage of reasoning which consists of a questionnaire. Noticeably enough, Cappelen et al (2010) observed a decrease in the proportion of participants referring to a strict egalitarian allocation. Whereas it does not constitute per se a proof for an equal split heuristics, this is consistent with this idea which has been suggested by many other work (Messik, 1993). This effect is also observed in our setting.

Our questionnaire follows the line of the empirical study of distributive judgments initiated by Yaari and Bar-Hillel (1984). In their approach, respondents are invited to express their judgments regarding the distribution of a good in a hypothetical situation.[3] Among the following studies in this line, our setting is closest to a recent study conducted by Schokkaert et al (2007). They study the judgments over the fair allocation of an extra harvest due to the allocation of a plough among three individuals. In particular, the

[3]This literature is reviewed by Konow (2003) and Gaertner and Schokkaert (2012).

authors check the consistency between the adhesion to general principles and the quantitative allocation chosen in particular cases. They do not find a clear correspondence. Recently, Hurley et al (2011) showed that the judgments expressed by the respondents can differ significantly over the verbal statements describing different allocation rules and their quantitative implications. They suggest that "some individuals do not understand the quantitative distributional implications of the various principles". We pursue this analysis by assessing the consistency of the adhesion to allocation rules with the judgments expressed in particular cases. These allocation rules are not able to account for the full range of pattern observed. For instance, most respondents deemed the welfare egalitarian allocation *rule* fairest. Yet, they do not seem to conform to it when expressing quantitative judgments in particular cases: a majority of respondents deem fair to give nothing to someone who would not be willing to perform the task (the no-dummy principle) and to split the benefit equally when both participants feature the same compensation requirement. Overall, few respondents are consistent with one of the four rules proposed.

The rest of the paper is organized as follows. Section 5.1 presents the experimental situation and motivates the choice of allocations rules and particular cases. Section 5.2 details and motivates the experimental protocol and formulates the predictions we intend to test. Results are presented in Section 5.3. Finally, we further discuss the results in Section 5.4.

5.1 Allocation rules

5.1.1 Presentation of the situation

The situation is chosen so as to feature a conflict between two norms of distributive justice. It is presented in the following vignette.

> Two persons, A and B, express separately the minimal amount they require for performing a given task. Person A declares that it costs her €6 to perform this task, and person B, 14€ for her. The task yields a benefit of €20 which can be shared. However, it may only be performed once. It will be performed by one of the persons who stated the lowest amount. Person A is therefore retained to perform the task.
> How should the €20 be shared?

The interest of this situation is that it features a conflict between two significant distributive norms. As the effort is actually undergone by the task performer, it seems desirable to recognize this individual entitled to the total value of the task. On the other hand, the existence of a single task is circumstantial. Therefore, it also seems desirable to take into account the lack of opportunities in this situation. Therefore, an answer to the question raised leaves room for reasoning and deliberation.

An informal example of a deliberation scheme could be the following. It can be claimed that *"the person who performs the task should get the value of it as it is her who makes the effort of performing the task"*. However, this would dismiss the fact that the other *"could also have performed the task in the absence of A"*. This suggests that *"he should get some amount as he did not get the opportunity of performing the task"*, but *"it is normal that the task performer gets more"*.[4] This leaves many possibilities. One can be to give the task performer her own compensation requirement and share the rest equally. Yet, in this case, B may get a part of value even in cases in which she would not be willing to do the task anyway. This may be deemed too high a compensation. So we may rather look for an allocation rule that depends on the difference between the value of the task and the compensation requirement of the deprived individual. One possibility is to give to the other the difference between the value of the task and her own compensation requirement, and the rest to the task performer. In this situation €6, and €14 to the task performer. This rule seems reasonable. Still, it would fail to give more to the task performer in the situation in which both compensation requirements amount to €6, this rule recommends to give €6 to the task performer and €14 for the other. This is not tenable. Another possibility is to gives to the deprived individual half the difference between the value of the task and her own compensation requirement, that is, in this situation to give €17 to the task performer and €3 to the other. This allocation rule seems to be a reasonable way to recognize a right of the task performer to the fruit of her labor while compensating the other from the deprivation of an opportunity. A more formal analysis is proposed in the following subsection.

5.1.2 Reasoning about fair allocation rules

In this section, we propose and motivate four allocation rules as alternative ways to prioritize or conciliate relevant norms applying to this situation. To reason out of general principle requires defining a level of generality. Then, we have to define a general class of problems, sufficiently similar to the problem at hand to require the application of the same principles. We propose to consider the class of situations $\mathcal{S}$ in which two individuals $i \in \{1, 2\}$ state their compensation requirements c_i to perform a task worth 20€. A situation S is then fully described by the pair (c_1, c_2).[5] This class of situations defines the

[4] Emphasis are quotations from the participants' justifications for their recommendation in this same situation collected at the end of the experiment (our translation). These statements and the data collected in this experiment are available upon request.

[5] Note that we do not allow the value of the task to vary across situations. This amounts to assume that the rules of judgments would not change depending on this value. This may be challenged, in which

level of generality at which reasoning is performed. The objective is to define general rules of judgment about allocations in all these situations. These are called allocation rules. Among this class, we already presented the situation $S_1 = (6, 14)$. We will also draw attention to the following particular situations that have a potential to question one's conclusions. In the first situation, the two individual are equally willing to perform the task. Two examples considered are $S_2 = (6, 6)$ and $S_3 = (14, 14)$. In the last situation, an individual expresses so high a compensation requirement that he would refuse to perform the task in the absence of the other. One example can be $S_4 = (6, 20)$.

We discuss fair allocation rules in this setting. An *allocation rule* specifies the assignment of the task to one of the two individuals (or, possibly, to no one) and, when the task is performed, in the allocation of the resulting benefit in all possible situations. Individual assignments and allocations are respectively referred to as *roles* and *transfers* hereafter.

We start by introducing two reasonable properties. First, it seems reasonable to consider allocation rules for which no other rule would be unanimously preferred given what is known about preferences. Therefore, we focus on *Pareto efficient* allocation rules. In this setting, they are rules that always assign the task to the individual with the lowest compensation requirement. The question then boils down to the allocation of the benefit of the task. Second, it seems reasonable to require that no individual is made worse off when involved in this situation. This is the requirement of *Individual rationality.* In this situation, it requires that the task performer never gets less than her own compensation requirement and that the other never gets a negative transfer. We now present more controversial principles. First, a direct application of the *accountability principle* (Konow, 2000)[6] to the actual effort undergone would lead to recommend to allocate the whole benefit of the task to the subject who performed it. This characterizes the *strong libertarian allocation rule* (SL).[7]

As argued earlier, such an allocation rule may be unsatisfactory in that it does not take into account the fact that the individual who did not perform the task has been deprived from the opportunity to perform it. This is particularly salient in the situation S_2 and S_3 in which the individual who performs the task turns out to be arbitrary chosen. One may be willing to acknowledge that no individual should benefit more than the other from the existence of the task in this particular situation. A way to capture this idea is the principle of no envy: "no participant should prefer the others' assignment to their own". When applied only to the particular situation when both compensation requirements are

case, we shall consider the more general domain.

[6]Konow's accountability principle "requires that a person's fair allocation (e.g., of income) vary *in proportion* to the relevant variable that he influence (e.g., work, effort) but not according to those that he cannot reasonably influence (e.g. a physical handicap)" (Konow, 2000; our emphasis).

[7]Given the artificial nature and constrained nature of the situation considered, the names of the allocations may not properly reflect the subtleties of underlying philosophies. We choose them in order to ease their interpretation and so as to echo to allocation names in Cappelen and Tungodden (2007).

the same, we will call this principle *No envy among equals.* As the former notion of accountability is too strong to accommodate this view, one may like to endorse a weaker principle. We propose the following *reward principle*: "Everything else equal, including the roles, an allocation rule should never grant a lower share to an individual as her compensation requirement decreases". As such this principle remains compatible with the strong libertarian allocation rule. Still, it is compatible with other rules. We shall now be able to clarify what we mean by expressing that the strong libertarian allocation rule goes too far in rewarding the task performer. We propose the following *weak solidarity principle*: "as an individual gets rewarded as his compensation requirement gets lower, this reward shall not penalize the other who's compensation requirement is greater". Along with no envy among equals, the reward and the weak solidarity principles characterize the *weak libertarian allocation rule* (WL).

A radically different view could consider that individual should not be held responsible for their preferences in this context. Therefore, they should also be compensated for the existing differences in their compensation requirements. We propose to capture this idea with the following *solidarity principle*: "for any change in the situation, both individual should either benefit or lose from it". Along with No envy among equals, this principle characterizes the *welfare egalitarian allocation rule* (WE). This rule and the weak libertarian allocation rule are the two extreme side of the set of allocation rules that meet the no-envy criterion. While the latter proposes to reward the individuals with low compensation requirements, the former sticks to the idea of compensating the individual who performs the task for her effort but rejects any further reward. In particular, it gives the stick to the same allocation in the situations S_1, S_2 and S_4.

Finally, the *strict egalitarian allocation rule* (SE) always recommends to split the benefit equally. A particularity of this allocation is that it does not satisfy the property of individual rationality. In the situation S_3 for instance, it leads to recommend to give to the task performer an amount which is inferior to her actual compensation requirement. It does not satisfy the property of No envy among equals either as in the situations S_2 and S_3, the task performer who rather not do the task for getting the same amount.

This discussion is summarized in Table 5.1. It singles out the four following allocations rules.

- The **strong libertarian (SL)** allocation rule recommends giving the whole benefit of the task to the individual who performs it.
- The **weak libertarian (WL)** allocation rule recommends giving half the difference between the benefit and her own compensation requirement to the individual who does not perform the task, and to give the rest to the participant who performs the task.

- The **welfare egalitarian (WE)** allocation rule recommends giving her own compensation requirement to the individual who performs the task, and to share the remaining benefit equally.
- The **strict egalitarian (SE)** allocation rule recommends sharing the total benefit equally among the two participants.

The recommendations of each of these four allocation rules in the four situations previously emphasized are presented in Table 5.1 and Figure 5.1. In the following experiment, a questionnaire is designed on the basis of this reasoning. It presents the four allocation rules introduced and asks for recommendations in the four particular cases that we identified as likely to score the limits of each of these rules. The recommendations presented will motivate the hypothesis.

	SE	WE	WL	SL
Properties				
Pareto efficiency	Yes[1]	Yes[2]	Yes[3]	Yes[4]
Individual rationality	No	Yes	Yes	Yes
No envy among equals	No	Yes	Yes[3]	No
Accountability principle	No	No	No	Yes[4]
Reward principle	Yes	No	Yes[3]	Yes
Weak Solidarity principle	Yes	Yes	Yes[3]	No
Solidarity principle	No	Yes[2]	No	No
Allocation to the task performer in particular situations				
$S_1 = (\underline{6}, 14)$	10	13	17	20
$S_2 = (\underline{6}, 6)$	10	13	13	20
$S_3 = (\underline{14}, 14)$	10	17	17	20
$S_4 = (\underline{6}, 20)$	10	13	20	20

Table 5.1: Properties of the four allocation rules considered.
The general properties of the allocation rules on the class of situations at hand are presented above and the allocations to the person who performs the task (underlined) as recommended by each allocation rule in four particular situations are presented below.

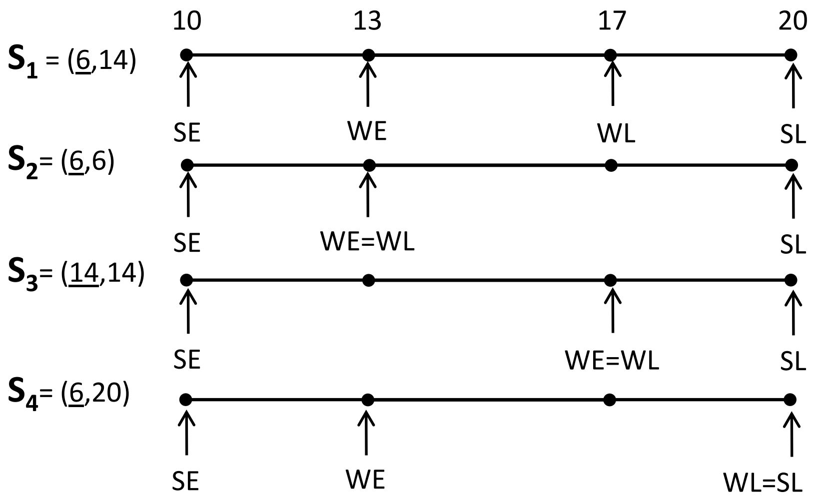


Figure 5.1: Allocation to the task performer recommended by each allocation rule for each particular case.

5.2 Experimental design

5.2.1 Organization

The experiment was programmed in PHP/SQL. A pilot session was conducted on November, 20th and eight sessions took place on November, 27th and 28th at the Toulouse School of Economics. Each session hosted between 13 and 20 participants and the sessions gathered 151 participants in total. Details about the experimental conditions are provided in Appendix 5.B.

Participants were allowed to leave the room as soon as they finish. The time spent on the experiment lasted from half an hour to an hour and a half. Payments were given in a separate room by a person who did not know the experiment. They were composed of a €5 show-up fee, an additional payment depending on their choices and the choices of others. They ranged from €5 to €35, with an average of €15.9.

5.2.2 Proceeding of the experiment

The experiment protocol proceeds in four parts. Its proceeding is summarized in Figure 5.1. Details and screenshots are provided in Appendix 5.C.

In the first part of the experiment, the participants are invited to perform an elementary task. This was intended to make them experience the final task that they may have to perform. It consists in counting the number of occurrence of the letter A in a randomly

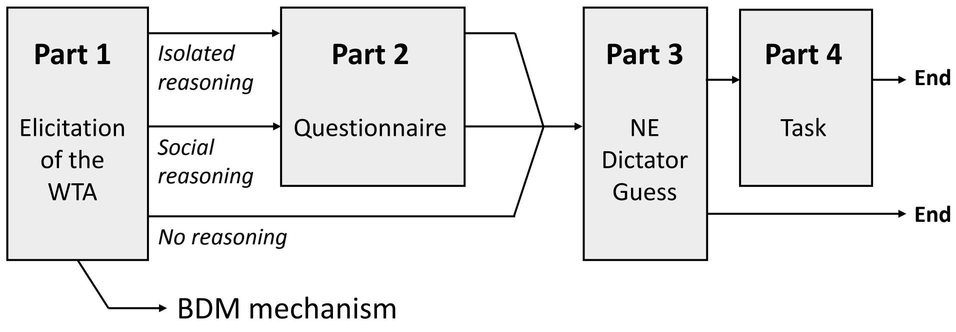


Figure 5.1: Proceeding of the experiment.

generated table of dimension 25x25. This task was chosen as it requires a significant effort while not requiring any particular skills. As all participants had to reach this stage before the experiment could proceed, and in order to avoid too long a waiting time, a maximum of 8 minutes was granted. After they finish or their time is exhausted, participants are invited to proceed further. In the following screen, they are invited to declare their willingness to accept to perform this task ten times at the end of the session, hereafter called their compensation requirements. In order to make their statement credible, they are explained the functioning of the Becker-De Groot-Marschak mechanism and told that they may be involved in this mechanism.[8] When all have finished this part, some are directed to the BDM mechanism and others are matched into pairs and allocated into the three different treatments. The matching process and the allocation into the treatments are explained in Appendix 5.D.

The treatments differ according to the second part. This part consists of a questionnaire, which is detailed in section 5.2.3. First, a third of the pairs is assigned to the *"no reasoning"* treatment. They do not take the questionnaire and are directed to the third part. The two other groups answer a questionnaire. In order to compare judgments related to individual moral norms and social norms, two treatments are designed. The questionnaire is the same for both groups. Only the instructions differ. In the *"isolated reasoning"* treatment, the participants are simply asked to answer the questionnaire whereas, in the *"social reasoning"* treatment, participants are paired and asked to coordinate in their answers. Coordination is not incentivized. However, the participants are informed that they will be shown how well they performed at the end of the session. We are interested in the effect of answering this questionnaire on behaviors and beliefs as measured in the next part.

[8]The BDM mechanism is the following: after the participant states her compensation requirement for performing the task, a random payment is drawn between 0 and 20€. If the stated compensation requirement is lower than this payment, the subject performs the task and gets this amount. If the stated compensation requirement is greater than this payment, the participant does not perform the task and gets no payment.

In the third part, participants are informed of the situation: the task is worth €20 and they are matched with another participant. The compensation requirements of both participants are disclosed and they are informed that only the participant with the lowest compensation requirement will be performing the task at the end of the session.[9] In a first screen, they declare the minimal offer they would find defendable from the other participant. This amount will be called the *stated normative expectation* hereafter. On a second screen, they chose an offer for the other participant. In each pair, one of the two offers is randomly selected and implemented. Participants were informed of this process. This amount will be called the *offer*. Finally, they were asked to guess the minimal offer declared by another participant in another pair. Correct guesses were paid €10. This amount will be called the *guessed normative expectation*. At this end of this part, the participants were asked about their perception of the other's compensation requirement, what would be a fair allocation in the case $S_1 = (6, 14)$ and some individual characteristics (gender, birth year, and political orientation).

Finally, the participants who were directed to the fourth part had to perform the task. Participants could leave the room and get paid as soon as they had finished.

5.2.3 Normative reasoning

The second part consisted in a questionnaire performed on a single screen and structured in two parts. Respondents first read a vignette describing the situation as the one presented in section 5.1.1.

The two parts of the questionnaire are the following:

1. **Judgments over allocation rules:** A verbal description of the four allocation rules introduced in section 5.1.2 is presented along with a justification for each of them. Their implications are illustrated in the situation $S_1 = (6, 14)$. The participants are asked, for each of them, whether they think they are "fair" or "unfair" and which one is fairest. This part aims at trigering deductive reasoning from allocation rules to particular allocations (Hurley et al, 2011).

2. **Judgments in particular cases:** the participants recommend a quantitative allocation in the 4 particular cases introduced in section 5.1.2. This part aims at encouraging inductive reasoning from particular allocations to allocation rules (Hurley et al, 2011).

The order of the two parts, of the elements within each part, and of choices was randomized. The participants are informed about this. The full questionnaire is presented in Appendix 5.C.

[9]In the case in which their compensation requirements are the same, the task performer is randomly chosen.

5.2.4 Tests

On top of informing us about how participants use the compensation requirements in forming their judgments, this experiment intends to test four points. First, whether the four allocation rules motivated in the first section satisfactorily account for the actual judgments observed. Second, we intend to test whether respondents actually judge according to rules of judgments, or rather whether they follow their intuitions. Third, we aim at testing whether moral and social judgments significantly differ. Finally, we test for the effect of reasoning on the stated fairness ideals. .

5.2.4.1 Relevance of the four allocation rules

In the questionnaire, were all the participants following one of the four allocation rules presented in section 5.1.2, we should expect the modes of the distribution to correspond to the recommendations of the four rules presented on Figure 5.1. This motivates the following assumptions.

Test 1. *For the particular cases presented in the questionnaire,*

1. *In the situation S_1, four modes are observed, that correspond to 10, 13, 17 and 20.*
2. *In the situation S_2, three modes are observed, that correspond to 10, 13 and 20. The mode at 13 is greater than in S_1 and similar for the other values.*
3. *In the situation S_3, three modes are observed, that correspond to 10, 17 and 20. The mode at 17 is greater than in S_1 and similar for the other values.*
4. *In the situation S_4, three modes are observed, that correspond to 10, 13 and 20. The mode at 20 is greater than in S_1 and similar for the other values.*

We may also be interested by the predicted effect of the compensation requirement of each participant on statements and choices. Note that the welfare egalitarian allocation rule is the only one that depends on the lower compensation requirement, c_1. Similarly, the weak libertarian allocation rule is the only one that depends on the greater compensation requirement, c_2. Assuming that a significant fraction of the participants follows these rules while the others follow the two others leads to the following assumptions.

Test 2. *For the particular cases presented in the questionnaire, on average, respondents give more to the task performer*

1. *in S_3 than in S_1,*
2. *in S_4 than in S_2, and*
3. *in S_2 than in S_1.*

Test 3. *The normative expectation*

1. *increases with c_1 for the task performer and decreases with c_1 for the other participant.*
2. *increases with c_2 for the task performer and decreases with c_2 for the other participant.*

Test 4. *The offer*

1. *decreases with c_1 for the task performer and increases with c_1 for the other participant.*
2. *decreases with c_2 for the task performer and increases with c_2 for the other participant.*

We also expect this to be reflected in the guessed normative expectations.

Test 5. *The normative expectations as guessed by the participants are consistent with Tests 3.1 and 3.2.*

5.2.4.2 Consistency of individual judgments

For the participants that took the questionnaire, we intend to test the consistency between the judgments expressed in particular situations and the adhesion to one of the four allocation rules introduced earlier.

Test 6. *For each respondent, the Euclidian distance between the answers in particular cases and the recommendations of the four allocation rules characterized is minimal for the allocation rule deemed fairest.*

We also check that respondents are consistent with their stated fairness ideals with the following test:

Test 7. *For each respondent, a fair allocation to the task performer as reflected in normative expectations and offers is all the greatest as her stated fairness ideal is in order the strict egalitarian, the welfare egalitarian, the weak libertarian and the strict libertarian allocation rule.*

5.2.4.3 Difference between isolated and social reasoning

Finally, we are also interested in testing the effect of reasoning and the difference between judgments expressed in isolation or when trying to coordinate with another participant. This leads us to the following predictions.

Test 8. *In the questionnaire,*

1. *judgments over allocation rules expressed in the isolated and social reasoning treatments differ, and*
2. *the average allocation chosen for the task performer in the isolated and social reasoning treatments differ.*

5.2.4.4 Effect of reasoning

We eventually test for the effect of reasoning on the average normative expectations and offers

Test 9. *The average normative expectation and the average offer differ between the "no reasoning" treatment and the two others.*

5.3 Results

We present the results in three sections. The first section presents the sample and provides some descriptive statistics. The following section focuses on the fairness judgments expressed in the questionnaire. We compare the answers in the “isolated reasoning” and “social reasoning” treatments, we present and discuss the answers to the particular cases referring to the framework presented in Section 5.1.2. In the last part, we present the answers and the choices made in the third part of experiment. We present the effects of the reasoning treatments, and we analyze how choices and answers depend on the features of the situation.

5.3.1 Descriptive statistics

The sample is composed of 151 students from Toulouse, mainly in Law, Management and Economics. 43% (65) were male. The average age was 21.5 year old. For the 136 participants having informed their political preferences from left to right on a scale from 1 to 10, the average index is 5.3.

No significant difference is observed across them in terms of gender, age or political preferences. It is however noticed that there are significantly fewer students in law, management and political sciences and more students in economics in the “no reasoning” treatment group than in the other two.[10]

The compensation requirements range from 0 to 19€, on average 9.3€. The participants spent on average 4.3 minutes on the first trial and 26 minutes on the final task (from 11

[10] A precise description is provided in Appendix.

to 46 min). Each pair consists in a participant with a lower compensation requirement, hereafter called the *task performer*, and a participant with a higher compensation requirement, hereafter called the *other participant*. In the event both participants in a pair have the same compensation requirement, the roles are allocated randomly.[11] Table 5.1 shows that a reasonable diversity of situations is achieved.

	Mean (sd)	min	max
c_1	6.3 (2.4)	0	10
c_2	12.3 (2.3)	8	19
$c_2 - c_1$	6.1 (3.5)	0	13

Table 5.1: Descriptive statistics of the compensation requirements in the pairs. c_1 denotes the lower compensation requirement in the pair, that is the compensation requirement of the task performer, and c_2 the compensation requirement of the other participant. $c_2 - c_1$ denotes the discrepancy between the two compensation requirements in the pair.

5.3.2 Fairness judgments in the questionnaire

The questionnaire corresponds to the second part of the experiment on Figure 5.1. As explained in subsection 5.2.3, it consists in two parts that were presented in a random order to the respondents. In this section, we first present and discuss their answers to each part separately. Then, we further discuss and relate them.

5.3.2.1 Judgments over allocation rules

In this part, the respondents were presented the four allocation rules motivated in Section 5.1.2. The recommendations of each allocation rule were illustrated in the situation $(6, 14)$. Respondents were asked to declare whether each of the four allocation rules was fair or unfair and which was the fairest. The order according to which the rules were presented was random.

The judgments expressed concerning each allocation rule are presented in Figure 5.1. It suggests that respondents reject the strict egalitarian allocation rule, are divided on the strict libertarian allocation rule and generally tend to judge the two intermediate allocation rules as fair.[12] The welfare egalitarian allocation rule appears to be the most favored rule overall.

As for Test 8.1, it seems that respondents judge the welfare egalitarian allocation rule as fair and fairest more often in the social reasoning treatment. Still, these observations are

[11] This was the case for 3 pairs out of 66.

[12] 72% of the participants deem more than a single allocation rule fair. The average number of allocation rules deemed fair is 1.9.

not statistically significant (respective 2-sided 2-sample test for equality of proportions, p-value = 0.364 and 0.135).

5.3.2.2 Answers to the particular cases

In this part, the respondents were asked to recommend a fair allocation to the task performer in the four particular situations introduced in section 5.1.2. The situations were presented in a random order.

Figure 5.2 shows the distribution of answers in each of the four cases. Consistently with Test 1, we observe the four modes corresponding to the four allocation rules in the situation $(6, 14)$ and, on each chart, the main mode corresponds to the recommendation of one of the four allocation rules: the strict egalitarian allocation is the most chosen alternative in the cases $(6, 6)$ and $(14, 14)$, the welfare egalitarian allocation in the case $(6, 14)$ and the libertarian allocations in the case $(6, 20)$. Still, part of the test fails. The mode on the strict egalitarian allocations is strikingly much greater in the situations $(6, 6)$ and $(14, 14)$ than predicted. Besides, we observe an additional mode on the allocation 14 in the situation $(14, 14)$ that is not predicted by any of the four allocation rules. One possible account could be the reliance on a constrained egalitarian rule that split the benefit equally under the constraint that the task performer gets at least her compensation requirement. Another unexpected observation is the high support for the equal split whenever the participants have the same compensation requirement, and the high support for the libertarian allocations in situations in which the second participants feature so high a compensation requirement that the task would have no value to her anyway.[13]

Figure 5.3 shows the average amount allocated to the task performer in each of these situations. Consistently with Tests 2.2 and 2.3, the amount granted to the task performer generally increases with the greater cost. It increases from the situation $(6, 6)$ to $(6, 14)$ (Wilcoxon rank-sum test, p-value$<$0.001) and from $(6, 14)$ to $(6, 20)$ (Wilcoxon rank-sum test, p-value$<$0.001). However, we do not observe that the average amount granted to the task performer increases from the situation $(6, 14)$ to $(14, 14)$. This result is surprising as this is what the welfare egalitarian allocation rule, the most chosen rule in the other part, would recommend. While this contradicts Test 2.1, this may be due to the specific treatment of case in which the participants feature the same compensation requirement.

Figure 5.3 further suggests that respondents tend to give more to the task performer in the isolated reasoning treatment. This difference is statistically significant in the cases $(6, 6)$, $(14, 14)$ and $(6, 14)$ (Wilcoxon rank-sum test, respective p-value are 0.0703, 0.00986 and 0.09349) but not in the case $(6, 20)$. While this seems to confirm Test 8., this may

[13]This is referred to as the *no-dummy* principle in axiomatic analyses.

be interpreted with care as this may be due to a different distribution of types in the two treatments that may not be significant given the size of the sample.

In summary, while part the previous observations are consistent with the tests intended, some surprising facts question the idea that respondents have a clear preference for a given *rule* among the four we consider. It rather suggests that the judgments observed in each particular situation vary in a way that none of the four allocation rule is able to account for.

In order to further assess the consistency of the individual answers with each of the allocation rules, Test 6 proposes to classify participants according to the average distance of their particular answers to the recommendations of each of the allocation rule presented in Table 5.1. For instance, consider a participant who would have respectively answered 10, 10, 20 and 13 to the cases $(6, 6)$, $(14, 14)$, $(6, 20)$, and $(6, 14)$. The respective distances to the recommendations of the four allocation rules SE, WE, WL and SL are 3.25, 4.25, 3.5 and 6.75. It is minimal for the strict egalitarian allocation rule so we would classify this respondent as a strict egalitarian. The classification obtained is presented in Table 5.2. Only 35 (40%) of the 88 respondents are classified consistently with the fairness ideal they choose in the other part of the questionnaire with this procedure. This contradicts Test 6 and confirms that a significant part of the respondents does not consistently apply one of the four allocation rules across all the cases. In order to get an idea of the proportion of respondents which are influenced by the features of the situation, we add a fifth possibility that consists in the choice of the most intuitive allocation in each of the situation, where the most intuitive allocation is defined as the main mode observed. This leads to the respectively recommend to allocate €10, €10, €20 and €13 to the task performer in the cases $(6, 6)$, $(14, 14)$, $(6, 20)$, and $(6, 14)$. Adding this possibility leads to capture 25% of the respondents that would have otherwise been classified as strict egalitarians or weak libertarians.

In summary, the choices made in particular situations give us a different picture on fairness judgments as the one suggested at the beginning. Consistently with the recommendation of the libertarian allocation rules, a majority of respondents deem fair to give nothing to someone who would not be willing to perform the task. Consistently with the strict egalitarian rule, a significant propotion of the respondents deem fair to split the benefit equally when both participants feature the same compensation requirement. However, a significant proportion of the respondents does not seek to make consistent recommendations across cases in the sense of the four allocation rules we proposed.

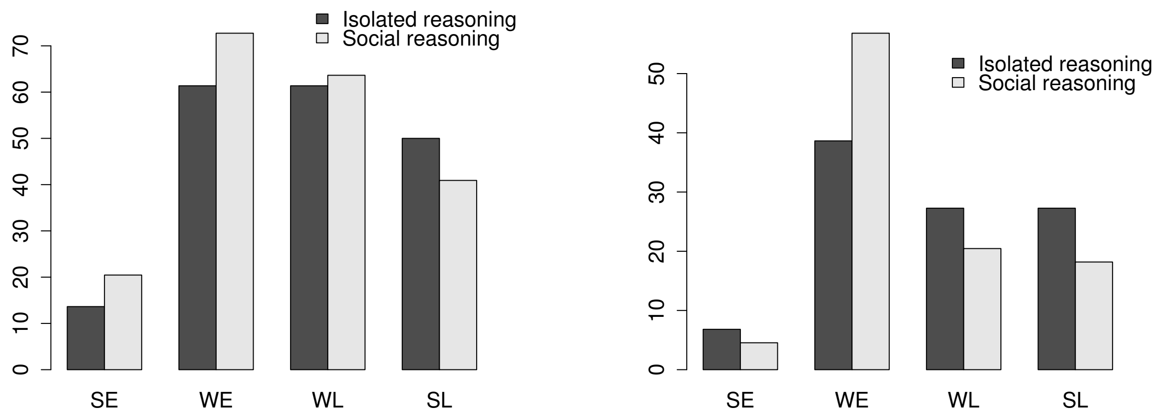


Figure 5.1: Proportion of respondents judging each allocation rule as “fair” against “unfair” (on the left) and proportion of respondents judging each rule as the “fairest” (on the right) in each treatment.

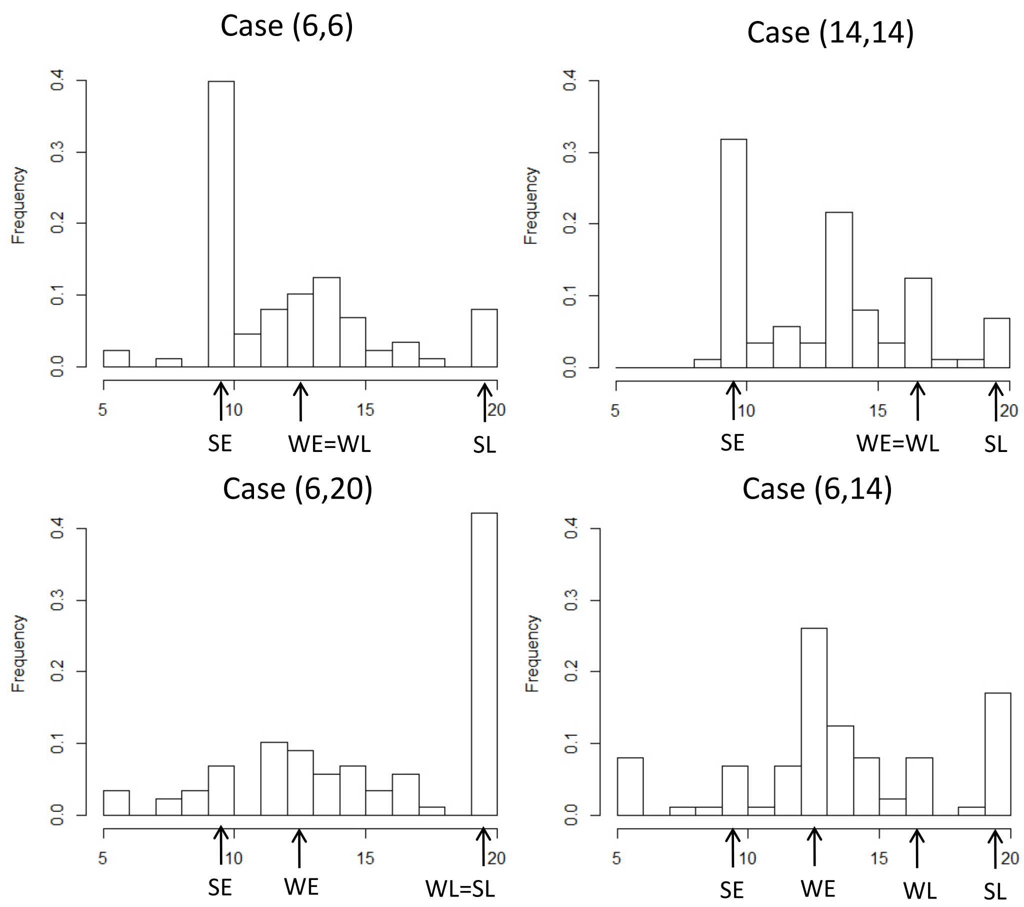


Figure 5.2: Distribution of the allocation recommended for the task performer in each case.

The recommendations associated with each allocation rule are indicated below.

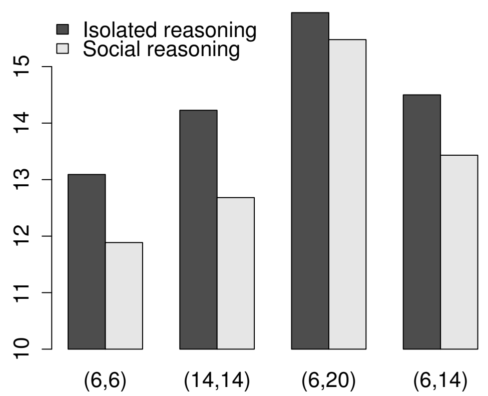


Figure 5.3: Average allocation to the task performer in each of the particular cases by treatments.

	SE	WE	WL	SL	Situation specific
Isolated reasoning	13	11	14	6	-
Social reasoning	17	11	14	2	-
Allocation rules first	12	13	14	6	-
Cases first	18	9	14	2	-
Total number	30	22	28	8	-
	34%	25%	32%	9%	-
Total number	18	22	18	8	22
	20%	25%	20%	9%	25%

Table 5.2: Classification of respondents according to their answers to the particular cases.

5.3.2.3 Further analysis and discussion

The reflective equilibrium procedure allows the judgments in particular cases and over allocation rules to affect each other both ways until an equilibrium is reached. However, in spite of the fact that participants were allowed to revise their answers, only 17 respondents (19%) did revise some answers to one part after having started answering the next one.[14] Besides, we observe in Figure 5.4 that the allocation recommended to the task performer in the particular case is systematically higher when respondents are invited to express their judgments on allocation rules first. This results is statistically significant for the cases (6,6), (14,14) and (6,14) (Wilcoxon rank-sum test, respective p-value are 0.03156, 0.0806 and 0.09064), but not for the case (6,20) (Wilcoxon rank-sum test, p-value=0.2341). Figure 5.5 suggests that this difference could be driven by the different distribution of fairness ideals in the two groups. Indeed we observe that the number of respondents judging the strict egalitarian allocation rule (resp. the strict egalitarian rule) as the fairest is higher (resp. lower) when rules are presented first. Even though, these differences are not statistically significant (2-sample test for equality of proportions respective p-values are 0.1237 and 0.1651), they could explain the higher average amount allocated to the task performer when rules are presented first. This may be mitigated by the fact that the weak libertarian allocation rule seems to be more chosen when cases are presented first. However, this results is not statistically significant either (2-sample test for equality of proportions, p-value=0.2677). This analysis suggests that judgments expressed in the questionnaire are not in equilibrium. However, these effects would have to be confirmed by a more focused and powered experiment.

5.3.3 Choices

After their situation is disclosed to them, the participants successively declared their normative expectation regarding how the other should split the revenue from the task, specified an offer to the other participant and guessed another participant's normative expectation. In this section, we present each of these measures and explore how they depend on the treatments and the features of the situation.

5.3.3.1 Normative expectations

Normative expectations correspond to the answer to the question "Would the other participant share the value of the task, what would be the minimal amount that he gives

[14] As observed in Amiel and Cowell (1999), participants seem reluctant to revise their answers even when they are invited to do so.

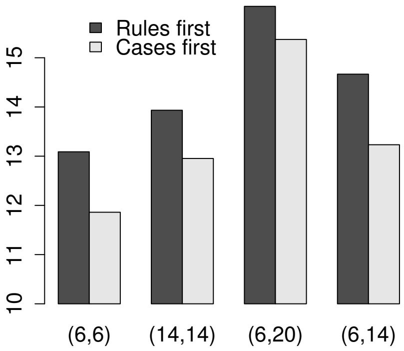


Figure 5.4: Average allocation to the task performer in each of the particular cases according to the order of the two parts.

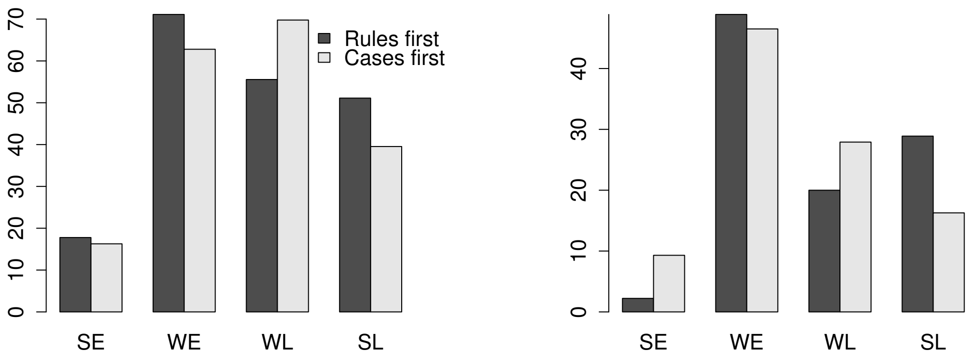


Figure 5.5: Proportion of respondents judging each allocation rule as "fair" (against "unfair") (on the left) and proportion of respondents choosing each allocation rule as "the fairest" (on the right) according to the order of the two parts.

you that you would find defendable?". The participants are informed that their answer to this question has no consequence on their final payment and that this amount will never be disclosed to the other participant. In the last situation, they are asked to guess the answer of another participant in the opposite role. Correct guesses are paid €10.

Figure 5.6 presents the average normative expectations by treatment and role. On average, the task performers declared to find defendable offers that are not lower than 12.9€ and the other participants, 7.9€. As for Test 9, the overall effect of the treatment is to increase the amount expected by the task performers (Wilcoxon rank-sum test, p-value=0.00407) and decrease the amount expected by the other participants (Wilcoxon rank-sum test, p-value=0.04704). The average guessed normative expectations are shown on Figure 5.7. They are on average of 12.6€ for the task performers and 8.5€ for the others. Test 5 is partly confirmed. The treatment effect on normative expectations is consistent with the effect observed on guesses: participants who will not perform the task are expected to expect lower amounts (Wilcoxon rank-sum test, p-value=0.04636). However, the guessed amount expected by the task performers cannot be said to be statistically lower for the no reasoning treatment as compared to the two others (Wilcoxon rank-sum test, p-value=0.153). The cumulative distribution of the stated normative expectations is presented on Figure 5.8 by treatment and role. The difference observed for the task performers seems to be mainly driven by a departure from the equal split whereas the origin of this difference is more diffuse for the other participants.

As suggested by the answers to the questionnaire, the situation influences the perceived fair amount in two ways. First, while the strict egalitarian and the strict libertarian position are insensitive to the costs, the welfare egalitarian allocation rule recommends a higher payment to the task performer as her own cost grows while the weak libertarian allocation rule recommends to give more to the task performer when as the other's cost grows. This motivates Test 3. Furthermore, the answers to the questionnaire suggest that the salient features of a situation tend to favor the application of different allocation rules. Therefore, we expect normative expectations to depend on the compensation requirements of the participants in the pair. Results are presented in Table 5.3. We first observe that the participants that will not perform the task do not expect significantly different payments depending on their situation. Contrary to what our framework predicts, results rather suggest that the higher their own cost, the more they expect. Task performers also feature surprising expectations. Regression 1 suggests that their normative expectations would increase with the other's cost but also that they decrease with their own cost. This last result cannot be explained by any of the allocation rules considered. It suggests that a better model would account for their normative expectations on the basis of the difference between the two compensation requirements of participants. This corresponds to Regression 2.

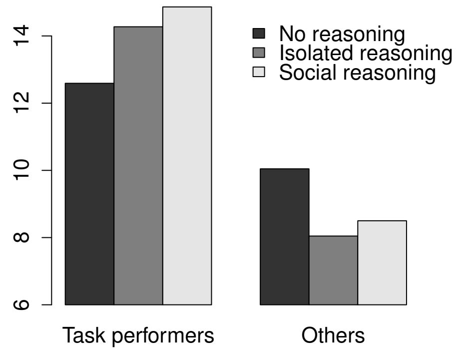


Figure 5.6: Average stated normative expectations by treatment and role.
Stated normative expectations correspond to the answer to the question “Would the other participant be to split the value of the task, what would be the minimal amount that he gives you that you would find defendable?”.

Figure 5.7: Average guess of the stated normative expectations of another participant in the opposite role by treatment and role.

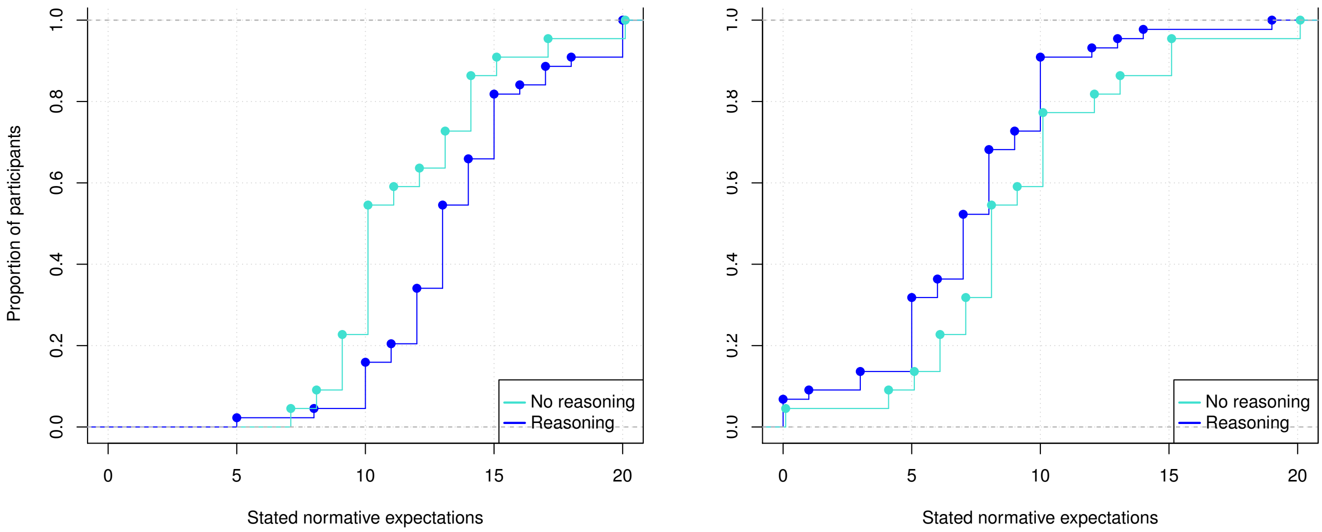


Figure 5.8: Cumulative distribution functions of stated normative expectations for task performers (on the left) and others (on the right).

	Task performers		Others	
	Regression 1	Regression 2	Regression 3	Regression 4
c1	-0.31^{*}		-0.04	
	(0.16)		(0.20)	
c2	0.30^{*}		0.36^{*}	
	(0.16)		(0.21)	
c2-c1		0.30^{***}		0.20
		(0.11)		(0.14)
No Reasoning	-2.24^{***}	-2.24^{***}	1.78^{*}	1.67
	(0.80)	(0.79)	(1.03)	(1.02)
Male	-0.46	-0.45	-0.25	-0.19
	(0.80)	(0.79)	(1.00)	(1.00)
Age	0.56^{**}	0.56^{**}	0.20	0.18
	(0.25)	(0.24)	(0.25)	(0.25)
(Intercept)	0.19	0.10	-1.16	2.23
	(5.58)	(5.25)	(6.24)	(5.37)
R^2	0.22	0.22	0.10	0.08
Adj. R^2	0.16	0.17	0.03	0.02
Num. obs.	65	65	66	66

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 5.3: Effect of the compensation requirements on declared normative expectations.

5.3.3.2 Offers

After having stated their normative expectations, the participants were invited to propose an allocation of the €20. They were informed that the other participant was also choosing one and that the offer that would eventually be implemented would be drawn randomly.

Figure 5.9 shows the average normative expectations by treatment and role. On average, the task performers offered 6.7€ and others, 9.8€. Test **??** is not observed: no statistically significant difference is observed between the "no reasoning" treatment and the two others (Wilcoxon rank sum test, p-value=0.1586). Test 4 is not observed either: the compensation requirements in the pair are not found to significantly affect the offers. The distribution of offers is presented in Figure 5.10.

As for Test 7, Table 5.4 presents regressions of the fair wage derived from the normative expectations and the offer on the fairness ideals deduced from the questionnaire for participants in the reasoning treatments. We observe that the answers that the fair wage tend to be higher for individuals that are classified as strong libertarians according to their answers to the particular cases. Still, the verbal statements fail to account for the observed offers.

5.3.3.3 Stated fairness judgments in the final questionnaire

In the final questionnaire, participants were invited to recommend an allocation for the task performer and justify it in the situation $(6, 14)$. Figure 5.11 present their final answers. This last question confirms the existence of a large diversity of views and the potential for the four allocation rules proposed to focus the attention. The justifications provided are presented in Appendix.

5.4 Discussion

Individual fairness judgments predict behavior. In some circumstances, we may expect them to be primary motive. Still, they may result from spontaneous feelings, heuristics, or customary association of ideas. They may also be sensitive to subtle cues. They may also significantly differ depending on the context in which they are expressed. Facing this, we propose to identify some patterns that are likely to be driven by the structure of this problem. Few studies have studied individual fairness judgments in this structure.[15]

[15] As mentioned in the introduction, Schokkaert et al (2007) is another study that focus on a similar structure.

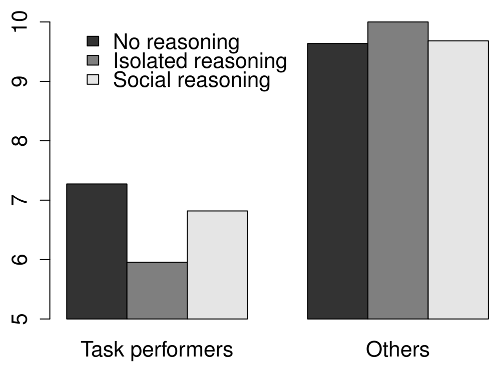


Figure 5.9: Average offer by treatment and role.

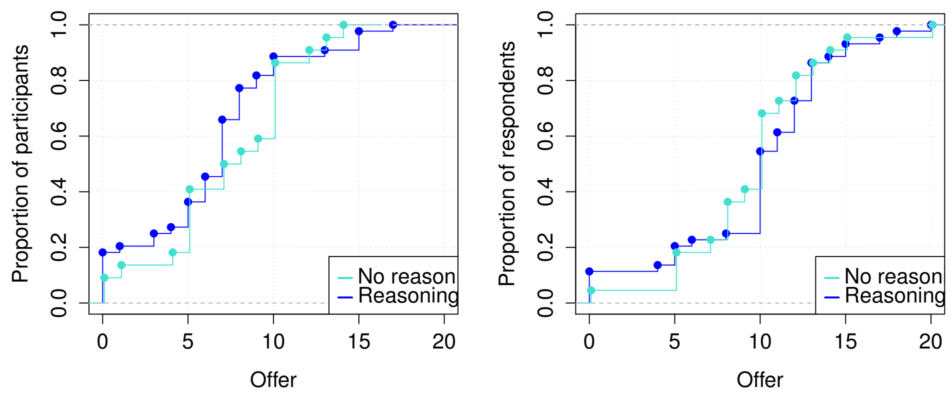


Figure 5.10: Cumulative distribution function of offers from task performers (on the left) and others (on the right).

	Stated fair wage	Chosen wage
Constant	10.25***	7.63***
	(1.36)	(2.06)
Consistent WE	0.49	1.53
	(0.96)	(1.45)
Consistent WL	2.87***	1.73
	(1.01)	(1.53)
Consistent SL	5.22***	4.96**
	(1.32)	(1.99)
Consistent situation spec.	1.65	−1.35
	(1.01)	(1.52)
Stated ideal WE	1.24	1.52
	(1.47)	(2.23)
Stated ideal WL	−0.04	1.03
	(1.54)	(2.33)
Stated ideal SL	1.96	0.23
	(1.64)	(2.48)
Task performer	0.44	4.07***
	(0.65)	(0.98)
R^2	0.34	0.27
Adj. R^2	0.27	0.19
Num. obs.	88	88

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 5.4: Determinants of the stated fair wage in the questionnaire and the chosen wage for participants in the reasoning treatments.

For the task performer, the stated fair wage is defined as her normative expectations. For the other, it corresponds to the difference between 20€ and the normative expectation of the other participant. The chosen wage is the difference between 20€ and her offer for the task performer and her offer for the other participant. The stated fairness ideals (above) corresponds to the preference expressed over the allocation rule in the questionnaire. The derived fairness ideals (below) correspond to the adhesion to the allocation rule inferred from the judgments expressed in the particular cases as detailed in subsection 5.3.2.2.

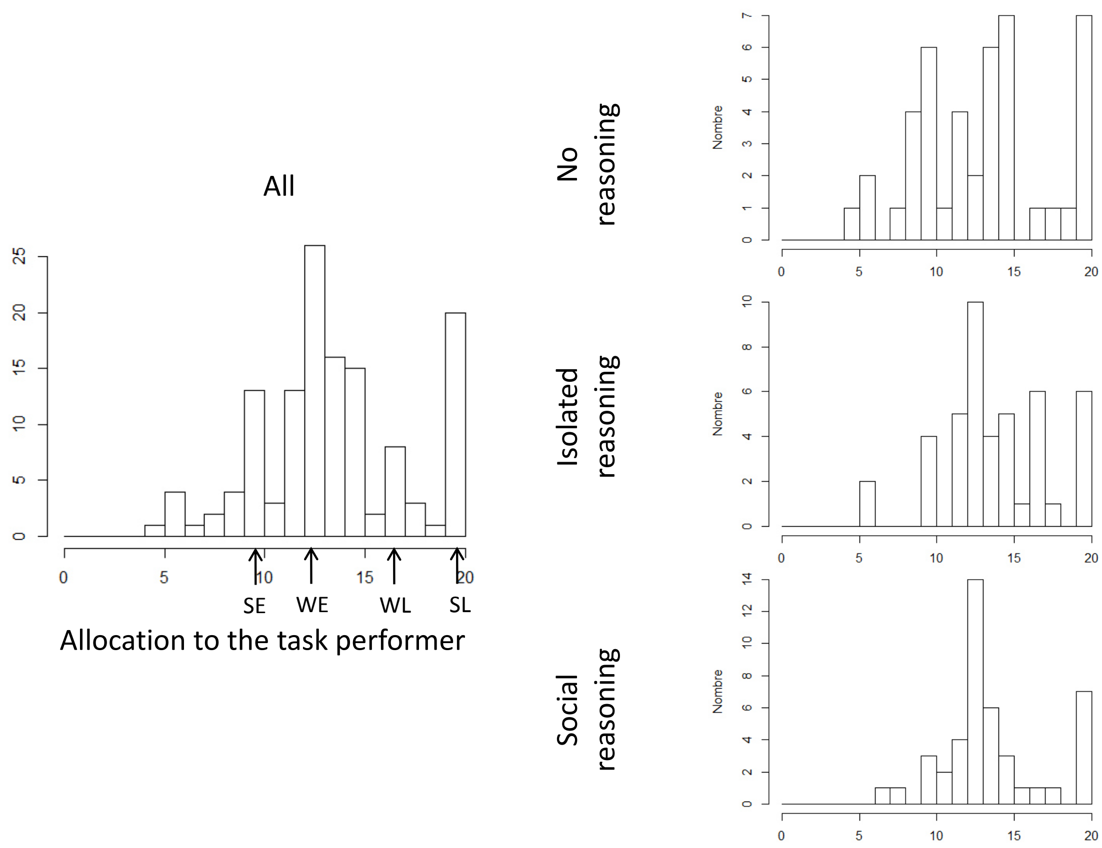


Figure 5.11: Allocation to the task performer in the situation $(6, 14)$ recommended in the final questionnaire.

Our three main findings are the following. First, this experiment confirms that fairness judgments are sensitive to reasoning. More precisely, we observed that answering the questionnaire has the effect of increasing the normative expectations of the task performer and decreasing the normative expectations of others. It is more pronounced for the task performers themselves and seems to be mainly explained by a move away from the equal split for the task performers. However, no significant effect is observed for the actual allocation choices in the dictator phase, which may be due to a lack of power. Actually, the effect of the treatment on the fraction of participants choosing the equal split is significant both for the stated normative expectations and the offers of the task performers. Second, we stress some apparently wide convergence on individual judgments. The first observation is that a majority of respondents deem fair to give nothing to someone who would not be willing to perform the task. This corresponds to the no-dummy principle in axiomatic studies. A second observation is that respondents tend to split the benefit equally when both participants feature the same compensation requirement. We suspect that this is driven by a heuristic, which would need to be explored further. Third, we observe that the normative expectations of the participants are sensitive to the differential in costs in a pair.

Despite these observations, we note that none of the proposed allocation rules is able to account for the variations of observed judgments across particular cases, nor for the fact that normative expectation increase with the differential in compensation requirements in a pair. This could be accounted for in two directions. A first possibility is that participants follow their first intuition and do not seek to be consistent across all these cases. In other word, the judgments observed are not in reflective equilibrium (Rawls, 1951). It is true that normative reasoning requires a high cognitive involvement, even in this stylized situation. Further experiment could question whether individuals value being consistent across cases and how participants would react when made aware of their "moral mistakes", understood as the inconsistencies in their judgments or the potentially undesirable consequences of their adopted principles. Social psychologists suggest that reasoning would best be incentivized in social interactions (Mercier et Landemore, 2012). It would be interesting to raise the question of whether individuals care about others being consistent when having to explicitly justify their choice to others. Another question is about whether individuals tend to accept a single or several allocations. In their final justification, many participants expressed a mere desire to give more to the task performer, while being very lax in justifying how much this difference should be.[16] The effect of reasoning on tolerance rather than on fairness ideals may actually be an interesting route to pursue for consensus building. There exists a second way to deal with the apparent inconsistencies observed in this experiment. It relies on the possibility that different allo-

[16] In the questionnaire, 72% of the respondents deemed several allocation rules as fair (against unfair) (on average 1.9).

cation rules than the one we considered may be consistent with the choices and judgments observed, while still meeting desirable principles. In the context at hand, this raises the question of whether allocation rules that would be sensitive to the differential between costs could be grounded in principles.

In the end, none of the four allocation rules proposed stands out. However, at least three of them have some attractive features that make them more relevant in different settings. First the strict egalitarian allocation rule seems to be spontaneously chosen by participants. This is consistent with the idea that the equal split is a widespread heuristic. It is also the allocation that attracted most answers to the guess. Therefore, it could be proposed in settings in which little time for reflection is available or when participants have to coordinate on their expectations with limited communication. However, we also observe that reasoning leads to a general departure from it and that it attracts little support as a general allocation rule in the questionnaire. In contrast, it is the welfare egalitarian allocation rule which seems to attract most support in the questionnaire, both as a general allocation rule and in the particular case $(6, 14)$. It would constitute a good candidate in the context of a deliberation about a specific situation. However, few actual choices seem to be consistent with this allocation rule in specific situations. In particular, we find no evidence that participants are willing to pay the task performer more, nor that the task performer expects to get more when his compensation requirement gets higher. Besides, this allocation rule fails to capture the judgments expressed in the situation $(6, 20)$. Finally, the weak libertarian solution seems to best conciliate the tendencies observed in particular cases. In all the four particular cases considered, it is the allocation rule that seems to constitute an attractive middle ground between the idea that one should get the outcome of her own work as soon as no one is deprived from an opportunity and redistribute to some extent. However, it does not account for the particular judgments expressed in the symmetric cases $(6, 6)$ and $(14, 14)$.

Conclusion

This experiment constitutes a preliminary investigation into the formation of individual fairness judgments for the allocation of an indivisible task and its benefit. This situation is particularly interesting as it requires some degree of normative reasoning and allows us to bring together the theory of fair allocations and the empirical methods for the observation of individual fairness judgments. Furthermore, this situation features the basic structure of many problems of current interest, such as the allocation of property rights over pollution (emission rights) or the extraction of natural resources (water or fishing quotas). Another policy context could be the allocation of locally undesirable land uses among communities.

The results presented confirm that judgments hold on two widely shared but conflicting norms. On the one hand, we find that many respondents deem the task performer entitled to the benefit of the task. In particular, we observe a wide adhesion to the no-dummy principle, which requires giving nothing to someone who would not be willing to perform the task. On the other hand, we also observe strong egalitarian motives, which manifest in different and sometimes contradictory ways. For instance, we observe that a vast majority of the respondents favor the welfare egalitarian allocation rule but that most of them would choose to split the benefit of the task equally when both participants feature the same compensation requirement. As such, the observation may call for the characterization of alternative allocation rules or question the internal consistency of individual judgments. In any case, the ways in which the respondents account for their judgments and adjust their beliefs would constitute a further step toward a characterization of reasoned judgments in this context.

Overall, and despite a relatively uniform sample in terms of age and background, a significant diversity of fairness views is observed. This lead us to suggest that this setting could be useful to address more fundamental questions about normative reasoning, the individual sense of justice and the possibility of an overlapping consensus (Rawls, 1993) in the presence of contrasted moral intuitions. Furthermore, we observe that self-declared political preferences did correlate with some answers. For instance, right-wing oriented participants tended to judge the welfare egalitarian allocation as the fairest less often. As one's political ideology may conflict with one's self interest in the context of this experiment, this situation may also prove interesting to study the relative role of deliberation, ideology and self-interest in shaping the judgments of actual stakeholders. In the end, we hope that the results presented here and the question raised by this analysis would motivate further inquiries in this specific setting.

Appendix

5.A Sample

Table 5.A.1 presents individual and situational characteristics in the three treatments.

	All (132)	NR (44)	IR (44)	SR (44)
Individual characteristics				
Male	59	20	19	20
Age	21.6	21.8	21.4	21.5
Political preference	5.3	5.2	5.5	5.2
Law, Management and Political science	72	**19**	**27**	26
Economics	34	**16**	**7**	11
Other	26	9	10	7
Situational characteristics				
c_1	6.3	6.0	6.5	6.5
c_2	12.3	12.2	12.5	12.3
$c_2 - c_1$	6.1	6.2	6.1	5.9

Table 5.A.1: Descriptive statistics of individual and situational characteristics across the treatments.

NR, IR and SR respectively denote the “No reasoning”, “Isolated reasoning” and “Social reasoning” treatments. The significant differences are reported in bold. There are significantly less student in law (NR-IR 2-sample test for equality of proportions, p-value=0.06759 and NR-SR 2-sample test for equality of proportions, p-value=0.10759) and more students in economics in the “no reasoning” treatment than in the other two.

5.B Experimental conditions

The experiment used TSE’s mobile lab. Pictures of the stations are provided by Figure 5.B.1.

Figure 5.B.1: Experimental room and an individual station.

5.C Detailed screens

5.C.1 Proceedings

The experiment proceeds as follows :

- Introduction

Screen 1: Welcome message, general instructions, and consent form

- **Part 1:** Elicitation of the compensation requirement

Screen 2: **(Introduction Part 1)**

Screen 3: **(Elementary task)** Participants perform a first elementary task: they count the number of occurrence of a letter in a square of randomly generated letters.

Screen 4: **(Elicitation)** Participants are asked their compensation requirements to perform the elementary task 10 times. Truthful revelation is incentivized by some possibility to play a Becker-De Groot-Marschak mechanism.[17]

Screen 5: **(Waiting screen)** Participants wait for all the others in their session to reach this stage. When it is the case, they are matched into pairs, pairs are allocated into treatments[18] and they are directed to the next page.

[17]The mechanism is the following: a random number is drawn in [0,20]. If the stated compensation requirement is lower than this number, the subject is offered to perform the task and paid this amount. If the stated compensation requirement is lower than this amount, the subject does not get the opportunity to perform the task. We do not want the revelation to be truthful but want this information to be credible.

[18]The matching process in described in Appendix.

- **Part 2:** Normative reasoning (treatments "isolated reasoning" and "social reasoning" only)

Screen 6: **(Introduction Part 2)**

Screen 7: **(Normative reasoning)** Participants fill in a questionnaire that depends on their treatment. This is the only screen which varies across treatments. In the treatments with normative reasoning, they are presented a vignette which is said to reflect the situation they will be involved in. Details are provided in a next subsection.

- **Part 3:** Elicitation of moral preferences and normative expectations

Screen 8: **(Introduction Part 3)** Participants are revealed the two compensation requirements in their pair, that the individual with the lower compensation requirement will be asked to perform the task.

Screen 9: **(Stated normative expectations)** Participants state the minimal offer they would find appropriate from the other participants. They are told that this statement would in no way influence they final payoff and not be observed by the other participant, nor the experimenter.

Screen10: **(Dictator game)** All participants specify an offer, knowing that one of the two offers in their pair will be randomly implemented.

Screen11: **(Elicitation of normative expectations)** Subjects are asked to guess the minimal offer another participant, in a different situation, deemed appropriate. Correct guess is incentivized.

- **Part 4:** Task (for the relevant subjects only)

Screen12: **(Introduction Part 4)**

Screen13: **(Task)** The subjects with the lower compensation requirements in their pair perform the task. The other subjects skip this stage.

Screen14: **(Waiting screen)**

- Concluding screens

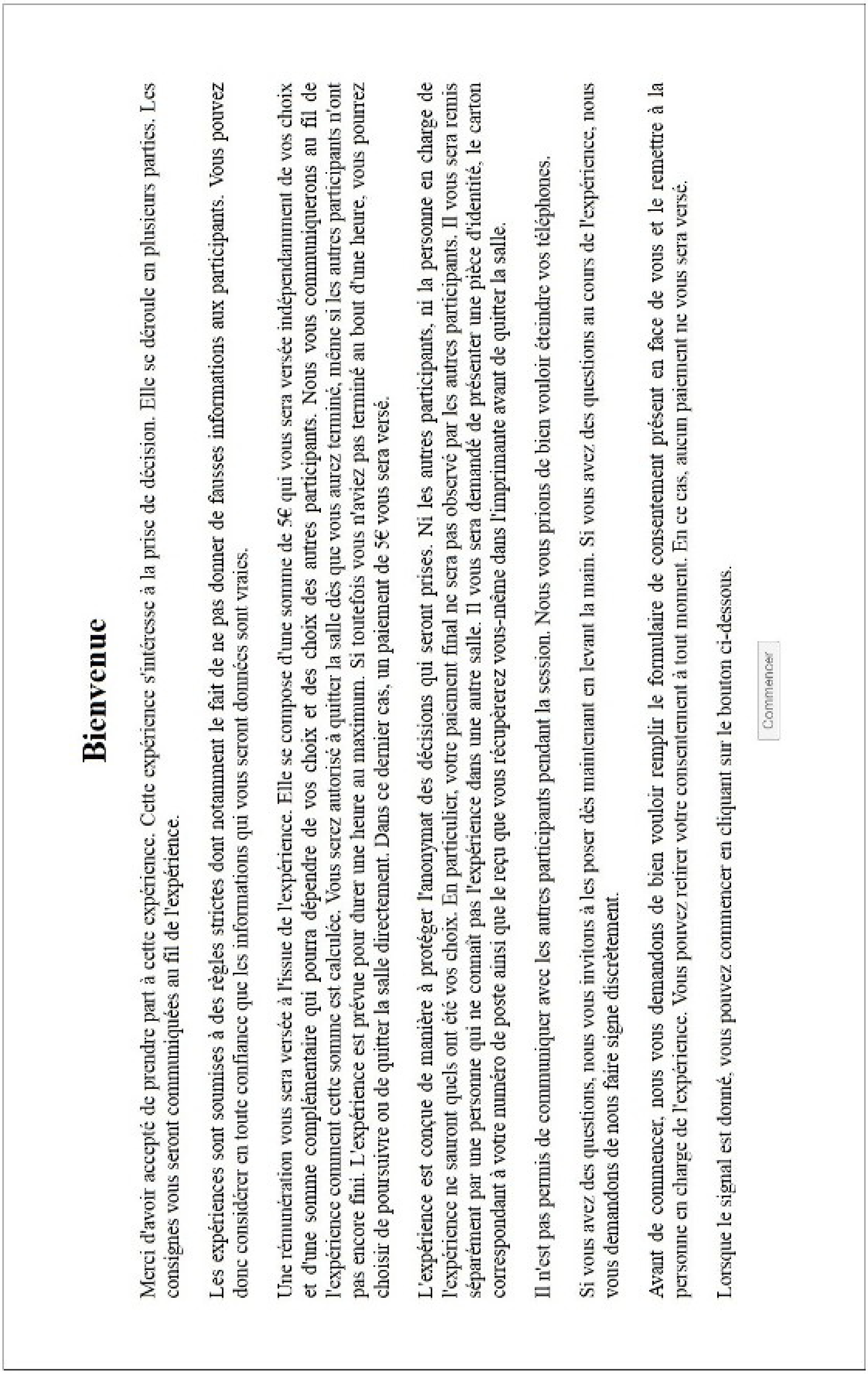
Screen15: **(Final questionnaire)** Questionnaire about socio-demographic characteristics.

Screen16: **(Results)** Final screen.

5.C.2 Screenshots

The following screenshots correspond to the "*Social reasoning*" treatment. Some comments are provided below each screen. For screens 7 and 15, all information is not visible on a single screen. The detailed text is added below.

Screen 1



Bienvenue

Merci d'avoir accepté de prendre part à cette expérience. Cette expérience s'intéresse à la prise de décision. Elle se déroule en plusieurs parties. Les consignes vous seront communiquées au fil de l'expérience.

Les expériences sont soumises à des règles strictes dont notamment le fait de ne pas donner de fausses informations aux participants. Vous pouvez donc considérer en toute confiance que les informations qui vous seront données sont vraies.

Une rémunération vous sera versée à l'issue de l'expérience. Elle se compose d'une somme de 5€ qui vous sera versée indépendamment de vos choix et d'une somme complémentaire qui pourra dépendre de vos choix et des choix des autres participants. Nous vous communiquerons au fil de l'expérience comment cette somme est calculée. Vous serez autorisé à quitter la salle dès que vous aurez terminé, même si les autres participants n'ont pas encore fini. L'expérience est prévue pour durer une heure au maximum. Si toutefois vous n'aviez pas terminé au bout d'une heure, vous pourrez choisir de poursuivre ou de quitter la salle directement. Dans ce dernier cas, un paiement de 5€ vous sera versé.

L'expérience est conçue de manière à protéger l'anonymat des décisions qui seront prises. Ni les autres participants, ni la personne en charge de l'expérience ne sauront quels ont été vos choix. En particulier, votre paiement final ne sera pas observé par les autres participants. Il vous sera remis séparément par une personne qui ne connaît pas l'expérience dans une autre salle. Il vous sera demandé de présenter une pièce d'identité, le carton correspondant à votre numéro de poste ainsi que le reçu que vous récupèrerez vous-même dans l'imprimante avant de quitter la salle.

Il n'est pas permis de communiquer avec les autres participants pendant la session. Nous vous prions de bien vouloir éteindre vos téléphones.

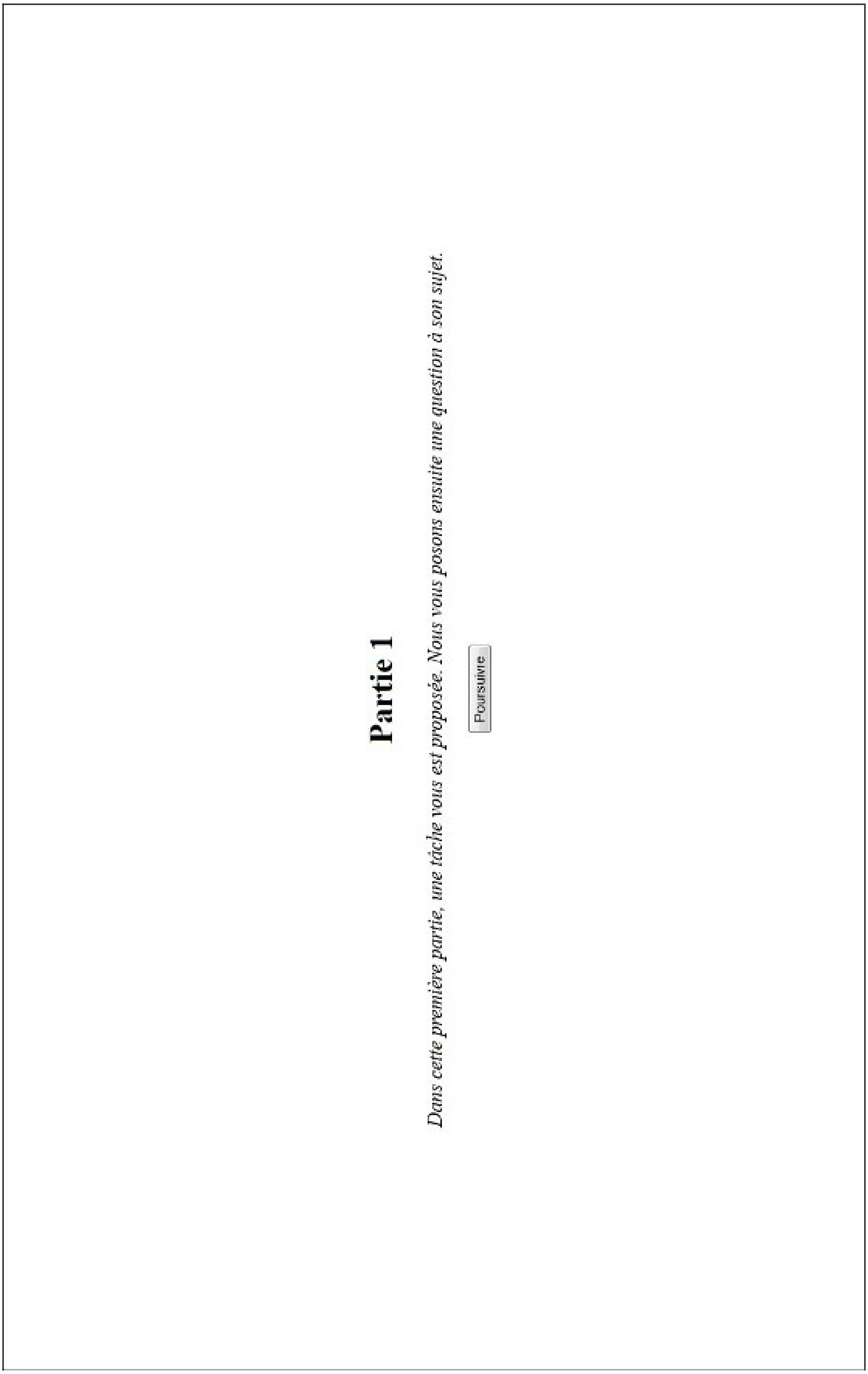
Si vous avez des questions, nous vous invitons à les poser dès maintenant en levant la main. Si vous avez des questions au cours de l'expérience, nous vous demandons de nous faire signe discrètement.

Avant de commencer, nous vous demandons de bien vouloir remplir le formulaire de consentement présent en face de vous et le remettre à la personne en charge de l'expérience. Vous pouvez retirer votre consentement à tout moment. En ce cas, aucun paiement ne vous sera versé.

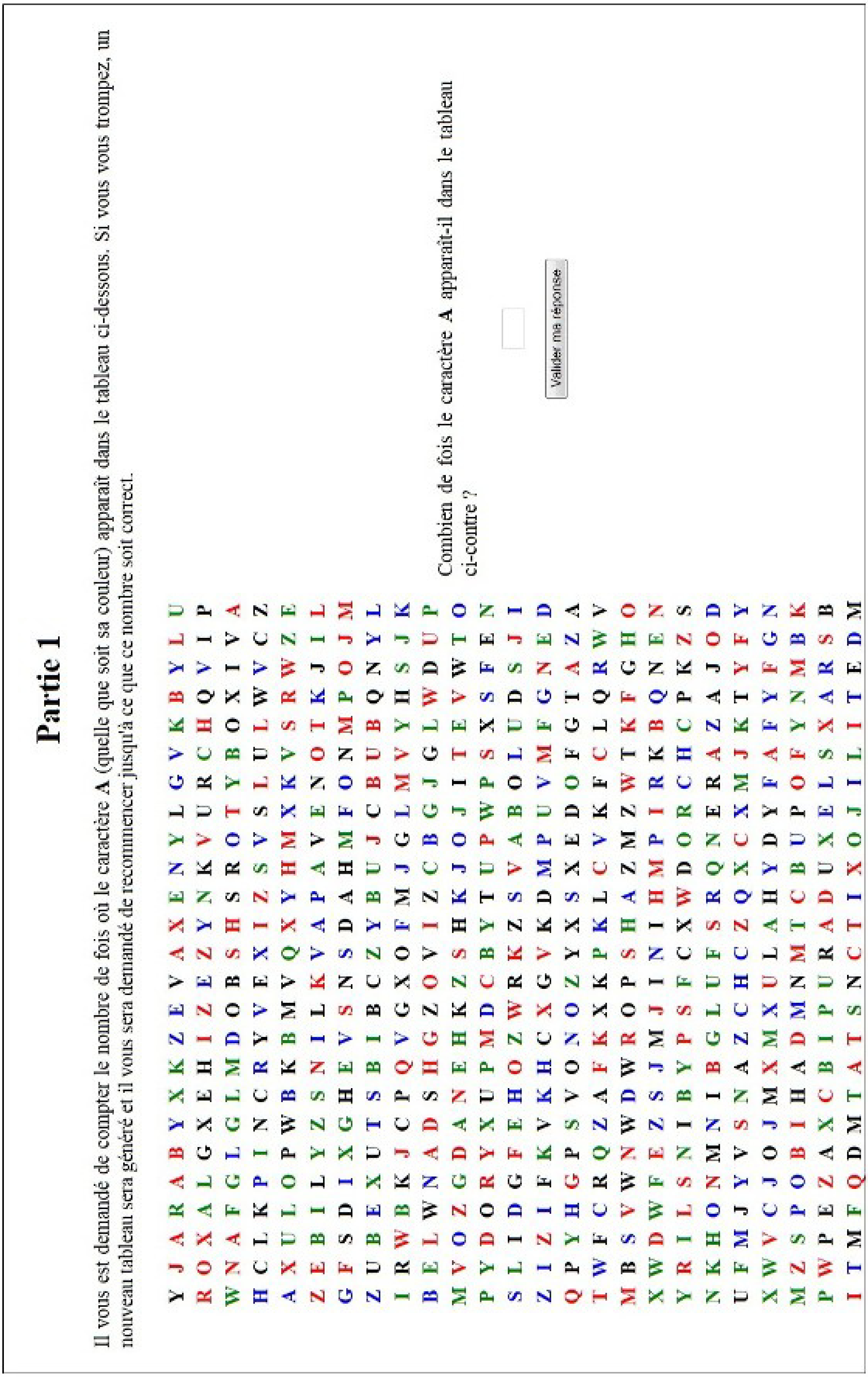
Lorsque le signal est donné, vous pouvez commencer en cliquant sur le bouton ci-dessous.

Commencer

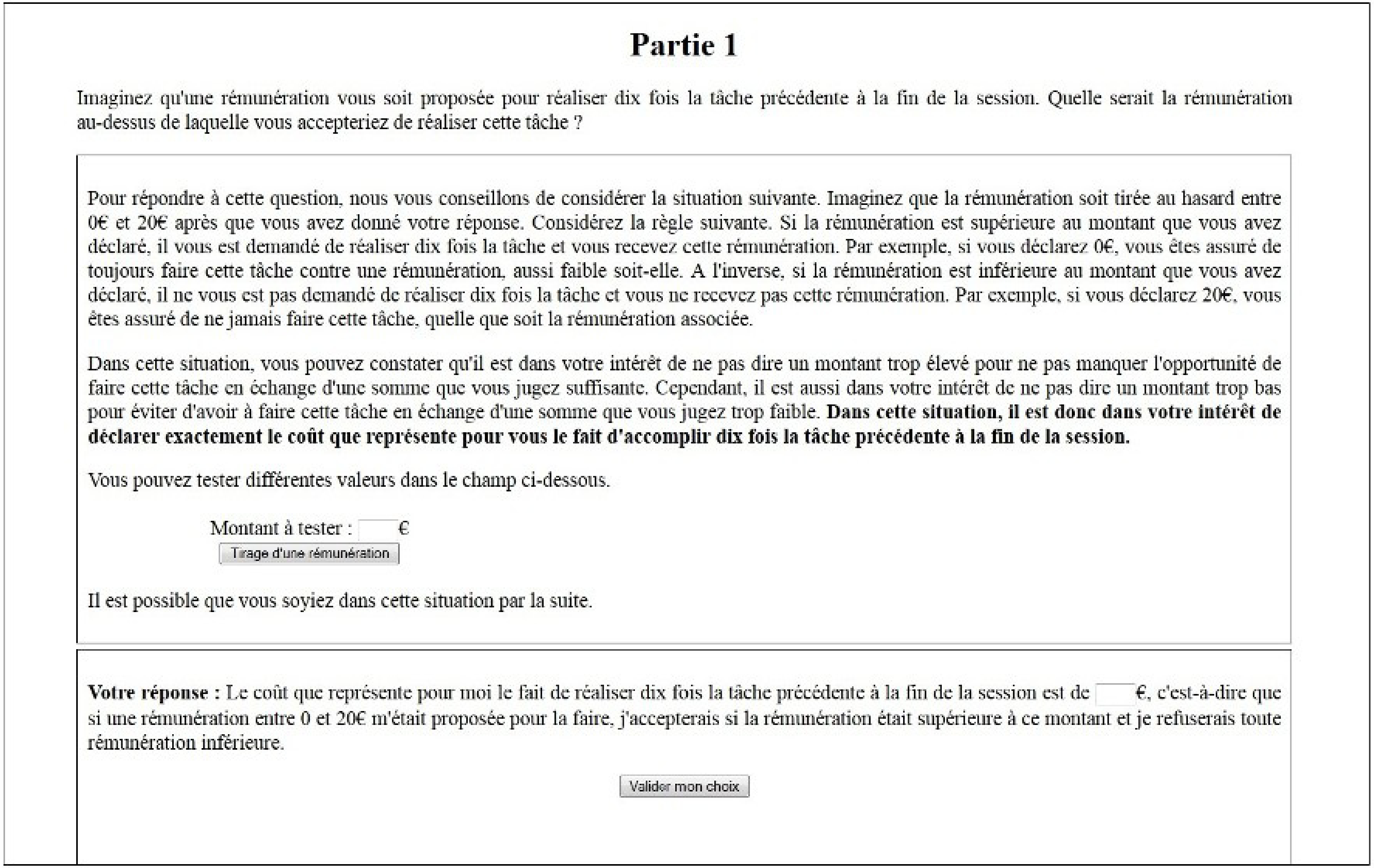
Screen 2


Partie 1
Dans cette première partie, une tâche vous est proposée. Nous vous posons ensuite une question à son sujet.
Poursuivre

Screen 3


Partie 1
Il vous est demandé de compter le nombre de fois où le caractère A (quelle que soit sa couleur) apparaît dans le tableau ci-dessous. Si vous vous trompez, un nouveau tableau sera généré et il vous sera demandé de recommencer jusqu'à ce que ce nombre soit correct.
Combien de fois le caractère A apparaît-il dans le tableau ci-contre ?
Valider ma réponse

Screen 4



Partie 1

Imaginez qu'une rémunération vous soit proposée pour réaliser dix fois la tâche précédente à la fin de la session. Quelle serait la rémunération au-dessus de laquelle vous accepteriez de réaliser cette tâche ?

Pour répondre à cette question, nous vous conseillons de considérer la situation suivante. Imaginez que la rémunération soit tirée au hasard entre 0€ et 20€ après que vous avez donné votre réponse. Considérez la règle suivante. Si la rémunération est supérieure au montant que vous avez déclaré, il vous est demandé de réaliser dix fois la tâche et vous recevez cette rémunération. Par exemple, si vous déclarez 0€, vous êtes assuré de toujours faire cette tâche contre une rémunération, aussi faible soit-elle. A l'inverse, si la rémunération est inférieure au montant que vous avez déclaré, il ne vous est pas demandé de réaliser dix fois la tâche et vous ne recevez pas cette rémunération. Par exemple, si vous déclarez 20€, vous êtes assuré de ne jamais faire cette tâche, quelle que soit la rémunération associée.

Dans cette situation, vous pouvez constater qu'il est dans votre intérêt de ne pas dire un montant trop élevé pour ne pas manquer l'opportunité de faire cette tâche en échange d'une somme que vous jugez suffisante. Cependant, il est aussi dans votre intérêt de ne pas dire un montant trop bas pour éviter d'avoir à faire cette tâche en échange d'une somme que vous jugez trop faible. **Dans cette situation, il est donc dans votre intérêt de déclarer exactement le coût que représente pour vous le fait d'accomplir dix fois la tâche précédente à la fin de la session.**

Vous pouvez tester différentes valeurs dans le champ ci-dessous.

Montant à tester : €

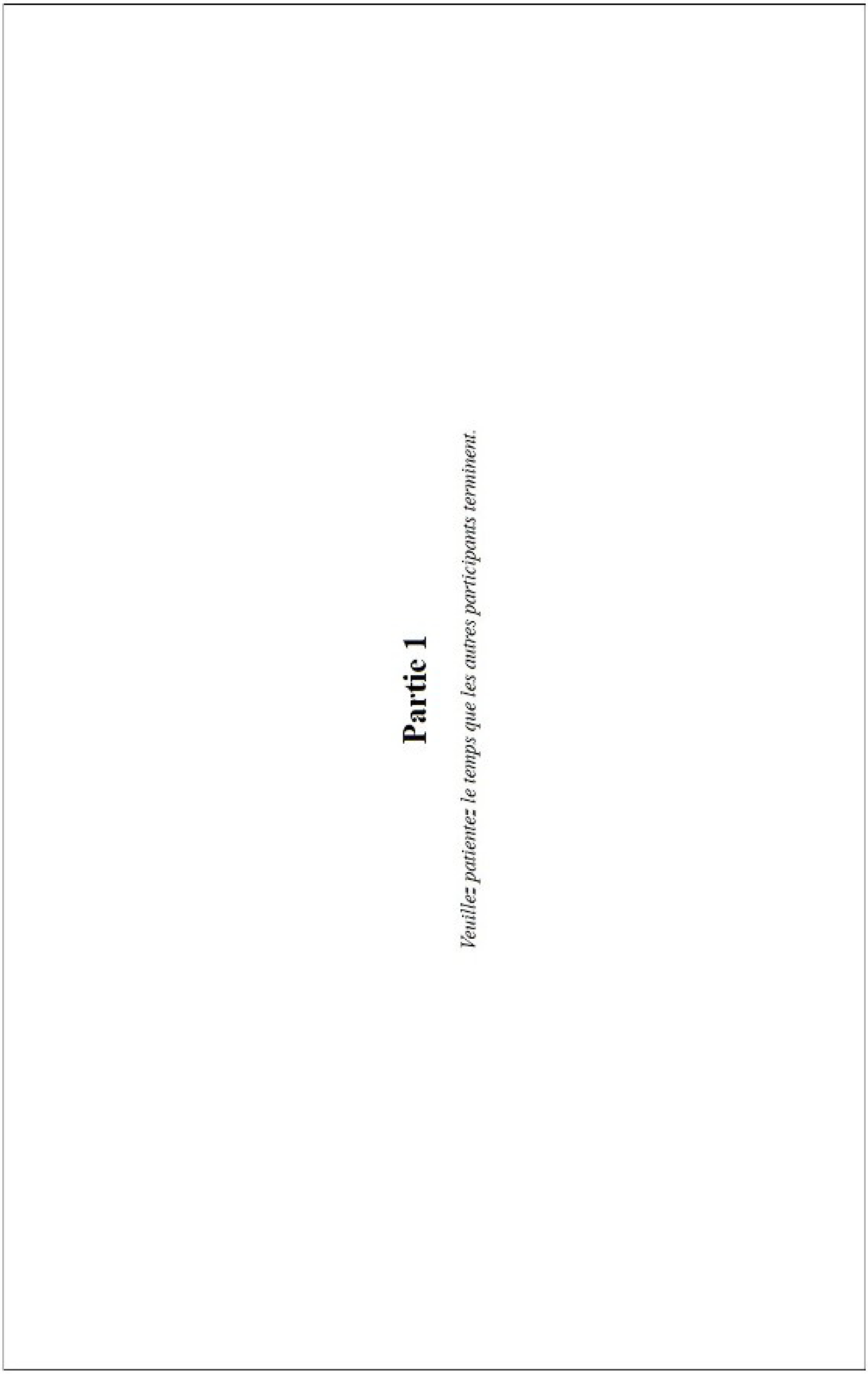
Tirage d'une rémunération

Il est possible que vous soyiez dans cette situation par la suite.

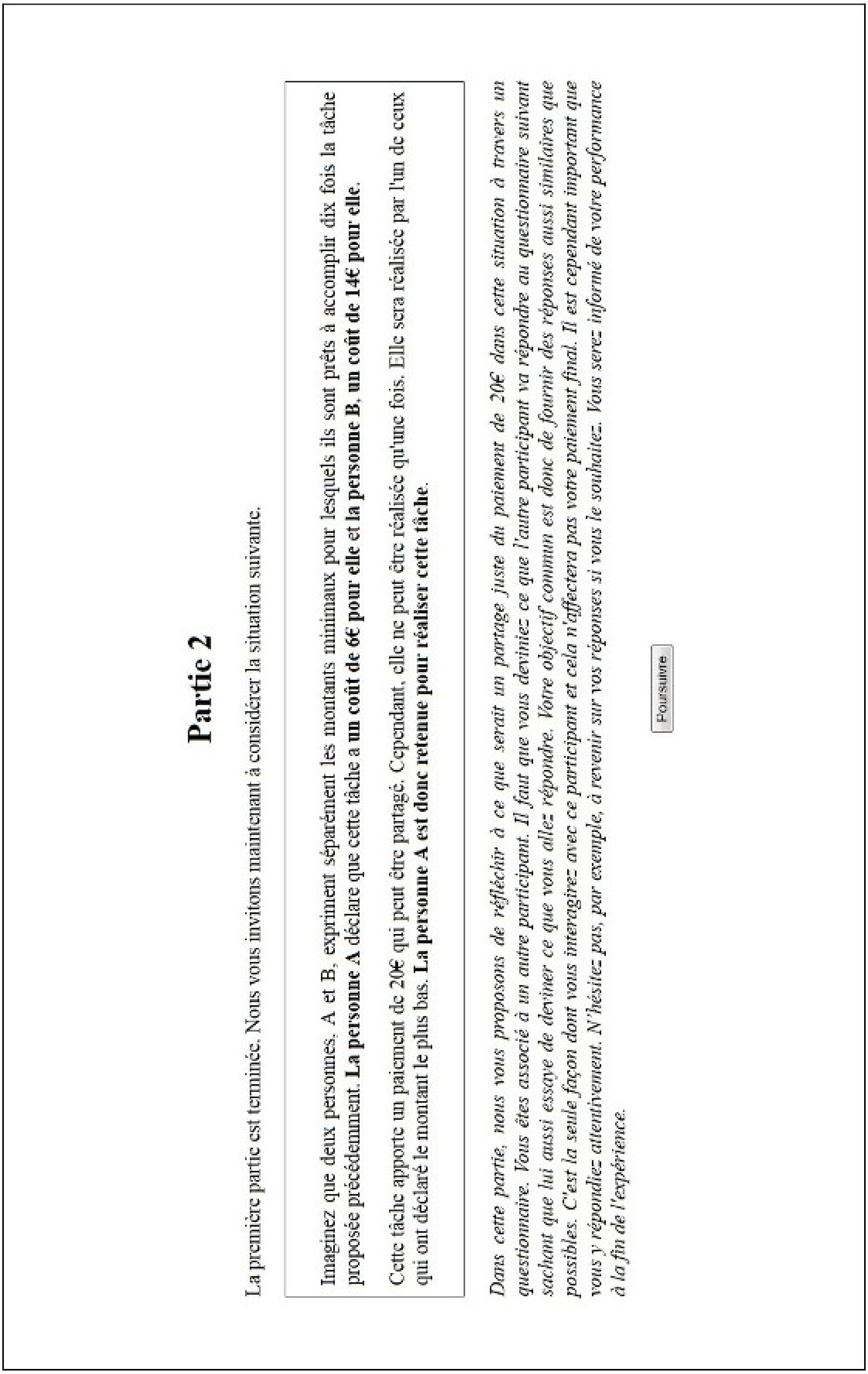
Votre réponse : Le coût que représente pour moi le fait de réaliser dix fois la tâche précédente à la fin de la session est de €, c'est-à-dire que si une rémunération entre 0 et 20€ m'était proposée pour la faire, j'accepterais si la rémunération était supérieure à ce montant et je refuserais toute rémunération inférieure.

Valider mon choix

Screen 5


Partie 1
Veuillez patientez le temps que les autres participants terminent.

Screen 6



Partie 2

La première partie est terminée. Nous vous invitons maintenant à considérer la situation suivante.

Imaginez que deux personnes, A et B, expriment séparément les montants minimaux pour lesquels ils sont prêts à accomplir dix fois la tâche proposée précédemment. **La personne A** déclare que cette tâche a **un coût de 6€ pour elle et la personne B, un coût de 14€ pour elle**.

Cette tâche apporte un paiement de 20€ qui peut être partagé. Cependant, elle ne peut être réalisée qu'une fois. Elle sera réalisée par l'un de ceux qui ont déclaré le montant le plus bas. **La personne A est donc retenue pour réaliser cette tâche**.

Dans cette partie, nous vous proposons de réfléchir à ce que serait un partage juste du paiement de 20€ dans cette situation à travers un questionnaire. Vous êtes associé à un autre participant. Il faut que vous deviniez ce que l'autre participant va répondre au questionnaire suivant sachant que lui aussi essaye de deviner ce que vous allez répondre. Votre objectif commun est donc de fournir des réponses aussi similaires que possibles. C'est la seule façon dont vous interagirez avec ce participant et cela n'affectera pas votre paiement final. Il est cependant important que vous y répondiez attentivement. N'hésitez pas, par exemple, à revenir sur vos réponses si vous le souhaitez. Vous serez informé de votre performance à la fin de l'expérience.

Poursuivre

Screen 7

Partie 2

Vous êtes associé à un autre participant. Il faut que deviniez ce que l'autre participant va répondre au questionnaire suivant sachant que lui aussi essaye de deviner ce que vous allez répondre. Votre objectif commun est donc de fournir des réponses aussi similaires que possibles. Vous serez informé de votre performance à la fin de l'expérience.

Le questionnaire est structuré en deux parties. L'ordre des parties et des éléments au sein des parties est aléatoire. Il n'est sans doute pas le même pour vous et lui.

Deux personnes, A et B, expriment séparément les montants minimaux pour lesquels ils sont prêts à accomplir dix fois la tâche proposée précédemment. **La personne A** déclare que cette tâche a **un coût de 6€ pour elle** et **la personne B, un coût de 14€ pour elle.**

Cette tâche apporte un paiement de 20€ qui peut être partagé. Cependant, elle ne peut être réalisée qu'une fois. Elle sera réalisée par l'un de ceux qui ont déclaré le montant le plus bas. **La personne A est donc retenue pour réaliser cette tâche**.

Cas particuliers

Chacun des cas suivants est identique à la situation décrite au départ, sauf que les coûts déclarés par les participants ne sont pas les mêmes. Dans chacun de ces cas, nous vous demandons de recommander un partage juste du paiement de 20€.

Avec un autre participant, votre objectif commun est de fournir les mêmes réponses.

Cas 1 : la personne A déclare que la tâche représente pour elle un coût de 6€ et B, un coût de 14€. La personne A est retenue pour réaliser la tâche.

Part de A : € *(veuillez entrer un montant compris entre 0 et 20€)*
Part de B :

Cas 2 : la personne A déclare que la tâche représente pour elle un coût de 6€ et B, un coût de 20€. La personne A est retenue pour réaliser la tâche.

The detailed text of the screen is the following:

Partie 2

Vous êtes associé à un autre participant. Il faut que deviniez ce que l'autre participant va répondre au questionnaire suivant sachant que lui aussi essaye de deviner ce que vous allez répondre. Votre objectif commun est donc de fournir des réponses aussi similaires que possibles. Vous serez informé de votre performance à la fin de l'expérience.

Le questionnaire est structuré en deux parties. L'ordre des parties et des éléments au sein des parties est aléatoire. Il n'est sans doute pas le même pour vous et lui.

Deux personnes, A et B, expriment séparément les montants minimaux pour lesquels ils sont prêts à accomplir dix fois la tâche proposée précédemment. La personne A déclare que cette tâche a un coût de 6€ pour elle et la personne B, un coût de 14€ pour elle. Cette tâche apporte un paiement de 20€ qui peut être partagé. Cependant, elle ne peut être réalisée qu'une fois. Elle sera réalisée par l'un de ceux qui ont déclaré le montant le plus bas. La personne A est donc retenue pour réaliser cette tâche.

Cas particuliers

Chacun des cas suivants est identique à la situation décrite au départ, sauf que les coûts déclarés par les participants ne sont pas les mêmes. Dans chacun de ces cas, nous vous demandons de recommander un partage juste du paiement de 20€. Avec un autre participant, votre objectif commun est de fournir les mêmes réponses.

Cas 1 : la personne A déclare que la tâche représente pour elle un coût de 6€ et B, un coût de 14€. La personne A est retenue pour réaliser la tâche.

- Part de A : € (veuillez entrer un montant compris entre 0 et 20€)
- Part de B :

Cas 2 : la personne A déclare que la tâche représente pour elle un coût de 6€ et B, un coût de 20€. La personne A est retenue pour réaliser la tâche.

- Part de A : € (veuillez entrer un montant compris entre 0 et 20€)
- Part de B :

Cas 3 : la personne A déclare que la tâche représente pour elle un coût de 14€ et B, un coût de 14€. La personne A est retenue pour réaliser la tâche.

- Part de A : € (veuillez entrer un montant compris entre 0 et 20€)
- Part de B :

Cas 4 : la personne A déclare que la tâche représente pour elle un coût de 6€ et B, un coût de 6€. La personne A est retenue pour réaliser la tâche.

- Part de A : € (veuillez entrer un montant compris entre 0 et 20€)
- Part de B :

Propositions

Vous êtes invité à considérer quatre propositions de partage du paiement de 20€. Ces propositions sont illustrées dans la situation présentée en début de page. Les propositions de partage suivantes peuvent-elles être considérées comme justes ? Avec un autre participant, votre objectif commun est de fournir les mêmes réponses.

Proposition 1 : "Les deux participants devraient recevoir la même somme après avoir compensé la personne qui réalise la tâche pour son effort. La personne qui réalise la tâche devrait donc être compensée à hauteur de son propre coût, ici 6€, et le reste, ici 20€- 6€=14€, devrait être partagé en parts égales. Dans le cas présenté, A devrait donc recevoir 13€ et B, 7€."

- Juste
- Injuste

Proposition 2 : "Le paiement de la tâche devrait revenir à la personne qui la réalise. Dans le cas présenté, A devrait donc recevoir 20€ et B, 0€. "

- Juste
- Injuste

Proposition 3 : "Les deux participants devraient recevoir la même somme. Dans le cas présenté, A devrait donc recevoir 10€ et B, 10€. "

- Juste
- Injuste

Proposition 4 : "Le paiement de la tâche devrait revenir à la personne qui la réalise après avoir compensé la personne qui ne la réalise pas pour le fait d'être privé de cette opportunité. Cette dernière devrait recevoir la moitié de la différence entre le paiement de 20€ et son propre coût, ici (20€-14€)/2=3€, et la personne qui réalise la tâche devrait recevoir le reste, ici 20€-3€=17€. Dans le cas présenté, A devrait donc recevoir 17€ et B, 3€."

- Juste
- Injuste

Laquelle de ces propositions peut être considérée comme la plus juste ?

- La proposition 1
- La proposition 2
- La proposition 3
- La proposition 4

Screen 8

Partie 3

La partie 2 est terminée. Dans cette partie, la situation est la suivante.

La tâche finale consiste à réaliser dix fois la tâche que vous avez réalisée au début de la session. Elle est rémunérée à hauteur de 20€. Vous avez déclaré être prêt à la réaliser pourvu qu'elle soit rémunérée à hauteur d'au moins 9€. Un autre participant a déclaré être prêt à la réaliser pourvu qu'elle soit rémunérée à hauteur d'au moins 16€. Cette tâche ne peut être réalisée qu'une fois. Il sera demandé au participant pour qui le coût de cette tâche est le plus bas de la réaliser à la fin de la session. Il s'agit de vous

Vous êtes tout les deux assurés d'obtenir 5€ du fait de votre venue. Cette partie vise à déterminer le partage du paiement additionnel de 20€ associé à la tâche.

Poursuivre

Screen 9

Partie 3

Nous vous rappelons la situation :

- La tâche finale est rémunérée à hauteur de 20€.
- Le coût que réprésente pour vous le fait de réaliser cette tâche est de 9€
- Le coût que réprésente pour l'autre participant le fait de réaliser cette tâche est de 16€
- La tâche finale sera réalisée par vous.

Si l'autre participant devait choisir comment partager le paiement de 20€, nous vous demandons ici quel est le montant minimal que vous devriez obtenir pour que vous jugiez son choix défendable.

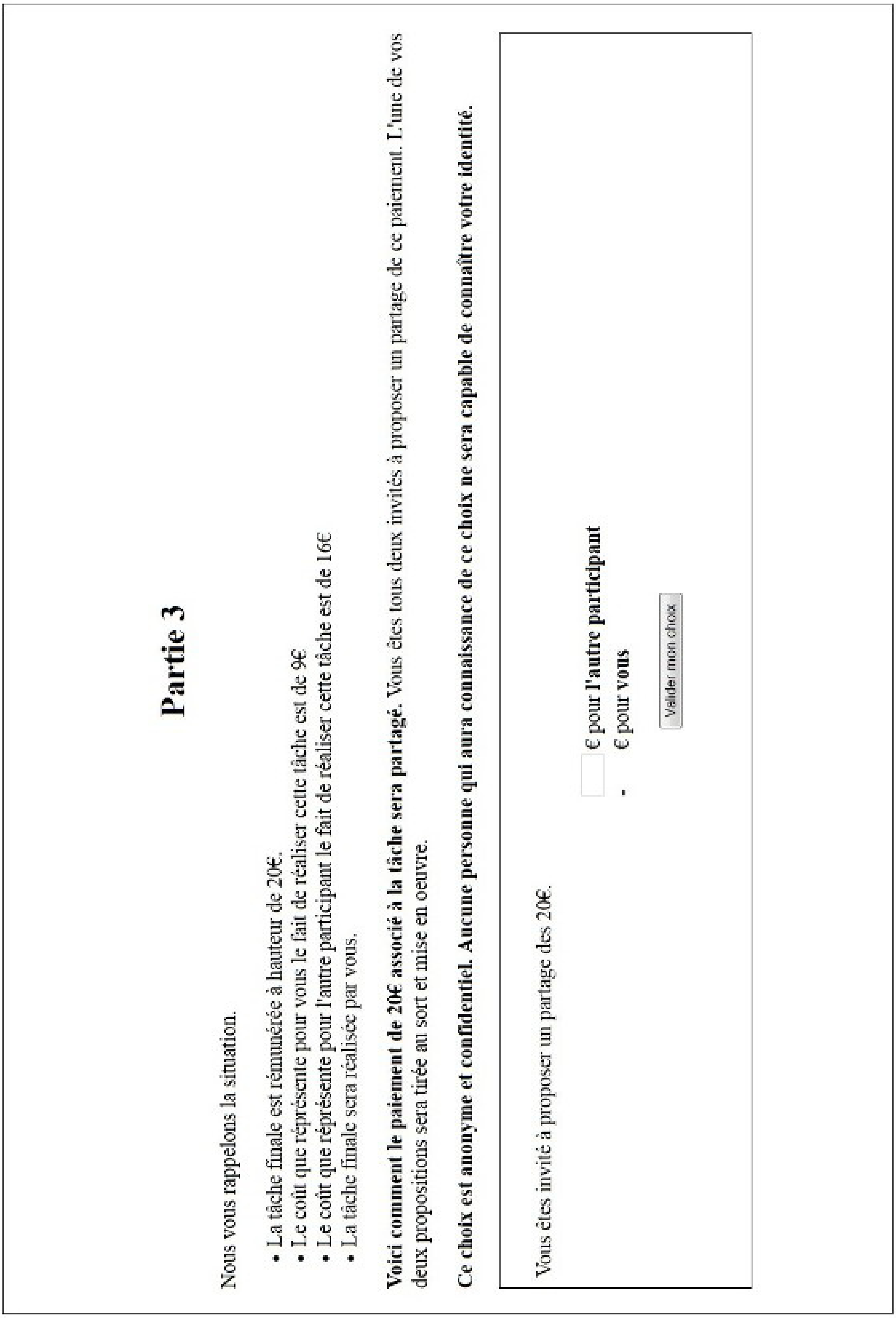
Votre réponse à cette question ne sera pas communiquée à l'autre participant. Elle n'aura, en aucune manière, d'impact sur le déroulement de la suite ni sur votre paiement final.

Si l'autre participant devait choisir comment partager le paiement de 20€, quel est le montant que vous devriez obtenir, au minimum, pour que vous jugiez son choix défendable ?

Votre réponse : Je juge qu'un choix de la part de l'autre participant est défendable si j'obtiens au moins ____ €, et lui, pas plus de - €.

Valider mon choix

Screen 10


Partie 3

Nous vous rappelons la situation.

- La tâche finale est rémunérée à hauteur de 20€.
- Le coût que représente pour vous le fait de réaliser cette tâche est de 9€
- Le coût que représente pour l'autre participant le fait de réaliser cette tâche est de 16€
- La tâche finale sera réalisée par vous.

Voici comment le paiement de 20€ associé à la tâche sera partagé. Vous êtes tous deux invités à proposer un partage de ce paiement. L'une de vos deux propositions sera tirée au sort et mise en oeuvre.

Ce choix est anonyme et confidentiel. Aucune personne qui aura connaissance de ce choix ne sera capable de connaître votre identité.

Vous êtes invité à proposer un partage des 20€.

€ pour **l'autre participant**

- € pour **vous**

Valider mon choix

Screen 11

Partie 3

Nous vous invitons à considérer la situation de deux autres participants.
Attention, cette situation ne correspond plus à votre situation.

- La tâche finale est rémunérée à hauteur de 20€.
- Le coût que réprésente pour **le premier participant** le fait de réaliser cette tâche est de 10€
- Le coût que réprésente pour **le second participant** le fait de réaliser cette tâche est de 3€
- La tâche finale sera réalisée par le second participant.

Ces participants ont répondu aux mêmes questions que vous précédemment. En particulier, **le premier participant**, celui qui ne réalisera pas la tâche, a déclaré l'offre minimum qu'il jugerait défendable de la part du second participant. Dans cette partie, nous vous invitons à deviner sa réponse.

Vous gagnerez 10€ supplémentaires si vous devinez sa réponse.

Quel est, selon vous, l'offre minimale que le premier participant juge défendable ?

Votre réponse : Je pense que le premier participant juge qu'un choix de la part du second est défendable si lui (le premier participant) obtient au moins ____ €, et le second participant, pas plus de - €.

Valider mon choix

Screen 12

Questionnaire final

Vos réponses aux 6 questions suivantes sont anonymes et confidentielles. Merci de bien vouloir y répondre avec attention. Il vous sera ensuite demandé de réaliser la tâche. Suite à cela, l'expérience sera terminée.

Question 1

L'autre participant a déclaré que le coût que représente pour lui le fait de réaliser dix fois la tâche précédente à la fin de la session est de 16€.
Que pensez-vous de ce montant ?

- Il est sans doute involontairement sous-évalué
- Il est sans doute volontairement sur-évalué
- Il est sans doute involontairement sur-évalué
- Il est sans doute volontairement sous-évalué
- Il reflète de manière crédible le coût de faire la tâche pour cette personne
- Aucune de ces propositions

Pouvez-vous nous préciser les raisons de votre réponse à cette question ?

Question 2

Nous vous invitons à considérer la situation suivante. La tâche consiste à réaliser dix fois la tâche que vous avez réalisée au début de la session. Imaginez que deux personnes, A et B, expriment séparément les montants pour lesquels ils sont prêts à accomplir cette tâche. **La personne A** déclare que la tâche a **un coût de 6€ pour elle** et **la personne B**, **un coût de 14€ pour elle**.

The detailed questionnaire is the following:

Questionnaire final

Vos réponses aux 6 questions suivantes sont anonymes et confidentielles. Merci de bien vouloir y répondre avec attention. Il vous sera ensuite demandé de réaliser la tâche. Suite à cela, l'expérience sera terminée.

Question 1

L'autre participant a déclaré que le coût que représente pour lui le fait de réaliser dix fois la tâche précédente à la fin de la session est de 16€. Que pensez-vous de ce montant ?

- Il est sans doute involontairement sous-évalué
- Il est sans doute volontairement sur-évalué
- Il est sans doute involontairement sur-évalué
- Il est sans doute volontairement sous-évalué
- Il reflète de manière crédible le coût de faire la tâche pour cette personne
- Aucune de ces propositions

Pouvez-vous nous préciser les raisons de votre réponse à cette question ?

Question 2

Nous vous invitons à considérer la situation suivante. La tâche consiste à réaliser dix fois la tâche que vous avez réalisée au début de la session. Imaginez que deux personnes, A et B, expriment séparément les montants pour lesquels ils sont prêts à accomplir cette tâche. La personne A déclare que la tâche a un coût de 6€ pour elle et la personne B, un coût de 14€ pour elle. Cette tâche apporte un paiement de 20€ qui peut être partagé. Cependant, elle ne peut être réalisée qu'une fois. Il est demandé à la personne pour qui le coût de la tâche est le plus bas, c'est-à-dire A, de la réaliser. Dans cette situation, quel partage du paiement de 20€ vous semble le plus juste ?

- € pour la personne A, qui réalisera cette tâche
- - € pour la personne B, qui ne réalisera pas cette tâche

Pouvez-vous nous préciser les raisons de votre réponse à cette question ?

Question 3

Vous êtes :

- un homme
- une femme

Question 4

Quelle est votre année de naissance ?

Question 5

Quel est votre domaine d'études ?

- Santé et Sport
- Sciences (Physiques, Chimie, etc.)
- Ingénieurie
- Art, Littérature et Sciences Humaines
- Droit, Sciences de Gestion et Sciences Politiques
- Economie
- Autres

Question 6

En politique, les gens parlent parfois de droite et de gauche. Où vous situriez-vous sur une échelle de 1 à 10, où 1 signifierait la gauche et 10 la droite ?

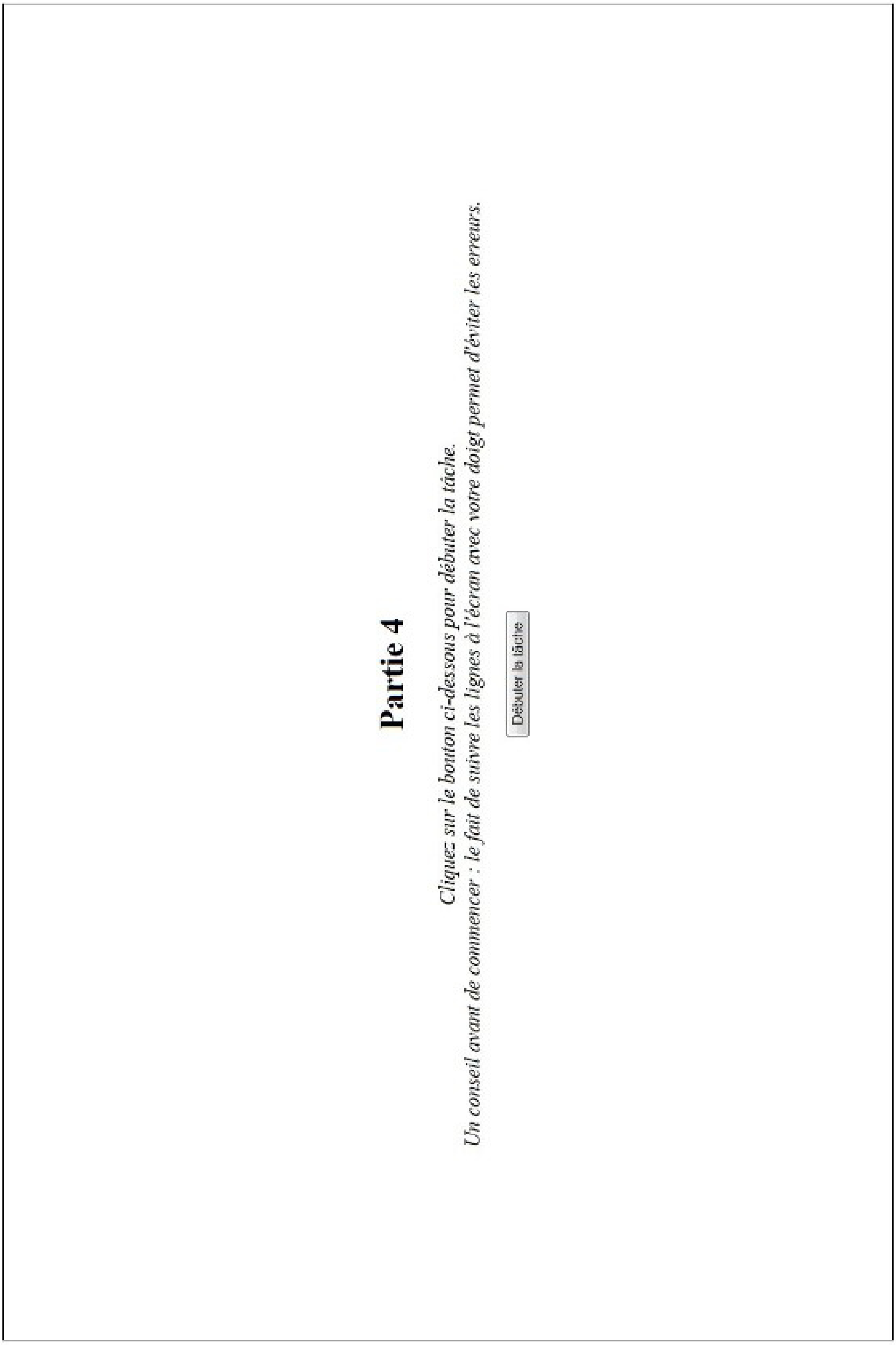
Gauche	1	2	3	4	5	6	7	8	9	10	Droite

- Ne souhaite pas répondre

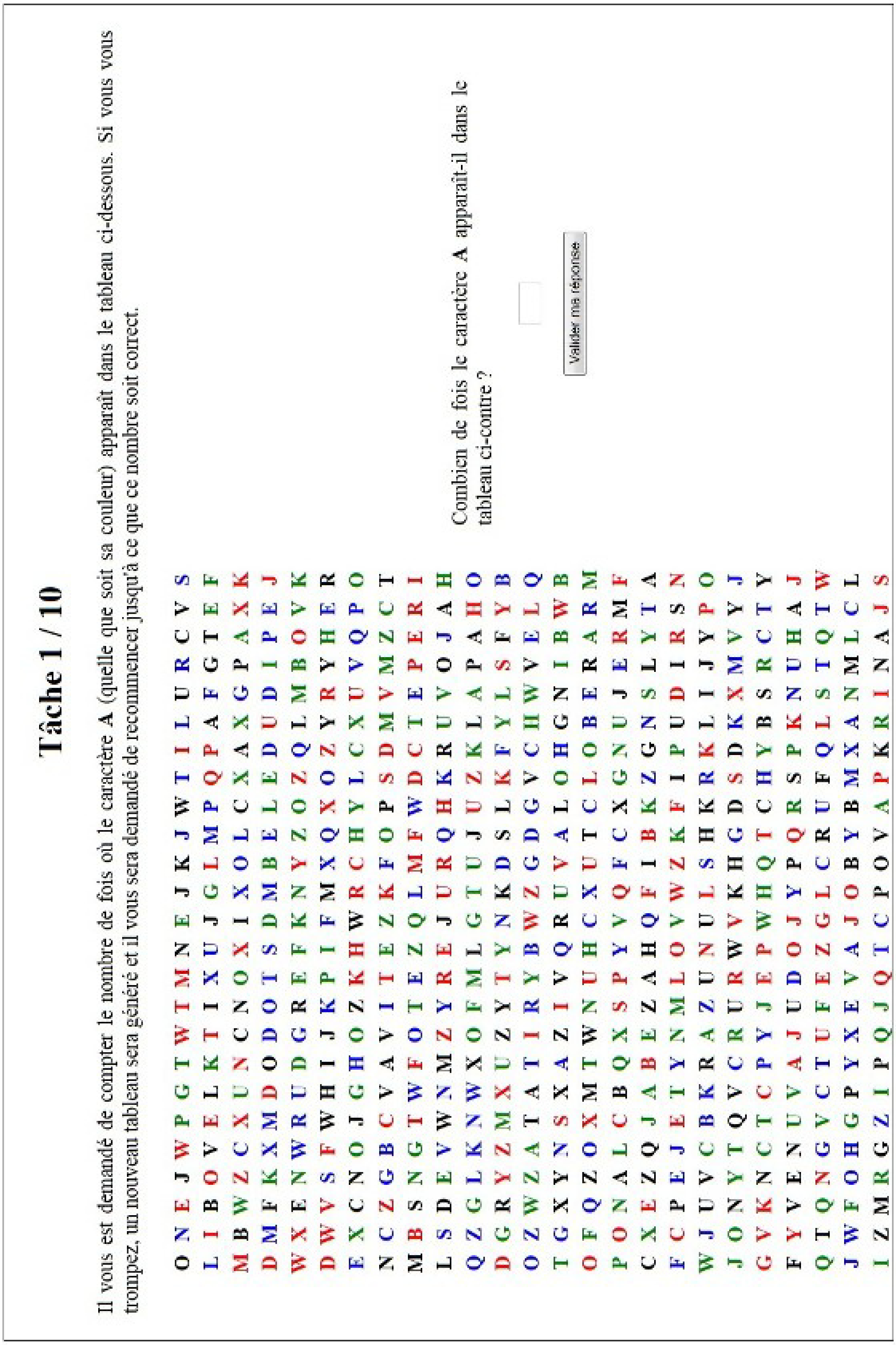
Si vous avez des commentaires à ajouter concernant l'expérience, vous pouvez le faire dans l'espace ci-dessous.

Vous pouvez cliquer sur le bouton suivant pour poursuivre.

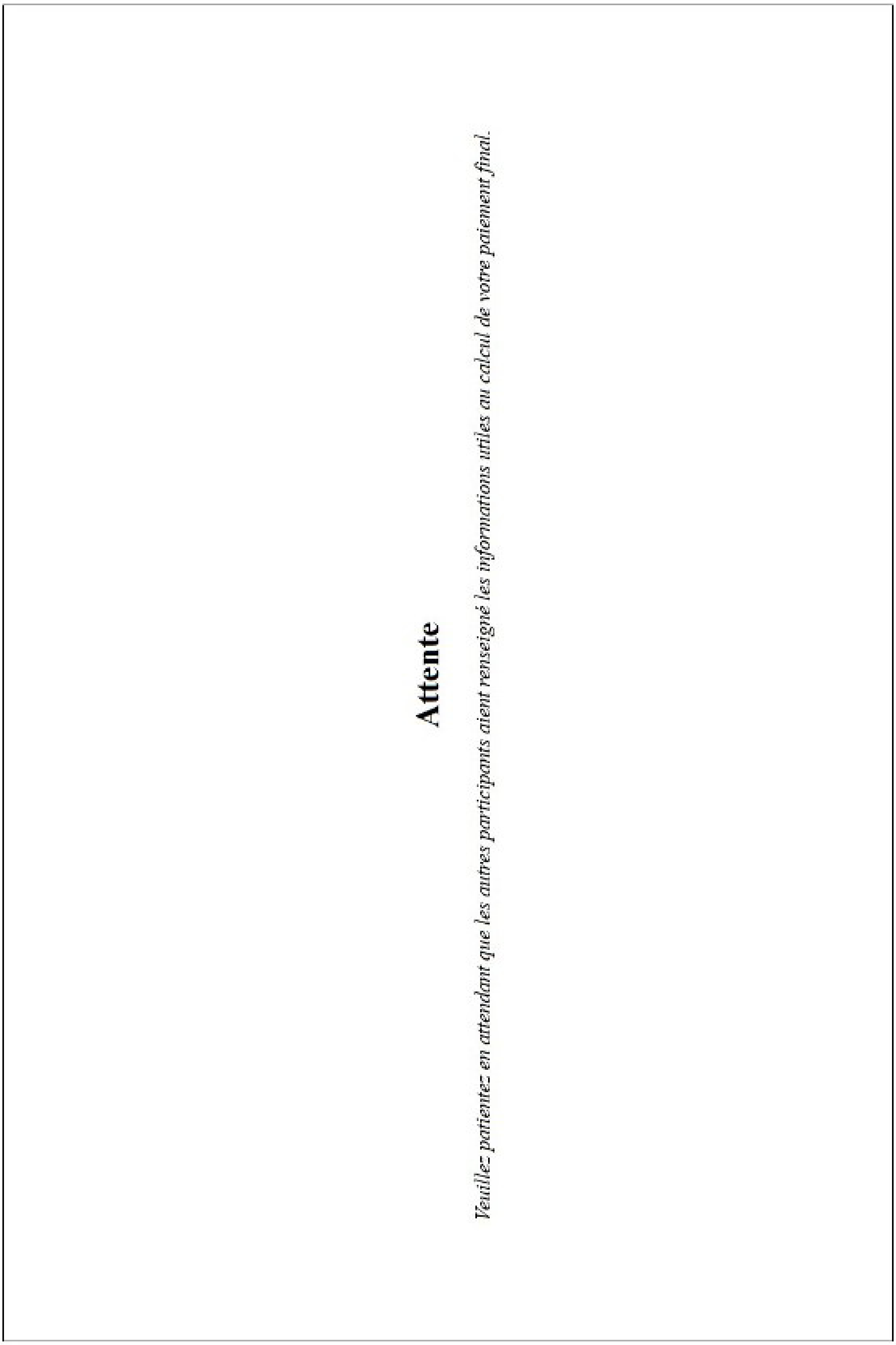
Screen 13


Partie 4
Cliquez sur le bouton ci-dessous pour débuter la tâche.
Un conseil avant de commencer : le fait de suivre les lignes à l'écran avec votre doigt permet d'éviter les erreurs.
Débuter la tâche

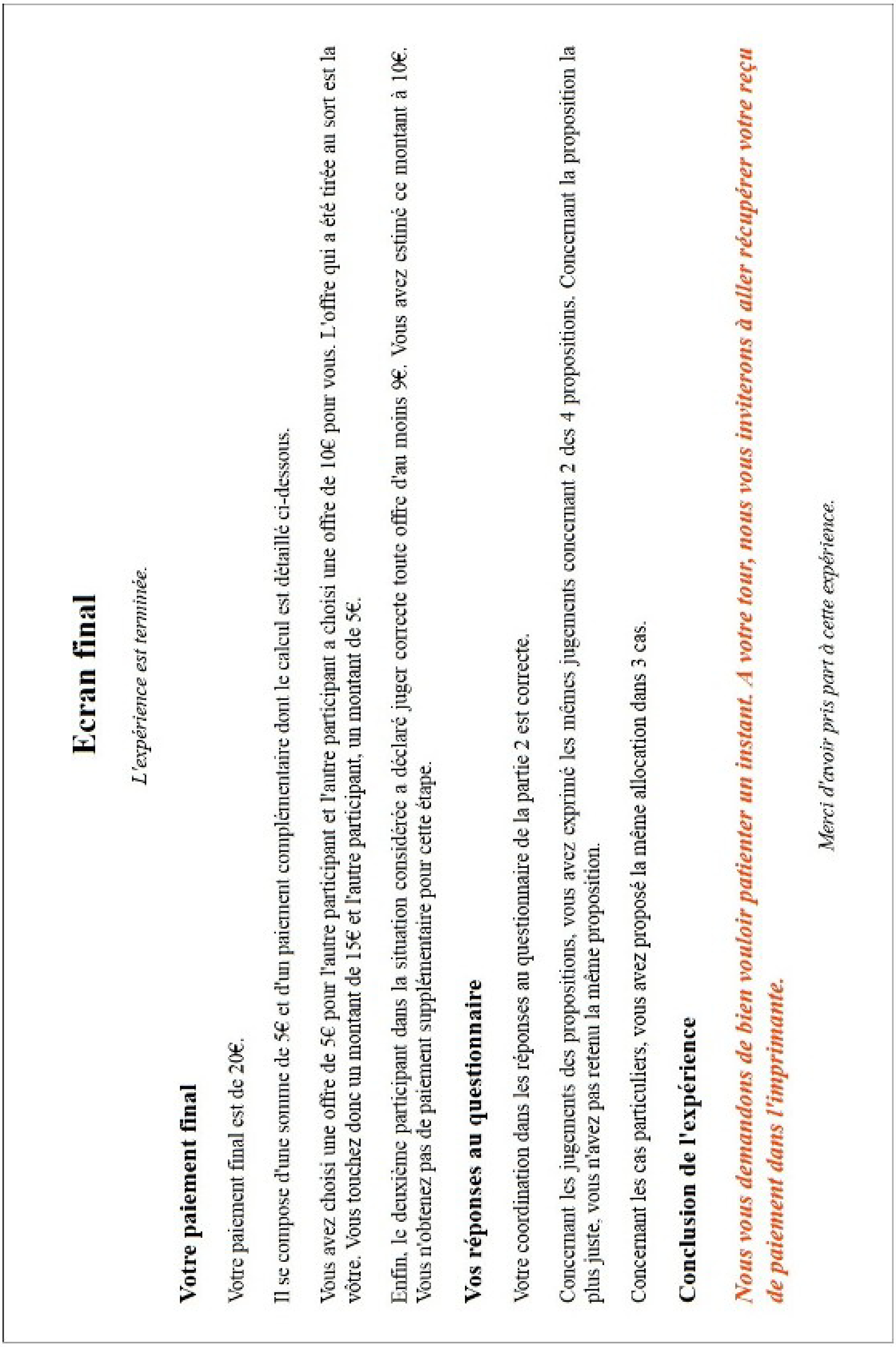
Screen 14


Tâche 1 / 10
Il vous est demandé de compter le nombre de fois où le caractère A (quelle que soit sa couleur) apparaît dans le tableau ci-dessous. Si vous vous trompez, un nouveau tableau sera généré et il vous sera demandé de recommencer jusqu'à ce que ce nombre soit correct.
Combien de fois le caractère A apparaît-il dans le tableau ci-contre ?
Valider ma réponse

Screen 15


Attente
Veuillez patientez en attendant que les autres participants aient renseigné les informations utiles au calcul de votre paiement final.

Screen 16



Ecran final

L'expérience est terminée.

Votre paiement final

Votre paiement final est de 20€.

Il se compose d'une somme de 5€ et d'un paiement complémentaire dont le calcul est détaillé ci-dessous.

Vous avez choisi une offre de 5€ pour l'autre participant et l'autre participant a choisi une offre de 10€ pour vous. L'offre qui a été tirée au sort est la vôtre. Vous touchez donc un montant de 15€ et l'autre participant, un montant de 5€.

Enfin, le deuxième participant dans la situation considérée a déclaré juger correcte toute offre d'au moins 9€. Vous avez estimé ce montant à 10€. Vous n'obtenez pas de paiement supplémentaire pour cette étape.

Vos réponses au questionnaire

Votre coordination dans les réponses au questionnaire de la partie 2 est correcte.

Concernant les jugements des propositions, vous avez exprimé les mêmes jugements concernant 2 des 4 propositions. Concernant la proposition la plus juste, vous n'avez pas retenu la même proposition.

Concernant les cas particuliers, vous avez proposé la même allocation dans 3 cas.

Conclusion de l'expérience

Nous vous demandons de bien vouloir patienter un instant. A votre tour, nous vous inviterons à aller récupérer votre reçu de paiement dans l'imprimante.

Merci d'avoir pris part à cette expérience.

5.D Matching process

We run 6 sessions in which pairs are formed and allocated into 3 treatments. Our aim is to obtain a fair diversity of pairs within each treatment but a similar distribution of the pairs between the treatments. Given the small size of the sample, a random process can be improved by directing allocation of pairs into treatments in order to keep the distribution of pairs comparable between the 3 treatments.

The problem is the following, for each session $s \in \{1, ..., 6\}$, we get a list of compensation requirements $\boldsymbol{c} = (c_1, ..., c_{k_s})$, where k_s is the number of subjects in session s (an even number below 18). We have to form $\frac{k_s}{2}$ pairs and allocate them into the 3 treatments sequentially. We intend to form a reasonable diversity of pairs and achieve a comparable distribution of the pairs within each treatment in a reasonable execution time.

Protocol

The matching protocol will consist in two stages, performed for each session:

- A protocol to allocate participants in the BDM mechanism
- A protocol to form pairs
- A protocol to allocate pairs into the treatments

For a pair of pairs $t = \{\{c_1^1, c_2^1\}, \{c_1^2, c_2^2\}\}$ the spread $S(t)$ between the pairs is defined as follows:

$$S(t) = \max(|c_1^2 - c_1^1|, |c_2^2 - c_2^1|, ||c_2^1 - c_1^1| - |c_2^2 - c_1^2||)$$

The spread of a triplet is defined as the maximal spread over all pairs of its elements. The scheme of the algorithm is the following: for each session rank the compensation requirements in increasing order and randomly match participants in the lower half with a participant in the upper half. Then compose triplets of all existing pairs in order to minimize the average spread of the triplets. Within each triplet, randomly allocate pairs into the treatments that are not already allocated. The minimization of the average spread over triplets is a NP-hard problem. It cannot be optimized to be performed in polynomial time. Even with a dozen of pairs the exploration of all possible matchings into treatments could potentially require intensive computation. Given that the matching is performed during the sessions, we want it to be executed fast. Then, we propose a greedy algorithm, less computationally intensive. This approach does not lead to the global optimum but would still reach satisfactory solutions.

- Pair formation: at each stage,

  1. Rank the compensation requirements in increasing order

2. Randomly match a subject in the lower half with a subject in the upper half

- Allocations of pairs into treatments: at each stage,

1. Find a matching among the newly generated pairs that minimizes the average spread (full optimization).
2. Sequentially explore and implement permutations between the new pairs and pairs in existing triplets that strictly reduces the average spread until no such opportunity exists.
3. Allocate treatment randomly with the constraint not to repeat already existing treatments within each matching

Chapter 6

Cooperative decision-making for the provision of a locally undesirable facility

This chapter is a joint work with Stefan Ambec.

It is accepted for publication in Social Choice and Welfare[1]

[1]DOI : 10.1007/s00355-015-0907-2

Abstract

We consider the decentralised provision of a global public good with local externalities in a spatially explicit model. Communities decide on the location of a facility that benefits everyone but exhibits costs to the host and its neighbors. They share the costs through transfers. We examine cooperative games associated with this so-called NIMBY (“Not In My Back-Yard") problem. We derive and discuss conditions for core solutions to exist. These conditions are driven by the temptation to exclude groups of neighbors at any potential location. We illustrate the results in different spatial settings. These results clarify how property rights can affect cooperation and shed further light on a limitation of the Coase theorem.

Keywords: NIMBY, externality, Coase theorem, pollution, waste, core, cooperative game, spatial model.

JEL codes: C71, D62, Q53, R53.

Ackowledgments: We wish to thank Olof Johansson-Stenman, Justin Leroux and Vera Zaporozhets for helpful comments. We are especially grateful to Michel Le Breton for his advice and suggestions. This paper also benefited from comments by participants at the 2013 conference of the European Association of Environmental and Resource Economics in Toulouse, the 2013 Public Economic Theory conference in Lisbon and the 2013 Spanish-Italian-Netherland Game theory conference in Vigo. We acknowledge financial support from ANR (France) through the project ANR-12-BSH1-0003-01 on “Political Economy of the Environment".

Introduction

The production of activities that are harmful to society is the source of a famous controversy between Arthur Pigou and Ronald Coase. In such a context, Pigou recommended that the producers pay the harmful damages inflicted on third parties. Using the example of the sparks from railway engines that set fire to woods surrounding the tracks, Pigou (1920) argued that the railways should be forced to compensate those whose woods are burnt. In the same example, Coase (1960) challenged the Pigouvian solution. He argued that the parties involved could resolve the problem themselves in the absence of transaction costs, provided the property rights on harmful externalities (or liability rules) were assigned to one of them. Such Coasean bargaining would lead to efficiency, regardless of the allocation of property rights. This is known as the "Coase Theorem".

The Coase theorem was subsequently invalidated in cooperative settings involving more than two players. A famous instance is a version of the Shapley and Shubik *garbage game* (1969), in which three neighbors decide on where to dump their garbage. For instance, let the disutility of having waste in one's backyard be -1 for one bag of garbage and -2 both for 2 and 3 bags of garbage. When utility is transferable, the efficient outcome is for the three agents to cooperate and to locate the garbage in the backyard of one of them. The total disutility so achieved is -2. Coasean bargaining predicts that, in the absence of transaction costs, they will exchange garbage and money to reach such a socially optimal outcome. Yet, if a player has the right to dispose of garbage as she or he likes, every group of two players will prefer to dump their garbage into the third player's garden without compensating her or him. In the previous example, one can easily check that whatever way they share the total cost, there will always be a couple of players willing to withdraw and coordinate their dumping on the third player: this game has an empty "core"'.[2] More generally, Starrett (1973) pointed out that economies with nonconvexities can have an empty core. Aivazian and Callen (1981) make a similar argument: they provide an example with a polluting facility for which the core in the cooperative game representing Coasean bargaining with a specific liability rule is empty.[3] Although the above examples do show that the Coase theorem cannot always be demonstrated, they do not tell us in which circumstances the Coase theorem is likely to hold. This paper fills that gap. Using quite a general model of production activities with negative externalities, we investigate what the driving economic parameters are that determine whether the core is empty.

[2]The total disutility that such a two-player coalition can guarantee to itself is at least -1: both members drop their garbage on the third player but may still get his or her garbage. Additionally, the disutility of the third agent, is -2, hence the total disutility is -3: social efficiency is not achieved. Hence, players may not be able to reach an efficient outcome.

[3]The argument is reproduced by Stearns (1993) with voting instead of bargaining as a collective decision process. In his example, a Condorcet cycle arises in a situation where three communities have to collectively decide where to site a nuclear waste repository.

To do so, we deal with a spatial model with externalities. It represents the problem of providing a locally undesirable but globally desirable facility. In Pigou's story on sparks from railway engines, all citizens connected to the railways benefit from it, but those who own woods along the track might suffer from the external cost. This is the well-known *Not In My Backyard* (NIMBY) problem. Examples include waste treatment plants, nuclear or coal power utilities, windmills, airports, or prisons.[4] Such facilities may sometimes be acknowledged to be socially beneficial in the sense of the Hicks-Kaldor criterion: the social benefits more than offset the social costs. Yet their provision could still face strong opposition from neighboring citizens who suffer from negative externalities such as air or water pollution, noise, or amenity losses.[5] This is why the localization of the facility is a sensitive issue. Some form of compensation may be offered to make it acceptable to the neighboring victims of external costs. In this paper we examine the feasibility of the decentralized provision of such facilities.

The NIMBY problem is first studied in its general form. Several communities plan to build a facility. The benefits from using the facility are excludable and non-rival. The costs are incurred by the host and its neighbors. The communities both agree on location and transfers: they decide on who is going to host the facility and how much the host and its neighbors must be compensated. Coalitions of communities block the agreement if they are better-off building and sharing their own facility (or not building at all). The outside option of coalitions defines a cooperative game associated with the NIMBY problem. Interestingly, the cooperative game exhibits specific properties. It is a cooperative game with externalities in the sense that the welfare that a group of communities can enjoy depends on the cooperative behavior of communities outside the group as well as the localization of the facility they build on their own. Yet externalities in the associated cooperative game can be negative or positive: a group of communities can benefit or suffer from the cooperation of others. Nevertheless, the best that can happen for a coalition of communities is that the other communities are not cooperating. We first define the value function of our cooperative game accordingly: a group of communities do not expect that the others will build their own facility when they oppose an agreement. We thus give maximal incentives for coalitions to deviate and block an agreement. Under some assumptions, we show that only two forces constrain the core: individual rationality and the exclusion of individual communities and communities in the neighborhood of any potential host of the facility. Individual rationality makes sure that all communities benefit from the facility. The motive for exclusion is similar to that in Shapley and Shubik's garbage game: communities are tempted to exclude those who suffer from the

[4]Some of these projects feature non-excludability of the benefits at the origin of free-riding behaviors; others not. We will emphasize here the garbage game dimension of such problems, which is common to all.

[5]Richman and Boerner (2006) define a NIMBY as follows "a socially desirable land use that broadly distributes benefits, yet is difficult or impossible to implement because of local opposition".

negative externality, to avoid compensating them. This restricts considerably the set of inequalities defining the core.

We next propose an original index for testing whether the core is empty or not. The core is non-empty if and only if this index is lower than one. We discuss some comparative statics. We show that the core is less likely to be empty when the costs which cannot be externalized increase for some communities, or when costs increase somewhere but at an optimal location. In particular, this means that the more harmful would the project be when located elsewhere than at the optimal location, the easier it is for the communities to reach an agreement. We also show that the problem is exacerbated when the number of communities increases in the linear case. However, the effect of the number of communities is ambiguous in general. Next, we generalize our results for other notions of the core. We finally provide illustrations on different spatial structures, first in the linear case and on simple graphs. Lastly, we compute the index for a French administrative unit to illustrate its potential applicability to real-world NIMBY problems.

Related litterature

Most of the theoretical papers in economics on the NIMBY problem rely on a mechanism design approach. A central planner designs a mechanism such as an auction to locate the undesirable facility optimally and to share its cost (O'Sullivan, 1993, Minehart and Neeman, 2002, Perez-Castrillo and Wettstein, 2002, Laurent-Lucchetti and Leroux, 2011). The central planner can impose the mechanism on the communities but does not know their costs. The implemented solution does not guarantee that some communities could not do better by providing the facility by themselves. In contrast, we assume that the costs associated with the facility are common knowledge and adopt a cooperative approach. Decisions are decentralized to communities that collectively negotiate and can make binding agreements about localization and compensations.

In a cooperative framework, Laurent-Lucchetti and Leroux (2010), Sakai (2012) and Dehez (2013) have analyzed core solutions of cooperative games associated with NIMBY problems.[6] They all implicitly rely on the assumption that externalities are concentrated within a jurisdiction. In practice, pollution (e.g. air or water pollution, risk of radioactive contamination) spreads out quite widely compared to the size of the communities (e.g. municipalities, countries). To the best of our knowledge, our paper is the first to explicitly introduce spatial externalities in a cooperative framework representing the NIMBY problem. It emphasizes the difficulties that arise when the costs are spread over more than a single community.

[6] Lejano and Davos (2001) also consider coalition formation in the NIMBY problem. In a numerical example, they argue that a compensation scheme that leaves the host indifferent may fail to be a core allocation.

Our paper is also related to the literature on public good provision which emphasizes the free-riding problem: users can benefit from public goods without contributing to their cost (Bergstrom and al., 1986). Free-riding arises when people cannot be excluded from consuming the good. We avoid free-riding by assuming that communities can be excluded from accessing the facility at no cost. However, the potential exclusion of the neighbors of any potential host can still compromise cooperation. Both the NIMBY and public good provision cooperative games are games with externalities. This raises interesting conceptual issues for the definition of the core and the representation of the game in partition form (Bloch and van den Nouweland, 2014). Such issues are discussed in Section 6.2.2.

Finally, our approach is also related to a literature on the core of cooperative facility location games. In Goemans and Skutella (2004), consumers have heterogeneous costs of being connected to the facility and differ on the benefit they enjoy using the facility, depending on its location. They provide conditions for the core to be non-empty. In Le Breton and Weber (2003), both the users and the facility are located along a line. The benefit of using the facility is proportional to the distance between the user and the facility called "transportation" cost. In this model, preferences are single-peaked, in the sense that the closer the facility, the better for the user. This hypothesis plays a crucial role for the existence of core allocations. In contrast, preferences can be single-dipped in the linear representation of our model. As a consequence, non-emptiness of the core is no longer guaranteed.[7]

The rest of the paper is organized as follows. Section 6.1 introduces the general NIMBY problem with excludable benefits. Section 6.2 presents the main results. We first deal with a case without cooperative externalities in Subsection 6.2.1 and we discuss the robustness of the results presented when cooperative externalities are considered in Subsection 6.2.2. Finally, Section 6.3 provides illustrations of the main results on explicit spatial structures.

6.1 The NIMBY problem

A set $N = \{1, ..., n\}$ ($n > 1$) of communities or agents (land owners, municipalities, cities, regions, countries, etc.) might decide to launch a facility such as a waste treatment plant, a utility (nuclear or coal power plant) or a polluting factory. Each community $i \in N$

[7] Barberà et al. (2012) and Manjunath (2014) have examined single-dipped preferences for the location of an indivisible bad. They deal with non-transferable utility (no money involved) whereas we assume transferable utility: players can transfer part of their welfare through side-payments. Their focus is on the localization of the public bad with strategy-proof rules. In contrast, we abstract for information problems so that the public bad can easily be efficiently located. In our setting, localization impacts the value that a deviating coalition can achieve. It thus determines the distribution of the welfare through side-payments.

enjoys an individual benefit $b_i \geq 0$ from using the facility. Benefits are non-rival and non-cumulative: once a community has access to a facility, it does not enjoy any additional benefit from accessing a second one. Yet the communities that launch a facility can exclude the others from using it. A facility also generates local nuisances to the host and its immediate neighbors (pollution, risk of accident or contamination, etc.). This is summarized by the cost matrix $\boldsymbol{C} = (c_{ij})_{(i,j)\in N^2}$ where $c_{ij} \geq 0$ denotes the cost incurred by community j from a facility located at i. In summary, a facility features non-rival and excludable benefits and rival costs: it is both a club good and a private bad. For every $i \in N$, c_{ii} will be called the host cost and c_{ij} $(i \neq j)$ an external cost. Community j is a neighbor of community i if and only if $c_{ij} > 0$. The matrix $\boldsymbol{C}$ provides a spatial representation of the problem. A NIMBY problem is defined as a triplet $(N, \boldsymbol{b}, \boldsymbol{C})$.

We will use some further notations. We denote by $h \in \arg\min_{i\in N} \sum_{j\in N} c_{ij}$ an optimal location in N. There may be several optimal locations or hosts h for a given problem. Let $\mathcal{H}$ denote the set of optimal locations

$$\mathcal{H} = \arg\min_{i\in N} \sum_{j\in N} c_{ij}$$

We denote by $\bar{\mathcal{N}}(i) = \{j \in N | c_{ij} > 0\}$ the *neighborhood* of i including i. Similarly, $\mathring{\mathcal{N}}(i) = \{j \in N\backslash\{i\} | c_{ij} > 0\}$ denotes the *strict neighborhood* of i. The set of the subsets of all strict neighborhoods $\mathring{\mathcal{N}}$ is defined by

$$\mathring{\mathcal{N}} = \{S \subseteq N | \exists i \in N, S \subseteq \mathring{\mathcal{N}}(i)\}$$

Finally, we define

$$\bar{\mathcal{N}} = \mathring{\mathcal{N}} \cup \{\{i\} | i \in N\}$$

Note that $\bar{\mathcal{N}}$ does *not* denote the set of the subsets of all neighborhoods but the set of the subsets of all strict neighborhoods plus the singletons. In the case, all communities pertain to a strict neighborhood, it coincides with $\mathring{\mathcal{N}}$. This is the case in the example that follows.

Example. Uniform linear NIMBY problems

Throughout the article, we will consider a particular illustrative NIMBY problem: the uniform linear case. A NIMBY is linear if it can be represented by a line in which a link between communities represents an external cost c_{ij}. In a line, each community has two neighbors, except the ones at the two ends. If we order communities according to their location from 1 to n, it means that the external costs are $c_{jj+1} > 0$ for $j = 1$ to $n - 1$. A uniform NIMBY problem is characterized by uniform benefits and costs. The benefit per community is denoted by b so that $\boldsymbol{b} = b\boldsymbol{e}$ where $\boldsymbol{e} = (1, ..., 1)$. The host cost is c and the external cost is δc for the neighbors of the host, where δ is a positive parameter reflecting the proportion of the host's cost that spreads to the neighboring communities with $0 \leq \delta$.[8] Uniform linear NIMBY problems are fully characterized by parameters (n, b, c, δ).

Figure 6.1: Distribution of the costs at an optimal location in the uniform linear case.

The cost matrix of a linear uniform NIMBY problem is:

$$\boldsymbol{C} = \begin{pmatrix} c & \delta c & 0 & \cdots & \cdots & \cdots & 0 \\ \delta c & \ddots & \ddots & \ddots & & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & \ddots & \delta c \\ 0 & \cdots & \cdots & \cdots & 0 & \delta c & c \end{pmatrix}$$

The strict neighborhoods are $\mathring{\mathcal{N}}(1) = \{2\}$, $\mathring{\mathcal{N}}(i) = \{i-1, i+1\}$ for $i = 2, ..., n-1$ and $\mathring{\mathcal{N}}(n) = \{n-1\}$. In the uniform linear setting, it is efficient to build the facility at one end of the line. The total welfare so achieved is $nb - (1+\delta)c$. Figure 1 provides a spatial representation of a linear uniform problem when the facility is optimally located at one end of the line.

For any set of communities $S \subseteq N$, let $b(S) = \sum_S b_i$ be the total benefit enjoyed by S from running a facility. While the total benefit does not depend on the location of the facility in S, total costs do. Let us denote by $c(S)$ the *lowest* total cost that the members of S incur by building and running a facility. We have $c(S) = min_{i \in S} \sum_{j \in S} c_{ij}$. A facility should be

[8] We insist on the interpretation of δ as the proportion of a neighbor's pollution cost as compared to the host's total cost. Formally, the latter may be the sum of a technical cost c_t (construction, management, etc.) and a pollution cost c_p. If α denotes the multiplicative change in the pollution cost for the immediate neighbors, the additional cost for each of them is αc_p. We then get $\delta = \alpha \frac{c_p}{c_t + c_p}$. So δ captures the change of pollution costs with distance, as well as the share of pollution costs in the host's total costs.

built if the total benefit exceeds the total cost when located optimally. We assume that $b(N) > c(N)$: it is efficient to build a facility in the grand coalition. Obviously, since a facility is non-rival and benefits are non-cumulative, it is efficient to build only one facility used by all communities in N. The total benefit from building a facility optimally located is thus $b(N) - c(N)$. In addition, we assume $b_i < c_{ii}$ for every $i \in N$ (in the uniform linear case, this means $b < c$): it is never efficient for a community to launch a facility alone. Therefore the cooperation of at least two communities is required to make a facility advantageous. A coalition $S \subset N$ is called a *building* coalition if $b(S) \geq c(S)$ and a *non-building* coalition otherwise.

The communities agree on a location of the facility h and a way to share the net benefit from using it. An *efficient allocation* is a vector $\boldsymbol{x} = (x_i)_{i=1,\dots,n}$ where x_i denotes community i's benefit with:

$$\sum_{i \in N} x_i = b(N) - c(N) = v(N).$$

An efficient allocation of the total net benefit $v(N)$ is induced by budget-balanced transfers $\boldsymbol{t} = (t_i)_{i=1,\dots,n}$ with $\sum_{i=1}^{n} t_i = 0$. The host h enjoys a welfare of $x_h = b_h - c_{hh} + t_h$ where t_h is the compensation received from hosting the facility. Its neighbors j obtain $x_j = b_j - c_{hj} + t_j$. They are thus paid t_j for the nuisances. Other communities $i \in N \backslash \bar{\mathcal{N}}(h)$ get $x_i = b_i + t_i$, thereby paying $-t_i$ to finance the compensations $t_h + \sum_{j \in \mathring{\mathcal{N}}(h)} t_j$.

An allocation is in the core of the NIMBY problem if it is not blocked by any coalition. We say a coalition $S \subset N$ blocks a distribution of the welfare if its members can achieve a higher welfare by themselves. We need to figure out what a coalition S can achieve by building and running its own facility. It depends on its own behavior and on the behavior of the communities in $N \backslash S$. Indeed, by agreeing to build a facility close to some members of S, the communities outside S can exert a negative externality on S, hence reducing its value. Technically, the cooperative game induced by the NIMBY problem exhibits *cooperative externalities*: the worth or value of a coalition S depends on the behavior of outside communities. For instance, if the communities outside S cooperate to build a facility, a member of S who adjoins the facility might suffer from a *negative externality* and S would experience a welfare loss. In the next section, we assume that if a coalition S builds a facility, communities outside S do not build any. Such an assumption is in line with the notion of γ-core whereby communities outside a coalition S play their individual best reply strategies (Chander and Tulkens, 1997). We show that the problem can be reduced under some assumptions. We propose a simple formula for an index related to the non-emptiness of the core.

6.2 Existence of core allocations

6.2.1 The NIMBY game without cooperative externalities

In first approach, we assume that if a coalition S builds a facility, remaining communities do not cooperate. Here, since it is too costly for a single community to build its own facility, its best individual strategy is not to build. This gives higher incentives for a coalition S to block a global agreement since, when doing so, it does not anticipate the potential negative externalities resulting from the cooperative behavior of remaining communities. We relax this assumption and extend our result in Section 6.2.2.

Under such an assumption, the value or worth of coalition $S \subset N$ is:

$$v(S) = max\{0, b(S) - c(S)\}$$

The NIMBY game without cooperative externalities thus defined belongs to the set of TU-games. It may not be superadditive.[9] Besides, not all TU-games can be represented as a NIMBY game without cooperative externalities, i.e. by picking benefits $\boldsymbol{b}$ and costs $\boldsymbol{C}$ to obtain its characteristic function: the NIMBY games without cooperative externalities form a strict subset of the set of TU-games.[10]

A core allocation is defined as follows:

Definition 6.1. An allocation $\boldsymbol{x}$ is in the core $\mathcal{C}$ if it satisfies $\sum_{i \in N} x_i = v(N)$ and the following core lower bounds:

$$\forall S \subset N, \sum_{i \in S} x_i \geq v(S).$$

The core is defined by a large number of inequality constraints. Some are binding, others are not. We introduce several assumptions which aim at simplifying the problem. The first one is related to the benefit achieved by a coalition formed by excluding a single community or members of a common neighborhood.

[9] For instance, the NIMBY problem with three players defined by $b_1 = b_2 = b_3 = 2$, $c_{11} = c_{22} = c_{33} = 1$, $c_{12} = c_{23} = c_{31} = 1$, and $c_{21} = c_{32} = c_{13} = 3$ does not lead to a superadditive TU-game. Indeed, we have $v(\{1,2,3\}) = 1 < v(\{1,2\}) + v(\{3\}) = 2 + 1 = 3$.

[10] For instance, in the case of TU-games with three players, the cooperative game induced by a NIMBY problem with three communities, benefit $\boldsymbol{b}$, and cost matrix $\boldsymbol{C}$ has the following characteristic function:

$$\begin{aligned} v(\{i\}) &= max(0, b_i - c_{ii}), i \in \{1,2,3\} \\ v(\{i,j\}) &= max(0, b_i + b_j - \min(c_{ii} + c_{ij}, c_{jj} + c_{ji})), i \neq j \\ v(\{1,2,3\}) &= b_1 + b_2 + b_3 - \min(c_{11} + c_{12} + c_{13}, c_{21} + c_{22} + c_{23}, c_{31} + c_{32} + c_{33}) \end{aligned}$$

Consider the TU-game represented by $v(\{i\}) = 1$, $v(\{i,j\}) = 0$ $(i \neq j)$, $v(\{1,2,3\}) = 2$. It is easy to check that no vector of benefits $\boldsymbol{b}$ and cost structure $\boldsymbol{C}$ can make the two characteristic functions coincide.

Assumption 1. $\forall S \in \bar{\mathcal{N}}, b(N \backslash S) \geq c(N \backslash S)$

Assumption 1 implies that coalitions formed by excluding some neighbors of the same community would always build a facility. It holds for *local* externalities, i.e. when few municipalities are negatively impacted in relation to the number of beneficiaries. In the uniform linear case, it holds when excluding the neighbors of communities at the extremities of the line (1 and n), or the two neighbors of a middle-community i (with $1 < i < n$), would not prevent the remaining communities from building a facility. The first requirement is met when $(n-1)b \geq c$ while the second holds true when $(n-2)b \geq c$.[11] The last inequality provides a condition on the parameters n, b and c such that Assumption 1 holds in the uniform linear case.

Assumption 2. *The optimal host is not unique:* $|\mathcal{H}| > 1$.

Assumption 2 might appear quite restrictive. Yet it holds in the linear case and can be replaced by a different one for the main result, as discussed in Appendix 6.D.

Relying on the above assumptions, we can significantly reduce the set of lower bounds defining the core.

Proposition 6.1. *Under Assumptions 1 and 2, an allocation* $\boldsymbol{x}$ *is in the core* $\mathcal{C}$ *if and only if*

$$\sum_{i \in N} x_i = v(N) \tag{6.1}$$

$$\forall i \in N, x_i \geq 0 \tag{6.2}$$

$$\forall i \in N, x_i \leq b_i \tag{6.3}$$

$$\forall S \in \mathring{\mathcal{N}}, \sum_{i \in S} x_i \leq b(S) - (c(N) - c(N \backslash S)). \tag{6.4}$$

Proposition 6.1 clarifies what constrains core allocations. Condition (6.1) is the efficiency condition. Condition (6.2) captures individual rationality: as we assume any single community anticipates no external cost from its withdrawing, it should be guaranteed 0 in core allocations. In regard to this lower bound on individual allocations, Condition (6.3) imposes a higher bound on individual allocations: the rationality of the coalitions of size

[11] In the first case, excluding a community at the extremity of the line allows a cost δc to be saved so that the total cost incurred by the coalition which excludes 1 or n is $(1+\delta)c$. Yet the coalition loses the benefit \$b\$ from the excluded community so that the total benefit is $(n-1)b$. In the second case, by excluding two communities that are neighbors of a middle-community i, the coalition can save the two external costs $2\delta c$ by locating the facility at i, although they loose the benefit of the two neighbors from using the facility so that the total benefit is $(n-2)b$.

$n-1$ ensures that no agent can be subsidized in the grand coalition. From Assumption 2, this requirement also holds for the host. Finally, Condition (6.4) reflects the possible exclusion of the host's neighbors: by excluding some neighbors of a potential host, a coalition disregards part of the external costs of the facility. Note that, even though it is not taken into account by the remaining coalition, the excluded communities could still suffer from the external costs. Due to such possibility of costless exclusion, the welfare of the host's neighbors is bounded from above. The point is that the upper bounds should not only hold for the *actual host*'s neighbors, but also for all *potential host*'s neighbors. Therefore, a coalition S of neighbors of a same community should contribute to the project at least to the extent of the cost saved by excluding them, that is, $c(N) - c(N \backslash S)$. We need this condition to hold for every coalition of neighbors of a same community, that is for all coalitions in $\mathring{\mathcal{N}}$.

We note that these constraints can be stringent enough to undermine the existence of core allocations: when exclusion is profitable enough, the allocation of the full value of the facility could be impossible in the grand coalition as such constraints would require the collection of more than the total cost of the project. The understanding we get from Proposition 6.1 leads us to a general statement about the existence of the core in NIMBY games. The following condition will be imposed.

Assumption 3. $\forall i \in N, b_i \geq max_{j \in N \backslash \{i\}} c_{ji}$

Assumption 3 states that the cost borne by a community when the facility is located at one of its neighbor's never surpasses its own benefit. In the uniform linear case, this means that $b \geq \delta c$: the external cost is bounded by the benefit of using the facility. In this specific case, we note that $b \geq \delta c$ is a necessary condition for the core not to be empty as we know that community 2 would have a maximum welfare of $b - \delta c$, which is negative when the condition is not met. This assumption is made in order to focus on cases for which individual rationality is not a source of emptiness of the core. We will see in Section 6.2.2 that it can be relaxed when considering more general notions of the core.

The following Proposition provides a simple test for non-emptiness of the core, involving an original index $I(\boldsymbol{C})$.

Proposition 6.2. *Under Assumptions 1 to 3, the core is non-empty if and only if* $I(\boldsymbol{C}) \leq 1$ *where*

$$I(\boldsymbol{C}) = \max_{\boldsymbol{\chi}} \left\{ \sum_{S \in \mathring{\mathcal{N}}} \chi_S \left(1 - \frac{c(N \backslash S)}{c(N)}\right) \middle| \forall i \in N, \sum_{S \in \mathring{\mathcal{N}} : i \in S} \chi_S = 1, \chi_S \geq 0 \right\}$$

The proof of Proposition 6.2 is provided in Appendix 6.H. A similar result can be obtained with a different assumption than Assumption 2. This is stated and proved in Appendix

6.D. The general scheme of the proof is the following: starting from the result of Proposition 6.1, we show that individual rationality constraints are never binding in a linear program related to non-emptiness of the core using Assumption 3. The expression of the dual of the resulting linear program then leads to Proposition 6.2.

Literally, the index $I(\boldsymbol{C})$ considers the savings induced by the exclusions of single agents and subsets of strict neighborhoods. It consists in the computation of the extent of the savings induced by such exclusion on all balanced collections of subsets of neighbors.[12] This is a combinatorial problem which is difficult to solve in general. Yet the computational complexity of $I(\boldsymbol{C})$ is greatly reduced as compared to the general problem of the existence of the core. As we will see in the uniform linear case and in Section 6.3, it can be computed for specific spatial structures.

It is well known that the core of a TU-game is non-empty if and only if it is balanced. Proposition 6.2 reformulates the balancedness condition for NIMBY games without cooperative externalities.[13] It emphasizes the role of the cost structure $\boldsymbol{C}$ in the difficulty of reaching an unanimously accepted solution to the NIMBY problem.

Note that the index $I(\boldsymbol{C})$ is invariant by the multiplication of $\boldsymbol{C}$ by a same positive number: what matter are the relative proportions of the different costs. The two following Corollaries clarify what features of the cost structure matter.

Corollary 6.1. *Let $t \in \mathbb{R}_+$ and define $\boldsymbol{C'} = \boldsymbol{C} + t\boldsymbol{I_n}$ the cost matrix obtained from the addition of t to all the costs c_{ii}, $i \in N$. If Assumptions 1 to 3 remain satisfied, we have $I(\boldsymbol{C'}) \leq I(\boldsymbol{C})$.*

More precisely, defining $\tau = \frac{t}{c(N)}$, we have :

$$I(\boldsymbol{C'}) = \frac{I(\boldsymbol{C})}{1+\tau}$$

The proof of Corollary 6.1 is presented in Appendix 6.E. It turns out that the index $I(\boldsymbol{C})$ can be interpreted as a measure of the proportion of the external costs as compared to the host costs. More precisely, it quantifies the minimal increase in the host costs at all locations that is required for core allocations to exist. An index of 0.8 indicates that at most 20% of the total cost $c(N)$ could be withdrawn to the host cost at any location and the core would remain non-empty. An index of 1.2 indicates that at least 20% of the total cost $c(N)$ should be added to the host cost at any location for the core to be non-empty.[14]

[12] A collection $\mathcal{B}$ of coalitions is said to be balanced if and only if there exist strictly positive weights $\boldsymbol{\chi^{\mathcal{B}}} = (\chi_S^{\mathcal{B}})_{S\in\mathcal{B}}$ such that, for any $i \in N$, $\sum_{S\in\mathcal{N}:i\in S} \chi_S^{\mathcal{B}} = 1$.

[13] As not all TU-games can be represented as NIMBY cooperative games without outside cooperation, this condition cannot be expressed for any TU-game.

[14] Other meaningful quantities could be defined in this context. For instance, Le Breton et al (2013) focus on the least core-value in problems of local public-project provision and financing. This value

Corollary 6.2. *Consider a decrease in $\boldsymbol{C}$ while the minimal cost in the grand coalition $c(N)$ is unchanged. If Assumptions 1 to 3 remain satisfied, the resulting index $I(\boldsymbol{C})$ weakly increases.*

The proof of Corollary 6.2 is presented in Appendix 6.F. It may appear surprising at first sight: when costs decrease everywhere but at an optimal location, the core shrinks. The reason is that a decrease in the costs which leaves the cost unchanged at an optimal location, weakly increases the profitability of the deviation for all coalitions. As a result, the core is more likely to be empty and $I(\boldsymbol{C})$ weakly increases. An illustration will be provided in Section 6.3.1 for the case of graphs.

Example.

An illustration of Proposition 6.2 can be provided in the uniform linear case. In such problems, an explicit computation of $I(\boldsymbol{C})$ (detailed in Appendix 6.G) leads to a higher bound on the parameter δ which depends only on the parameter n, as stated in the following Corollary.[15]

Corollary 6.3. *Under Assumptions 1 and 3, the core of the uniform linear NIMBY problem (n, b, c, δ) is non-empty if and only if*

$$\delta \leq \bar{\delta}(n) = \begin{cases} \frac{2}{n-2} & \text{if } n = 4k, k \in \mathbb{N} \\ \frac{2}{n-1} & \text{if } n = 4k+1, k \in \mathbb{N} \\ \frac{2}{n} & \text{if } n = 4k+2, k \in \mathbb{N} \\ \frac{2}{n-1} & \text{if } n = 4k+3, k \in \mathbb{N} \end{cases}$$

In the linear case, a higher number of communities n causes the index $I(\boldsymbol{C})$ to increase. Hence, for any number of communities there exists a critical level of δ above which the core is empty or, for any δ there exists a critical number of communities above which the core is empty. This result does not generalize to all NIMBY problems. A counterexample will be provided in Section 6.3.1.

Remark.

The Shapley value is an attractive solution to define compensations when the core is non-empty. Unfortunately, the Shapley value might not belong to the core. It can indeed be

quantifies the minimal tax required on deviating coalitions for stabilizing the grand coalition. In this line, the cost of stability (Bachrach et al, 2009), quantifies the minimal subsidy to the grand coalition required to stabilize it. However, neither of them has a clear explicit form in the NIMBY cooperative game.

[15] Note that Assumption 2 is always satisfied in the uniform linear case.

the case as shown by Dehez (2013) in a slightly different problem without externalities. We reproduce this result in our framework with $n = 3$ players. Assume that benefits are uniform ($\forall i \in N, b_i = b$) and that the cost of hosting the facility are $c_{11} = c_{22} < c_{33} < b$ while $c_{ij} = 0$ for $i \neq j$ (no cost externality). Assumptions 1 to 3 trivially hold, communities 1 or 2 should host the facility and all coalitions would build a facility. The core is not empty: for instance the equal sharing solution $x_i = b - \frac{c_{11}}{3}$ for $i \in \{1, 2, 3\}$ belongs to the core. The Shapley value $\boldsymbol{\phi}$ assigns:

$$\begin{aligned} \phi_1 &= b - \frac{c_{11}}{2} + \frac{c_{33}}{6} = \phi_2 \\ \phi_3 &= b - \frac{c_{33}}{3} \end{aligned}$$

Since $v(1, 3) = 2b - c_{11}$, $\phi_1 + \phi_3 < v(1, 3)$ if $3c_{11} < c_{33}$: the coalition $\{1, 3\}$ can be made better-off if it builds its own facility rather than share the total welfare according to the Shapley value. The same applies for the coalition $\{2, 3\}$. The reason why the Shapley value does not always belong to the core is that it may recommend a positive transfer to the community 1 or 2 when it does not host the project. In these cases, remaining communities are better off leaving this community aside.

6.2.2 The NIMBY game with cooperative externalities

The results of the previous section stand for a notion of the core which relies on a coalition's anticipation that outside members will not build any project. In some cases, it may be unrealistic to assume such behaviors. For instance, a single community or a small coalition would more likely expect outside members to cooperate and build a facility. In this section, we examine alternative and plausible expectations on the behavior of outsiders that a coalition might form. In doing so, we generalize the result to other notions of the core.

As discussed above, this game features *cooperative externalities*: the value of a coalition S depends on the cooperative behavior of communities outside S and their related facility-building decisions. We now formalize the problem in partition form (Thrall and Lucas, 1963). Let $\mathbb{P}(N)$ be the set of all partitions of N. The cooperative behavior of communities is summarized by an element $\mathcal{P}$ of $\mathbb{P}(N)$ where each element S of $\mathcal{P}$ is a coalition. The members of S jointly decide on whether to build a facility and on its location. Let us denote S's building decision by its location choice $l \in S \cup \{0\}$ where $l = 0$ if no facility is built. In a partition $\mathcal{P} = \{S_1, ..., S_m\}$, each coalition of communities $S_i \in \mathcal{P}$ picks one of its best location decisions l_i. A rational location vector in partition $\mathcal{P}$ is a vector $\boldsymbol{l} = (l_1, ..., l_m)$ where each decision l_i minimizes the cost of the facility (the cost of hosting the facility and the external costs *within* S_i). Let us denote by $\mathcal{L}(\mathcal{P})$ the set of rational location decision vectors in the partition $\mathcal{P}$.[16] They can be multiple due to potential indifference. For instance, in the linear homogeneous NIMBY problem with $n = 5$ communities and $2b < c \leq 3b$, the partition $\{\{1,2\},\{3,4,5\}\}$ might implement two different location vectors $(0,3)$ and $(0,5)$. They are both rational. However, as coalitions do not internalize the effect of their decisions on outside communities, some rational location will not be efficient. In this case, $(0,5)$ is preferred to $(0,3)$ by coalition $\{1,2\}$ since in the first case one of its members, namely community 2, will incur the negative external cost δc.

In a standard approach, the value function of a game with externalities depends on the coalition S and the partition in which the coalition is embedded. In the NIMBY game, it also depends on the rational location vector. Hence, we define the value in partition function form v as a function that assigns to every coalition S, partition $\mathcal{P}$ of N such that $S \in \mathcal{P}$, and rational location vector $\boldsymbol{l} \in \mathcal{L}(\mathcal{P})$, a real number $v(S, \mathcal{P}, \boldsymbol{l})$. It is the welfare achieved by coalition S embedded in the partition $\mathcal{P}$ with the rational location vector $\boldsymbol{l}$ on $\mathcal{P}$. The value is then defined for any potential configuration in terms of partition and rational location decisions.Yet some of those configurations might still appear irrelevant. We account for this by introducing exogenous expectation formation rules. An *expectation formation rule* R is a mapping which assigns to each coalition S a pair $R(S)$ of a partition $\mathcal{P}$ of N including S with a rational location vector $\boldsymbol{l}$ for $\mathcal{P}$. Such a function represents the expectations of a deviating coalition regarding the cooperative behavior of outside members and the non-cooperative behavior of the resulting coalitions. The value of S under the rule R is denoted $v^R(S) = v(S, R(S))$. Note that expectations are taken as exogenous here. They could be endogenized following the literature on dynamic coalition formation (Bloch and van den Nouweland, 2014).

We now define the core based on exogenous expectation formation rules. A coalition S in a partition $\mathcal{P}$ blocks a global agreement $\boldsymbol{x}$ under the rule R if it can achieve a higher welfare under such a rule. An allocation belongs to the R-core of the NIMBY game with cooperative externalities, denoted $\mathcal{C}^R$, if it is not blocked by any coalition of N. Formally:

Definition 6.2. Let R be an exogenous expectation formation rule. An agreement $\boldsymbol{x}$ is in the R-core $\mathcal{C}^R$ if it satisfies $\sum_{i \in N} x_i = v(N)$ and the following core lower bounds:

$$\forall S \subset N, \sum_{i \in S} x_i \geq v^R(S)$$

[16]Two comments are called for here. First, our restricting the attention to the set of rational decisions is in contrast with the standard approach of the α-core and β-core which respectively consider what a coalition can achieve regardless of the behavior of outside members or when having the possibility to adjust to others actions. Consistently with a remark by Laffont (1977) in the context of the garbage game, the α-core would never be empty in our context. Second, in our case, location decisions are independent. Yet, in the case of non-excludable benefits, strategic interactions would arise among coalitions for the provision of facilities.

As an illustration, we propose to discuss two specific rules relying on polar assumptions on the behavior of outside members.[17]

We call the first expectation formation rule *Collapse In Outside Cooperation* (CIOC). When deviating from a global agreement by blocking an allocation, a coalition S expects that the remaining communities will not cooperate to build facilities. It is formally defined by: $\forall S \subset N$, $R(S) = (\{S, \{i\}_{i \in N \setminus S}\}, \boldsymbol{l})$. Note that since communities outside S are singletons and that we assume that no community would build on its own, they never build. The location decision vector $\boldsymbol{l}$ boils down either to no-building at all, or to a single facility located inside S. If S builds a facility, there might be multiple optimal localizations of the facility in S. Yet, all these localizations lead to a single value $v^c(S) = b(S) - c(S)$. This leads us to the following definition:

Definition 6.3. The CIOC value function is defined as

$$v^c(S) = max(0, b(S) - c(S))$$

The CIOC-core $\mathcal{C}^c$ is the R-core associated with the CIOC expectation formation rule.

This is the notion of the core that we have examined in the previous section. It corresponds to the notion of the γ-core introduced in the context of public good games (see e.g. Chander and Tulkens, 1997).

The second expectation formation rule we consider corresponds to the case of full cooperation. We call this second rule *Rational Hostile Outside Cooperation* (RHOC). It is formally defined by $\forall S \subset N, R(S) = (\{S, N\setminus S\}, \boldsymbol{l})$. The expectations of a coalition S when considering blocking an allocation is that the remaining communities will cooperate and (potentially) build a facility. Moreover, the coalition S expects that if the coalition $N\setminus S$ is indifferent between different locations, it will locate it at the worst place from S's point of view. This leads us to the following definition:

Definition 6.4. The RHOC value function for a coalition S is defined as

$$v^r(S) = min_{\boldsymbol{l} \in \mathcal{L}(\{S, N\setminus S\})} v(S, \{S, N\setminus S\}, \boldsymbol{l})$$

The RHOC-core $\mathcal{C}^r$ is the R-core associated with the RHOC expectation formation rule.

We first investigate the cooperative externalities in the NIMBY problem. A cooperative game exhibits positive (resp. negative) externalities if coalitions benefit (resp. suffer) from the cooperative behavior of players outside (De Clippel and Serrano, 2008). It turns

[17]On the cooperative aspects, these rules respectively correspond to the $\underline{N}$-exogenous and the $\bar{N}$-exogenous rules in Bloch and van den Nouweland (2014).

out that externalities in the cooperative game induced by the NIMBY problem can be either positive or negative. The following proposition links the value functions of the game in partition form, and the CIOC and the RHOC rules.

Proposition 6.3. *For any $S \subset N$, $\mathcal{P} \ni S$ and $\boldsymbol{l} \in \mathcal{L}(\mathcal{P})$,*

1. $v^c(S) \geq v(S, \mathcal{P}, \boldsymbol{l})$
2. *We might have $v^r(S) > v(S, \mathcal{P}, \boldsymbol{l})$ or $v^r(S) < v(S, \mathcal{P}, \boldsymbol{l})$ depending on $\mathcal{P}$*
3. *We might have $v(S, \mathcal{P}, \boldsymbol{l}) > v(S, \mathcal{P}', \boldsymbol{l})$ or $v(S, \mathcal{P}, \boldsymbol{l}) < v(S, \mathcal{P}', \boldsymbol{l})$ when $\mathcal{P}'$ is a finer partition of N including S.*

First, the CIOC value is the highest possible value that a coalition can obtain by deviating from the global agreement. This is because a coalition can only be bothered by nuisances generated by the facilities built by outsiders. So the best that can happen for a coalition is that the outsiders do not build any facility which holds under CIOC. Therefore, the CIOC rule can also be seen as the optimistic expectation formation rule in this context (Shenoy, 1979).

Second, the RHOC value can be lower or higher than the value with other partitions. This can be shown in the 5-player uniform linear NIMBY problem. For $2b \geq c$ the lowest value for coalition $S = \{2\}$ would be achieved with $\mathcal{P} = \{\{2\}; \{1, 4\}; \{3, 5\}\}$ and location decisions $(0, 1, 3)$ because S would undergo the externalities linked to 2 facilities instead of a single one in the case $\mathcal{P} = \{\{2\}; \{1, 4, 3, 5\}\}$. This remark emphasizes the fact that full cooperation of outsiders is not the worst that can happen to a coalition.

Third, a coalition does not necessarily benefit from the merger of other coalitions. For instance, S could experience a negative externality when two former non-building coalitions merge and build next to it. This would be the case in the homogeneous linear case with 5 communities when $c \leq 4b < 2c$ and $\mathcal{P} = \{\{2\}; \{1, 4\}; \{3, 5\}\}$. The merger to $\mathcal{P} = \{\{2\}; \{1, 4, 3, 5\}\}$ would induce the construction of a facility at 1 and make the worth of $\{2\}$ decrease. Therefore, the RHOC rule does not correspond to the pessimistic expectation formation rule that leads to the notion of the α-core (Aumann, 1967).

As a consequence of Proposition 6.3, we know that for any exogenous expectation formation rule R, $\mathcal{C}^c \subseteq \mathcal{C}^R$. The CIOC-core is the most restrictive notion of a core as it amounts to considering that coalitions do not take into account the negative externality that outside members could exert on them. Hence, the emptiness of the CIOC-core does not imply the emptiness of any core. This leads us to question the generality of Proposition 6.2, and more especially to doubt whether the necessary character of the identified condition would extend to any core. We actually show that this condition carries forward to any exogenous expectation formation rule, provided we additionally assume that neighborhoods are small enough so that they never build when excluded. This is the sense of Assumption 4.

Assumption 4. $\forall S \in \bar{\mathcal{N}}, b(S) < c(S)$

Assumption 4 suffices for any coalition of neighbors of a same community not to build a facility on their own.[18] It holds when the number of direct neighbors is limited as compared to the minimal number of communities for which building a facility is efficient. For instance, it does hold in the uniform linear case when at least three communities are needed to build a facility, because each community has at most two neighbors. Formally, it requires that $2b < c$, so that no neighborhood of a community would build on its own.[19]

Assumption 3 will also be needed. Yet, it can be weakened to the following assumption.

Assumption 5. $\forall S \in \bar{\mathcal{N}}, b(S) \geq \sum_{i \in S} \max_{j \in N \setminus \{i\}} c_{ji} + v^R(S)$

Unfortunately, Assumption 5 has no direct interpretation. Along with Assumptions 1, 2 and 4, it generalizes Proposition 6.2 to any R-core.

Proposition 6.4. *Under Assumptions 1, 2, 4 and 5, for any exogenous expectation formation rule R, the R-core is non-empty if and only if $I(\boldsymbol{C}) \leq 1$.*

Proposition 6.2 is generalized in games with externalities for two reasons. First, if a coalition excludes some communities in a common neighborhood, then, due to Assumption 4, it is not rational for the excluded communities to build their own facility. The value of the former coalition is therefore independent from the expectation formation rule. Thus, we end up with the same core lower bounds as in the game without cooperative externalities when it comes to the exclusion of communities in a common neighborhood. Second, we show that, under Assumption 5, the core lower bounds of other non-building coalitions are never binding in a linear program related to the non-emptiness of the core as in Proposition 6.2.

Note that Proposition 4 can be extended to more general representations of expectation formation rules. More generally, communities could form non-deterministic expectations about the cooperative behavior of others defined as a coalition structure (in which the coalition is embedded) and the rational location vector for the facilities. We could then build the expectation of a coalition S as a probability distribution over partitions $\mathcal{P} \ni S$ and location vectors $\boldsymbol{l} \in \mathcal{L}(\mathcal{P})$. The worth of S would be its expected value. Proposition 4 would remain valid with such rules.[20]

Before moving to the illustrative examples, we briefly discuss the case of non-excludable benefits. It includes for instance NIMBY problems such as shale gas wells or nuclear waste

[18] Along with Assumption 1, it emphasizes a crucial feature for our results to hold: neighborhoods should be sufficiently smaller than their complementary to induce different building decisions. For this reason, our results apply to local pollution at the scale of N.

[19] If two communities neighboring a community i with $1 < i < n$ share a facility, they incur the hosting cost but no external cost for a benefit of $2b$.

[20] We thank a referee for suggesting this point.

repositories. When benefits are non-excludable, coalitions of communities are tempted to block a global agreement because they can benefit from the facility without paying its cost. Both free-riding and the exclusion of potential neighbors compromise the existence of core allocations. However, free-riding is mitigated by two forces. First, large coalitions would not rationally expect remaining communities to build a project by themselves. Hence, free-riding would only increase the core lower bounds for small coalitions. Second, small deviating coalitions may expect the project to be located at their borders when withdrawing from the grand coalition as their interests would no longer be taken into account. At first sight, this would mitigate free-riding incentives. The RHOC-core can offer interesting insights on this problem as it presupposes cooperation among remaining communities and the associated credible threat. We again focus on the linear case in the following example.

Example

Let us consider the uniform linear case with $n \geq 6$. Under Assumptions 1 and 4, the RHOC-core of a uniform linear NIMBY problem with at least six communities and non-excludable benefits is empty. Indeed, for any RHOC-core allocation $\boldsymbol{x}$, the core lower bound of the coalition $N \backslash \{2\}$ (respectively $N \backslash \{n-1\}$) requires $x_2 \leq b_2 - \delta c$ (resp. $x_{n-1} \leq b_{n-1} - \delta c$). On the other hand, we know that the coalition $N \backslash \{2, n-1\}$ is a building coalition due to Assumption 4. Hence the coalition $\{2, n-1\}$ can free-ride and its core lower bound is written $x_1 + x_{n-1} \geq b_1 + b_{n-1} - \delta c$. The latter condition is not compatible with the other two conditions identified. Hence the RHOC-core is empty.

We conclude this part by insisting that the emptiness of the core in the non-excludable case is likely to stem from the interplay between free-riding incentives and the garbage-game dimension of the problem. The latter dimension puts a higher bound on the welfare of the neighborhoods. When benefits are excludable, these small coalitions will often not get more in a core allocation than what they would achieve if they withdrew, even if they have to bear the threat imposed by the remaining communities. Therefore, even if credible threats exist, they would often fail to stabilize the grand coalition. In the following section we redirect the focus to the CIOC-core in explicit spatial structures with excludable benefits.

6.3 Illustrations on explicit spatial structures

6.3.1 Uniform NIMBY problems on graphs

The linear example developed previously put the emphasis on the effect of the number of individuals on the core. We propose here a natural extension of the linear case which allows us to investigate the effect of the spatial structure on the core. For a given NIMBY problem $(N, \boldsymbol{b}, \boldsymbol{C})$, we restrict our attention to NIMBY problems on graphs. As in the linear case, the cost of building the project at a community i is the same for all. However, it entails an identical additional cost δc on each of i's neighbors. Hence, the matrix of costs can be written $\boldsymbol{C} = cI_n + \delta c G$ where I_n is the identity matrix and G is the adjacency matrix of a simple graph (with values 0 on the diagonal). A NIMBY problem on a graph is fully characterized by parameters $(N, \boldsymbol{b}, c, \delta, G)$.

Example

Figure 6.1 below represents the graph associated with the following cost matrix for $n = 6$:

$$\boldsymbol{C} = \begin{pmatrix} c & \delta c & \delta c & 0 & 0 & 0 \\ \delta c & c & \delta c & 0 & 0 & 0 \\ \delta c & \delta c & c & \delta c & 0 & 0 \\ 0 & 0 & \delta c & c & \delta c & \delta c \\ 0 & 0 & 0 & \delta c & c & \delta c \\ 0 & 0 & 0 & \delta c & \delta c & c \end{pmatrix}$$

On this graph, a facility built at 1 would yield an external cost δc at 2 and 3. The minimal cost in the grand coalition is $c(N) = c + 2\delta c$. The efficient locations are 1, 2, 5 and 6.

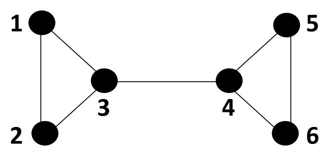


Figure 6.1: A graph with 6 communities

As in the linear case, the condition for non-emptiness can be stated as an upper bound on the parameter δ.

Corollary 6.4. *Under Assumptions 1, 2 and 3, the core of a NIMBY problem on a graph* $(N, \boldsymbol{b}, c, \delta, \boldsymbol{G})$ *is non-empty if and only if* $\delta \leq \bar{\delta}(\boldsymbol{G})$ *where* $\bar{\delta}(\boldsymbol{G}) > 0$.

The proof and an explicit expression for the critical value $\bar{\delta}(\boldsymbol{G})$ are provided in Appendix 6.I, which is directly related to $I(\boldsymbol{C})$: for a given δ, the more $I(\boldsymbol{C})$, the less $\bar{\delta}(\boldsymbol{G})$. The expression of the critical value $\bar{\delta}(\boldsymbol{G})$ involves the value of a linear program that can be computed for specific examples.

Figure 6.2 provides an example of the ambiguous effect of the number of communities. Indeed, from the linear graph A with 5 communities, the addition of a community on the extremity of the line to form graph B implies a decrease in $\bar{\delta}(\boldsymbol{G})$. For a given δ, this is associated with a decrease in $I(\boldsymbol{C})$. However, the further addition of a community to form graph $\boldsymbol{C}$ implies an increase in $\bar{\delta}(\boldsymbol{G})$.

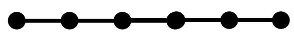
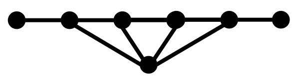
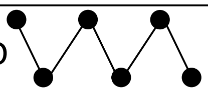
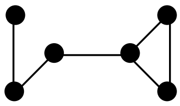
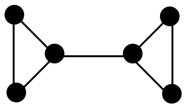
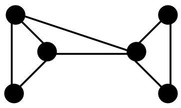
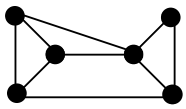
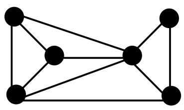
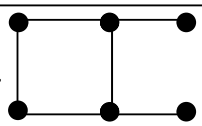
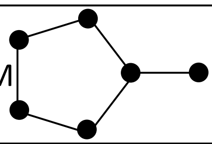
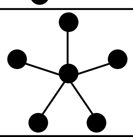
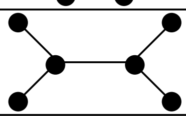
Graph G		n	$\bar{\delta}$(G)
A		5	1/2
B		6	1/3
C		7	1/2

Figure 6.2: Critical value of δ for different graphs with different number of communities. These values are obtained from the explicit computation of $\bar{\delta}(\boldsymbol{G})$ according to the expression derived in Appendix 6.I. The code used is provided in Appendix 6.J.1.

Figure 6.3 presents the critical value $\bar{\delta}(\boldsymbol{G})$ associated with different graphs, all involving 6 communities. Corollary 6.2 is illustrated on graphs D to K and P to S: we observe that, when a link is added while keeping the minimum degree constant, requirements on δ can only be relaxed. In particular, the lax condition obtained for the complete graph X can easily be extended to all complete graphs. This further emphasizes that our argument mainly stands for local pollutions. Yet, in this case, since neighborhoods and their complements are no longer asymmetric, Assumptions 1 and 4 cannot be met at the same time. Finally, this assessment shows that the spatial structure is in itself an important source of variability for the set of core agreements. We propose to carry further the exercise on a real administrative unit.

6.3.2 A tentative assessment on real geographies

We now introduce a hypothetical problem in a real administrative geographical division and compute the associated index $I(\boldsymbol{C})$. The main purpose of this exercise is to illustrate

Graph G		$\underline{d}(N)$	$\overline{\delta}(G)$
D		1	1/3
E		1	1/2
F		2	1/4
H		2	1/3
J		2	1/2
K		2	1
L		1	1
M		1	1/2
N		1	1
O		1	1

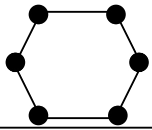
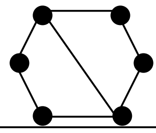
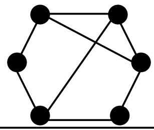
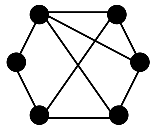
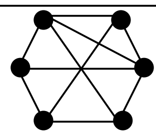
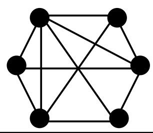
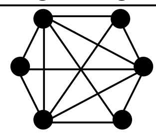
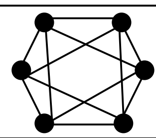
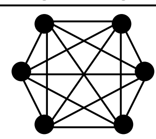
Graph G		$\underline{d}(N)$	$\overline{\delta}(G)$
P		2	1/4
Q		2	1/4
R		2	1/3
S		2	1/2
T		3	1/3
U		3	1/2
V		3	1
W		4	1/2
X		5	1

Figure 6.3: Critical values for different graphs with 6 communities. $\underline{d}(N)$ denotes the minimal degree of the graph. These values are obtained from the explicit computation of $\bar{\delta}(\boldsymbol{G})$ according to the expression derived in Appendix 6.I. The code used is provided in Appendix 6.J.1.

how the previous analysis could motivate further empirical analyses.

More precisely, let us consider a hypothetical negotiation among municipalities for locating and funding a facility in the French *département* of Haute-Garonne. This project would yield a potentially heterogeneous benefit to each municipality. As Proposition 6.2 emphasizes, we do not need precise knowledge of the benefits, provided that Assumptions 1 and 3 are met. However, Assumption 2 does not hold in this environment. Hence, we rely on the modification of Proposition 6.2 established in Appendix 6.D. It allows us to drop this assumption at the cost of a different, more realistic one and a slight modification of the expression of $I(\boldsymbol{C})$. What matters most is the structure of costs. Assume that a facility could only be located at the centroid of each municipality.[21] Assume additionally that, wherever it is built, a facility yields a uniform pollution cost within a fixed radius from the site. For instance, Figure 6.4 shows, in red, an impacted area of 3 km around a facility located at the centroid of a given municipality i. Given our assumption of uniformity within the impacted area, the total pollution cost is directly proportional to the red area. Moreover, as Proposition 6.2 emphasizes, we do not need to specify absolute values to compute the index: only relative values matter. Hence, c_{ii} can be normalized to the area of the intersection of the red plain circle and i's territory, and c_{ij}, to the area of the intersection between the red plain circle and j's territory. The matrix $\boldsymbol{C}$ is obtained by computing all such areas. Neighborhood sets and the optimal location are derived from the matrix $\boldsymbol{C}$.

We compute the index $I(\boldsymbol{C})$ for different radii. The code used to perform this computation is provided in Appendix 6.J.2. It yields the following results.

Radius (in km)	1	2	3
$I(\boldsymbol{C})$	4.46	34.2	50.2

We note that the index $I(\boldsymbol{C})$ is always higher than 1, so the core is empty in all these cases. This illustrative exercise can be improved by including a better estimate of the costs of such a facility. The main difficulty here lies in the use of plausible values for the perceived pollution costs. Any step in this direction would rely on a good understanding of the monetary as well as the non-monetary costs of such facilities. The resulting index would in particular be sensitive to the cost at the optimal site, which, in this illustration, was to be at some indentation of the boundary.

[21] This exercise emphasizes a limitation in the model: in order to compute the cost matrix $\boldsymbol{C}$, a hypothesis has to be made on where the facility would be located within a given municipality regardless of the coalition it belongs to. In this example, we chose the centroids of the municipalities. In a more general framework, we could expect coalitions to have some flexibility in the location choice. By increasing the value of all coalitions, such flexibility would strengthen requirements for non-emptiness. It would yield complications but, in our view, few more insights.

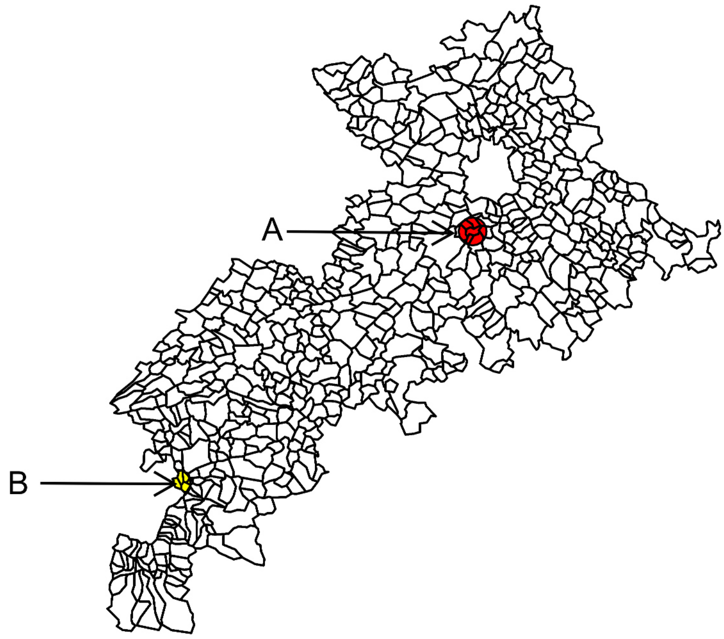


Figure 6.4: Municipalities composing Haute-Garonne.
Area A corresponds to the impacted area when located at an arbitrary municipality for a radius of 3 km. Area B corresponds to the impacted area at the optimal location for the same radius.

Conclusion

In this paper, we analyzed the cooperative provision of economic activities that are globally beneficial but locally harmful in an explicit spatial model. Examples include facilities such as landfills, waste treatment plants or polluting utilities. When communities can be excluded from using the facilities, free-riding is not a problem like in standard localized public-good provision problems. Yet, this may not suffice to warrant cooperation. We show that the exclusion of the neighboring communities of potential hosts can be a significant obstacle to cooperation. It sets upper bounds on compensations, which together with the participation constraints, determine whether a global cooperative solution exists. That is, if the core of the cooperative game is non-empty. An index is computed to test the existence of a core solution. Its definition is robust to several assumptions on the value function of induced cooperative games. It can be estimated in practice.

As mentioned in the introduction, our investigation of the NIMBY problem using cooperative game theory formalizes Coasian bargaining in economies with externalities. If, when the core is empty, the parties involved fail to implement the project, the "Coase theorem" does not hold. In this work, property rights were implicitly assigned to the polluters because a facility could be built without the consent of the neighboring communities. Under other assignments of property rights, the core might be non-empty. In particular, it is easy to show that the core is never empty under the polluter-pays principle: if the communities building the facility are forced to compensate all neighboring communities for the damages, a global and efficient cooperative agreement can always be reached. Therefore, in contradiction with another interpretation of the Coase theorem, the assignment of property rights could matter for achieving efficient outcomes, even in the absence of transaction costs.

Appendix

6.A Notations

N	Set of communities
$\boldsymbol{b}$	Vector of individual benefits b_i derived from using the facility
$\boldsymbol{C} \in \mathcal{M}_n(\mathbb{R}_+)$	Cost matrix
$c_{ij} \in \mathbb{R}_+$	Cost for community j of a facility located at i
$\mathcal{H} \subseteq N$	Set of optimal locations
$\mathring{\mathcal{N}}(i)$	Strict neighborhood of i (without i)
$\mathring{\mathcal{N}}$	Set of subsets of all strict neighborhoods
$\bar{\mathcal{N}}(i)$	Neighborhood of i, including i
$\bar{\mathcal{N}}$	Set of subset of all strict neighborhoods plus the singletons
$\boldsymbol{x}$	An allocation of the net benefit
$\boldsymbol{t}$	Transfers between communities
$S \subseteq N$	A coalition
$v(S)$	Worth of a coalition, assuming no outside cooperation
$b(S)$	Benefit of a facility for a coalition
$c(S)$	Optimal cost of a facility for a coalition
$\mathcal{P}$	A partition of N
$\boldsymbol{l}$	A rational location vector
R	An exogeneous expectation formation rule
$\mathcal{C}^R$	Set of core allocations, given the exogeneous expectation formation rule R
$\mathcal{C} = \mathcal{C}^c$	Set of core allocations, assuming no outside cooperation
$\mathcal{C}^r$	Set of core allocation, assuming rational hostile outside cooperation

6.B Proof of Proposition 6.1

Let $\boldsymbol{x}$ be an allocation which meets the conditions stated in Proposition 6.1, that is, the efficiency condition (6.1), individual rationality conditions (6.2), and the following lower bounds for every $S \in \mathcal{Y} = \{N\backslash S | S \in \mathring{\mathcal{N}}\} \cup \{N\backslash\{i\} | i \in N\}$,

$$\sum_{i \in S} x_i \geq v(S) \tag{6.5}$$

We first show that it satisfies the core lower bounds (6.5) for any arbitrary coalition. Let $T \subseteq N$.

- If T is a non-building coalition, we have $v(T) = 0$. $\forall i \in T$, $\boldsymbol{x}$ meets the individual rationality constraint $x_i \geq 0$. The sum of these constaints yields Condition (6.5) for T. \item If T is a building coalition, we have $v(T) = b(T) - c(T)$. Let us consider $j^* \in argmin_{j \in T} \sum_{i \in T} c_{ij}$ an optimal site in T and $S^* = \mathring{\mathcal{N}}(j^*)\backslash T$, the set of strict neighbors of j^* that are not in T and $\bar{T} = N\backslash S^*$. Since $\bar{T} = N\backslash S^* \in \mathcal{Y}$,

$$\sum_{i \in \bar{T}} x_i \geq b(\bar{T}) - c(\bar{T})$$

  Besides, $c(\bar{T}) \leq c(T)$ so:

$$\sum_{i \in \bar{T}} x_i \geq b(\bar{T}) - c(T) \tag{6.6}$$

  As for every $i \in N$, $N\backslash\{i\} \in \mathcal{Y}$, we have $\sum_{j \in N\backslash\{i\}} x_j \geq v(N\backslash\{i\})$. This inequality can be rewritten, using the efficiency condition (6.1), as $x_i \leq v(N) - v(N\backslash\{i\})$. We have $\forall i \in N\backslash\mathcal{H}, v(N) - v(N\backslash\{i\}) \leq b_i$ and Assumption 2 additionally implies $\forall h \in \mathcal{H}, v(N) - v(N\backslash\{h\}) \leq b_h$. Thus, $\forall i \in N, -x_i \geq -b_i$. From the summation of the latter inequalities for all agents in $\bar{T}\backslash T$ to inequality (6.6), we obtain $\sum_{i \in T} x_i \geq b(T) - c(T) = v(T)$. Hence condition (6.5) holds for T.

We have shown that the core lower bounds can be restricted to coalitions in $\mathcal{Y}$. From Assumption 1, coalitions in $\{N\backslash S | S \in \mathring{\mathcal{N}}\}$ are all building coalitions so the constraints associated with them are: $\sum_{i \in N\backslash S} x_i \geq b(N\backslash S) - c(N\backslash S)$. Combining them with the efficiency constraints yields conditions (6.4) in Proposition 6.1.

6.C Proof of Proposition 6.2

From Proposition 1, the core can be defined as:

$$\left\{ \boldsymbol{x} \in \mathbb{R}^n_+ \middle| \sum_N x_i = v(N) \text{ and } \forall S \in \bar{\mathcal{N}}, \sum_S x_i \leq b(S) - (c(N) - c(N\backslash S)) \text{ and } \forall i \in N, x_i \geq 0 \right\} \tag{6.7}$$

A necessary and sufficient condition for the non-emptiness of this set involves the linear program $(LP1)$:

$$\max_{\boldsymbol{x}} \left\{ \sum_{i \in N} x_i \middle| \forall S \in \bar{\mathcal{N}}, \sum_{i \in S} x_i \leq b(S) - (c(N) - c(N \backslash S)), \forall i \in N, x_i \geq 0 \right\} \geq v(N) \quad (6.8)$$

To show this equivalence, first note that if the set involved in condition (6.7) is non-empty. Therefore any element of it constitutes a feasible solution in the linear program defined in condition (6.8). A fortiori, its optimal solution also satisfies the condition. Therefore (6.7) $\Rightarrow$ (6.8). For the converse, consider a solution $\boldsymbol{x}^*$ to the linear program and assume it satisfies condition (6.8). Consider the allocation $\boldsymbol{x}^\epsilon$ defined by $x_i^\epsilon = \max(0, x_i^* - \epsilon)$, for all $i \in N$ and some $\epsilon > 0$. For any ϵ, the resulting allocation still pertains to the feasible set of the linear program and, by a continuity argument, we can always find ϵ such that $\sum_N x_i^\epsilon = v(N)$. This allocation pertains to the set defined in condition (6.7), which, therefore, is non-empty.

Assumption 3 implies that the saving induced by the withdrawal of a community will never exceed its benefit so for any $S \subseteq T \subseteq N$, we have $v(S) \leq v(T) \leq v(N)$. In particular, for every $S \in \bar{\mathcal{N}}$, $v(N \backslash S) \leq v(N)$ which implies $b(S) - (c(N) - c(N \backslash S)) \geq 0$. Hence, this linear program is feasible when Assumption 3 is met (take, for all $i \in N, x_i = 0$). Besides, it is bounded (by $\sum_N b_i$ for instance) so it admits a finite value.

We now show that the individual rationality constraints $x_i \geq 0$ are non-binding in $(LP1)$. We start to show that, for any optimal solution, no community, $i_0 \in N$, pays more than $\max_{j \in N \backslash \{i_0\}} c_{ji_0}$, the highest external cost it can bear. Let $\boldsymbol{x}^*$ be an optimal solution to $(LP1)$ and $i_0 \in N$. Assume that:

$$x_{i_0}^* < b_{i_0} + \min_{T \in \bar{\mathcal{N}}: i_0 \in T} \{c(N \backslash T) - c((N \backslash T) \cup \{i_0\})\} \quad (6.9)$$

We can then increase $x_{i_0}^*$ by some $\epsilon > 0$ such that:

$$x_{i_0}^* + \epsilon < b_{i_0} + \min_{T \in \bar{\mathcal{N}}: i_0 \in T} \{c(N \backslash T) - c((N \backslash T) \cup \{i_0\})\} \quad (6.10)$$

Such an increase improves the objective. We shall show that it also leads to a feasible solution. Let $S \in \bar{\mathcal{N}}$ such that $i_0 \in S$. Because $S \in \bar{\mathcal{N}}$, $S \backslash \{i_0\}$ also pertains to $\bar{\mathcal{N}}$ (except for the case $S = \{i_0\}$, in which the result is direct). By feasibility of $\boldsymbol{x}^*$, we have:

$$\sum_{i \in S \backslash \{i_0\}} x_i^* \leq b(S \backslash \{i_0\}) - (c(N) - c((N \backslash S) \cup \{i_0\})) \quad (6.11)$$

Summing inequalities (6.10) and (6.11), we get:

$$\sum_{i \in S} x_i^* + \epsilon < b(S) - (c(N) - c((N \backslash S) \cup \{i_0\})) + \min_{T \in \bar{\mathcal{N}}: i_0 \in T} \{c(N \backslash T) - c((N \backslash T) \cup \{i_0\})\}$$

Since $\min_{T \in \bar{\mathcal{N}}: i_0 \in T} \{c(N \backslash T) - c((N \backslash T) \cup \{i_0\})\} \leq c(N \backslash S) - c((N \backslash S) \cup \{i_0\})$, we have:

$$\sum_{i \in S} x_i^* + \epsilon < b(S) - (c(N) - c(N \backslash S))$$

All the constraints involving x_{i_0} are met. This contradicts the optimality of $\boldsymbol{x}^*$. Hence, inequality (6.9) cannot hold by contradiction. We have:

$$x_{i_0}^* \geq b_{i_0} + \min_{T \in \bar{\mathcal{N}}: i_0 \in T} \{c(N \backslash T) - c((N \backslash T) \cup \{i_0\})\} \tag{6.12}$$

Besides, for any $T \in \bar{\mathcal{N}}$ such that $i_0 \in T$,

$$c((N \backslash T) \cup \{i_0\}) - c(N \backslash T) = \min_{j \in (N \backslash T) \cup \{i_0\}} \sum_{k \in (N \backslash T) \cup \{i_0\}} c_{jk} - \min_{j \in N \backslash T} \sum_{k \in N \backslash T} c_{jk}$$

Let us denote by j^* an optimal host in $N \backslash T$. Since $\min_{j \in (N \backslash T) \cup \{i_0\}} \sum_{k \in (N \backslash T) \cup \{i_0\}} c_{jk} \leq \sum_{k \in (N \backslash T) \cup \{i_0\}} c_{j^* k}$, by definition of the minimum, we have:

$$c((N \backslash T) \cup \{i_0\}) - c(N \backslash T) \leq \sum_{k \in (N \backslash T) \cup \{i_0\}} c_{j^* k} - \sum_{k \in N \backslash T} c_{j^* k} = c_{j^* i_0}$$

Hence:

$$c((N \backslash T) \cup \{i_0\}) - c(N \backslash T) \geq - \max_{j \in N \backslash \{i_0\}} c_{j i_0} \tag{6.13}$$

From conditions (6.12) and (6.13), we get $x_{i_0}^* \geq b_{i_0} - \max_{j \in N \backslash \{i_0\}} c_{j i_0}$ and, from Assumption 3, $x_{i_0}^* \geq 0$. Thus, individual rationality constraints can be discarded from $(LP1)$ without altering the value of the objective. This leads us to consider the linear program $(LP2)$:

$$\max_{\boldsymbol{x}} \left\{ \sum_{i \in N} x_i \middle| \forall S \in \bar{\mathcal{N}}, \sum_{i \in S} x_i \leq b(S) - (c(N) - c(N \backslash S)) \right\}$$

Again, this linear program is bounded and feasible. Therefore, it admits a finite value and so its dual $(LP2^*)$:

$$\min_{\boldsymbol{x}} \left\{ \sum_{S \in \bar{\mathcal{N}}} \chi_S (b(S) - (c(N) - c(N \backslash S))) \middle| \forall i \in N, \sum_{S \in \bar{\mathcal{N}}: i \in S} \chi_S = 1, \chi_S \geq 0 \right\}$$

Which can be further simplified to:

$$\min_{\boldsymbol{x}} \left\{ b(N) - \sum_{S \in \bar{\mathcal{N}}} \chi_S (c(N) - c(N \backslash S)) \middle| \forall i \in N, \sum_{S \in \bar{\mathcal{N}} : i \in S} \chi_S = 1, \chi_S \geq 0 \right\}$$

A necessary and sufficient condition for non-emptiness of the core is that the value of $(LP2^*)$ is lower than $v(N) = b(N) - c(N)$. This leads to the following condition:

$$\max_{\boldsymbol{\chi}} \left\{ \sum_{S \in \bar{\mathcal{N}}} \chi_S (1 - \frac{c(N \backslash S)}{c(N)}) \middle| \forall i \in N, \sum_{S \in \bar{\mathcal{N}} : i \in S} \chi_S = 1, \chi_S \geq 0 \right\} \leq 1$$

6.D Discarding Assumption 2 in Proposition 6.2

If $|\mathcal{H}| > 1$, the proof of Proposition 6.2 holds. Here we assume $|\mathcal{H}| = 1$ and show that a similar result to Proposition 6.2 can still be obtained. The difference lies in the fact that the host can get more than b_h in core allocations, which prevents an immediate focus on neighborhoods. However, we show that, as soon as an additional assumption is met, requiring that $x_h \leq b_h$ does not alter the value of the linear program. This allows a focus on neighborhoods. The proof proceeds as the proof of Propositions 6.1 and 6.2: we first discard redundant constraints and simplify non-binding constraints in a linear program related to the emptiness of the core.

Let us denote by h the unique optimal host in N and let $\boldsymbol{x}$ be an allocation which meets the efficiency condition (6.1), individual rationality constraints (6.2) and the following core lower bounds for every $S \in \mathcal{Y}' = \mathcal{E} \cup \mathcal{E}_h \cup \{N \backslash \{i\} | i \in N\}$,

$$\sum_{i \in S} x_i \geq v(S) \tag{6.14}$$

Where $\mathcal{E} = \{N \backslash S | S \in \mathring{\mathcal{N}} \text{ and } c(N \backslash S) \leq c(N)\}$ and $\mathcal{E}_h = \{S \subset N | h \notin S\}$.

We first show that it satisfies the core lower bounds (6.14) for any arbitrary coalition. Let $T \subseteq N$.

- If T is a non-building coalition, we have $v(T) = 0$. $\forall i \in T$, $\boldsymbol{x}$ meets the individual rationality constraint $x_i \geq 0$. The sum of these constraints yields Condition (6.14) for T.
- If T is a building coalition, we have $v(T) = b(T) - c(T)$. Let us consider $j^* \in argmin_{j \in T} \sum_{i \in T} c_{ij}$ an optimal site in T and $S^* = \mathring{\mathcal{N}}(j^*) \backslash T$, the set of strict neighbors of j^* that are not in T. We define $\bar{T} = N \backslash S^*$. If $c(N \backslash S^*) \leq c(N)$, then $\bar{T} \in \mathcal{E} \subset \mathcal{Y}'$. If $c(N \backslash S^*) > c(N)$, then it must be that h is not in $N \backslash S^*$ hence

$\bar{T} \in \mathcal{E}_h \subset \mathcal{Y}'$. Therefore:

$$\sum_{i \in \bar{T}} x_i \geq b(\bar{T}) - c(\bar{T})$$

Besides, $c(\bar{T}) = \min_{j \in \bar{T}} \sum_{k \in \bar{T}} c_{jk} \leq \sum_{k \in \bar{T}} c_{j^*k} = \sum_{k \in T} c_{j^*k} = c(T)$, where the third equality comes from the fact that communities in $\bar{T} \backslash T$ do not belong to the neighborhood of j^* by construction. Hence:

$$\sum_{i \in \bar{T}} x_i \geq b(\bar{T}) - c(T) \tag{6.15}$$

The rationality of coalitions $N \backslash \{i\}$ yields $\forall i \in N \backslash \{h\}, -x_i \geq -b_i$. From the summation of the latter inequalities for all agents in $\bar{T} \backslash T$ to inequality (6.15), we obtain $\sum_{i \in T} x_i \geq b(T) - c(T) = v(T)$. Hence condition (6.14) holds for T.

We have shown that the core lower bounds can be restricted to coalitions in $\mathcal{Y}'$. Combining them with the efficiency constraints and defining $\bar{\mathcal{E}} = \{T | T \in \mathring{\mathcal{N}} \text{ and } c(N \backslash T) \leq c(N)\}$ and $\bar{\mathcal{E}}_h = \{T | h \in T\}$, the respective complementary of $\mathcal{E}$ and $\mathcal{E}_h$, the core is non-empty if and only if:

$$\max_{\boldsymbol{x}} \left\{ \sum_{i \in N} x_i \middle| \forall S \in \bar{\mathcal{E}} \cup \bar{\mathcal{E}}_h \cup \{i | i \in N\}, \sum_{i \in S} x_i \leq b(S) - (c(N) - c(N \backslash S)), \forall i \in N, x_i \geq 0 \right\} \geq v(N)$$

We now eliminate constraints in $\bar{\mathcal{E}}_h$. Let us denote by $(LP3)$ the former linear program. Let us consider $\boldsymbol{x^*}$ as an optimal solution to $(LP3)$ and let us assume $x_h^* > b_h$ so that we can write $x_h^* = b_h + \epsilon$, $\epsilon > 0$. At this stage, an additional assumption is required:

Assumption 6. $\exists S \in \mathring{\mathcal{N}}$ *such that* $h \notin S$ *and* $c(N \backslash S) \leq c(N)$

This assumption implies that it is always possible to exclude some agents different from h and save on the cost of the project. We will show there always exists another optimal solution, $\boldsymbol{x'}$, such that $x'_h \leq b_h$. From Assumption 6, there exists $S \in \mathring{\mathcal{N}}$ such that $h \notin S$ and $c(N \backslash S) \leq c(N)$. Let us consider $S \cup \{h\} \in \bar{\mathcal{E}}_h$. We have, by feasibility of $\boldsymbol{x^*}$ in $(LP3)$:

$$\sum_{i \in S} x_i^* + x_h^* \leq \sum_{i \in S} b_i + b_h$$

Hence,

$$\sum_{i \in S} x_i^* \leq \sum_{i \in S} b_i - \epsilon$$

Besides the rationality of coalitions $N \backslash \{i\}$ requires $\forall i \in S$, $x_i^* \leq b_i$. Hence, there exists

$(\epsilon_i)_{i\in S} \in \mathbb{R}_+^{|S|}$ such that, $\sum_{i\in S}\epsilon_i = \epsilon$ and, for all $i \in S$, $x_i^* \leq b_i - \epsilon_i$. Let us define $\boldsymbol{x'}$ as follows:

$x'_h = x^*_h - \epsilon = b_h$

$x'_j = x^*_j + \epsilon_i$for all $j \in S$

$x'_i = x^*_i$for all $i \notin S \cup \{h\}$

By construction this solution yields the same objective. We want to show it is feasible as well. Let T_S be such that $T_S \cap S \neq \emptyset$ and $T_S \in \bar{\mathcal{E}} \cup \bar{\mathcal{E}}_h \cup \{i|i \in N\}$, an arbitrary coalition of $\bar{\mathcal{E}} \cup \bar{\mathcal{E}}_h \cup \{i|i \in N\}$ containing elements of S. Three cases arise:

- If $T_S \in \{i|i \in N\}$, then the associated constraint $x_i \leq b_i$ is met by construction.
- If $T_S \in \bar{\mathcal{E}}_h$, we have, where the first inequality comes from the fact that $\sum_{i\in T_S\cap S}\epsilon_i - \epsilon \leq 0$ and the second is the feasibility of $\boldsymbol{x}^*$ in $(LP4)$:

$$\sum_{i\in T_S} x'_i \leq \sum_{i\in T_S} x^*_i \leq b(T_S) - (c(N) - c(N\backslash T_S))$$

- If $T_S \in \bar{\mathcal{E}}$, $T_S \cup \{h\} \in \bar{\mathcal{E}}_h$ and by feasibility of $\boldsymbol{x}^*$ in $(LP4)$:

$$\sum_{i\in T_S} x^*_i + x^*_h \leq b(T_S) + b_h - (c(N) - c(N\backslash(T_S \cup \{h\})))$$

  Simplifying b_h and because $\sum_{i\in T_S\cap S}\epsilon_i \leq \epsilon$,

$$\sum_{i\in T_S} x'_i \leq \sum_{i\in T_S} x^*_i + \epsilon \leq b(T_S) - (c(N) - c(N\backslash(T_S \cup \{h\})))$$

  Because $c(N\backslash T_S) \leq c(N)$, the optimal location in $N\backslash T_S$ cannot be h. Hence, the withdrawal of h can only lead to a decrease in cost, so that $c(N\backslash(T_S \cup \{h\})) \leq c(N\backslash T_S)$. Finally, we have, for any constraint T_S involving elements of S:

$$\sum_{i\in T_S} x'_i \leq b(T_S) - (c(N) - c(N\backslash T_S))$$

This establishes that $\boldsymbol{x'}$ is feasible. Hence it is an optimal solution as well. Finally, we can require that $x_h \leq b_h$ without altering the value of the linear program. This defines the linear program $(LP4)$:

$$\max_{\boldsymbol{x}} \left\{ \sum_{i\in N} x_i \middle| \forall S \in \bar{\mathcal{E}} \cup \bar{\mathcal{E}}_h, \sum_{i\in S} x_i \leq b(S) - (c(N) - c(N\backslash S)), \forall i \in N, x_i \leq b_i \text{ and } x_i \geq 0 \right\}$$

It is straightforward to show that, following the introduction of the additional constraint $x_h \leq b_h$, all constraints in $\bar{\mathcal{E}}_h$ are redundant in $(LP4)$. Hence $(LP4)$ can be rewritten:

$$\max_{\boldsymbol{x}} \left\{ \sum_{i \in N} x_i \middle| \forall S \in \bar{\mathcal{E}}, \sum_{i \in S} x_i \leq b(S) - (c(N) - c(N \backslash S)), x_h \leq b_h, \forall i \in N, x_i \geq 0 \right\}$$

And, adding some redundant constraints to simplify the notations:

$$I(\boldsymbol{C}) = \max_{\boldsymbol{x}} \left\{ \sum_{i \in N} x_i \middle| \forall S \in \bar{\mathcal{N}}, \sum_{i \in S} x_i \leq b(S) - (c(N) - c(N \backslash S)), \forall i \in N, x_i \leq b_i \text{ and } x_i \geq 0 \right\}$$

We eventually get an expression similar to the one introduced in Proposition 6.2: Assumption 2 can be replaced by Assumption 6 provided we impose the additional condition $x_h \leq b_h$ in the former linear program. Hence, an expression of $I(\boldsymbol{C})$ can be obtained by defining the function c' such that $c'(N \backslash \{h\}) = \min\{c(N \backslash \{h\}), c(N)\}$ and, for all $S \subset N$ different from $N \backslash \{h\}$, $c'(S) = c(S)$. Then:

$$I(\boldsymbol{C}) = \max_{\boldsymbol{\chi}} \left\{ \sum_{S \in \bar{\mathcal{N}}} \chi_S (1 - \frac{c'(N \backslash S)}{c'(N)}) \middle| \forall i \in N, \sum_{S \in \bar{\mathcal{N}} : i \in S} \chi_S = 1, \chi_S \geq 0 \right\}$$

6.E Proof of Corollary 6.1

Consider the NIMBY problem $\sigma = (N, \boldsymbol{b}, \boldsymbol{C})$ and let $t \in \mathbb{R}_+$. Define $\sigma' = (N, \boldsymbol{b}, \boldsymbol{C'})$, where $\boldsymbol{C'} = \boldsymbol{C} + t\boldsymbol{I_n}$ and $\boldsymbol{I_n}$ denotes the identity matrix. For any building coalition in σ', we have $c'(S) = c(S) + t$ where $c(S)$ and $c'(S)$ denote the cost of the project for coalition S in σ and σ' respectively.

The linear programs defining $I(\boldsymbol{C'})$ writes:

$$I(\boldsymbol{C'}) = \max_{\boldsymbol{\chi}} \left\{ \sum_{S \in \bar{\mathcal{N}}} \chi_S \left(1 - \frac{c(N \backslash S) + t}{c(N) + t}\right) \middle| \forall i \in N, \sum_{S \in \bar{\mathcal{N}} : i \in S} \chi_S = 1, \chi_S \geq 0 \right\}$$

Defining $\tau = \frac{t}{c(N)}$ and substituting t, we get:

$$I(\boldsymbol{C'}) = \frac{1}{1+\tau} \max_{\boldsymbol{\chi}} \left\{ \sum_{S \in \bar{\mathcal{N}}} \chi_S \left(1 - \frac{c(N \backslash S)}{c(N)}\right) \middle| \forall i \in N, \sum_{S \in \bar{\mathcal{N}} : i \in S} \chi_S = 1, \chi_S \geq 0 \right\}$$

Therefore

$$I(\boldsymbol{C'}) = \frac{I(\boldsymbol{C})}{1+\tau}$$

6.F Proof of Corollary 6.2

Let $\sigma = (N, \boldsymbol{b}, \boldsymbol{C})$ and $\sigma' = (N, \boldsymbol{b}, \boldsymbol{C'})$ be two NIMBY problems meeting Assumptions 1 to 4. Define c and c' the cost function in the problem σ and σ' respectively, and assume

1. $c(N) = c'(N)$;
2. $\boldsymbol{C} \geq \boldsymbol{C'}$.

Let (LP) and (LP') be the linear programs defining respectively $I(\boldsymbol{C})$ and $I(\boldsymbol{C'})$ and let $\boldsymbol{\chi}$ be an optimal solution to (LP). In case, some additional coalitions appears in the set $\bar{\mathcal{N}}$ in σ, extend $\boldsymbol{\chi}$ by assigning a weight of 0 to them. This defines a feasible solution $\bar{\boldsymbol{\chi}}$ in (LP'). Because $\boldsymbol{C} \geq \boldsymbol{C'}$ and $c(N) = c'(N)$, we have $\forall S \in \bar{\mathcal{N}}, 1 - \frac{c(N \backslash S)}{c(N)} \leq 1 - \frac{c'(N \backslash S)}{c'(N)}$, so the objective of (LP') at $\bar{\boldsymbol{\chi}}$ is not lower than $I(\boldsymbol{C})$. Therefore, the value of (LP'), $I(\boldsymbol{C'})$, cannot be lower than $I(\boldsymbol{C})$.

6.G Proof of Corollary 6.3

In the linear case, Assumption 2 holds, and we explicitly compute the value of $I(\boldsymbol{C})$. In this section we will use the notion of balanced collections. A collection $\mathcal{B}$ of subsets of N is said to be balanced if and only if there exist strictly positive weights $\boldsymbol{\chi}^{\mathcal{B}} = (\chi_S^{\mathcal{B}})_{S \in \mathcal{B}}$ such that, for any $i \in N$, $\sum_{S \in \mathcal{B}: i \in S} \chi_S^{\mathcal{B}} = 1$. Denoting by $\mathbb{B}(\bar{\mathcal{N}})$ the set of balanced collections over N composed of elements of $\bar{\mathcal{N}}$ only, we can write:

$$I(\boldsymbol{C}) = \frac{1}{c(N)} \max_{\mathcal{B} \in \mathbb{B}(\bar{\mathcal{N}})} \left\{ \sum_{S \in \mathcal{B}} \chi_S^{\mathcal{B}} (c(N) - c(N \backslash S)) \right\}$$

We compute the costs saved by excluding a set of neighboring communities from the grand coalition $c(N) - c(N/S)$ for every $S \in \bar{\mathcal{N}}$. In the linear case, S is of size 1 or 2.

- Case $|S| = 1$. Some cost is saved by excluding a single community only if the community excluded is neighbor of one of the optimal hosts: 1 or n. The external cost δc is then saved: $c(N) - c(N \backslash S) = \delta c$ for $S \in \{\{2\}, \{n-2\}\}$.
- Case $|S| = 2$. Let S be a coalition of two communities $S = \{j, j+2\}$ neighbor of a community not located at the extreme of the line $j + 1 \in \{1, ..., n-2\}$. The cost saved by excluding S is δc because the optimal host becomes $j+1$ with $c(N \backslash S) = c$ while it is 1 or n in the grand coalition with $c(N) = c + \delta c$. For all other coalitions of size 2 neighbor of the same community, no cost is saved: $c(N) - c(N \backslash S) = 0$.

Therefore, for any $S \in \bar{\mathcal{N}}$, we have the corresponding values:

$$c(N) - c(N \backslash S) = \begin{cases} \delta c \text{ if } S \in \{2, n-1\} \\ \delta c \text{ if } S \in \{\{j, j+2\} | j \in \{1, ..., n-2\}\} \\ 0 \text{ otherwise} \end{cases}$$

$\bar{\mathcal{N}}$is a set of coalitions of no more than two players. Hence, for any balanced collection $\mathcal{B}$ of elements of $\bar{\mathcal{N}}$, there exists a partition of N into pairwise disjoint sets $N_1, ..., N_l, l = 0...L$ where each N_l with $l > 0$ is a coalition of at least three communities such that $\mathcal{B}$ consists of full cycles on each N_l and a partition of N_0 (Balinski, 1970, as stated in Le Breton and Weber, 1995: 316). Because no cycle can be formed out of elements of $\bar{\mathcal{N}}$ in the linear case, all balanced collections over $\bar{\mathcal{N}}$ are partitions. In summary, we are interested in finding partitions $\mathcal{P}$ of N, composed with elements of $\bar{\mathcal{N}}$ which maximize $\sum_{S \in \mathcal{P}}(c(N) - c(N \backslash S))$. We now explain how to find such optimal partitions.

First, for any partition involving coalitions in which 2 or $n-1$ belongs to a two-agent coalition, we weakly improve on the objective by splitting such coalitions into singletons. Hence, we can restrict our attention to coalitions in which such communities appear as singletons. The construction of an optimal partition then consists in maximizing the number of coalitions of the form $\{\{j, j+2\} | j \in \{1, ..., n-2\}\}$. In the case $n \in \{4, 5, 6, 7\}$, such optimal partitions are trivial as soon as communities 2 and $n-1$ appear as singletons. Figure 6.G.1 presents optimal partitions and the corresponding value of $\sum_{S \in \mathcal{P}}(c(N) - c(N \backslash S))$.

For any $n > 7$, we know that n can be decomposed as $n = 4k + i$, $k \in \mathbb{N}$ and $i \in 0, 1, 2, 3$. According to this decomposition, an optimal partition can be found by combining the initial patterns above and the iterative pattern presented in Figure 6.G.2 which maximizes the value that can be obtained by adding 4 communities to the initial pattern.

We eventually find the following optimal partitions:

- If $n = 4k, k \in \mathbb{N}$, $\mathcal{P} = \{\{1\}, \{2\}, \{n-1\}, \{n\}\} \cup_{j=1}^{k-1} \{\{4j-1, 4j+1\}, \{4j, 4j+2\}\}$
- If $n = 4k+1, k \in \mathbb{N}$, $\mathcal{P} = \{\{1, 3\}, \{2\}, \{n-1\}, \{n\}\} \cup_{j=1}^{k-1} \{\{4j, 4j+2\}, \{4j+1, 4j+3\}\}$
- If $n = 4k+2, k \in \mathbb{N}$, $\mathcal{P} = \{\{1, 3\}, \{2\}, \{n-1\}, \{n-2, n\}\} \cup_{j=1}^{k-1} \{\{4j, 4j+2\}, \{4j+1, 4j+3\}\}$
- If $n = 4k+3, k \in \mathbb{N}$, $\mathcal{P} = \{\{1, 3\}, \{2\}, \{4\}, \{n-1\}, \{n-2, n\}\} \cup_{j=1}^{k-1} \{\{4j+1, 4j+3\}, \{4j+2, 4j+4\}\}$

And the associated values are:

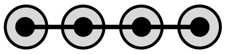
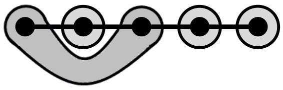
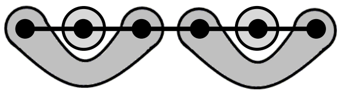
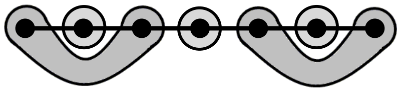
Case	Optimal partition	Value
n=4	1 2 3 4 	2δc
n=5	1 2 3 4 5 	3δc
n=6	1 2 3 4 5 6 	4δc
n=7	1 2 3 4 5 6 7 	4δc

Figure 6.G.1: Initial patterns.

The reasoning adopted for finding the optimal partitions consists in considering all possible cases. We detail the case $n = 7$. First, we know that there is always an optimal partition containing $\{2\}$ and $\{6\}$ as singletons. The value associated with each is δc. The value associated with any other single individual is 0 whereas the value associated with any pair of $\bar{\mathcal{N}}$ is δc. An optimal partition thus contains as many pairs of $\bar{\mathcal{N}}$ as possible. This is achieved with the partition $\mathcal{P} = \{\{1,3\},\{2\},\{4\},\{6\},\{5,7\}\}$.

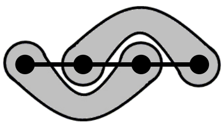


Figure 6.G.2: Iterative pattern.

$$I(\boldsymbol{C}) = \begin{cases} \frac{n}{2}\frac{\delta}{1+\delta}\text{if } n = 4k, k \in \mathbb{N} \\ \frac{n+1}{2}\frac{\delta}{1+\delta} \text{ if } n = 4k+1, k \in \mathbb{N} \\ \frac{n+2}{2}\frac{\delta}{1+\delta}\text{if } n = 4k+2, k \in \mathbb{N} \\ \frac{n+1}{2}\frac{\delta}{1+\delta}\text{if } n = 4k+3, k \in \mathbb{N} \end{cases}$$

The condition on δ expressed in Corollary 6.3 directly follows from the comparison of $I(\boldsymbol{C})$ with 1.

6.H Proof of Proposition 6.4

Let R be an exogenous expectation formation rule and v^R its associated characteristic function. We want to show that under Assumptions 1, 2, 4 and 5, the R-core is non-empty if and only if $I(\boldsymbol{C}) \geq 1$. We extend the proof of Propositions 6.1 and 6.2.

First, we eliminate redundant constraints in the system defining the core. We distinguish between building and non-building coalitions. $NB = \{T \subset N | b(T) < c(T)\}$ is the set of non-building coalitions. Replicating the proof of Proposition 6.1, the constraints for building coalitions can be restricted to $\{N \backslash S | S \in \bar{\mathcal{N}}\}$. However, the constraints for non-building coalitions cannot be reduced to individual rationality: an allocation $\boldsymbol{x}$ is in the R-core $\mathcal{C}^R$ if and only if

$$\sum_{i \in N} x_i = v(N) \tag{6.16}$$

$$\forall S \in NB, \sum_{i \in S} x_i \geq v^R(S) \tag{6.17}$$

$$\forall i \in N, x_i \leq b_i \tag{6.18}$$

$$\forall S \in \mathring{\mathcal{N}}, \sum_{i \in S} x_i \leq b(S) - (c(N) - c(N \backslash S)) \tag{6.19}$$

where the constraints (6.17) contain the individual rationality constraints. We consider the linear program $(LP5)$:

$$\max_{\boldsymbol{x}} \left\{ \sum_{i \in N} x_i \middle| \forall S \in \bar{\mathcal{N}}, \sum_{i \in S} x_i \leq b(S) - (c(N) - c(N \backslash S)) \text{ and } \forall S \in NB, \sum_{i \in S} x_i \geq v^R(S) \right\}$$

The R-core $\mathcal{C}^R$ is non-empty if and only if $(LP5)$ is feasible and reaches a value higher than $v(N)$. We first note that such a program would always be feasible under Assumption 5.

Second, as in the proof of Proposition 6.2, we can show that the constraints (6.17) are never binding under Assumption 5.

Let $\boldsymbol{x}^*$ be an optimal solution to the above linear program and assume there exists $i_0 \in N$ such that:

$$x^*_{i_0} < b_{i_0} + \min_{T \in \bar{\mathcal{N}}: i_0 \in T} \{c(N \backslash T) - c((N \backslash T) \cup \{i_0\})\} \tag{6.20}$$

Then we can increase $x^*_{i_0}$ by some $\epsilon > 0$ such that:

$$x^*_{i_0} + \epsilon < b_{i_0} + \min_{T \in \bar{\mathcal{N}}: i_0 \in T} \{c(N \backslash T) - c((N \backslash T) \cup \{i_0\})\} \tag{6.21}$$

Such an increase improves on the objective. We shall show that it also leads to a feasible solution. First, it is straightforward to see that the constraints (6.17) are met. We concentrate on the remaining constraints.

Let $S \in \bar{\mathcal{N}}$ with at least two communities, such that $i_0 \in S$. Because $S \in \bar{\mathcal{N}}$, $S \backslash \{i_0\}$ also pertains to $\bar{\mathcal{N}}$. By feasibility of $\boldsymbol{x}^*$, we have:

$$\sum_{i \in S \backslash \{i_0\}} x^*_i \leq b(S \backslash \{i_0\}) - (c(N) - c((N \backslash S) \cup \{i_0\})) \tag{6.22}$$

Summing inequalities (6.21) and (6.22), we get:

$$\sum_{i \in S} x^*_i + \epsilon < b(S) - (c(N) - c((N \backslash S) \cup \{i_0\})) + \min_{T \in \bar{\mathcal{N}}: i_0 \in T} \{c(N \backslash T) - c((N \backslash T) \cup \{i_0\})\}$$

Therefore,

$$\sum_{i \in S} x^*_i + \epsilon < b(S) - (c(N) - c(N \backslash S))$$

Therefore, all the constraints involving x_{i_0} are met. This contradicts the optimality of $\boldsymbol{x}^*$. Hence, inequality (6.20) cannot hold by contradiction. We have:

$$x^*_{i_0} \geq b_{i_0} + \min_{T \in \bar{\mathcal{N}}: i_0 \in T} \{c(N \backslash T) - c((N \backslash T) \cup \{i_0\})\}$$

Besides, as established in the proof of Proposition 6.2:

$$\forall S \in \bar{\mathcal{N}} : i_0 \in S, c(N \backslash S) - c((N \backslash S) \cup \{i_0\}) \geq - \max_{j \in N \backslash \{i_0\}} c_{j i_0}$$

so $x^*_{i_0} \geq b_{i_0} - \max_{j \in N \backslash \{i_0\}} c_{j i_0}$ and $\forall S \in \{T \subset N | b(T) < c(T)\}$, $\sum_{i \in T} x^*_i \geq b(T) - \sum_{i \in T} \max_{j \in N \backslash \{i\}} c_{ji}$. Hence, using Assumption 5, $\sum_{i \in T} x^*_i \geq v^R(T)$.

The constraints (6.17) can then be removed from the linear program (LP5) without changing its value. This leads us back to the linear program ($LP2$) and the proof of Proposition 6.2 applies.

6.I Proof of Corollary 6.4

The cost of the project on a graph depends on the minimal degree of this graph. For any $S \subseteq N$, we denote by $\underline{d}(S)$ the minimal degree of the graph induced by S on G. Rewriting the condition $I(\boldsymbol{C}) \geq 1$, we get the following condition on δ:

$$\delta \leq \bar{\delta}(\boldsymbol{G}) = \frac{1}{max_{\boldsymbol{\chi}} \{\sum_{S \in \bar{\mathcal{N}}} \chi_S(\underline{d}(N) - \underline{d}(N \backslash S)) | \forall i \in N, \sum_{S: i \in S} \chi_S = 1, \chi_S \geq 0\} - \underline{d}(N)}$$

We want to show $\bar{\delta}(\boldsymbol{G}) > 0$. Let $h \in \mathcal{H}$ be an optimal host in N and $j \in \mathring{\mathcal{N}}(h)$.[22] Consider the following partition: $\{\mathring{\mathcal{N}}(h), S_j, N \backslash (\mathring{\mathcal{N}}(h) \cup S_j)\}$, where $S_j = \mathring{\mathcal{N}}(j) \backslash \mathring{\mathcal{N}}(h)$ is the strict neighborhood j from which we withdraw members of $\mathring{\mathcal{N}}(h)$. A feasible solution $\boldsymbol{\chi'}$ associated with this partition is defined as follows:

- $\chi'_{\mathring{\mathcal{N}}(h)} = 1$;
- $\chi'_{S_j} = 1$;
- $\chi'_{N \backslash (\mathring{\mathcal{N}}(h) \cup S_j)} = 1$;
- $\chi'_S = 0$ for all other coalitions

We compute the value of this linear program at this feasible solution. First, we know that $|\mathring{\mathcal{N}}(h)| = \underline{d}(N)$. Hence community j has at most $\underline{d}(N) - 1$ neighbors in $\mathring{\mathcal{N}}(h)$. The withdrawal of its neighbors in S_j therefore leads to a graph with a degree of at least $\underline{d}(N) - 1$. Hence, $\underline{d}(N) - 1 \geq \underline{d}(N \backslash S_h)$, which implies that $\underline{d}(N) - \underline{d}(N \backslash S_h) \geq 1$. Second, we have $\underline{d}(N \backslash \mathring{\mathcal{N}}(h)) = 0$; hence, $\underline{d}(N) - \underline{d}(N \backslash \mathring{\mathcal{N}}(h)) = \underline{d}(N)$. Finally, as we have $h \in S_j$ by construction, the minimal degree of $\mathring{\mathcal{N}}(h) \cup S_h$ is at most $\underline{d}(N)$; hence $\underline{d}(N) - \underline{d}(\mathring{\mathcal{N}}(h) \cup S_h) \geq 0$. The value associated with the feasible solution $\boldsymbol{\chi'}$ is $\underline{d}(N) + 1$, hence the optimal value of the linear program defining $\bar{\delta}(\boldsymbol{G})$ can only be higher than it. Therefore, $\bar{\delta}(\boldsymbol{G}) > 0$.

6.J Codes

These are the codes used with the software R.

[22] We assume here that $|\mathring{\mathcal{N}}(h)| > 0$. If it is not, the core is always non-empty ($\bar{\delta}(\boldsymbol{G}) = +\infty$).

6.J.1 NIMBY problems of graphs

```
#######################################################
# Finding the critical value of delta on a graph #
#######################################################
rm(list=ls())
library(linprog)
#This function returns the critical value deltac
#Input: Adjacency matrix
#Output: deltac
deltac<-function(M){
n<-dim(M)[1]
A<-NULL
b<-NULL
for(i in 1:dim(M)[1]){
neighbors<-which(M[i,]>0,arr.ind=TRUE)
di<-length(neighbors)
for(k in 0:(min(rowSums(M))-1)){
ExcludableCoalitions<-matrix(neighbors[combn(1:di,di-k)],ncol=choose(di,di-k))
for(l in 1:choose(di,di-k)){
constraint<-rep(0,dim(M)[1])
constraint[ExcludableCoalitions[,l]]<-1
A<-rbind(A,constraint)
b<-cbind(b,min(rowSums(M))-k)
}
}
}
rownames(A)<-NULL
A<-rbind(A,diag(1,n))
b<-c(t(b),rep(0,n))
chi<-solveLP(b,rep(1,2*n),rbind(t(A),t(A)),maximum = TRUE, const.dir =
c(rep("<=",n),rep(">=",n)),lpSolve=FALSE,solve.dual = FALSE)$solution
1/(chi%*%b-min(rowSums(M)))
}

#Case A of Figure 6.3
MA<-matrix(c(
0,1,0,0,0,0,
1,0,1,0,0,0,
0,1,0,1,0,0,
0,0,1,0,1,0,
0,0,0,1,0,1,
```

```
0,0,0,0,1,0), nrow = 6, ncol = 6)
deltac(MA)
```

6.J.2 NIMBY problem on a French administrative unit

The GIS data used is the GEOFLA® Communes database. It is publicly available at http://professionnels.ign.fr/geofla.

```
######################################
# Computation of I(C) #
######################################
rm(list=ls())
memory.size(8000)
require("rgdal")
require("rgeos")
require("spdep")
require("linprog")

# Importation of the GIS data
mun <- readOGR(dsn="COMMUNES", layer="COMMUNE")
numdep<-"31"
dref<-1000
mundep<-mun[as.character(mun@data$CODE_DEPT)==numdep,]
dep<-gUnaryUnion(mundep)
centroids<-SpatialPoints(cbind(mundep@data$X_CENTROID*100,mundep@data$Y_CENTROID*100),
mundep@proj4string)
impactArea<-gBuffer(centroids,width=dref,byid=TRUE,id=rep(" ",length(centroids)))

#Derivation of the matrix C
M<-rep(0,length(mundep))%*%t(rep(0,length(mundep)))
colnames(M)<-(mundep@data[,1]-1)
for(i in 1:length(mundep)){
neighbors<-gArea(gIntersection(impactArea[i],mundep,byid=TRUE),byid=TRUE)
M[i,as.character(as.numeric(names(neighbors)))]<-as.numeric(neighbors)
}
cN<-min(rowSums(M))

#Computation of I(C)
storage.mode(M) <- "integer"
A<-rep(0L,800000)%*%t(rep(0L,dim(M)[1]))
```

```
b<-rep(0L,800000)
compteur<-1
for(i in 1:dim(M)[1]){
neighbors<-which(M[i,]>0, arr.ind=TRUE)
neighbors<-neighbors[neighbors!=i]
di<-length(neighbors)
ci<-sum(M[i,])
if(di>0){
for(k in 1:di){
ExcludedAgents<-matrix(neighbors[combn(1:di,k)],,ncol=choose(di,k))
for(l in 1:dim(ExcludedAgents)[2]){
if(ci-sum(M[i,ExcludedAgents[,l]])<=cN){
A[compteur,ExcludedAgents[,l]]<-1
b[compteur]<-cN-ci+sum(M[i,ExcludedAgents[,l]])
compteur<-compteur+1 } } } } }
A<-A[1:compteur-1,]
b<-b[1:compteur-1]
A<-rbind(A,diag(1,dim(M)[1]))
b<-c(b,rep(0,dim(M)[1]))
I(C)<-solveLP(b,rep(1,dim(M)[1]),t(A),maximum = TRUE,const.dir = rep("=",dim(M)[1])
,maxiter=300000,lpSolve=TRUE,solve.dual = FALSE,verbose=1)$opt/cN
```

Conclusion

Main results

Let us recall our original and overarching questioning which regards the assessment of the fair character of public policies whose distributive impacts are well understood. I first argued that a relevant assessment cannot avoid paying attention to the political community concerned. To illustrate this, I proposed three broad perspectives on the democratic political community and argued that a relevant approach would significantly differ in these perspectives. In particular, it appeared that the degree of ethical reasoning involved, the underlying empirical material and the formulation of the assessment would differ significantly depending on the perspective taken. I next focused on the deliberative perspective and discussed how two particular branches of the social choice literature are susceptible of bringing insights and proposals that could be relevant to public deliberations: the empirical social choice literature and the theory of fair allocations. In a nutshell, the latter provides a theory of the formation of reasoned judgments in the course of deliberations, while the former can be used to assess the relevance of the theories proposed.

The rest of the thesis is intended to illustrate this approach in the context of the NIMBY problem. The results presented in Chapter 4 further suggest a some fair mechanisms along with a justification for them. A selection among these may be justified based on considerations of compensation or reward. Observing the judgments expressed or revealed on these problems and others with a similar structure in Chapters 4 and 5, I conclude that the welfare egalitarian allocation rules is the rule that attracts most support, despite some clues that these judgments may still rely on heuristics and then be suceptible of change. The main conclusion that I may propose out of these results is that, absent externalities, mechanisms such as the first or second-price auctions may be considered as reasonably fair policy options, in the sense that can be justified on the basis of widely shared principles, but they are not the only ones. However, I insisted on two main limits. First the analysis did not consider the possibility of externalities, which brings additional complications as suggested in Chapter 6. Second, observation suggests that the judgments reported are not in reflective equilibrium, which limits their relevance for motivating a social judgment.

Applied and interdisciplinary perspectives

In the end, an application of the proposed framework to actual policymaking would have to tackle two challenges. The first is a need for structuring the existing information in a way that is useful to the fairness assessment of actual public policies. In Chapter 2, I claimed that the theory of fair allocations can assist the formation of reasoned judgments over actual policies and I proposed to organize the results of these theories in a comprehensive framework that encompasses reflections on the ethical relevance of different features of a problem. While necessary, this also requires the consideration of many possible problems and a lot of information. The best way to structure and present such information remains open.

The second challenge relates to the need for designing protocols for the observation of judgments that are relevant to the deliberative perspective. In Chapter 1, I argued that judgments may vary according to several dimensions. More knowledge is required regarding the dimensions along which judgments are most dependent. This also requires a discussion of the conditions which are most relevant in a given perspective. In the deliberative perspective, I proposed that reasoned judgments shall be preferred over spontaneous judgments. In practice, however, protocols for the observation of reasoned judgments would have to conciliate the complexity of axiomatic analyses with the limited willingness of individuals to perform complicated reasoning. In the attempts presented, I proposed to simplify the phrasing of the axioms, which eventually hardened the interpretation of the results. Bridging the gap between the actual judgments and the theories of fairness represents a significant challenge. As individual motivations for justification may be strongest in social interactions, I stressed the potential interest to consider the results and methods in other disciplines.

Beyond the many possible perspectives on the political community and from a pragmatic viewpoint, I eventually put forward how this approach can contribute to policy making in practice. A first possibility is to contribute to the design of robust justification for policies. In a given context, contemplating the many possible arguments and their logical relationship allows to clearly identify the normative underpinnings of a given policy choice. In particular, I argued that these results may ease the argumentation through the identification of appropriate particular cases or principles for counterarguments.

A second possibility is to contribute to the design of innovative justification schemes which, in turn allows for the considerations of broader options in policy making. This approach may actually correspond to what has been achieved on some particular problems such as school choice or organ donation. But the potential application of these approaches may extend far beyond these problems. Referring to environmental policies, Amartya Sen (1995) emphasized that "the threats that we face call for organized international action as well as changes in national policies, particularly for better reflecting social costs in prices and incentives. But they are also dependent on value formation, related to public discussions, both for their influence on individual behavior and for bringing about policy changes through the political process." As such, the approach proposed can contribute to identify significant policy changes from the status quo, and bring a justification for them and evidence of its ability to convince.

A third possibility that is related to the implementation of these relates to the design of communication strategies. In particular, the potential of the empirical observation of the judgments in conditions of impartiality to contribute to value change could be hypothesized from the intuition that knowledge of other's view may constitute an important input in shaping individuals' judgments. This remains to investigate.

In all these case, the potential contributions presented mainly lie in increasing the acceptance of public policies and reducing conflicts in society. Despite the widespread skepticism toward the possibility and interest of knowledge on normative matters, I hope to have convinced the reader that a joint understanding of the logics of norms and the judgments held in society is a relevant knowledge basis for policymaking. This being said, I wish to conclude by recollecting what the resulting assessments do *not* seek to replace. At an individual level, a full-fledged moral inquiry and at the collective level, the constitution of a political consensus, both will remain required regardless of the level of understanding achieved on a particular problem, through this approach or any other.

Bibliography

[1] David Aadland and Van Kolpin. Shared irrigation costs: an empirical and axiomatic analysis. *Mathematical Social Sciences*, 35(2):203–218, 1998.

[2] Fernando Aguiar, Alice Becker, and Luis Miller. Whose impartiality? an experimental study of veiled stakeholders, involved spectators and detached observers. *Economics and Philosophy*, 29(02):155–174, 2013.

[3] Varouj A. Aivazian and Jeffrey L. Callen. The coase theorem and the empty core. *Journal of Law and Economics*, 24(1):175–181, 1981.

[4] Yoram Amiel and Frank Cowell. Measurement of income inequality experimental test by questionnaire. *Journal of Public Economics*, 47:3–26, 1999.

[5] Elizabeth Anderson. *Value in ethics and economics*. Harvard University Press, 1995.

[6] James Andreoni and John Miller. Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, pages 737–753, 2002.

[7] Anthony Appiah. *Experiments in ethics*. Harvard University Press, 2008.

[8] Kenneth J Arrow. *Social choice and individual values, 3rd edition, 2012*, volume 12. Yale university press, 1951.

[9] R. J. Aumann. A survey of cooperative games without side payments. In Martin Shubik, editor, *Essays in Mathematical Economics in Honor of Oskar Morgenstern*, page 3–27. Princeton University Press, 1969.

[10] Linda Babcock and George Loewenstein. Explaining bargaining impasse: the role of self-serving biases. *Journal of Economic Perspectives*, 11(1):109–126, 1997.

[11] Linda Babcock, George Loewenstein, and Samuel Issacharoff. Creating convergence: Debiasing biased litigants. *Law & Social Inquiry*, 22(4):913–925, 1997.

[12] Yoram Bachrach, Elkind Edith, Reshef Meir, Dmitrii Pasechnik, Michael Zuckerman, Jorg Rothe, and Jeffrey S. Rosenschein. The cost of stability in coalitional games. In *SAGT' 09: 2nd international symposium on algorithmic game theory*, pages 122–134, 2009.

[13] M. Balinski. On the maximum matching set, minimal covering. In *Proceedings of the Princeton symposium on mathematical programming*, 1970.

[14] Salvador Barberá, Dolors Berga, and Bernardo Moreno. Domains, ranges and strategy-proofness: the case of single-dipped preferences. *Social Choice and Welfare*, 39(2-3):335–352, 2012.

[15] Steven R Beckman, John P Formby, W James Smith, and Buhong Zheng. Envy, malice and pareto efficiency: An experimental examination. *Social Choice and Welfare*, 19(2):349–367, 2002.

[16] Theodore Bergstrom, Lawrence Blume, and Hal Varian. On the private provision of public goods. *Journal of Public Economics*, 29:25–49, 1986.

[17] Cristina Bicchieri. *The grammar of society The nature and dynamics of social norms*. Cambridge University Press, 2006.

[18] Cristina Bicchieri and Alex Chavez. Behaving as expected: Public information and fairness norms. *Journal of Behavioral Decision Making*, 23:161–178, 2010.

[19] Francis Bloch and Anne Van den Nouweland. Expectation formation rules and the core of partition function games. *Games and Economic Behavior*, 88:339–353, 2014.

[20] Gary E Bolton and Axel Ockenfels. Erc: A theory of equity, reciprocity, and competition. *American Economic Review*, pages 166–193, 2000.

[21] Tilman Borgers, Daniel Krahmer, and Roland Strausz. *An introduction to the theory of mechanism design*. Oxford University Press, 2015.

[22] Samuel Bowles. Endogenous preferences: The cultural consequences of markets and other economic institutions. *Journal of Economic Literature*, pages 75–111, 1998.

[23] Kjell Arne Brekke, Snorre Kverndokk, and Karine Nyborg. An economic model of moral motivation. *Journal of public economics*, 87(9):1967–1983, 2003.

[24] Colin Camerer and Ernst Fehr. Foundations of human sociality economic experiments and ethnographic evidence from fifteen small-scale societies. In J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, and H. Gintis, editors, *Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists*. Oxford University Press, 2004.

[25] Colin F. Camerer. *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press, 2003.

[26] Colin F Camerer and Ernst Fehr. Measuring social norms and preferences using experimental games: A guide for social scientists. 2002.

[27] Alexander W. Cappelen, Astri Drange Hole, , Erik O. Sorensen, and Bertil Tungodden. The pluralism of fairness ideals: An experimental approach. *American Economic Review*, 97(3):818–827, 2007.

[28] Alexander W. Cappelen, James Konow, Erik O. Sorensen, and Bertil Tungodden. Just luck: An experimental study of risk-taking and fairness. *American Economic Review*, 103(4):1398–1413, 2013.

[29] Alexander W. Cappelen, Erik O. Sorensen, and Bertil Tungodden. Responsibility for what? fairness and individual responsibility. *European Economic Review*, 54:429–441, 2010.

[30] Alexander W. Cappelen and Bertil Tungodden. Distributive interdependencies in liberal egalitarianism. *Social Choice and Welfare*, 36:35–47, 2011.

[31] Fredrik Carlsson, Mitesh Kataria, Alan Krupnick, Elina Lampi, Åsa Löfgren, Ping Qin, and Thomas Sterner. A fair share: Burden-sharing preferences in the united states and china. *Resource and Energy Economics*, 35(1):1–17, 2013.

[32] Parkash Chander and Henry Tulkens. The core of an economy with multilateral environmental externalities. *International Journal of Game Theory*, 26:379–401, 1997.

[33] Todd L. Cherry, Peter Frykblom, and Jason F. Shogren. Hardnose the dictator. *American Economic Review*, 92(2):1218–1221, 2002.

[34] Ronald Harry Coase. The problem of social cost. *Journal of Law and Economics*, 3:1, 1960.

[35] Norman Daniels. *Justice and justification: Reflective equilibrium in theory and practice*, volume 22. Cambridge Univ Press, 1996.

[36] Geoffroy De Clippel and Roberto Serrano. Marginal contributions and externalities in the value. *Econometrica*, 76(6):1413–1436, 2008.

[37] Pierre Dehez. Cooperative provision of indivisible public goods. *Theory and Decision*, 74:13–29, 2013.

[38] Udo Ebert. Taking empirical studies seriously: the principle of concentration and the measurement of welfare and inequality. *Social Choice and Welfare*, 32(4):555–574, 2009.

[39] Gerald Eisenkopf, Urs Fischbacher, and Franziska Föllmi-Heusi. Unequal opportunities and distributive justice. *Journal of Economic Behavior & Organization*, 93:51–61, 2013.

[40] Jon Elster. *Social norms and the explanation of behavior*, pages 195–217. The Oxford Handbook of Analytical Sociology. Oxford University Press, 2009.

[41] Christoph Engel. Dictator games: a meta study. *Experimental Economics*, 14(4):583–610, 2011.

[42] Armin Falk and Urs Fischbacher. Distributional consequences and intentions in a model of reciprocity. *Annales d'Economie et de Statistique*, 63-64:111–129, 2001.

[43] Ernst Fehr and Urs Fischbacher. Third-party punishment and social norms. *Evolution and human behavior*, 25(2):63–87, 2004.

[44] Ernst Fehr and Klaus M Schmidt. A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3):817–868, 1999.

[45] James Gordon Finlayson. *Habermas - A very short introduction*. Oxford University Press, 2005.

[46] James S Fishkin and Robert C Luskin. Experimenting with a democratic ideal: Deliberative polling and public opinion. *Acta Politica*, 40(3):284–298, 2005.

[47] Raymond Fisman, Shachar Kariv, and Daniel Markovits. Individual preferences for giving. *The American Economic Review*, 97(5):1858–1876, 2007.

[48] Marc Fleurbaey. *Fairness, Responsibility and Welfare*. Oxford University Press, 2008.

[49] Marc Fleurbaey and François Maniquet. Handbook of social choice and welfare. volume 2, chapter 22 - Compensation and Responsibility, pages 507–604. Elsevier, 2011.

[50] Marc Fleurbaey and François Maniquet. *A theory of fairness and social welfare*. Cambridge University Press, 2011.

[51] Marc Fleurbaey, Kotaro Suzumura, and Koichi Tadenuma. The informational basis of the theory of fair allocation. *Social Choice and Welfare*, 24(2):311–341, 2005.

[52] Duncan K Foley. Resource allocation and the public sector. *Yale Economic Essays*, 7(1):45–98, 1967.

[53] Michel Forsé and Maxime Parodi. *Une théorie empirique de la justice sociale*. Hermann, 2010.

[54] Vito Fragnelli and Stefano Gagliardo. Cooperative models for allocating an object. *Economics Letters*, 117(1):227–229, 2012.

[55] Bruno S. Frey and Felix Oberholzer-Gee. The cost of price incentives: An empirical analysis of motivation crowding-out. *The American Economic Review*, 87(4):746–755, 1997.

[56] Norman Frohlich and Joe A Oppenheimer. *Choosing justice: An experimental approach to ethical theory*, volume 22. Univ of California Press, 1992.

[57] Simon Gächter and Arno Riedl. Dividing justly in bargaining problems with claims. *Social Choice and Welfare*, 27(3):571–594, 2006.

[58] Wulf Gaertner and Erik Schokkaert. *Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice*. Cambridge University Press, 2012.

[59] Michel X. Goemans and Martin Skutella. Cooperative facility location games. *Journal of Algorithms*, 50:194–214, 2004.

[60] Natalie Gold, Briony D Pulford, and Andrew M Colman. Do as i say, don't do as i do: Differences in moral judgments do not translate into differences in decisions in real-life trolley problems. *Journal of Economic Psychology*, 47:50–61, 2015.

[61] Jerry R. Green and Jean-Jacques Laffont. *Incentives in Public Decision-Making*. Amsterdam: North-Holland, 1979.

[62] Jonathan Haidt. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, 108(4):814–834, 2001.

[63] J Hamilton. Environmental equity and the siting of hazardous waste facilities in oecd countries. In Y. Serret and N. Johnstone, editors, *The International Yearbook of Environmental and Resource Economics 2005/2006*, chapter 3, pages 227–285. USA and Paris: Edward Elgar Publishing and OECD, 2006.

[64] Rögnvaldur Hannesson. *The privatization of the oceans*. Mit Press, 2004.

[65] John C. Harsanyi. Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy*, 61:434–435, 1953.

[66] Hélène Hermansson. The ethics of nimby conflicts. *Ethical theory and moral practice*, 10(1):23–34, 2007.

[67] Carmen Herrero, Juan D. Moreno-Ternero, and Giovanni Ponti. On the adjudication of conflicting claims: An experimental study. *Social Choice and Welfare*, 34:145–179, 2010.

[68] Jeremiah Hurley, Neil J. Buckley, Catherine Cuff, Mita Giacomini, and David Cameron. Judgments regarding the fair division of goods: the impact of verbal versus quantitative descriptions of alternative divisions. *Social Choice and Welfare*, 37:341–372, 2011.

[69] Herbert Inhaber. *Slaying the NIMBY dragon*. Transaction Publishers, 1998.

[70] Andrea Isoni, Anders Poulsen, Robert Sugden, and Kei Tsutsui. Efficiency, equality, and labeling: An experimental investigation of focal points in explicit bargaining. *The American Economic Review*, 104(10):3256–3287, 2014.

[71] Dan M Kahan. The logic of reciprocity: Trust, collective action, and law. In Herbert Gintis, Samuel Bowles, Robert Boyd, and Ernst Fehr, editors, *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*, chapter 12, pages 339–378. MIT Press, 2005.

[72] Daniel Kahneman, Jack L Knetsch, and Richard Thaler. Fairness as a constraint on profit seeking: Entitlements in the market. *The American economic review*, pages 728–741, 1986.

[73] L Robin Keller and Rakesh K Sarin. Fair processes for societal decisions involving distributional inequalities. *Risk Analysis*, 15(1):49–59, 1995.

[74] Hyungjun Kim. Population monotonic rules for fair allocation problems. *Social Choice and Welfare*, 23(1):59–70, 2004.

[75] Paul R. Kleindorfer and Murat R. Sertel. Auctioning the provision of an indivisible public good. *Journal of Economic Theory*, 64:20–34, 1994.

[76] James Konow. Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review*, 90(4):1072–1091, 2000.

[77] James Konow. Which is the fairest one of all? a positive analysis of justice theories. *Journal of Economic Literature*, 41:1188–1239, 2003.

[78] James Konow. Is fairness in the eye of the beholder? an impartial spectator analysis of justice. *Social Choice and Welfare*, 33:101–127, 2009.

[79] James Konow. Adam smith and the modern science of ethics. *Economics and Philosophy*, 28(03):333–362, 2012.

[80] Michal Krawczyk. A glimpse through the veil of ignorance: Equality of opportunity and support for resdistribution. *Journal of Public Economics*, 94:131–141, 2010.

[81] Stephan Kroll, Todd L. Cherry, and Jason F. Shogren. The impact of endowment heterogeneity and origin on contributions in best-shot public good games. *Experimental Economics*, 10:411–428, 2007.

[82] Erin L. Krupka and Roberto A. Weber. Identifying social norms using coordination games: why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3):495–524, 2013.

[83] Howard Kunreuther, Paul Kleindorfer, and Rudy Knez, Peter J.and Yaksick. A compensation mechanism for siting noxious facilities: Theory and experimental design. *Journal of Environmental Economics and Management*, 14:371–383, 1987.

[84] Howard Kunreuther and Paul R. Kleindorfer. A sealed-bid auction mechanism for siting noxious facilities. *The American Economic Review*, 76(2):295–299, 1986.

[85] Jean-Jacques Laffont. *Effets externes et théorie économique.* Collection Monographies du Séminaire d'Econométrie, 1977.

[86] Julian Lamont. The concept of desert in distributive justice. *The philosophical quaterly*, 44(174), 1994.

[87] Andreas Lange, Andreas Löschel, Carsten Vogt, and Andreas Ziegler. On the self-interested use of equity in international climate negotiations. *European Economic Review*, 54(3):359–375, 2010.

[88] Jeremi Laurent-Lucchetti and Justin Leroux. Choosing and sharing. *Games and Economic Behavior*, 73(1):296–300, 2011.

[89] Jeremy Laurent-Lucchetti and Justin Leroux. Lindhal prices solve the nimby problem. *Economic Bulletin*, 30(3):2457–2463, 2010.

[90] Michel Le Breton, Juan D Moreno-Ternero, Alexei Savvateev, and Shlomo Weber. Stability and fairness in models with a multiple membership. *International Journal of Game Theory*, 42:673–694, 2013.

[91] Michel Le Breton and Shlomo Weber. Stability of coalition structures and the principle of optimal partitioning. In W. Barnett, H. Moulin, M. Salles, and N. Schofield, editors, *Social Choice, Welfare and Ethics*, pages 301–319. 1995.

[92] Michel Le Breton and Shlomo Weber. The art of making everybody happy: How to prevent a secession. *IMF Staff Papers*, 50:p. 403–435, 2003.

[93] Raul P. Lejano and Climis A. Davos. Siting noxious facilities with victim compensation: n-person games under transferable utility. *Socio-Economic Planning Sciences*, 35:109–124, 2001.

[94] Steven D Levitt and John A List. What do laboratory experiments measuring social preferences reveal about the real world? *The journal of economic perspectives*, pages 153–174, 2007.

[95] John A List. On the interpretation of giving in dictator games. *Journal of Political Economy*, 115(3):482–493, 2007.

[96] Wolfgang J. Luhan, Odile Poulsen, and Michael W. M. Roos. Unstructured bargaining over an endogenously produced surplus and fairness ideals an experiment. Ruhr Economic Papers 438, 2013.

[97] V Manjunath. Efficient and strategy-proof social choice when preferences are single-dipped. *International Journal of Game Theory*, 43(3):579–597, 2014.

[98] Hugo Mercier and Helen Landemore. Reasoning is for arguing: Understanding the success and failures of deliberation. *Political Psychology*, 33(2):243–258, 2012.

[99] Hugo Mercier and Dan Sperber. Why do humans reason? arguments for an argumentative theory. *Behavioral and brain sciences*, 34(02):57–74, 2011.

[100] David M Messick. Equality as a decision heuristic. In Barbara A. Mellers and Jonathan Baron, editors, *Psychological Perspectives on Justice*, Cambridge series on judgment and decision making, chapter 2, page 11. Cambridge University Press, 1993.

[101] David Miller. Distributive justice: What the people think. *Ethics*, 102(3):555–593, 1992.

[102] Deborah Minehart and Zvika Neeman. Effective siting of waste treatment facilities. *Journal of Environmental Economics and Management*, 43:303–324, 2002.

[103] Philippe Mongin. L'axiomatisation et les théories économiques. *Revue économique*, 54(1):99–138, 2003.

[104] Hervé Moulin. Fair division under joint ownership: Recent results and open problems. *Social Choice and Welfare*, 7:149–170, 1990.

[105] Hervé Moulin. Uniform externalities: Two axioms for fair allocation. *Journal of Public Economics*, 43(3):305–326, 1990.

[106] Hervé Moulin. Welfare bounds in the cooperative production problem. *Games and Economic Behavior*, 4:373–401, 1992.

[107] Hervé Moulin. *Fair division and collective welfare*. MIT press, 2003.

[108] Hervé Moulin and John Roemer. Public ownership of the external world and private ownership of self. *Journal of Political Economy*, 97(2):347–367, 1989.

[109] Arthur O'Sullivan. Voluntary auctions for noxious facilities: incentives to participate and the efficiency of siting decisions. *Journal of Environmental Economics and Management*, 25(1):S12–S26, 1993.

[110] David Perez-Castrillo and David Wettstein. Choosing wisely: A multibidding approach. *The American Economic Review*, 92(5):1577–1587, 2002.

[111] Arthur Cecil Pigou. *The Economics of Welfare*. MacMillan, London, 1920.

[112] John Rawls. Outline of a decision procedure for ethics. *The Philosophical Review*, 60(2):177–197, 1951.

[113] John Rawls. *A theory of justice*. Harvard University Press, 1971.

[114] John Rawls. *Political liberalism*. Columbia University Press, 1993.

[115] Leigh Stafford Raymond. *Private rights in public resources: equity and property allocation in market-based environmental policy*. Resources for the Future, 2003.

[116] Ernesto Reuben and Arno Riedl. Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior*, 77(1):122–137, 2013.

[117] Barak D. Richman and Christopher Boerner. A transaction cost economizing approach to regulation: Understanding the nimby problem and improving regulatory responses. *Yale Journal on Regulation*, 23:29–76, 2006.

[118] Alvin E. Roth. Bargaining experiments. In John H. Kagel and Alvin E. Roth, editors, *The handbook of experimental economics*, pages 253–348. Princeton University Press, 1995.

[119] Toyotaka Sakai. Fair waste pricing: An axiomatic analysis to the nimby problem. *Economic Theory*, 50(2):499–521, 2012.

[120] Michael J. Sandel. Market reasoning as moral reasoning: why economists should reengage with political philosophy. *Journal of Economic Perspectives*, 27(4):121–140, 2013.

[121] Leland F Saunders. Reason and intuition in the moral life: A dual-process account of moral justification. In Jonathan Evans and Keith Frankish, editors, *In two minds: Dual processes and beyond*, chapter 15, pages 335–354. Oxford University Press Oxford, 2009.

[122] Thomas Schelling. *The strategy of conflict.* Harvard University Press, 1960.

[123] Erik Schokkaert, Kurt Devoogt, Bart Capeau, and Sara Lelli. Allocating an indivisible good. a questionnaire-experimental study of intercultural differences. *Leuven Center for Economic Studies - Discussion Paper Series 07.16*, June 2007.

[124] Amartya Sen. The impossibility of a paretian liberal. *The journal of political economy*, pages 152–157, 1970.

[125] Amartya Sen. Rationality and social choice. *The American Economic Review*, 85(1):1–24, 1995.

[126] Amartya Sen. *The Idea of Justice.* Harvard University Press, 2009.

[127] Lloyd S. Shapley. A value for n-person games. In H. W. Kuhn and W. Tucker, editors, *Contributions to the theory of games II.* Princeton University Press, 1953.

[128] Lloyd S. Shapley and Martin Shubik. On the core of an economic system with externalities. *The American Economic Review*, 4(1):678–684, 1969.

[129] Prakash P Shenoy. On coalition formation: a game-theoretical approach. *International Journal of Game Theory*, 8(3):133–164, 1979.

[130] Adam Smith. Théorie des sentiments moraux. *Presses Universitaires de France. Paris.*, 1999.

[131] David Starett. A note on externalities and the core. *Econometrica*, 41(1):179–183, 1973.

[132] Maxwell L. Stearns. The misguided renaissance of social choice. *The Yale Law Journal*, 103:1219–1292, 1993.

[133] Robert Sugden. The behavioural economist and the social planner: to whom should behavioural welfare economics be addressed? *Inquiry*, 56(5):519–538, 2013.

[134] Cass R Sunstein. Moral heuristics. *Behavioral and brain sciences*, 28(4):531–541, 2005.

[135] William Thomson. The replacement principle in economies with indivisible goods. *Social Choice and Welfare*, 15(1):57–66, 1997.

[136] William Thomson. On the axiomatic method and its recent applications to game theory and resource allocation. *Social Choice and Welfare*, 18:327–386, 2001.

[137] William Thomson. Fair allocation rules. In Kenneth J. Arrow, Amartya Sen, and Kotaro Suzumura, editors, *Handbook of Social Choice and Welfare*, volume 2, chapter 21, pages 393–506. Elsevier, 2011.

[138] R. M. Thrall and W. F. Lucas. N-person games in partition function form. *Naval Research Logistics Quarterly*, 10(1):281–298, 1963.

[139] Bertil Tungodden. Some reflections on the role of moral reasoning in economics. *Nordic Journal of Political Economy*, 30:49–59, 2004.

[140] Paloma Ubeda. The consistency of fairness rules: An experimental study. *Journal of Economic Psychology*, 41:88–100, 2014.

[141] Ferdinand M Vieider. Separating real incentives and accountability. *Experimental Economics*, 14(4):507–518, 2011.

[142] Michael Walzer. *Spheres of justice: A defense of pluralism and equality.* Basic Books, 1983.

[143] Menahem E Yaari and Maya Bar-Hillel. On dividing justly. *Social Choice and Welfare*, 1(1):1–24, 1984.

[144] H Peyton Young. Cost allocation. *Handbook of game theory*, 2:1193–1235, 1994.

[145] H Peyton Young. *Equity: in theory and practice.* Princeton University Press, 1995.

Index