



**HAL**  
open science

# A game-theoretical and experimental approach to expert-based knowledge

Philippe Colo

► **To cite this version:**

Philippe Colo. A game-theoretical and experimental approach to expert-based knowledge. Economics and Finance. Université Panthéon-Sorbonne - Paris I, 2021. English. NNT : 2021PA01E038 . tel-03539631

**HAL Id: tel-03539631**

**<https://theses.hal.science/tel-03539631>**

Submitted on 21 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**UNIVERSITÉ PARIS I PANTHÉON SORBONNE**  
**UFR 02**

**Paris-Jourdan Sciences Économiques - Paris School of Economics**

**THÈSE**

Pour l'obtention du titre de Docteur en Economie

Présentée et soutenue publiquement

le 8 Juillet 2021 par

**Philippe Colo**

**A game-theoretical and experimental approach to  
expert-based knowledge**

**Sous la direction de Jean-Marc Tallon & Stéphane Zuber, Directeurs de recherche  
au CNRS**

**Membres du Jury :**

<b>Berger</b>	<b>Loic</b>	Chargé de recherche au CNRS, IESEG	Rapporteur
<b>Kellner</b>	<b>Christian</b>	Professeur associé à l'Université de Southampton	Rapporteur
<b>Koessler</b>	<b>Frédéric</b>	Directeur de recherche au CNRS, PSE	Examineur
<b>Ottaviani</b>	<b>Marco</b>	Professeur à l'Université Bocconi	Examineur



*Sapere aude !*



# Remerciements

Il y a bien longtemps, quand je suis arrivé à l'Ecole d'économie de Paris, Jean-Marc m'a dit qu'écrire une thèse, c'était comme courir marathon. J'ai bien essayé d'arguer qu'un marathon était finalement une épreuve assez courte à l'intensité élevée et que trois ans à plancher sur un sujet (hé oui, à l'époque je pensais encore que cela durerait trois ans) ce n'était sous doute pas la même chose. A vrai dire, grâce à sa grande expérience sportive, Jean-Marc n'avait pas tort. La thèse est comme un marathon, car c'est une épreuve où il faut trouver son rythme, un équilibre mental et savoir faire face aux hauts et aux bas. Mais surtout, comme sur un marathon, vous ne tiendriez pas la moitié de la course sans les fanfares de supporters le long du parcours. Je dois dire que cela fait cinq ans maintenant (hé oui, cinq ans) que les encouragements s'accumulent le long de la route. A quelques encablures de la ligne d'arrivée, dans les 195 derniers mètres qui font qu'un marathon dure plus de 42km, c'est toujours à ces encouragements que l'on pense. C'est presque autant pour eux que pour nous qu'on arrive jusque là.

Me voilà dans les 195 derniers mètres et, formidable tradition de la thèse, on m'offre une tribune pour remercier la fanfare. Evidemment, la plus présente, la plus patiente mais aussi la plus exigeante était celle de mes directeurs, Jean-Marc et Stéphane. Ils ont été pour moi comme ces entraîneurs qui sont toujours là pour leur athlète, qu'il pleuve, qu'il neige ou qu'il vente. Je ne les remercierai jamais assez de m'avoir suivi du début à la fin de cette aventure avec la même attention. Ils m'ont appris par l'exemple que la liberté intellectuelle de poursuivre ses propres idées n'est possible qu'en s'imposant rigueur et clarté. Bien-sûr, ces remerciements vont tout autant à mon directeur officieux, Frédéric, qui m'a suivi

sur toute la deuxième partie de ma thèse. Ce sont les après-midis passés ensemble à plonger intensément dans les méandres de mes modèles qui m'ont appris ce qu'était de s'approprier véritablement un problème. Je voudrais aussi remercier Francis et Franz, qui ont tous deux eu la patience de participer à mon comité de thèse. Grâce à eux, j'ai découvert des manières totalement différentes d'aborder des questions, avec une clarté et une profondeur que je ne connaissais pas.

Dans le cheminement d'un athlète, ce sont aussi les conseils et le soutien des coureurs plus expérimentés qui font sa progression. Tout au long de mon chemin, les encouragements, le soutien, la confiance et les conseils de Loïc et Brian ont été cela pour moi et, à eux aussi, je dois énormément. Travailler avec eux, c'est avoir pris la foulée de ceux qui m'ont donné envie de participer à la course. Je dois particulièrement à Loïc de m'avoir permis de finir ma thèse dans les meilleures conditions possible, en me finançant jusqu'au bout. Je dois particulièrement à Brian de m'avoir montré qu'il était possible (mais pas facile) de faire le pont entre les questions qui me fascinent en philosophie et les formidables outils de l'économie. Ces questions, ce sont Anouk Barberousse, Isabelle Drouet et Marion Vorms qui me les ont faites découvrir. Si j'ai eu un jour le courage de m'y atteler, elles y sont pour beaucoup. Je n'ai jamais été aussi fier d'un travail que j'ai fait que du mémoire que je leur ai rendu. La philosophie avait toujours été pour moi une course si attirante et impressionnante à la fois – une sorte d'ultra-triathlon montagneux – que je pensais ne pouvoir jamais en prendre le départ.

Evidemment, on ne court pas seul une course comme celle que je termine. Dans mon groupe d'allure, nous étions plusieurs compagnons d'infortune : Guillaume, Shaden, Aurore, Thomas D., Marion, Malka, Elias, Quentin ont déjà tous passé la ligne d'arrivée. Alexis, Thomas Z. et Lydia sont partis bien après moi, mais sont déjà sur mes talons. Avec eux tous, la route a été bien plus belle. Je souhaite qu'elle nous mène encore loin ensemble.

Arrivé à ce point de ma tribune, le lecteur aura sans doute saisi mes travers sportifs. C'est donc le moment de remercier ceux qui ont fait ma vie parallèle, celle qui m'a donné l'équilibre nécessaire pour parvenir au bout de cette thèse,

celle du PUC. Je dois à Laura, Thomas et Geneviève de m'avoir appris comment construire, jour après jour, un projet ambitieux, qu'il soit sportif ou intellectuel. Comme mes entraîneurs de thèse, ils m'ont montré que pour cela, peu importent les capacités de base; c'est la persévérance qui est la première des qualités. J'ai une pensée pour mes compagnons de thèse du PUC, Tristan, Félicien, Marine, Hassan et Antoine, avec qui j'ai eu en commun les mêmes vies parallèles. Je pense aussi à Marc, grand sage, qui m'aura souvent permis de relativiser mes échecs (« Ta preuve ne marche pas ? C'est ça la vraie vie ! »).

La dernière année de cette thèse a évidemment été particulière, et difficile, notamment en raison de l'isolement physique et social. Je dois beaucoup à Catherine et Georges-Henri de m'avoir accueilli en Bretagne dans les moments les plus difficiles – sans doute était-ce le fameux mur du 30ème. Je voudrais aussi remercier Christine, chez qui, lors du premier confinement, j'ai fait une rencontre quasi-amoureuse avec la montagne. Sans nul doute était-ce l'euphorie du 20ème kilomètre.

Mes supporters les plus fidèles, les plus inconditionnels, seront toujours mes parents. Même si j'ai fait tout à l'inverse de leurs conseils, personne ne criera jamais plus fort qu'eux au bord de la route. Cette route, je n'aurais même pas osé la regarder si Sophie n'avait pas été là. Si je passe la ligne d'arrivée, c'est avant tout à elle que je le dois. Et à Nala, bien évidemment.



# Résumé

Cette thèse est constituée de trois essais indépendants, chacun portant sur un aspect théorique ou empirique relatif à la question des connaissances provenant d'experts.

Dans le premier chapitre de cette thèse, j'étudie la transmission de connaissances scientifiques entre un expert et un décideur. Un modèle scientifique est formalisé par une distribution de probabilité sur un ensemble de scénarios possibles. L'expert est supposé connaître le modèle le plus probable parmi un ensemble possible et cherche à le communiquer au décideur. Toutefois, parce que ces modèles sont trop complexes, l'expert ne peut pas certifier cette information au décideur. Je montre que s'il y a une différence d'intérêt entre les deux partis, à l'équilibre, la transmission de l'information est toujours partielle. L'expert ne pourra jamais communiquer de manière crédible quel modèle est le plus probable. Toutefois, il pourra désigner un ensemble de modèles contenant celui-ci. La taille de cet ensemble, et donc le degré d'information qu'il pourra communiquer, dépend à la fois de la différence d'intérêt entre les partis mais aussi du consensus entre les modèles scientifiques. Si la science n'est pas suffisamment consensuelle, il y a une asymétrie dans la transmission de l'information. Si le modèle le plus probable est parmi les plus optimistes, la transmission d'information dépend uniquement de la différence d'intérêt entre les partis. Mais s'il est parmi les plus pessimistes, aucune transmission d'information n'est possible.

Dans le second chapitre de cette thèse, mes co-auteurs et moi mesurons expérimentalement les croyances de sujets sur des événements dont ils sont plus ou

moins familiers. Pour ce faire, nous proposons une méthode novatrice d'identification des croyances des sujets qui s'appuie sur l'utilisation d'intervalles de probabilités objectives. Pour chaque événement, notre approche nous permet d'éliciter des ensembles de distribution de probabilités majoritairement non dégénérés. De plus, plus les événements sont familiers, plus les intervalles élicités sont restreints. Ainsi, plus nos sujets se sentent experts sur une question, plus leurs croyances sont précises. Notre approche nous permet également d'estimer la manière dont ces sujets agissent en correspondance avec leurs croyances. Ce faisant, nous parvenons à la première estimation du coefficient de mixture  $\alpha$  dans le modèle de décision  $\alpha$ -maxmin EU de Hurwicz, en contrôlant par les croyances des sujets.

Dans le troisième chapitre de cette thèse, je reprends l'hypothèse que la connaissance scientifique est trop complexe pour être certifiée à des non-experts. J'étudie les conséquences de celle-ci dans un cas appliqué : celui du changement climatique. Je modélise le problème de la sur-émission de gaz à effet de serre (GES) comme un jeu de contribution à un mal public. Dans ce jeu, tous les contributeurs gagnent individuellement à émettre, car les GES sont corrélés à la consommation de biens, mais tous les contributeurs souffrent du total des émissions car celles-ci sont responsables de dommages climatiques. A l'équilibre, le niveau d'émission est toujours trop élevé, car chaque contributeur ne tient pas compte des externalités négatives dont il est responsable. Les contributeurs ne sont pas des experts du climats, et leurs connaissances sur les dommages auxquels ils s'exposent s'appuient uniquement sur un expert. Ce dernier tient compte des externalités des contributeurs, et voudrait toujours un niveau d'émission plus bas que celui obtenu à l'équilibre par les contributeurs. Il y a donc toujours une différence d'intérêt entre l'expert et les non-experts. Dans ce chapitre, je prouve qu'aucune transmission d'information ne peut avoir lieu à l'équilibre. Ce résultat montre que la seule parole de l'expert, sans pouvoir de certification, ne suffit pas lorsqu'il s'agit de communiquer sur le risque climatique.

# SUMMARY

This thesis consists of three independent essays, each of which focuses on a theoretical or empirical aspect related to expert-based knowledge.

In the first chapter of this thesis, I study the transmission of scientific knowledge between an expert and a decision maker. A scientific model is formalised by a probability distribution over a set of possible scenarios. The expert is assumed to know the most likely model among a set of possible models and tries to communicate it to the decision maker. However, because these models are too complex, the expert cannot certify this information to the decision maker. I show that if there is a difference of interest between both parties, at equilibrium, the transmission of information is always partial. The expert will never be able to credibly communicate which model is the most likely. However, he will be able to designate a set of models containing it. The size of this set, and thus the degree of information that can be communicated, depends both on the difference of interest between the parties and on the consensus among scientific models. If the science is not sufficiently consensual, there is an asymmetry in the transmission of information. If the most likely model is among the most optimistic, the transmission of information depends solely on the difference in interest between the parties. But if it is among the most pessimistic, no information transmission is possible.

In the second chapter of this thesis, my co-authors and I experimentally measure subjects' beliefs about events with which they are more or less familiar. To do so, we propose a novel method for identifying subjects' beliefs that relies on

the use of objective probability intervals. For each event, our approach allows us to elicit mostly non-degenerate probability distribution sets. Moreover, we find that the more familiar the events, the smaller the elicited intervals. Thus, the more our subjects believe themselves to have expertise on a question, the more accurate their beliefs are. Our approach also allows us to estimate how these subjects act in correspondence with their beliefs. In doing so, we arrive at the first estimate of the mixture coefficient  $\alpha$  in Hurwicz's  $\alpha$ -maxmin EU decision model, controlling for the subjects' beliefs.

In the third chapter of this thesis, I once again assume that scientific knowledge is too complex to be certified to non-experts. I study the consequences of this assumption in the case of climate change mitigation. I model the problem of greenhouse gas (GHG) over-emission as a game of contribution to a public bad. In this game, all contributors individually gain from emitting, because GHGs are correlated with the consumption of goods, but all contributors suffer from the total level of emissions because these emissions are responsible for climate damage. In equilibrium, the level of emissions is always too high, because each contributor does not take into account the negative externalities for which it is responsible. The contributors are not climate experts, and their knowledge of the damage they are exposing themselves to is based solely on an expert. The expert takes into account the externalities of the contributors, and would always want a lower emission level than the one obtained in equilibrium by the contributors. Thus, there is always a difference of interest between the expert and the non-experts. In this chapter I prove that no transmission of information can take place at equilibrium. This result shows that the word of the expert alone, without certification power, is not enough when it comes to communicating about climate risk.

# Contents

<b>Remerciements</b>	<b>iii</b>
<b>Résumé</b>	<b>vii</b>
<b>Summary</b>	<b>ix</b>
<b>General Introduction</b>	<b>1</b>
<b>1 Communicating over Models</b>	<b>11</b>
1.1 Introduction . . . . .	12
1.2 Setup . . . . .	19
1.2.1 Primitives . . . . .	19
1.2.2 Strategic interaction . . . . .	21
1.2.3 Consensus and division in Science . . . . .	24
1.3 Equilibrium analysis . . . . .	29

1.4	General results . . . . .	31
1.5	The linear-quadratic example . . . . .	36
1.6	$\alpha$ -MEU receiver . . . . .	38
1.6.1	Optimal actions and structure of equilibria . . . . .	38
1.6.2	Comparative ambiguity aversion . . . . .	42
1.7	Conclusion . . . . .	45
1.8	Appendix . . . . .	47
1.8.1	Supplementary Assumptions . . . . .	47
1.8.2	Proofs of the results in the main text . . . . .	50
<b>2</b>	<b>Eliciting Multiple Prior Beliefs</b>	<b>79</b>
2.1	Introduction . . . . .	80
2.2	Theoretical Background . . . . .	83
2.2.1	Preliminaries . . . . .	83
2.2.2	Upper and lower probabilities and CDFs . . . . .	85
2.2.3	Decision model . . . . .	86
2.2.4	Matching Probability Benchmark . . . . .	87
2.2.5	Theory: Matching Probability Intervals . . . . .	89

2.2.6	Implementation 1: 2D Choice lists . . . . .	91
2.2.7	Implementation 2: Binary-choice procedure . . . . .	95
2.3	Experimental Methods . . . . .	97
2.3.1	Subjects . . . . .	98
2.3.2	Stimuli and treatments . . . . .	98
2.3.3	Elicitation techniques . . . . .	101
2.3.4	Incentivizing subjects . . . . .	104
2.4	Results . . . . .	105
2.4.1	Descriptive Statistics and Performance . . . . .	105
2.4.2	Bayesian analysis . . . . .	107
2.4.3	Imprecision . . . . .	108
2.4.4	Imprecision and familiarity . . . . .	113
2.4.5	Matching probabilities and the $\alpha$ -maxmin EU mixture coefficient . . . . .	114
2.5	Discussion . . . . .	115
2.6	Conclusion . . . . .	122
2.7	Theoretical Appendix . . . . .	123
2.7.1	Proofs . . . . .	123

2.7.2	Binary-choice procedure . . . . .	124
2.7.3	Robustness of the elicitation method . . . . .	130
2.8	Supplementary Statistics . . . . .	140
2.8.1	Descriptive Statistics . . . . .	140
2.8.2	Bayesian estimation . . . . .	147
2.8.3	Matching Probability data and analysis of $\alpha$ . . . . .	154
2.8.4	Elicitation-free check of $\alpha > \frac{1}{2}$ . . . . .	154
2.9	Experimental design and displays . . . . .	157
2.9.1	Probability interval elicitation: displays . . . . .	157
2.9.2	MP elicitation: displays . . . . .	160
2.10	Lemmas for the proof of Proposition 4 . . . . .	162
<b>3</b>	<b>Communicating over climate risk</b>	<b>167</b>
3.1	Introduction . . . . .	168
3.2	Setup . . . . .	171
3.3	Emission stage . . . . .	174
3.4	Welfare analysis . . . . .	176
3.5	Communication stage . . . . .	179

3.5.1	Strictly positive energy consumption . . . . .	179
3.5.2	Null energy consumption . . . . .	184
3.6	Conclusion . . . . .	186
3.7	Appendix . . . . .	187
	<b>References</b>	<b>193</b>



# INTRODUCTION GENERALE<sup>1</sup>

Alors que nous vivons à une époque où la somme des travaux scientifiques n'a jamais été aussi importante, il semble, paradoxalement, que notre compréhension individuelle du monde soit chaque jour plus partielle. Comprendre le fonctionnement d'une simple montre GPS demande par exemple de maîtriser les principes de l'électronique, de la programmation informatique mais aussi de la relativité générale. Fort peu d'entre nous maîtrisent suffisamment de pans de la science pour en comprendre pleinement le fonctionnement. Cet apparent paradoxe provient du fait que l'avancement collectif de la science n'a été possible qu'au prix d'une division croissante de la compréhension des théories scientifiques. Alors que le fonctionnement du monde est toujours plus complexe et que le savoir scientifique sous-jacent à son organisation s'éparpille chaque jour davantage entre une multitude d'experts, le degré de connaissance *du monde* nous possédons *individuellement* n'a jamais été aussi incertain. Autrement dit, on peut se demander dans quelle mesure nous pouvons dire que nous savons où nous nous trouvons uniquement en nous appuyant sur une montre GPS, alors que nous ne sommes experts ni en électronique, ni en programmation informatique, ni en relativité générale.

Bien que les travaux développés dans cette thèse relèvent de la science économique, leur ambition est de contribuer à la question de la connaissance, au temps de la division du savoir. L'accès à des croyances sur le monde tel qu'il est, ou "croyance vraies", est à la fois l'ambition fondamentale de l'approche scientifique et l'objet d'une perpétuelle remise en cause. Toute l'entreprise scientifique consiste

---

<sup>1</sup>Due to legal restrictions, this introduction had to be written in french. An english version can be found [here](#).

à proposer des méthodes de *justification* de croyances sur le monde suffisamment robustes pour pouvoir être tenues pour des connaissances. Traditionnellement, en philosophie de la connaissance, cette justification ne peut provenir que de la perception ou d'un raisonnement logique que le sujet a formulé lui-même. Cette position, notamment défendue par Kant (1787) dans *Critique de la raison pure*, est celle de l'indépendance épistémique. Elle impose que pour accéder à la connaissance, le sujet doit avoir une maîtrise directe des justifications sous-tendant la proposition considérée afin de pouvoir s'appuyer uniquement sur son propre jugement.

Or, il apparaîtra sans doute au lecteur qu'en approchant la connaissance sous cet angle, nous ne pouvons pas dire que nous savons où nous nous trouvons, uniquement grâce à une montre GPS. Autrement dit, dans cet exemple, et dans un grand nombre de cas similaires, l'exigence d'indépendance épistémique est trop restrictive pour qu'on puisse parler de connaissance. Plus généralement, ce constat nous ramène vers notre paradoxe initial: si l'on se tient à la seule position kantienne, la très vaste majorité des propositions que nous tenons pour des connaissances perdrait ce statut. Il faudrait alors dire que plus la science avance collectivement, moins nous en savons sur le monde individuellement. Pour autant, on peut penser que c'est l'inverse qui se produit: avec l'avancée de la science nous avons le sentiment d'en savoir plus sur le monde. Il faudrait alors dire que plus nous avançons dans l'Histoire, plus les sociétés sont irrationnelles car elles tiennent pour des connaissances de plus en plus de choses dont la justification n'est pas épistémiquement autonome.

C'est ce constat paradoxal qui a été fait par John Hardwig dans *Epistemic Dependence* (Hardwig, 1985). Pour le dépasser, Hardwig propose dans cet article, et dans *The role of trust in knowledge* (Hardwig, 1991), une approche alternative de la connaissance : on peut être justifié à tenir une proposition provenant d'un autre individu pour connaissance si on est justifié à croire ceux qui la soutiennent et que ceux-ci sont des autorités intellectuelles en la matière. Autrement dit, il est possible de fonder nos connaissances sur des experts. Par cette proposition, Hardwig s'affranchit de l'exigence kantienne d'indépendance épistémique en introduisant

une troisième source de justification pour la connaissance : le témoignage.

La proposition de Hardwig a notamment contribué au développement récent de l'épistémologie du témoignage. Pour les épistémologues du témoignage<sup>1</sup>, la déférence épistémique à autrui ne peut être justifiée que si son bénéficiaire est à la fois plus compétent que soi-même et que le déférent est rationnellement justifié à accorder sa confiance. Dans cette thèse, je me place à la suite de cette approche de la connaissance et je propose d'approfondir l'étude du lien de confiance entre expert et non-expert. Pour qui veut fonder la connaissance sur le témoignage, il est d'une importance capitale d'étudier le lien de confiance entre expert et non-expert même lorsque les intérêts de chaque parti sont différents. Dans un grand nombre de cas, il serait en effet naïf de croire que ces intérêts sont alignés. Pour prendre l'exemple initial de la montre GPS, il est raisonnable de penser que son concepteur a intérêt à nous convaincre que c'est bien notre position que l'appareil va afficher. En supposant qu'il est un expert concernant les montre GPS, dans quelle mesure pouvons-nous nous fier à lui concernant les propriétés de sa montre ?

Pour étudier les déterminants de la confiance que nous pouvons avoir dans les experts, il a été souvent reproché à la philosophie des sciences de ne pas suffisamment tenir compte de la dimension sociale et politique de l'activité scientifique (Barberousse et al., 2011).<sup>2</sup> C'est notamment sur ces aspects que se concentrent les premier et troisièmes chapitres de cette thèse. Pour ce faire, l'usage de la théorie des jeux m'a paru prometteuse, en raison des possibilités qu'elle offre pour représenter les comportements individuels dans un cadre stratégique. En m'engageant dans cette voie, j'ai donc pris le parti de proposer la notion d'*équilibre stratégique* comme justification à une dépendance épistémique. Ainsi, un non-expert sera justifié à croire une proposition que lui transmet un expert si ce message est une action d'équilibre. Un avantage de cette approche est le formalisme mathématique qui l'accompagne. Celui-ci permet de décrire le degré

---

<sup>1</sup>Voir notamment Goldman (2001), Coady (1992) et Burge (1993)

<sup>2</sup>Cette critique a notamment été formulée par les fondateurs des science studies, discipline entièrement empirique qui s'est notamment développée à partir des du programme proposé dans Bloor (1976) et Collins (1981).

d'information qu'un non-expert est stratégiquement justifié à croire. Autrement dit, même si un expert a une connaissance parfaite de son sujet, un non-expert, s'il est son dépendant épistémique, ne sera pas nécessairement justifié au même degré de connaissance. Cette thèse se propose d'étudier ce degré de justification comme étant le résultat d'une interaction stratégique entre un émetteur (sachant) et un receveur (dépendant). A la suite de [Crawford and Sobel \(1982\)](#), tenir une proposition provenant d'un expert pour vrai sera justifiée pour le dépendant si, à l'équilibre, l'expert a intérêt à transmettre cette proposition.

Dans le premier chapitre de cette thèse, *Expert-based scientific knowledge: communicating over models*, j'étudie la transmission de connaissances scientifiques dans ce cadre. Un modèle scientifique est formalisé par une distribution de probabilité sur un ensemble de scénarios possibles. Un expert est supposé connaître le modèle le plus probable parmi un ensemble possible et cherche à le communiquer à un décideur. Le décideur est le dépendant épistémique de l'expert : il ne peut s'appuyer que sur ce dernier pour justifier ses croyances sur les modèles. Je montre que s'il y a une différence d'intérêt entre les deux partis, à l'équilibre, la transmission de l'information est toujours partielle. L'expert ne pourra jamais faire accéder le décideur au même degré de connaissance que lui. Il ne pourra jamais communiquer de manière crédible quel modèle est le plus probable. Toutefois, il pourra désigner un ensemble de modèles contenant celui-ci. La taille de cet ensemble, et donc le degré d'information qu'il pourra communiquer, dépend à la fois de la différence d'intérêt entre les partis mais aussi du consensus entre les modèles scientifiques. Si la science n'est pas suffisamment consensuelle, il y a une asymétrie dans la transmission de l'information. Si le modèle le plus probable est parmi les plus optimistes<sup>3</sup>, la transmission d'information dépend uniquement de la différence d'intérêt entre les partis. Mais s'il est parmi les plus pessimistes, aucune transmission d'information n'est possible. Ces résultats suggèrent qu'en matière de modèles scientifiques, la dépendance épistémique est possible. Mais celle-ci implique nécessairement une forme de dilution des connaissances. S'il y a

---

<sup>3</sup>Optimisme et pessimisme sont ici pensés en terme d'espérance d'utilité maximale pour le décideur. Dans l'exemple du GPS, le modèle qui prédit que la montre donne avec certitude la position d'un décideur qui cherche à se localiser est le plus optimiste possible. Un modèle qui prédit que la montre n'a aucune capacité de localisation est plus pessimiste pour le décideur.

une différence d'intérêt entre expert et non-expert, ce dernier n'est jamais justifié au même degré de connaissance que l'expert. Autrement dit, les connaissances du dépendant seront toujours plus imprécises.

Il semble donc que, par la voie du témoignage, nous ne soyons justifiés qu'à des croyances imprécises sur le monde. Je ne suis généralement pas justifié à croire que je connais le modèle scientifique le plus probable pour expliquer le fonctionnement de ma montre GPS, mais seulement que je le connais avec une certaine probabilité. Se pose alors la question de savoir dans quelle mesure on peut encore parler de connaissance. Nous avons jusqu'ici tenu une connaissance pour une croyance vraie et justifiée sur le monde. Comment une croyance peut-elle être vraie et seulement probable à la fois ? On attribue souvent à David Hume la paternité de ce débat. Avant Hume, la définition qui était donnée à la vérité est celle que l'on nomme désormais correspondantiste<sup>4</sup>. Au sens correspondantiste, un énoncé est vrai si et seulement s'il correspond à l'objet auquel il se réfère dans la réalité. Dans la section X de *l'Enquête sur l'entendement humain* (Hume, 1748), Hume fait valoir que cette approche de la vérité est inaccessible à l'expérience. Par l'expérience nous n'avons accès qu'à des conjonctions d'événements, à leurs connexions habituelles, mais jamais aux lois certaines de la nature. La position de Hume ouvre la voie à une position plus pragmatique concernant la notion de vérité dans le concept de connaissance: seule l'inférence, au sens des corrélations factuelles, nous est accessible et peut remplir le rôle laissé vide par la vérité. Avec Hume, il n'est plus nécessaire de chercher à accéder à la vérité pour fonder une connaissance. Il suffit de s'appuyer sur la notion de conjonction des événements. On peut tenir une proposition pour connue si on observe un grand nombre de conjonctions d'événements en sa faveur. Autrement dit, pour ceux qui prendront la suite de Hume, une connaissance est une croyance suffisamment probable et suffisamment justifiée sur le monde. Nous avons vu dans le premier chapitre que la manière dont nous justifions nos croyances peut influencer sur leur probabilité. Pour déterminer si nous pouvons fonder nos connaissances sur des experts il nous faut donc déterminer quand une croyance est suffisamment probable pour pouvoir parler de connaissance.

---

<sup>4</sup>Voir par exemple le *Théétète* de Platon

Cette question est encore largement débattue dans la philosophie de la connaissance contemporaine. Une des réponses proposé est celle du courant pragmatiste et est notamment défendue par les philosophes Fantl et McGrath dans *Knowledge in an Uncertain World* 2009. Considérant une proposition  $p$ , ils estiment que " votre probabilité de  $p$  est *knowledge-level* si et seulement si la probabilité de  $\text{non-}p$  ne s'oppose pas à ce que  $p$  soit mise en pratique comme base pour la croyance ou l'action ". Autrement dit,  $p$  peut être tenue pour connaissance si et seulement si aucune raison qui semble rationnellement justifiée de croire que  $p$  est fausse ne s'oppose à agir en s'appuyant sur  $p$ . Ainsi, si je suis justifié à croire que la position que me donne ma montre a au moins 90% de chance d'être la bonne, je peux dire que je connais ma position si les actions que je pourrais entreprendre en m'appuyant sur cette croyance peuvent s'accommoder d'une probabilité de 10% pour que je me trompe. Le lecteur non initié trouvera sans doute que Fantl et McGrath opèrent un retournement inhabituel pour définir la connaissance<sup>5</sup>. Connaître la proposition  $p$  n'est plus un état mental indépendant qui peut déboucher sur une action, c'est une action qui peut faire de la proposition  $p$  un état mental que l'on qualifiera de connaissance.

Dans le second chapitre de cette thèse, *Eliciting Multiple Prior Beliefs*, mes co-auteurs et moi mesurons expérimentalement les croyances de sujets sur des événements dont ils sont plus ou moins familiers. Notre approche s'appuie sur le champs de la théorie de la décision, qui étudie la décision individuelle lorsque l'incertitude est élevée. En particulier, nous nous appuyons sur l'approche des a priori multiples introduit par [Gilboa and Schmeidler \(1989a\)](#), qui permettent de rendre compte d'imprécision dans les croyances. Dans cette expérience, nous proposons une méthode novatrice d'identification des croyances des sujets qui s'appuie sur l'utilisation d'intervalles de probabilités objectives. Pour chaque événement, notre approche nous a permis d'éliciter des ensembles de distribution de probabilités. La méthode trouve une prédominance d'intervalles de probabilité non dégénérés parmi les sujets dans tous les cas explorés : nos sujets ont des croyances imprécises. Plus les événements sont familiers, plus ces intervalles

---

<sup>5</sup>Cet angle est en réalité propre au courant pragmatiste selon [Weatherson \(2012\)](#), qui le qualifie de *pragmatic encroachment*.

sont restreints. Ainsi, plus nos sujets se sentent experts sur une question, plus leurs croyances sont précises. Notre approche nous permet également d'estimer la manière dont ces sujets agissent en correspondance avec leurs croyances. Nous trouvons alors que, étant donnée une croyance qui est un intervalle de probabilité, les sujets sont disposés à agir en s'appuyant sur elle lorsque, même pour une probabilité basse de cet intervalle, cette action leur paraît justifiable. On peut donc dire que dans ce chapitre nous mesurons les propositions que nos sujets tiennent pour des connaissances, au sens de Fantl et McGrath. La méthode que nous avons employée peut servir de mécanisme d'identification des croyances pour des experts, par exemple sur des sujets complexes en science du climat. Ce type d'identification est, par exemple, très important dans la calibration de modèles de prospectives sur les impacts du changement climatique. Les prédictions donnés par ces modèles reposent en effet souvent sur la calibration de paramètres de départ qui ne peut se faire qu'à *dire d'expert*. La manière dont le jugement de ces derniers est mesuré a un impact déterminant sur l'estimation de ces paramètres.<sup>6</sup> En suivant Fantl et McGrath, on peut en réalité dire que plus que des croyances, cette méthode nous permet également de mesurer les propositions que ces experts tiennent pour des connaissances. En particulier, dans le cadre de la famille de modèle de décision  $\alpha$ -MEU (Ghirardato et al., 2004b), pour une proposition donnée, notre approche permet à la fois de mesurer l'ensemble d'a priori d'un sujet et le paramètre  $\alpha$  qui va guider son action en fonction de ses croyances. Ainsi, on pourra dire par exemple qu'un sujet estime qu'il y a entre 80% et 90% de chances pour que la position donnée par sa montre GPS soit juste. Mais aussi que, si son  $\alpha = 0.9$ , il considère connaître sa position si l'action qu'il doit prendre en s'appuyant sur cette connaissance peut admettre une probabilité d'erreur d'au plus 11 points de pourcentage.

En résumé, les deux premiers chapitres de cette thèse ont cherché à explorer les raisons pour lesquelles nous pouvons fonder nos connaissances sur les experts. Dans mon premier chapitre j'ai défendu l'idée qu'en tant que non-experts nous n'étions généralement justifiés qu'à des croyances sur le monde plus imprécises que celles des experts. A la suite de Fantl et McGrath, j'ai adhéré à l'idée que

---

<sup>6</sup>Le lecteur pourra trouver plus de détail dans Morgan et al. (1992) ou Cooke (1991).

celles-ci sont légitimes à être tenues pour des connaissances dans la mesure où les actions que nous pourrions entreprendre en nous appuyant sur elles peuvent s'accommoder de la probabilité que nous nous trompions. Il est ainsi possible qu'une même croyance soit suffisamment justifiée pour être tenue pour connaissance en ce qui concerne une action, mais pas en ce qui concerne une autre. Pour les décisions de tous les jours, nous pensons généralement que nous connaissons notre position grâce à notre montre GPS. Mais si cette position doit nous servir à aiguiller les secours venus nous sauver en montagne, et qu'une erreur causerait notre perte, dirons-nous toujours que nous connaissons notre position ? Les résultats empiriques obtenus dans mon second chapitre suggèrent que la réponse sera oui si, même en étant pessimiste, je choisis de m'appuyer sur ma montre pour donner ma position aux secours.

Dans le troisième chapitre de cette thèse, *Communicating over climate risk*, j'explore les conséquences de cette approche de la dépendance épistémique aux experts dans un cas pratique : celui du changement climatique. Je modélise le problème de la sur-émission de gaz à effet de serre (GES) comme un jeu de contribution à un mal public. Dans ce jeu, tous les contributeurs gagnent individuellement à émettre, car les GES sont corrélés à la consommation de biens, mais tous les contributeurs souffrent du total des émissions car celles-ci sont responsables de dommages climatiques. A l'équilibre, le niveau d'émission est toujours trop élevé, car chaque contributeur ne tient pas compte des externalités négatives dont il est responsable. Les contributeurs ne sont pas des experts du climat, et leurs connaissances sur les dommages auxquels ils s'exposent s'appuient uniquement sur un expert. Ce dernier tient compte des externalités des contributeurs, et voudrait toujours un niveau d'émission plus bas que celui obtenu à l'équilibre par les contributeurs. Il y a donc toujours une différence d'intérêt entre l'expert et les non-experts. Dans ce chapitre je montre qu'aucune transmission d'information ne peut avoir lieu à l'équilibre. L'expert n'est pas une source de connaissance pour les contributeurs dans ce contexte. Autrement dit, le témoignage n'est pas une source épistémique suffisamment forte pour motiver à lui seul l'action en faveur du climat.

Ce dernier résultat est sans doute celui qui illustre le plus clairement l'importance de poursuivre notre réflexion sur la connaissance, au temps de la division du savoir. L'approche proposée dans cette thèse aide à comprendre pourquoi pour tant d'entre nous, les décisions que nous impose le changement climatique sont encore trop difficiles à prendre. Si l'on définit la connaissance par l'action, à la manière des pragmatistes, et si l'on comprend que la division de la connaissance implique une dépendance épistémique qui réduit la précision de nos croyances, on prend une nouvelle mesure de la difficulté des choix environnementaux.



# 1. Expert-based Scientific Knowledge: Communicating over Models <sup>1</sup>

## Abstract

This paper studies the transmission of complex scientific knowledge. Scientific models are formalised as probability distributions over possible scenarios. An expert is assumed to know the most likely model and seeks to communicate it to a decision maker, but cannot certify it. As a result, communication of scientific knowledge is a cheap talk game over models. The decision maker is in a situation of model-uncertainty and is ambiguity sensitive. I show that information transmission depends on both the strategic misalignment of players and the consensus among scientific models. When science is divided, there is an asymmetry in information transmission when the receiver has maxmin expected utility preferences. Types below a certain threshold are necessarily pooled, regardless of the misalignment. All equilibria of the game are outcome equivalent to a partitional equilibria and, unlike similar models in the literature, the most informative one is interim Pareto dominant.

---

<sup>1</sup>I would like to thank Jean-Marc Tallon, Stéphane Zuber, Frédéric Koessler, Massimo Marinacci, Marco Ottaviani, Valentina Bossetti, Mark Le Quement, Francis Bloch, Larry Samuelson, Loïc Berger and Guillaume Pommey for helpful discussions. I also thank seminar and conference participants at PSE (TOM, SRE), Bocconi and FUR 2018 and Bocconi University for its hospitality. Financial support through ANR CHOp (ANR-17-CE26-0003), ANR ADE (ANR-18-ORAR-0005-01), ANR INDUCED (ANR-17-CE03-0008) and EUR PGSE is gratefully acknowledged.

## 1.1. Introduction

We are laymen on most of the knowledge we claim to possess. Most of today's science is too complex for an individual to understand it at first hand. It is often so complex that even for the experts who do understand it, it is hard to convey the evidence supporting their claims in a convincing way. In such cases, our knowledge relies much more on the word of these experts than on the evidence they can convincingly provide. The importance of our confidence in experts is even higher if one considers that, on topics such as climate change or the COVID-19 pandemic, science is highly uncertain. This is not only because data is scarce. The mechanism through which tobacco causes lung cancer is well understood and can give rise to a precise estimate of the chances of getting cancer upon smoking a packet a day. Issues like climate change give rise to various models predicting widely different probabilistic scenarios. The uncertainty over the models themselves is of a different nature than the mere randomness over outcomes that one has to face even with a single well-established model.

In this paper, I test the expert-laymen bound of trust by studying the transmission of information in the context of this complex, uncertain science. Information is about models, which I represent as probability distributions over states of the world. The transmission is strategic, as the sender (the expert) does not necessarily have the same interests as the receiver (the decision maker). For instance, the expert can be concerned with externalities among decision makers on issues such as global warming or the spread of a deadly virus. The expert reviews a set of scientific models and decides which is the most accurate. This model is the expert's type. He then communicates its findings to the decision maker who acts upon it. Given the strategic nature of the communication, the expert is typically not able to truthfully reveal which model is the most accurate. The resulting uncertainty over models creates a situation which has been extensively studied and designated as ambiguity. It calls for the use of specific ambiguity-sensitive preferences for the decision maker.

The game I study is in the tradition of Crawford and Sobel (1982)'s (hereafter CS) cheap talk game. The main difference lies in the fact that communication bears on models, seen as probability distributions over states of the world, rather than on states themselves. I mainly focus on the cases where the receiver displays maxmin expected utility preferences (MEU) or subjective expected utility (SEU) ones. The sender knows the state-generating model which is his type. At equilibrium, the sender designates a set of models as containing his type. In the SEU case, the equilibria of the game are similar to CS. But in the MEU case, the change in the nature of information has a major impact on the outcome of the game. Because of ambiguity aversion, the most pessimistic model is a strong point of attraction for the receiver. When the sender's preferred action leans towards the recommended one in this model, his influence is extremely high. When his interest is to induce an action in the opposite direction though, his influence is nonexistent.

Two cases may arise. In the first case, no univocally worst state, in terms of utility for the receiver, can be identified. Then, I say that science is divided because models putting a lower probability on one state do not necessarily decrease the expected welfare. In this case, an ambiguity averse behaviour is a hedging process against uncertainty. The receiver acts as if the worst possible combination of probabilities over states, an interior element of the set of types, would realise. For types of senders on one side of that element, whatever the difference of interest between both parties, the sender is unable to convey any information. For types in the other side, the precision of information transmission depends on the difference of interest between them. In the second case science is consensual and ambiguity aversion consists in acting as if the state-generating model was the one giving more weight to the univocally worst state. Communication then falls in one of two previous categories. In both cases, it is always in the sender's interest to be as precise as he can, even after he learns his type. This second result also strongly contrasts with the SEU case, where such incentives do not exist. Under SEU preferences, the precision of information transmission depends only on the difference of interest between both parties. When the latter is small, information transmission can be almost perfect. But once the sender learns his type, he does not always have an interest in being as informative as he can.

Restricting attention to the popular linear-quadratic example of CS, I characterise all equilibria when the receiver has both MEU and SEU preferences. In that context I show that when the receiver is MEU, less alignment in the players' interest is required for an equilibrium with a given number of cut-off types to be possible. Finally, I extend the model to the case of  $\alpha$ -MEU decision-making, a preference relation that allows to continuously vary the level of ambiguity aversion. I show that, whatever the misalignment, there is a degree of ambiguity aversion such that an absence of consensus in science leads to the asymmetry in information transmission I observed in the MEU case. This result suggests a form of robustness of what I observe in the MEU case. In addition, I show that ambiguity aversion always eases information transmission.

My results apply when, in any state of the world, trade-offs are to be made. In a pessimistic global-warming scenario, a high level of green house gases (GHG) abatement should be chosen, at the expense of economic growth. In an optimistic one, the opposite choice should be made. When decision-making is of this nature, two situations are to be noticed. Either one state of the world is worse than the others, whatever the decision maker's action. Then, science is necessarily consensual. Or conversely, the decision maker is caught between a rock and a hard place, and science can be divided. Climate change is a good example of the former situation. Scenarios such as the melting of the Antarctic ice sheet or the collapse of the Atlantic thermohaline circulation have been called "tipping elements" [Lenton et al. \(2008\)](#) because they imply a radical change in the climate system. In other words, they are worse than any other. Paradoxically, it is the existence of these catastrophic threats that creates the conditions for information transmission over all models. Conversely, current decisions regarding the COVID-19 pandemic are a choice between the lesser of two evils. One possibility is that the virus lethality is limited and that sparse sanitary measures are enough to contain casualties, while safeguarding the economy. But the opposite possibility cannot be discarded, triggering decisions which would lead to a much more rigorous limitation of social life ([Hollingsworth et al., 2011](#)). Arguably, both scenarios have a comparable profile of consequences. In this situation, science can be divided and my results show that information can only be conveyed regarding

the models that consider the optimistic scenario (epidemiologically speaking) as the most likely. More generally, models that call for more efforts in the provision of a public good, such as social distancing or reduction in GHG emissions, will be on the inaudible side of science.

The first specificity of my paper is to model complex science as non-certifiable information. Complex science underlies some of the greatest challenges our societies have to face. Consider the estimation of the effects of GHG on global temperature, which relies heavily on *black box* prospective computer simulations. Firstly, the process through which these simulations provide predictions is obscure; as pointed out by [Pidgeon and Fischhoff \(2011\)](#), black box simulations are hardly considered as convincing supporting evidence, even for scientists whose disciplines use observational methods. It is also extremely difficult for an expert of this field to justify why a given prospective simulation was chosen, a given methodology implemented or given assumptions made. What distinguishes an expert is precisely his direct understanding of the scientific foundation supporting existing models and of their relative quality. The modelling choice made by epidemiologists in order to evaluate the impact of sanitary measures on the COVID-19 pandemic is another good example. Two main approaches exist: process-based models, that try to capture the mechanisms by which diseases spread and curve-fitting approaches that aim at mathematically approximate the shape of the growth of the epidemic ([Ferguson et al., 2003](#)). The latter class of models does not attempt to characterise the underlying transmission process. As argued by [Berger et al. \(2020\)](#), choosing among these models is a fine art. It requires balancing between simplicity and comprehensiveness as a function of available data and the general understanding of the underlying mechanism. A task which makes the expert who he is. A final appropriate example, is the one of economists when they represent the social world through models. Constantly we have to navigate among modelling choices for the sake of tractability, compatibility with the rest of the literature or empirical testability. Economic modelling is complex because it requires this expertise. The resulting choices can be extremely hard to justify outside of the profession, a difficulty that has and still does attract a lot criticism.

The second specificity of my paper is to identify model uncertainty as the specific uncertainty surrounding scientific knowledge. Consider again the case of the effects of GHG emissions on the global temperature. Its estimation differs widely among the numerous existing models, largely because they rely on very different modelling choices. Predicting the impact of the rise of global temperature involves, for instance, modelling the socioeconomic response of our societies. As argued by [Heal and Millner \(2014\)](#), this can be done in a great variety of ways, leading to model uncertainty. The same challenge is present in the management of the COVID-19 pandemic. At the beginning of the pandemic, there was uncertainty about some of the more fundamental characteristics of the virus, such as its transmission channel, assuming one or another medium did, naturally, highly impact the resulting public policy recommendation (see [Hellewell et al. \(2020\)](#) or [Anderson et al. \(2020\)](#)). Following a tradition in statistics and decision theory dating back at least to [Wald \(1949\)](#) (see [Marinacci \(2015\)](#) for a survey), I represent scientific models as probability distributions over states of the world. Scientific models are simplified representations of reality, or models, capturing the main effects at stake in a given situation. For different courses of actions, a model predicts consequences as a function of the state of the environment, and uses probabilities to estimate their likelihood. When various models exist to represent the same phenomena, and offer different predictions, we face *model uncertainty*. Under model uncertainty, preferences generally fail to satisfy the expected-utility requirements, as famously pointed out by [Ellsberg \(1961a\)](#). In particular, decision makers may display ambiguity aversion. An individual is exposed to ambiguity when the expected payoff to his strategy varies with the probabilities over which he is uncertain. An ambiguity-averse individual will tend to favour strategies that reduce that exposure. MEU preferences, introduced by [Gilboa and Schmeidler \(1989a\)](#), which I will focus on throughout most of the paper, are the more popular example of preferences capturing that trend. [Ghirardato et al. \(2004b\)](#)'s  $\alpha$ -MEU preference is the most tractable extension of MEU, where  $\alpha$  captures the degree of ambiguity aversion of the decision maker. When  $\alpha = 1$ , the two criteria coincide.

The third specificity of my paper is to take into account the consequences of

the often overlooked difference of interest between experts and decision makers regarding science's relative lack of persuasive power. The case of climate change is again an example of utmost importance. Consider a receiver as a country deciding on its GHG emissions in the widely studied context of a game of contribution to a public bad ([Barrett, 1994](#)). Such a receiver benefits from her own emissions but fails to internalise the externalities she produces on others, leading to inefficiently high levels of emission. An expert communicating on the corresponding scientific knowledge might be expected to take this inefficiency into account. As a result, if he cares about all countries' welfare, there is always an asymmetry of interests between the two. The rate of vaccination in a population is a similar example. Vaccination is a public good: the higher it is, the better the population is protected from the disease. Yet, individuals face a private cost in doing so and may be assumed to care only about their own welfare. As pointed out by [Geoffard and Philipson \(1997\)](#), in this context, the overall rate is inefficient. Again, there is an asymmetry of interest between a public health authority acting as a social planner and the individual members of the population.

This study contributes to the recent literature on cheap talk communication with ambiguity sensitive preferences. [Kellner and Le Quement \(2017\)](#) were the first to study this question. In their model, communication is on states of the world. They allow for Ellsbergian communication strategies which are a kind of mixed strategy of the sender, where the mixing probability is ambiguous. They show that the use of these strategies reduces misalignment between players, creating equilibria which ex-ante Pareto dominates the corresponding ones in CS. [Kellner and Le Quement \(2018\)](#) explore a simple two actions two states setting, with only standard mixed strategies allowed, but an ambiguous prior over the states. They show that the optimal communication strategy of the sender is a randomisation over partitions. Because my communication is over probability distributions and only pure strategies are allowed in my model, these results differ from mine. In addition, as pointed out by [Hanany et al. \(2020\)](#), because communication is over states, the updating assumed in these papers violate sequential optimality. This is an issue I don't face when communication is over models.

[Hansen and Sargent \(2001\)](#) and [Hansen et al. \(2006\)](#) have explored the effect

of ambiguity aversion on model uncertainty in the context of dynamic decision making, showing its connection with robust control. [Millner et al. \(2013\)](#) and [Berger et al. \(2016\)](#) have argued for the relevance of model uncertainty and ambiguity aversion in the context of climate change management, where knowledge is scarce. They show that under ambiguity averse preferences, model disagreement is the main driver of GHG abatements. This paper belongs to that line of thought, highlighting the informational and decisional consequences of this type of uncertainty when the source of information is explicitly modelled.

This paper also relates to a continuing debate in epistemology regarding the role of testimony in the foundation of knowledge. In classical epistemology, beliefs qualify as knowledge only if by perception or inference one can verify their truth. This position has been called *reductionist* and has notably been defended by [Hume \(1740\)](#) and [Chisholm et al. \(1989\)](#). But then, why can we say, for instance, that we know that GHG emissions are responsible for global warming? For most of us, this comes neither from perception nor from logical inference. As argued by [Burge \(1993\)](#), perception and inference cannot be seen as *warrants* for most of what we collectively designate as knowledge. An alternative *anti-reductionist* approach argues in favour of adding testimony to the list of primary warrants of knowledge ([Hardwig, 1985](#)). For supporters of this view, it is the confidence in an expert's testimony which rationally entitles the layman to hold the expert's judgement for knowledge ([Goldman, 2001](#)). It is the strength of this bond of trust that epistemologically *entitles* the layman to knowledge. This paper's contribution is to formally model the relationship of trust between expert and layman as a strategic interaction. How much expert-based knowledge the layman is entitled to possess is the information he holds *at equilibrium*. For instance, in the absence of misalignment, the layman is *entitled* to the same knowledge as the expert, as the former has no strategic reason to manipulate his information. The study of this game's equilibria thus contributes to the study of the foundation of expert based-knowledge.

Section [1.2](#) introduces the base model and provides important preliminary results (Proposition [1](#)). Section [1.3](#) establishes general results regarding the structure of equilibria. No full revelation can happen at equilibrium: the sender never

discloses his private information on models. Yet, when the interest of both parties is not too distant, partial information transmission can happen. All equilibria are outcome equivalent to those where the sender credibly points out an interval of models containing the most accurate one (Proposition 2). In general, multiple equilibria may exist in which the designated intervals are more or less broad (Proposition 3). In section 1.4, I show that when the receiver has MEU preferences, information transmission can only be conveyed for models below a given threshold, even if misalignment is arbitrarily small (Theorem 1). I show that the sender always prefers to convey as much information as possible. That is, I show that all equilibria can be ranked by informativeness and that the sender is always interim better off (i.e. after having learned his type) playing the most informative one (Theorem 2). This does not hold when the receiver has SEU preferences. In section 1.5, adapting from the CS linear-quadratic example with uniform prior I characterise equilibrium cut-off types and show that cells are constant in size when the receiver is MEU (Corollary 1). Section 1.6 extends to  $\alpha$ -MEU preferences and shows that whatever the misalignment, there is a degree of ambiguity aversion for which no information transmission is possible for types above a given threshold (Proposition 9 and 10). The appendix contains all generalisations and proofs.

## 1.2. Setup

### 1.2.1. Primitives

I consider a game of communication between an expert acting as a sender S (he), and a decision maker acting as a receiver R (she). Let  $\mathcal{A} = \mathbb{R}$  be the set of actions of R and let  $\Omega = \{0, 1\}$  be the set of possible states of nature. For  $i = S, R$ , let  $u_i : \mathcal{A} \times \Omega \rightarrow \mathbb{R}$  be the von Neumann-Morgenstern utility function of player  $i$ , that maps her actions and the state into her welfare. I start by making the following assumptions:

**Assumption 1 (Utilities - Crawford and Sobel (1982)).**  $u_i$  is assumed twice continuously differentiable and strictly concave in  $a$ . For every  $\omega \in \Omega$ , there is  $a \in \mathbb{R}$  such that  $\frac{\partial u_i(a, \omega)}{\partial a} = 0$ . For all  $a \in \mathbb{R}$ ,  $\frac{\partial u_i(a, \omega)}{\partial a}$  is strictly increasing in  $\omega$ .

This assumption implies that  $u_i$  admits a unique maximum for each state. Define  $a_i(\omega) = \arg \max_{a \in \mathcal{A}} u_i(a, \omega)$  this maximum. It is the optimal action of player  $i$  under perfect information that the state is  $\omega$ . Assumption 1 ensures that  $a_i(\omega)$  is strictly increasing in  $\omega$ . For instance, in the context of our climate application, think of  $\mathcal{A}$  as the level of GHG abatement and of  $\Omega$  as the set of climate scenarios. I call  $\omega = 1$  ( $\omega = 0$ ) the high (low) state as it is the one where the optimal action is the highest (lowest). Taking the climate example again, where  $\mathcal{A}$  represents GHG abatements, a level of contribution to a public good, the high state is the one where climate damage is the highest. Conversely, if  $\mathcal{A}$  captures a level of social distancing, as in the COVID-19 example, the high state is the one where the mortality of the virus is the highest. The choice of an abatement level is the result of a trade-off between economic growth (positively correlated with abatements) and potential damages created by global warming. In the climate example, assumption 1 states that for any climate scenario, there is a single optimal abatement level. A higher abatement level  $a > a_i(\omega)$  is not optimal for  $i$  because it might create too much climate damage. A lower abatement level  $a < a_i(\omega)$  is neither optimal for  $i$  as it implies to reduce economic growth too much.

There is model uncertainty in the sense that, ex-ante, it is not known according to which distribution the state is drawn. Instead, there is a family of Bernoulli distributions  $\mathcal{D} = \{p_\theta | \theta \in [\underline{\theta}, \bar{\theta}]\}$ , where  $\underline{\theta}, \bar{\theta} \in [0, 1]$ , that potentially generates the true state, where  $p_\theta$  is the probability mass function of a Bernoulli distribution of parameter  $\theta$ :

$$p_\theta(\omega) = \begin{cases} \theta & \text{if } \omega = 1 \\ 1 - \theta & \text{if } \omega = 0 \end{cases}$$

There is a bijection between the sets  $\mathcal{D}$  and  $\mathcal{C} = [\underline{\theta}, \bar{\theta}]$ . In the rest of the paper,

for simplicity, I will specify all the communication strategy on the set  $\mathcal{C}$  which will be referred to as the set of *models*. Let  $A_i(\theta) = \operatorname{argmax}_{a \in \mathcal{A}} \mathbb{E}_\theta(u_i(a, \omega))$  be the optimal action in the eyes of player  $i$  under model  $\theta$ , where  $\mathbb{E}_\theta(u_i(a, \omega)) = (1 - \theta)u_i(a, 0) + \theta u_i(a, 1)$ .

**Assumption 2 (Model misalignment).** *For any model, the optimal actions of S and R are always misaligned:*

$$A_S(\theta) > A_R(\theta) \text{ for all } \theta \in \mathcal{C}$$

Assumption 2 states that regardless of the model, there is always a difference of interest between S and R such that optimal actions are ordered in the same way<sup>1</sup>. Note that excluding the case where  $A_S(\theta) < A_R(\theta)$  for all  $\theta \in \mathcal{C}$  is without loss of generality, as all results are symmetrical.

Finally, notice that the sorting condition over states of Assumption 1 implies a sorting condition over models.

**Lemma 1.** *Assumption 1 implies that:*

$$\frac{\partial^2 \mathbb{E}_\theta(u_i(a, \omega))}{\partial a \partial \theta} > 0$$

Lemma 1 states that the marginal utility of actions is increasing with  $\theta$ . As, for a given model, the expect utility of actions is single-peaked, it implies that the optimal action of players,  $A_i(\theta)$ , is a strictly increasing function of  $\theta$ .

## 1.2.2. Strategic interaction

Ex-ante, both players are in a situation of model uncertainty, also called ambiguity. In order to model the way R acts under model uncertainty, I will consider two

---

<sup>1</sup>In appendix 1.8.1, I show that Assumption 2 is implied by the equivalent assumption made on optimal actions as a function of the state (as in CS) plus an assumption on the ordering of the marginal utility of actions of both players.

separate cases. First, I will consider the case where they evaluate actions under uncertainty through the maxmin decision criteria (MEU) proposed by [Gilboa and Schmeidler \(1989a\)](#). According to [Gilboa and Schmeidler \(1989a\)](#), in addition to their utility function, players are characterised by a set of priors over  $\Omega$ , which I will assume to be  $\mathcal{C}$ .  $R$  evaluates action  $a \in \mathcal{A}$  by:

$$V_R^{MEU}(a) = \min_{\theta \in \mathcal{C}} \mathbb{E}_\theta(u_R(a, \omega))$$

Second, I will also consider the case where the receiver's decision making coincides with [Savage \(1972\)](#)'s subjective expected utility (SEU), often identified as a case of ambiguity neutrality. In that case,  $R$ 's preferences are represented by a utility function and a subjective prior over models  $\mu \in \Delta(\mathcal{C})$  admitting a probability distribution function  $g$ . In order to study a case of communication over models which is similar to CS, I will assume that in this case,  $R$  knows the objective distribution according to which the model is drawn. Thus,  $\mu$  is an objective distribution and I also assume that  $\text{supp}(\mu) = \mathcal{C}$ .  $R$  then evaluates action  $a$  under uncertainty through:

$$V_R^{SEU}(a) = \int_{\theta \in \mathcal{C}} g(\theta) \mathbb{E}_\theta(u_R(a, \omega)) d\theta$$

In the following, the MEU case (respectively SEU case) is the one where  $R$ 's evaluation of action coincides with the MEU (respectively SEU) decision criteria.

The timing of the game is as follows:

1. Nature draws the state generating distribution  $\theta_0$ , according to  $\mu$ .  $S$  is privately informed.
2.  $S$  sends a message regarding his type.
3.  $R$  updates her beliefs and chooses an action.

Having learned the state generating distribution<sup>2</sup>  $\theta_0 \in \mathcal{C}$  from nature, S sends a message  $m \in \mathcal{M}$ , where  $\mathcal{M} = [0, 1]$  to R. A signalling strategy for S is the strategy  $\sigma : \mathcal{C} \rightarrow \mathcal{M}$ . An action rule for R is a strategy  $y : \mathcal{M} \rightarrow \mathcal{A}$ . Notice that I will focus only on pure strategies. Let  $\sigma^{-1}(m) \subseteq \mathcal{C}$ , be the set of potential types of S, having received message  $m$ , when S follows strategy  $\sigma$ . An equilibrium  $(\sigma^*, y^*)$  is defined such that:

1. A sender of type  $\theta$  evaluates message  $m$  by:

$$V_S^\theta(m) = \mathbb{E}_\theta(u_S(y^*(m), \omega))$$

$\forall \theta \in \mathcal{C}$ , any  $\sigma^*(\theta) \in \mathcal{M}$  solves  $\max_{m \in \mathcal{M}} V_S^\theta(m)$ .

2. Having received an equilibrium message  $m \in \text{supp}(\sigma^*)$ , an MEU receiver updates her belief such that she evaluates action  $a$  by:

$$V_R^{MEU}(a, \sigma^{-1}(m)) = \min_{\theta \in \sigma^{-1}(m)} \mathbb{E}_\theta(u_R(a, \omega))$$

An SEU receiver is able to update her prior using Bayes' rule such that:

$$g(\theta|m) = \begin{cases} \frac{g(\theta)}{g(\sigma^{*-1}(m))} & \text{if } \theta \in \sigma^{*-1}(m) \\ 0 & \text{if not} \end{cases}$$

R then evaluates action  $a$  by:

$$V_R^{SEU}(a, \sigma^{-1}(m)) = \int_{\theta \in \mathcal{C}} g(\theta|m) \mathbb{E}_\theta(u_R(a, \omega)) d\theta$$

In both cases, R chooses action  $y^*(m)$  which solves  $\max_{a \in \mathcal{A}} V_R^{SEU}(a, \sigma(m))$  (respectively  $\max_{a \in \mathcal{A}} V_R^{MEU}(a, \sigma(m))$ )

---

<sup>2</sup>In appendix 1.8.1 I show that this assumption can be replaced by the one that the sender receives a noisy signal regarding models' likelihood

As usual, any message  $m$  such that  $m \notin \text{supp}(\sigma^*)$  is interpreted as some equilibrium message  $m_* \in \text{supp}(\sigma^*)$ .

### 1.2.3. Consensus and division in Science

Before directing our attention to the equilibria of this game and their specificity, it is useful to take a moment to study the players' decision making under uncertainty. This will help us understanding the specificities of the SEU and MEU decision criteria in my model and provide us with the intermediate result we need for the study of the equilibria. Given the assumptions and lemma 1, when the receiver evaluates action according to SEU, the model is very similar to CS. In that case, one can identify each model as a state in CS setting, where the payoff is the expected utility under that model and  $\mu$  is the prior over states. This case can thus be used as a benchmark. Yet, when the receiver evaluates action according to MEU, the game dramatically changes. In order to see why, let me introduce the following definitions. Regarding models, two special cases arise: either the maximal expected welfare of the receiver is always increasing (or decreasing) with the model, or this is not the case. I call the former case *consensual science* and the latter *divided science*.

**Definition 1.** *Science is consensual if any model that puts a higher probability to the state giving the lowest maximal utility to the receiver decreases his maximal expected welfare.*

$$\forall \theta, \theta' \in \mathcal{C}, \theta < \theta', \begin{cases} u_R(a_R(0), 0) \geq u_R(a_R(1), 1) \Rightarrow \mathbb{E}_\theta(u_R(A_R(\theta), \omega)) > \mathbb{E}_{\theta'}(u_R(A_R(\theta'), \omega)) \\ u_R(a_R(0), 0) < u_R(a_R(1), 1) \Rightarrow \mathbb{E}_\theta(u_R(A_R(\theta), \omega)) < \mathbb{E}_{\theta'}(u_R(A_R(\theta'), \omega)) \end{cases}$$

*Science is divided if it is not consensual.*

Consensus in science is a monotonicity condition on the maximal expected

welfare of the receiver  $\mathbb{E}_\theta(u_R(A_R(\theta), \omega))$ : the more likely the state that gives the lowest maximal utility the lower the maximal expected welfare of the receiver.

Take the climate example of the introduction where the high state is the one where a catastrophic event happens and in the low state the one where it does not. Then, all models agree that the more likely the catastrophic state, the lower the maximal expected pay-off of the receiver. Now consider the converse case where science is divided. Then, the monotonicity assumed above does not hold any more: increasing the likelihood of a given state does not necessarily lower the maximal expected welfare of the receiver. For this to be the case it must be that no state fully dominates the other in terms of utility for the receiver. Because of the single crossing assumption I made on utilities, both states can give the same utility for a given action in  $(a_R(0), a_R(1))$  at most once.

**Definition 2.** Define  $\tilde{a} = \operatorname{argmax}_{a \in \mathcal{A}} \min_{\omega \in \Omega} u_R(a, \omega)$  as the precautionary action and  $\tilde{\theta} \in [0, 1]$  such that  $A_R(\tilde{\theta}) = \tilde{a}$  as the cautious model.

$\tilde{a}$  is the action that maximises the function that gives the worst possible utility to the receiver. I call it the precautionary action because it is the optimal action anticipating that the worst state will always realise.  $\tilde{\theta}$  is the model - not necessarily in  $\mathcal{C}$  - for which the precautionary action is the optimal action.<sup>3</sup> Notice that for the utility functions I have assumed, if both states give that same utility for a given action in  $(a_R(0), a_R(1))$ , this action is the precautionary action  $\tilde{a}$ . Then  $\tilde{\theta} \in (\underline{\theta}, \bar{\theta})$  and the maximal expected utility of the receiver is decreasing for models putting a lower weight on the high state than the cautious one ( $\theta < \tilde{\theta}$ ) and increasing for the others ( $\theta > \tilde{\theta}$ ).

To illustrate, consider the following parametric example:

---

<sup>3</sup>The fact that  $\tilde{\theta}$  exists and is unique is proven in Lemma 3

**Linear-quadratic example:**

- $u_S(a, \omega) = -(a - \omega - b)^2 - c\omega$  where  $b > 0$  and  $c \in \mathbb{R}$
- $u_R(a, \omega) = -(a - \omega)^2 - c\omega$
- $\mathcal{C} = [0, 1]$  and  $\mu \sim U(\mathcal{C})$

Then:

$$\begin{cases} A_S(\theta) = \theta + b \\ A_R(\theta) = \theta \end{cases}$$

The example above is similar to CS's linear-quadratic one, to the difference of the  $-c\omega$  term. In function of the value  $c$ , the maximal utility in a given state is either higher or lower than in the other. When  $c = 0$ , both states are comparable, in the sense that under perfect information, the receiver could achieve exactly the same pay-off in both of them. To the contrary, if for instance  $c > 0$ , state 0 gives a higher maximum pay-off to the receiver than state 1.

As figure 1.1 shows, when  $c \in (-1, 1)$  both states give that same utility to  $R$  for  $\tilde{a} = \frac{1+c}{2} \in (0, 1)$  and we are in a case of divided science and  $\tilde{\theta} = \frac{1+c}{2}$ . Consider the special case where  $c = 0$ . Models on both side of  $\tilde{\theta} = \frac{1}{2}$  increase the maximal expected welfare of the receiver. For instance,  $\theta = 0.2$  and  $\theta = 0.8$  both improve the receiver's maximal expected utility compared with  $\tilde{\theta}$ , as illustrated by figure 1.1. Thus, there is no strict ordering over models with respect to  $R$ 's expected utility but a division of  $\mathcal{C}$  in two sets of models with opposite impact on  $R$ 's maximal expected utility. For models on one side of  $\tilde{\theta}$ , the higher the probability of one state, the better  $R$ 's welfare. But for models on the other side of  $\tilde{\theta}$ , the higher the probability of that same state, the worse  $R$ 's welfare.

Conversely, when  $c \geq 1$  we are in a case of consensual science. As illustrated by figure 1.2, the precautionary action then is the optimal action in the worst

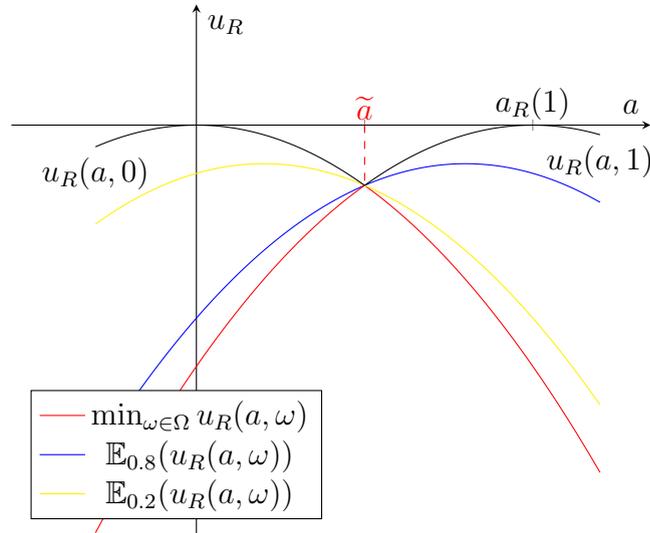


Figure 1.1:  $c = 0$  a case of divided science: for models above 0.5, the receiver's welfare is increasing with the probability of the high state. For models below 0.5, the opposite happens.

state:  $\tilde{a} = a_R(1) = 1$  and  $\tilde{\theta} = 1$ . The maximal expected utility of the receiver is strictly decreasing in  $\theta$ , the probability of the high state. In other words, for any model of  $\mathcal{C}$ , the higher the probability of the high state, the lower R's welfare. Similarly, when  $c \leq -1$  science is also consensual because the maximal expected utility of the receiver is strictly increasing in  $\theta$ , making the precautionary action  $\tilde{a} = a_R(0) = 0$ .

It is clear that the behavioural response of an MEU decision maker will be of a different nature, would science be divided or consensual. In the latter case, the precautionary action consists in anticipating the fully dominated state, thus acting as if the cautious model was the one putting the highest probability on that state. In the former case, the precautionary action consists in hedging against uncertainty, thus acting as if the cautious model was balancing odds between both states in the exact manner that leads to  $\tilde{a}$  as an optimal action. For  $B \subset \mathcal{C}$ , define  $A_R(B) \subset \operatorname{argmax}_{a \in \mathcal{A}} \min_{\theta \in B} \mathbb{E}_{\theta}(u_R(a, \omega))$  the set of optimal actions of a MEU receiver given the set of priors  $B$ . Given these definitions, we can now state the following result:

**Proposition 1.** Define  $B = [\theta_1, \theta_2] \subset \mathcal{C}$  the set of priors of the receiver. Given that

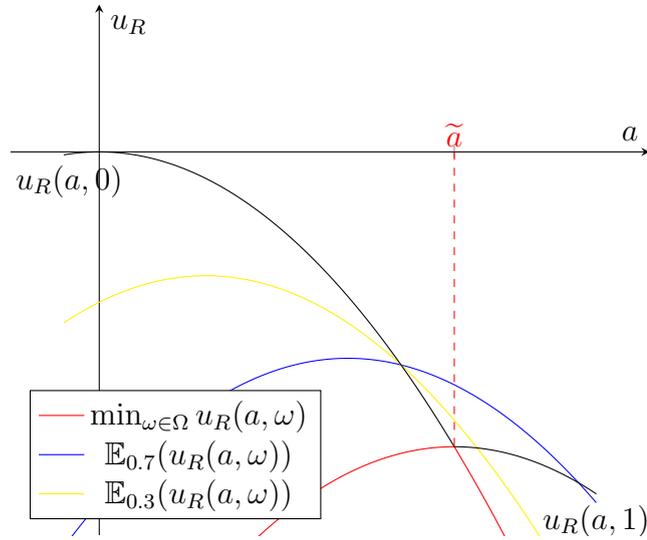


Figure 1.2:  $c = 1$ , a case of consensual science : for all models,  
the receiver's welfare is decreasing with the probability of the high state.

$\theta_0 \in B$ , an MEU receiver has a unique optimal action which is given by:

$$A_R(B) = \begin{cases} A_R(\theta_2) & \text{if } \theta_2 < \tilde{\theta} \\ A_R(\tilde{\theta}) & \text{if } \tilde{\theta} \in B \\ A_R(\theta_1) & \text{if } \theta_1 > \tilde{\theta} \end{cases}$$

Proposition 1 states that an MEU receiver has a unique optimal action for any belief  $\theta_0 \in B$  where  $B$  is an interval of  $\mathcal{C}$ . When she further believes that all models are below  $\tilde{\theta}$  ( $\theta_0 \in [\theta_1, \theta_2]$  and  $\theta_2 < \tilde{\theta}$ ) she optimally acts as if the probability of the high state were maximal. When she believes that all models are above  $\tilde{\theta}$  ( $\theta_1 > \tilde{\theta}$ ) she optimally acts as if the probability of the high state were minimal. Thus, when science is consensual, whatever  $R$ 's belief, she will always act as if the probability of the high state was maximal, or minimal. But when science is divided  $R$  will always act as if the probability of the high state was maximal for beliefs below  $\tilde{\theta}$  ( $\theta_2 < \tilde{\theta}$ ) and minimal for beliefs above  $\tilde{\theta}$  ( $\theta_1 > \tilde{\theta}$ ). Finally, when  $R$  believes that the cautious model could be the state generating model ( $\tilde{\theta} \in [\theta_1, \theta_2]$ ),

she optimally acts as if it were the case. This behavioural change described by Proposition 1, directly caused by the change in the monotonicity of R's maximal expected welfare at  $\tilde{\theta}$ , will be essential in the upcoming results. To the contrary, notice that in the SEU case the cautious model plays no particular part.

### 1.3. Equilibrium analysis

Let us now turn to the study of the equilibrium structure. First, I introduce the following definition:

**Definition 3.** Set  $\{\theta_0, \dots, \theta_q\} \subseteq \mathcal{C}$  such that:

- $\underline{\theta} = \theta_0 < \dots < \theta_q = \bar{\theta}$  where  $\theta_k$ , for  $0 \leq k \leq q$ , is called the  $k$ -th cut-off.
- $\cup_{k=1}^q [\theta_{k-1}, \theta_k] = [\underline{\theta}, \bar{\theta}]$ , where  $[\theta_{k-1}, \theta_k)$ , for  $1 \leq k < q - 1$ , is called the  $k$ -th cell and  $[\theta_{q-1}, \bar{\theta}]$  the  $q$ -th cell.

A  $q$ -cut-off partition equilibrium is an equilibrium of the game where the signaling strategy of  $S$  is uniform on every cell. That is, for  $\theta \in [\theta_{k-1}, \theta_k)$ ,  $\sigma^*(\theta) = m_k$ , for  $1 \leq k \leq q - 1$  and for  $\theta \in [\theta_{q-1}, \bar{\theta}]$ ,  $\sigma^*(\theta) = m_{q-1}$ .

A  $q$ -cut-off partition equilibrium is an equilibrium where there is a partition of the set of types in  $q$  cells. For any cell of this partition, any sender who is in that cell credibly sends the same message to the receiver. Having received that message, the receiver learns in what cell the sender is and acts optimally.

**Proposition 2.** In every equilibrium of the game, there is a partitioning of  $\mathcal{C}$  in a finite number of cells where every cell induces a distinct action. Thus, any equilibrium is outcome equivalent to a partition equilibrium.

The proof of Proposition 2 starts by showing that the number of actions induced at equilibrium is finite. The argument is similar to the one given in CS's

Lemma 1 and follows from both the concavity of  $S$ 's evaluation of actions and the fact that the optimal actions of  $R$  for a given belief  $B \subset \mathcal{C}$  is in the convex hull of the optimal actions for every element of  $B$ . Then I show that types that induce a given action must form an interval. This is a consequence of the concavity of  $S$ 's evaluation of actions.

Proposition 2 shows that there is a finite partition of  $\mathcal{C}$  where types in every cell induce a given action from the receiver. Notice that this does not imply that types in every cell send the same message, as it is possible that different messages induce the same action. As a result, every equilibrium is not necessarily a partition equilibrium, but must be outcome equivalent to one. In the following, we focus only on partition equilibria. Notice that there is always at least one partition equilibrium: the babbling equilibrium, where all types send the same message.

In the following, I give a characterisation of all partition equilibria of the game.

**Proposition 3.** *In any partition equilibrium of the game  $(\sigma_q^*, y^*)$ , the cut-off types  $\theta_0^q, \dots, \theta_q^q$  are defined such that for  $k \in 1, \dots, q$ :*

$$V_S^{\theta_k^q}(y^*(m_{k-1}^q)) = V_S^{\theta_k^q}(y^*(m_k^q)) \quad (1.1)$$

where  $m_k^q$  is the equilibrium message of types  $\theta \in [\theta_k^q, \theta_{k+1}^q]$ .

Figure 1.3 represents the interim utility of  $S$  when his type is  $\theta_k$ . As a convex combination of concave and single peaked functions, it is concave and maximal at  $A_S(\theta_k)$ . Figure 1.3 illustrates that  $m_{k-1}$  and  $m_k$  are equilibrium messages because they induce actions that give the same level of welfare to  $S$ . As a result,  $\theta_k$  is a cut-off type.

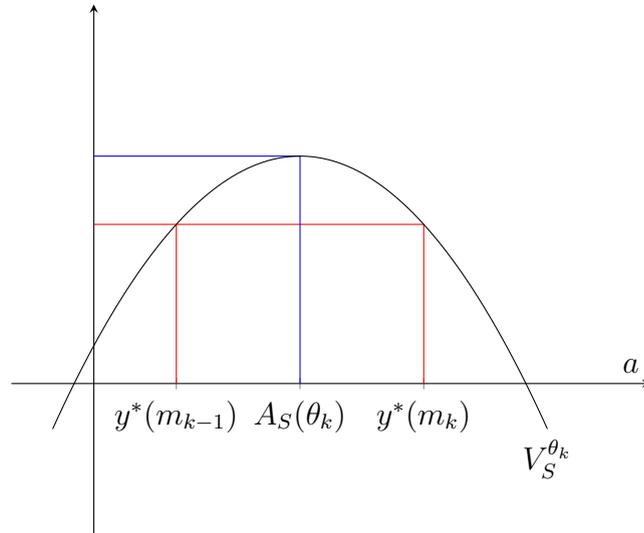


Figure 1.3: Identifying cut-offs

## 1.4. General results

In the following, I go one step further in the characterisation of the game's equilibria. Recall that in the MEU case, R evaluates action  $a$  by:

$$V_R^m(a) = \min_{\theta \in \sigma^{*-1}(m)} \mathbb{E}_\theta(u_R(a, \omega))$$

For any  $m \in \mathcal{M}$ , call  $\sigma^{*-1}(m) = [\theta_1, \theta_2]$ . Recall from Proposition 1 that:

$$A_R(\sigma^{*-1}(m)) = \begin{cases} A_R(\theta_2) & \text{if } \theta_2 < \tilde{\theta} \\ A_R(\tilde{\theta}) & \text{if } \tilde{\theta} \in B \\ A_R(\theta_1) & \text{if } \theta_1 > \tilde{\theta} \end{cases} \quad (1.2)$$

**Theorem 1.** *When the receiver has MEU preferences, all cut-offs in  $(\underline{\theta}, \bar{\theta})$  are below  $\tilde{\theta}$ .*

An upward misaligned sender is never capable of conveying information over

$[\tilde{\theta}, \bar{\theta}]$  when the receiver has MEU preferences. Assume that there is a  $q + 1$  cut-off equilibrium. Recall the characterisation result of partition equilibria given by Proposition 3. For  $\theta_q$  to be a cut-off type, it must be that the message sent by types in the cell below and above  $\theta_q$  induce actions that gives the same utility to a sender of type  $\theta_q$ . Would  $\theta_q$  be a cut-off type, following (1.2) we would have that:

$$\begin{cases} y^*(m_{q-1}) = A_R(\sigma^{*-1}(m_{q-1})) = A_R([\theta_{q-1}, \theta_q)) = A_R(\tilde{\theta}) = \tilde{a} \\ y^*(m_q) = A_R(\sigma^{*-1}(m_q)) = A_R([\theta_q, \theta_{q+1}]) = A_R(\theta_q) \end{cases}$$

Yet, as illustrated by Figure 1.4, the utility of the sender induced by  $m_{q-1}$  is always lower than the one induced by  $m_q$ . This is a direct consequence of the change in the monotonicity of R's maximal expected welfare at  $\tilde{\theta}$ . When R believes that the cautious model could be the state generating model, she optimally acts as if it were the case. When she believes that  $\theta_0 \in [\theta_q, \theta_{q+1})$  and  $\theta_q > \tilde{\theta}$  she will act as if the model was  $\theta_q$ . As a result, because S is upwards misaligned we have that:

$$\tilde{a} < A_R(\theta_q) < A_S(\theta_q)$$

and as  $V_S^{\theta_q}$  is strictly increasing for  $a \leq A_S(\theta_q)$  it is impossible that messages sent by types in the cell below and above  $\theta_q$  induce actions that gives the same utility to a sender of type  $\theta_q$ . As a result, the indifference between actions induced by messages  $m_{q-1}$  and  $m_q$  needed for  $\theta_q$  to be a cut-off type (as displayed in figure 1.3) is impossible.

A consequence of Theorem 1 is that when science is consensual such that  $\tilde{\theta} \leq \underline{\theta}$ , the only equilibrium is the babbling equilibrium. That is, whatever the sender's type, whatever the message he sends, the induced action is always the same. In other words, in this situation, the sender is inaudible. The next proposition shows

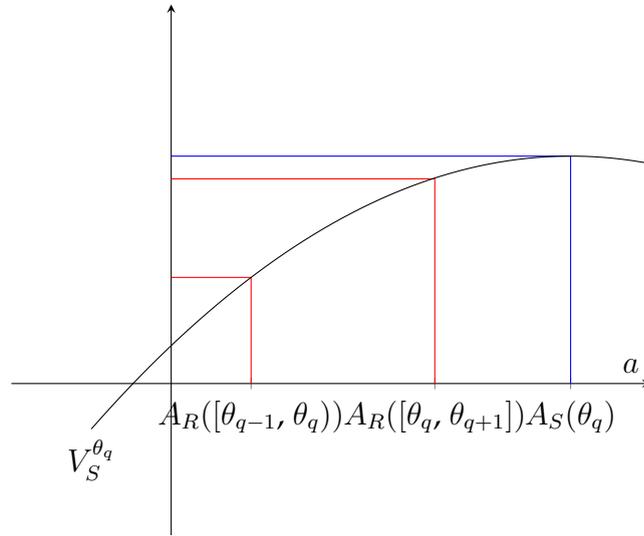


Figure 1.4: MEU best responses for  $\theta_{q-1} < \tilde{\theta} < \theta_q$

that all partitional equilibria of our game can be built from the same finite set of cut-off types.

**Proposition 4.** *When the receiver evaluates actions following the MEU criteria there are  $M > 0$  partition equilibria. Call  $\theta_0 < \dots < \theta_M$  the cut-offs of the equilibrium with most cut-offs. Then the  $q$  cut-off partition equilibrium is defined by cut-offs  $\theta_0 < \theta_{M-q} < \dots < \theta_M$ , for  $0 \leq q \leq M$ .*

As illustrated by Figure 1.5, in the context of MEU preferences, the cut-off types of every equilibrium are the same. That is: the interior cut-off of the one cut-off equilibrium is the same as the first interior cut-off of the two cut-off equilibrium. Similarly, the second cut-off of the two cut-off equilibrium is the same as the second cut-off of the three cut-off equilibrium. The same holds for all cut-offs of the existing equilibria. In particular, there is always an equilibrium with two cut-offs -  $\underline{\theta}$  and  $\bar{\theta}$  - corresponding to the babbling equilibrium.

Given Theorem 1, all interior cut-offs are in  $[\underline{\theta}, \bar{\theta}]$ . As a result, when S points out an interval of models, R only cares about its upper bound. As a result, cut-offs types will not be determined by an indifference between pairs of expectations of intervals of models (the two adjacent cells) but by an indifference between pairs of expectations of models (the lower bounds of the two adjacent cells). In the former case, each indifference condition depends on three distinct types and

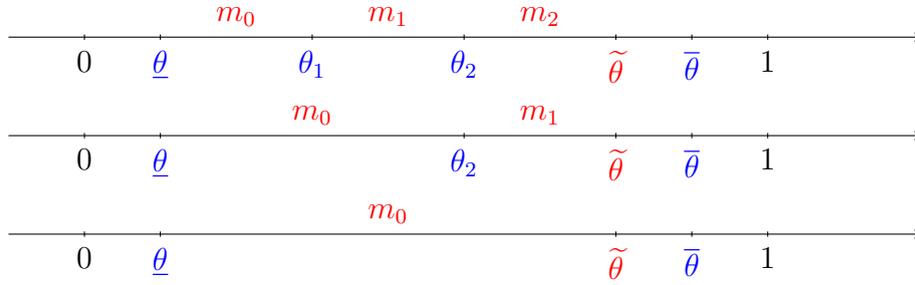


Figure 1.5: MEU equilibria for  $\underline{\theta} < \tilde{\theta} < \bar{\theta}$

the prior. Thus, in order to determine the cut-off types, the entire sequence of indifference conditions is needed. In the latter case, each indifference condition depends on two distinct types only. Given that  $[\underline{\theta}, \tilde{\theta}]$  is a closed set, it is then possible to find the first cut-off starting from  $\underline{\theta}$  and then to iterate the process to find the following ones. In doing so, I derive the cut-off types of the equilibrium that has the most cut-offs. Then, all other equilibria are characterised by the  $q$  first terms ( $1 \leq q \leq M$ ) of that sequence, assuming it has  $M$  elements.

A direct consequence of Proposition 4 is that all equilibria of the game can be ranked by informativeness, something which is never possible in the SEU case.<sup>4</sup> The following result can thus be established regarding interim Pareto dominance among equilibria.

**Theorem 2.** *When the receiver has MEU preferences, the sender is always interim weakly better off by playing the most informative equilibrium strategy*

The intuition of the proof, for S downwards misaligned, is the following. Consider the equilibria described in figure 1.5. Whatever the equilibrium considered, types in  $[\theta_1, \tilde{\theta}]$  will induce the same action  $\tilde{\theta}$ . But types in  $[\underline{\theta}, \theta_1]$  will induce action  $\tilde{\theta}$  in the babbling equilibrium, and  $\theta_1$  in the 3 cut-off equilibrium. Yet, by construction of the latter equilibrium, all types in  $[\underline{\theta}, \theta_1]$  prefer to induce action  $\theta_1$  than  $\tilde{\theta}$ . It follows that the 3 cut-off equilibrium interim Pareto dominates the babbling equilibrium. The same reasoning can be applied regarding types in  $[\underline{\theta}, \theta_2]$  to show

<sup>4</sup>The informativeness ranking comes from the fact that when receiving  $m \in \mathcal{M}$  from a type in  $[\theta_1, \theta_2]$  with  $\theta_2 < \bar{\theta}$  an MEU receiver acts exactly as when receiving  $m' \in \mathcal{M}$  from a type in  $[\theta'_1, \theta_2]$ , for any  $\theta'_1 < \theta_1$ . For an SEU receiver, this behavioural pattern is impossible, the optimal action would necessarily shift to the left.

that the 4 cut-off equilibrium Pareto dominates the 3 cut-off one.

**Discussion.** Theorem 2 gives that  $S$  is always interim better-off communicating following the most informative equilibrium strategy. This result differs significantly from those obtained in CS's framework. Under their monotonicity condition (M), CS show that the ex ante expected payoffs for both Sender and Receiver is maximal for the equilibrium with most cut-off. Condition (M) is satisfied if for any two sequence of cut-off types the  $k$ -th cut-off of each sequence can be ordered in the same direction, for any  $k \geq 1$ . This assumption is in particular verified by the linear-quadratic example. The resulting selected equilibrium is often the one studied in applications. Yet, as already pointed out in CS, ex-ante Pareto dominance is a questionable equilibrium-selection criterion, since once having learned their type, different sender types will necessarily have opposed preferences. CS suggests that ex-ante Pareto dominance could be retained only if there is an equilibrium selection agreement made ex-ante between players or if it can be seen as a convention maintained over repeated plays with several opponents. An alternative approach regarding equilibrium selection has been proposed by [Chen et al. \(2008\)](#). In this paper, a condition on utility functions, NITS, has been proposed. Under this condition, combined with Assumption (M), only the equilibrium with most cut-offs survives in CS's framework. An equilibrium satisfies NITS if the Sender of the lowest type weakly prefers the equilibrium outcome to the outcome induced by credibly revealing his type (if he could). In my case, one could adopt interim Pareto dominance as a selection criterion, which is immune to the limitations of ex-ante Pareto dominance and does not require supplementary assumptions as for NITS. Yet, it brings out the same (most informative) equilibrium and provides a foundation for the attention it receives in applications.

Assume one sticks to the interpretation of  $\mathcal{C}$  as the set of objective possible models, as it is for instance the case for the climate and COVID-19 examples of the introduction. The size of  $\mathcal{C}$  captures the degree of objective imprecision in scientific knowledge. For instance, if  $\mathcal{C} = [0, 1]$  the objective probability of the high state is between 0 and 1. It is then possible to analyse the effects of a change in this objective imprecision. In particular, an increase in precision can move sci-

ence from a state of division (where  $\tilde{\theta}$  is an interior point of  $\mathcal{C}$ ) to one of consensus (where  $\tilde{\theta}$  is at the boundary of  $\mathcal{C}$ ). As exposed until now, such an increase in precision could result in a major change in communication possibilities, making  $R$  fully influential over  $\mathcal{C}$  or to the contrary, completely inaudible.

Take the linear-quadratic example introduced page 10 in the case when  $c = 0$ . Then science is divided because  $\tilde{\theta} = 0.5$  and all interior cut-offs are in  $[0, 0.5]$ , whatever the misalignment  $b > 0$  of the sender. Assume the objective imprecision  $\mathcal{C}$  shifts from  $[0, 1]$  to  $\mathcal{C}' = [0, 0.5]$ . Then science becomes consensual and the partitioning of the set of types is possible over the entire  $\mathcal{C}'$ . Conversely, objective imprecision could shift to  $\mathcal{C}'' = [0.5, 1]$ . Science would then be consensual as well, but partitioning impossible.

## 1.5. The linear-quadratic example

In order to give a further insight of the results in the MEU case, I characterise all partitional equilibria in the context of the parametric example introduced page 10 and adapted from the widely used linear-quadratic example of CS. I also provide the same characterisation for the SEU case.

**Proposition 5.** *In the context of our linear-quadratic example for any  $c \in \mathbb{R}$ :*

- *When  $R$  has SEU preferences, a  $n$ -cut-off equilibrium exists if and only if:*

$$0 < b < \frac{1}{2n(n+1)} \quad (1.3)$$

*and, for  $k \in 1, \dots, n$ , cut-offs are:*

$$\theta_k = \frac{k}{n+1} - 2kb(n-k+1)$$

- *When  $R$  has MEU preferences, a  $n$ -cut-off equilibrium exists if and only if  $c > -1$*

and

$$0 < b < \frac{1}{2n} \quad (1.4)$$

and, for  $k \in 1, \dots, n$ , cut-offs are:

$$\theta_k = 1 - 2b(n - k)$$

Proposition 1.5 shows that  $c$  has no influence on communication in the SEU case. Yet, in the MEU one, when  $c \leq -1$ , the maximal pay-off in state 0 is always lower than in state 1. As a result, the attraction exerted by ambiguity aversion plays against the sender's communication possibilities and no non-babbling equilibrium is possible. Conversely, when  $c \geq 1$  the attraction exerted by ambiguity aversion plays in favour of the sender's communication possibilities in cut-offs can be on the entire set  $\mathcal{C}$ .

A corollary of Proposition 1.5 is that it is possible to characterise each equilibrium's cell sizes.

**Corollary 1.** *Consider a  $q$ -cut-off partition equilibrium. When  $R$  is SEU, for any  $c \in \mathbb{R}$ , cells are increasing in size. For all  $k \in 1, \dots, q - 1$ :*

$$\theta_{k+1} - \theta_k = \theta_k - \theta_{k-1} + 4b$$

*When  $R$  has MEU preferences and  $c > -1$ , non-terminal cells are of constant size. For all  $k \in 1, \dots, q - 2$ :*

$$\theta_{k+1} - \theta_k = 2b$$

*where the cell containing  $\underline{\theta}$  is called the terminal cell.*

In the MEU case non-terminal cells have always the same size ( $2b$ ), whatever the considered equilibrium. In the SEU case, it depends on the considered equilibrium. This illustrates why the general result proved in Proposition 4 holds. If all non-terminal cells have the same size in any given equilibrium, and that in addition the first cut-off is always the same (as proven in Proposition 5), it is straightforward that those equilibria can be ranked by informativeness in the Blackwell sense. Corollary 1 also states that in the SEU case, cells are at least of size  $4b$  and are thus always strictly larger.

The sender is able to induce a finer partition of types when the receiver is MEU. Consider a given positive bias such that it is possible to get a  $n$  cut-off equilibrium with a MEU receiver; then it is not always possible to sustain an  $n$  cut-off equilibrium with an SEU receiver. More precisely : call the supremum of the bias for which a  $n$ -cut-off equilibrium is possible in the MEU case  $b_M(n) = \frac{1}{2n}$ . Call the equivalent value of the bias in the SEU case  $b_S(n) = \frac{1}{2n(n+1)}$ . Both functions are increasing in  $n$ . In addition, for  $n \geq 2$ ,  $b_S(n) = b_M(n(n+1))$ . Thus, there is a  $n$  cut-off equilibrium between an SEU receiver of bias  $b$  and the sender if and only if there is a  $n(n+1)$  cut-off equilibrium between an MEU receiver of bias  $b$ .

## 1.6. $\alpha$ -MEU receiver

### 1.6.1. Optimal actions and structure of equilibria

In this section, I consider the case where R evaluates actions under uncertainty through the  $\alpha$ -maxmin decision criteria proposed by [Ghirardato et al. \(2004b\)](#) ( $\alpha$ -MEU). According to [Ghirardato et al. \(2004b\)](#), in addition to their utility function, players are characterised by two more elements. First, a set of priors over  $\Omega$ , which I will assume to be  $\mathcal{C}$ . Second, a parameter  $\alpha_i \in [0, 1]$  which captures their attitude towards ambiguity. As all the analysis will be conducted at the interim stage,  $\alpha_S$  is irrelevant. Thus, in the following, I will erase the subscript.  $R$

evaluates action  $a \in \mathcal{A}$  by :

$$V_R^\alpha(a) = \alpha \min_{\theta \in \mathcal{C}} \mathbb{E}_\theta(u_R(a, \omega)) + (1 - \alpha) \max_{\theta \in \mathcal{C}} \mathbb{E}_\theta(u_R(a, \omega)) \quad (1.5)$$

Here, the behavioural consequences of ambiguity aversion are captured by  $\alpha$ . It translates the decision maker's weighting between optimistic and pessimistic models regarding his expected utility. When  $\alpha = 1$ , the  $\alpha$ -MEU decision criteria coincides with MEU<sup>5</sup>. Adapting [Ghirardato et al. \(2004b\)](#)'s proposition 20 to our model one can state the following :

**Definition 4** ([Ghirardato et al. \(2004b\)](#)). *Receiver  $i$ , evaluating actions through  $V_R^{\alpha_i}$ , is said to be more ambiguity averse<sup>6</sup> than receiver  $j$ , evaluating actions through  $V_R^{\alpha_j}$ , if and only if*

$$\alpha_i > \alpha_j$$

Thus, for a fixed set of priors and utility function, increasing ambiguity aversion leads the receiver to anticipate an increasingly worst model in terms of expected utility.

As for the MEU case, having received an equilibrium message  $m \in \text{supp}(\sigma^*)$ , an  $\alpha$ -MEU receiver updates her belief such that she evaluates action  $a$  by:

$$V_R^\alpha(a, \sigma(m)) = \alpha \min_{\theta \in \sigma^{-1}(m)} \mathbb{E}_\theta(u_R(a, \omega)) + (1 - \alpha) \max_{\theta \in \sigma^{-1}(m)} \mathbb{E}_\theta(u_R(a, \omega))$$

By a natural extension of the notations introduced above, for  $B \subset \mathcal{C}$ , define

---

<sup>5</sup>In general, for non-symmetric utility functions,  $\alpha$ -MEU does not have SEU as a special case here. For instance, both criteria coincide if the set of models is a singleton. Yet, this set depends of the information conveyed by the sender, which, at equilibrium, is never a singleton.

<sup>6</sup>In the sense of [Ghirardato and Marinacci \(2002\)](#)

$A_R(B) = \operatorname{argmax}_{a \in \mathcal{A}} V_R^\alpha(a, B)$  the set of optimal actions of the  $\alpha$ -MEU receiver when his belief is  $B$ .

Figure 1.6 illustrates the ex-ante evaluation of actions of the receiver in the context of the linear-quadratic example. All valuation functions are located in the blue area and are a convex combination between  $\min_{\theta \in \mathcal{C}} \mathbb{E}_\theta(u_R(a, \omega))$  (in red) and  $\max_{\theta \in \mathcal{C}} \mathbb{E}_\theta(u_R(a, \omega))$  (in black). Then, notice that for a given  $\alpha$   $V_R^\alpha(a)$  is not necessarily single-peaked. For instance, for  $\alpha = 0.3$ ,  $V_R^{0.3}(a)$  is maximal at 0.3 and 0.7<sup>7</sup>.

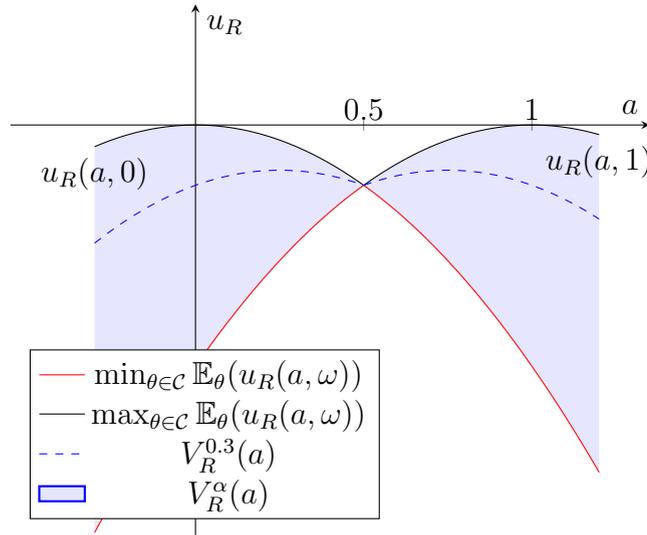


Figure 1.6:  $\alpha$ -MEU ex-ante valuation

I now characterise the set of optimal actions of R for a given set of priors.

**Proposition 6.** *Define  $B \subset \mathcal{C}$  the set of priors of the receiver with minimal element  $\theta_1$  and maximal element  $\theta_2$ . Given this belief, her optimal set of actions  $A_R(B) \subset [A_R(\theta_1), A_R(\theta_2)]$ . In the context of the linear-quadratic example, the set of optimal action of a  $\alpha$ -MEU receiver is given by:*

<sup>7</sup>In the linear-quadratic example for  $c = 0$  there are two optimal actions for any  $\alpha \leq 0.5$ . The fact that this threshold is the one separating love and aversion for ambiguity is non-generic. For sharper utility functions this threshold would be above 0.5. A formal definition is given in assumption 4 of the appendix

$$A_R(B) = \begin{cases} \alpha A_R(\theta_2) + (1 - \alpha)A_R(\theta_1) & \text{if } \theta_2 < \tilde{\theta} \\ \{\alpha A_R(\tilde{\theta}) + (1 - \alpha)A_R(\theta_M) \mid \theta_M \in \operatorname{argmax}_{\theta \in \{\theta_1, \theta_2\}} \mathbb{E}_\theta(u_R(a, \omega))\} & \text{if } \tilde{\theta} \in B \\ \alpha A_R(\theta_1) + (1 - \alpha)A_R(\theta_2) & \text{if } \theta_1 > \tilde{\theta} \end{cases}$$

A direct consequence of Proposition 6 is that ex-ante, for any  $a \in \mathcal{A}$ ,  $\min_{\theta \in \mathcal{C}} \mathbb{E}_\theta(u_R(a, \omega)) = \mathbb{E}_{\bar{\theta}}(u_R(a, \omega))$  and that  $\max_{\theta \in \mathcal{C}} \mathbb{E}_\theta(u_R(a, \omega)) = \max(\mathbb{E}_{\underline{\theta}}(u_R(a, \omega)), \mathbb{E}_{\bar{\theta}}(u_R(a, \omega)))$ . Thus,  $A_R(\mathcal{C}) = A_R(\alpha\tilde{\theta} + (1 - \alpha)\bar{\theta})$  when  $\mathbb{E}_{\underline{\theta}}(u_R(a, \omega)) < \mathbb{E}_{\bar{\theta}}(u_R(a, \omega))$ ,  $A_R(\mathcal{C}) = A_R(\alpha\tilde{\theta} + (1 - \alpha)\underline{\theta})$  when  $\mathbb{E}_{\underline{\theta}}(u_R(a, \omega)) > \mathbb{E}_{\bar{\theta}}(u_R(a, \omega))$  and  $A_R(\mathcal{C}) = \{A_R(\alpha\tilde{\theta} + (1 - \alpha)\underline{\theta}), A_R(\alpha\tilde{\theta} + (1 - \alpha)\bar{\theta})\}$  when  $\mathbb{E}_{\underline{\theta}}(u_R(a, \omega)) = \mathbb{E}_{\bar{\theta}}(u_R(a, \omega))$ . This explains the fact that in the example we considered before, where  $\mathbb{E}_0(u_R(a, \omega)) = \mathbb{E}_1(u_R(a, \omega))$ , optimal actions were not unique. Notice also that, when  $\alpha$  increases,  $A_R(\mathcal{C})$  gets closer to  $\tilde{a}$ , in the euclidean sense. Thus, an increase in ambiguity aversion gets R's ex-ante optimal action closer to the precautionary action.

I now prove that under  $\alpha$ -MEU preferences, all equilibria are still outcome equivalent to a partition equilibria.

**Proposition 7.** *In every equilibrium of the game, there is a partitioning of  $\mathcal{C}$  in a finite number of cells where every cell induces a distinct action. Thus, any equilibrium is outcome equivalent to a partition equilibrium.*

As for the proof of Proposition 2, I start by showing that the number of actions induced at equilibrium is finite. The argument is similar to the one given in CS's Lemma 1 and follows from both the concavity of  $S$ 's evaluation of actions and the fact that the optimal actions of R for a given belief  $B \subset \mathcal{C}$  is in the convex hull of the optimal actions for every element of  $B$ . This is also true when R has  $\alpha$ -MEU preferences, as one can deduce from Proposition 6. Then I show that types that induce a given action must form an interval. This is a consequence of the concavity of  $S$ 's evaluation of actions.

### 1.6.2. Comparative ambiguity aversion

In the following, I'm interested in the effect that ambiguity aversion has on the structure of partitional equilibria. I compare the equilibria of two versions of the game, where the only difference is the degree of ambiguity aversion of the receivers identified to their degree of ambiguity aversion  $\alpha_1$  and  $\alpha_2$ . Notice that the ex-post optimal action  $A_{R_i}(\theta)$  is unaffected by ambiguity aversion, thus, I will erase the subscript. I will only consider the linear-quadratic example introduced before. Recall that in the MEU case, when  $c = 1$  science is consensual and communication is possible over the entire  $\mathcal{C}$ , when  $c = 0$  science is divided and communication is only possible over  $(0, \frac{1}{2})$  and finally when  $c = -1$  science is consensual but no information transmission is possible.

I start by considering the consensual science cases:  $c = 1$  or  $-1$ . In the following, I characterise all the cut-offs of the corresponding partition equilibrium.

**Proposition 8.** *In the linear quadratic example, when  $R$  is  $\alpha$ -MEU, for  $\alpha \notin \{0, \frac{1}{2}, 1\}$ :*

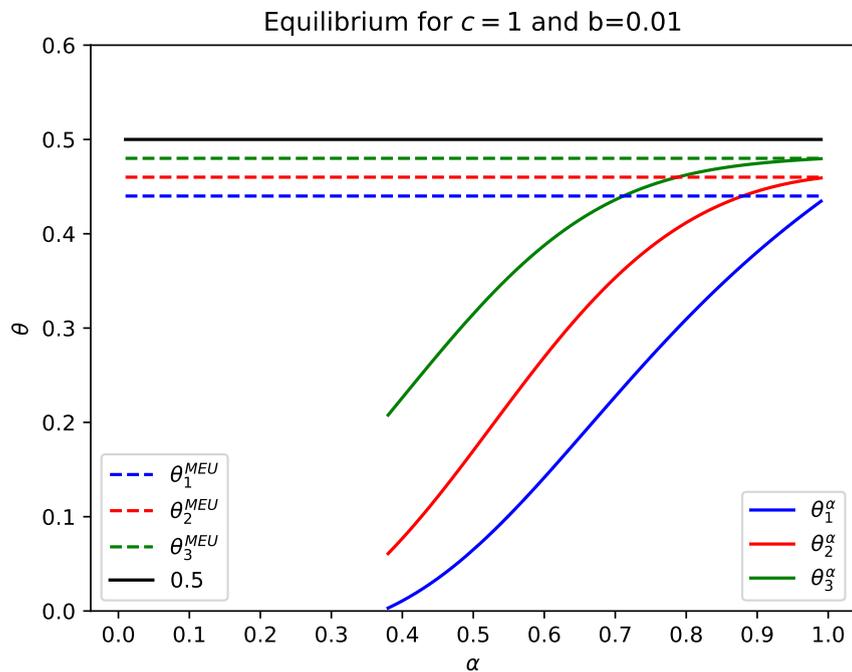
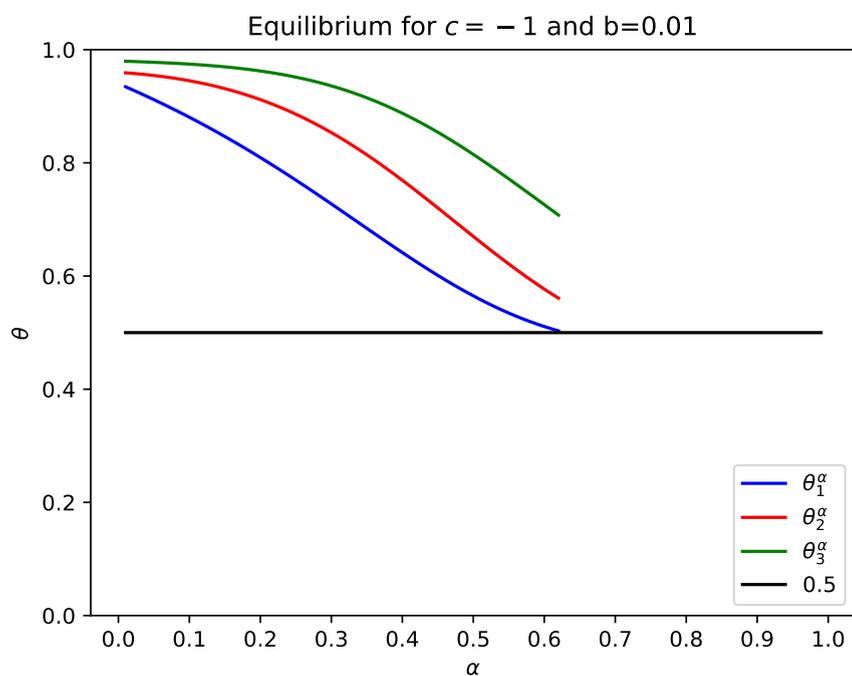
- *When  $c = 1$ , there are  $N > 0$  cut-off equilibria, one by number of cut-offs, and the  $k$ -th cut-off of the  $1 \leq n \leq N$  cut-off equilibrium is given by :*

$$\underline{\theta}_k^n(\alpha) = \left(\frac{1}{2} - \frac{2bn}{2\alpha - 1}\right) \left(\frac{1 - \left(\frac{1-\alpha}{\alpha}\right)^k}{1 - \left(\frac{1-\alpha}{\alpha}\right)^n}\right) + \frac{2bk}{2\alpha - 1}$$

- *When  $c = -1$ , there are  $M > 0$  cut-off equilibria, one by number of cut-offs, and the  $k$ -th cut-off of the  $1 \leq n \leq M$  cut-off equilibrium is given by :*

$$\bar{\theta}_k^n(\alpha) = \left(\frac{1}{2} - \frac{2bn}{2\alpha - 1}\right) \left(\frac{1 - \left(\frac{\alpha}{1-\alpha}\right)^k}{1 - \left(\frac{\alpha}{1-\alpha}\right)^n}\right) - \frac{2bk}{2\alpha - 1} + \frac{1}{2}$$

Figure 1.7 and 1.8 compute the cut-offs of the 3-cut-offs equilibrium as a function of  $\alpha$  for a fixed positive bias  $b = 0.01$ . Notice that for a given level of misalignment, information transmission is possible in both cases, for given levels of

Figure 1.7: 3-cut-offs equilibria for  $c = 1$ Figure 1.8: 3-cut-offs equilibria for  $c = -1$ 

ambiguity aversion. Thus, the asymmetry of the MEU case does not survive at any level of ambiguity aversion. Yet, simulations suggest that, when  $c = -1$ , cut-

off values decrease with  $\alpha$  towards  $\tilde{\theta}$ . Thus for a given bias, there is a level of ambiguity aversion from which all types in  $\mathcal{C}$  must pool. Conversely, simulations suggest that, when  $c = 1$ , cut-off values continuously increase with  $\alpha$  towards their MEU values.

I now formally prove that in the case of consensual science, no communication is possible in  $c = -1$  when  $\alpha$  is above a given threshold. In addition, I show that for a given bias, ambiguity aversion eases the existence of a  $n$  cut-off equilibrium, for  $n \geq 2$ .

**Proposition 9.** *In the context of the linear-quadratic example, when  $c = -1$ ,  $b > 0$  and  $\alpha \in (\frac{1}{2}, 1)$ :*

1. *There is  $\alpha(b) \in (1/2, 1)$  such that for  $\alpha \geq \alpha(b)$ , no information transmission is possible in  $[0, 1]$ . Moreover,  $\alpha(b)$  is a decreasing function.*
2. *For two receivers  $\alpha_1$  and  $\alpha_2$  such that  $\alpha_1 < \alpha_2$ , if there is a  $n \geq 2$  cut-off equilibrium between  $S$  and  $\alpha_1$ , there is a  $n$  cut-off equilibrium between  $S$  and  $\alpha_2$*

Thus, as suggested by the simulations, when  $c = -1$ , for a given bias, there is a level of ambiguity aversion from which all types in  $\mathcal{C}$  must pool. This follows from the fact that for any  $N \geq 2$ ,  $\bar{\theta}_{N-1}^N(\alpha)$  is a strictly decreasing and continuous function and that  $\lim_{\alpha \rightarrow +\infty} \bar{\theta}_{N-1}^N(\alpha) < \frac{1}{2}$ . As a result, there must be  $\alpha \in (\frac{1}{2}, 1)$  such that no partitioning of  $\mathcal{C}$  is possible.

In addition, when there is an equilibrium with at least 3 cut-offs, ambiguity aversion eases the existence of a  $n$  cut-off equilibrium. Recall bounds of  $\mathcal{C}$  are included in the count, which means that we are looking at every equilibrium which is a non-babbling one. In other words, for a given bias, increasing ambiguity aversion might enable the existence of a  $k$ -cut-off equilibrium which was not sustainable for a lower level of ambiguity aversion. In that sense, ambiguity aversion eases information transmission, when science is consensual. This second result follows from the fact that for any  $N \geq 2$ ,  $\underline{\theta}_{N-1}^N(\alpha)$  is a strictly increasing function and that  $\lim_{\alpha \rightarrow +\infty} \underline{\theta}_{N-1}^N(\alpha) < \frac{1}{2}$ .

I know further prove that the first result of Proposition 9 extends to the case where  $c = 0$  and science is divided:

**Proposition 10.** *In the context of the linear-quadratic example when  $c = 0$ ,  $b > 0$  and  $\alpha \in (\frac{1}{2}, 1)$ , there is  $\alpha(b) \in (1/2, 1)$  such that for  $\alpha \geq \alpha(b)$ , only one action can be induced by types in  $[\frac{1}{2}, 1]$ . Moreover,  $\alpha(b)$  is a decreasing function.*

As for the consensual science case, for a given bias, there is a level of ambiguity aversion from which all types in  $[\frac{1}{2}, 1]$  must pool. This suggests that there is a form of continuity in the division of the set of types - on both sides of the hedging model - that we have observed in the MEU case. For any level of misalignment of  $S$ , there is degree of ambiguity aversion of  $R$  in  $(\frac{1}{2}, 1)$  such that all models above  $\tilde{\theta}$  must pool. The proof of Proposition 10 builds on the one of proposition Proposition 9. I show that for any  $N \geq 2$ ,  $\bar{\theta}_{N-1}^N(\alpha)$  is a strictly decreasing and continuous function and that  $\lim_{\alpha \rightarrow +\infty} \bar{\theta}_{N-1}^N(\alpha) < \frac{1}{2}$ . As a result, there must be  $\alpha \in (\frac{1}{2}, 1)$  such that no partitioning of  $[\frac{1}{2}, 1]$  is possible.

## 1.7. Conclusion

This paper models the transmission of expert-based scientific knowledge as cheap-talk communication over models, in a framework similar to Crawford and Sobel (1982). Because models can be represented as probability distributions, a receiver of this game can naturally be assumed to be ambiguity sensitive. For every preferences I considered, I showed that all equilibria are outcome equivalent to a partition equilibria. When the receiver is MEU, information transmission can only happen for models below a given threshold, even if misalignment is arbitrarily small. In addition, the sender always prefers to convey as much information as possible as the most informative equilibrium is interim Pareto dominant. This is not true when the receiver has SEU preferences, a case which is equivalent to the model of communication over states proposed in Crawford and Sobel (1982). In the linear-quadratic example with a uniform prior over models, more cut-offs can exist in the MEU case than in the SEU one, for a given bias. This special case

shows that when the expert's preferred action is aligned with the effect of ambiguity aversion, his influence is extremely high; but in the opposite case, it is nonexistent. Assuming that the receiver has  $\alpha$ -MEU preferences allows to show the robustness of these results : whatever the misalignment, there is a degree of ambiguity aversion for which no information transmission is possible for types above a given threshold.

Coming back to the COVID-19 pandemic as an iconic case of conflicting science, my results suggest that because of the epistemic nature of expert-based knowledge, when faced with a choice between the lesser of two evils, decision makers will tend to be unconvinced by scientific results which recommend more collective efforts. To the contrary, when the state which demands more collective efforts is also always worst than all others - as in the case of climate change - ambiguity aversion is a powerful ally for the transmission of expert-based scientific knowledge. For a scientific authority, the optimal communication policy is thus opposed in these two cases. In the latter one, ambiguity aversion serves information transmission and communication over models should be encouraged. In the former, the asymmetry resulting from conflict in science is an important bias in information transmission. Therefore, simple announcements over complex science that can be seen as cheap-talk for a non-expert audience should be avoided.

There are, of course, limitations to this work. First, it remains to show that results extend to a framework with multiple states. Second, it would be interesting to introduce the possibility of delegating the action to the sender, as introduced by [Dessein \(2002\)](#)—a possibility which is often welfare improving for the receiver. Finally, in my view, the main avenue for future research is the extension of the current framework to multiple experts. Many situations where expert-based scientific knowledge plays an important part involve multiple senders, communicating both simultaneously or sequentially. Existing work in the context of communication over states, such as [Battaglini \(2002\)](#) or [Krishna and Morgan \(2001\)](#) show that significant differences can appear.

## 1.8. Appendix

### 1.8.1. Supplementary Assumptions

#### Assumptions on states

In the following I show that Assumption 2 is implied by the two following assumptions.

**Assumption 3 (Misalignment - Crawford and Sobel (1982)).** *The optimal actions of S and R are always misaligned:*

$$a_S(\omega) > a_R(\omega) \text{ for all } \omega \in \Omega$$

Assumption 3 states that whatever the state, there is always a difference of interest between S and R such that optimal actions are ordered the same way.

**Assumption 4 (Sharpness).** *Whatever the state, the sender has sharper preferences than the receiver, for every action  $a \in \mathcal{A}$*

$$\forall a \in \mathcal{A}, \frac{\partial u_R(a, \omega)}{\partial a} < \frac{\partial u_S(a, \omega)}{\partial a}$$

Assumption 4 is a more technical assumption on the players utility function. I assume that the player with highest optimal action in a given state has a more concave utility function in that state, as illustrated by Figure 1.9. I call that property sharpness, in the sense that it translates a sharper preference for the optimal action.

Given Assumptions 3 and 4, I now show that both players optimal actions are never aligned, whatever the model.

**Lemma 2.** *Assumptions 3 and 4 imply that:*

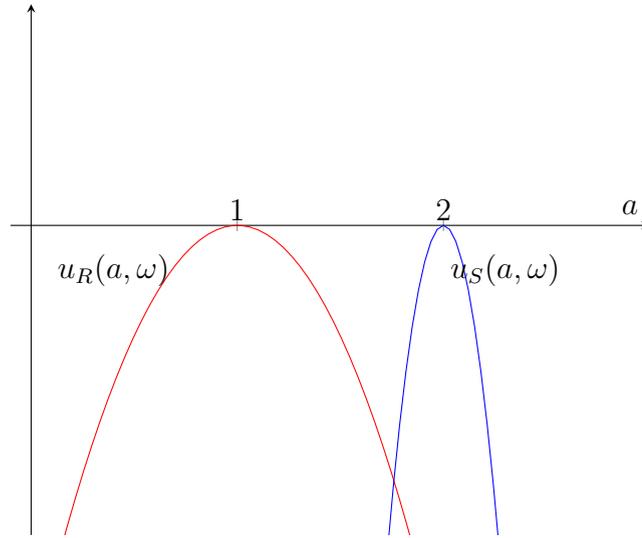


Figure 1.9: Sharpness Assumption

$$A_S(\theta) < A_R(\theta) \text{ for all } \theta \in \mathcal{C} \text{ or } A_S(\theta) > A_R(\theta) \text{ for all } \theta \in \mathcal{C}$$

**Proof of lemma 2:**

For player  $i$  and any  $\theta \in \mathcal{C}$ , define  $f_i^\theta : a \rightarrow (1 - \theta) \frac{\partial u_i(a, 0)}{\partial a} + \theta \frac{\partial u_i(a, 1)}{\partial a}$ .  $f_i^\theta$  is a continuous and decreasing function crossing the x-axis only once, at  $A_i(\theta)$ . We want to prove that for all  $\theta \in \mathcal{C}$ ,  $A_R(\theta) < A_S(\theta)$ . In order to do so, it is enough to prove that for any  $\theta \in \mathcal{C}$ ,  $f_R^\theta(a) < f_S^\theta(a)$ . Set  $h^\theta : a \rightarrow f_R^\theta(a) - f_S^\theta(a)$ .

$$h^\theta(a) = (1 - \theta) \left( \frac{\partial u_R(a, 0)}{\partial a} - \frac{\partial u_S(a, 0)}{\partial a} \right) + \theta \left( \frac{\partial u_R(a, 1)}{\partial a} - \frac{\partial u_S(a, 1)}{\partial a} \right)$$

Thus, by Assumption 4, for all  $a \in \mathcal{A}$ ,  $h^\theta(a) < 0$ .

□

Lemma 2 states that whatever the realised model,  $R$  and  $S$  optimal actions' are always ordered in the same direction. Notice that Assumption 3 isn't enough for this result. When Assumption 4 is violated, there can be  $\theta \in \mathcal{C}$  such that  $A_S(\theta) = A_R(\theta)$ .

## Imperfect knowledge of the model

In the following I show that the assumption that the sender observes the state generating distribution - the *true* model - can be replaced without significant change in the result. Instead, I will assume that  $S$  observes a noisy signal regarding the state generating distribution. I focus on the case where both players have MEU preferences. Yet, results regarding the linear-quadratic example differ. The noise decreases the precision of information transmission (cell sizes), acting as an additional bias.

Following [Gajdos et al. \(2008\)](#), I assume that  $S$  does not know the true model  $\theta_0$  but only observes an interval of models  $[\theta_0 - \epsilon, \theta_0 + \epsilon]$  of size  $2\epsilon > 0$ .

Assume that  $S$ 's preference under uncertainty are MEU. Then, having observed  $[\theta_0 - \epsilon, \theta_0 + \epsilon]$ ,  $S$  evaluates action  $a$  through:

$$V_S^{MEU}(a) = \min_{\theta \in [\theta_0 - \epsilon, \theta_0 + \epsilon]} \mathbb{E}_\theta(u_S(a, \omega))$$

Then, notice that the structure of equilibria is unaffected by those changes. Proposition 2 which guarantees that all equilibrium are outcome equivalent to a partition equilibrium only depends on the sender's type, and not the state generating distribution.

The fact that information transmission can only take place below  $\tilde{\theta}$  (Theorem 1) is also unaffected under my assumptions. Recall that there can not be a cut-off  $\theta_k$  above  $\tilde{\theta}$  because  $A_S(\theta_k) > A_R(s \in [\theta_k, \theta_{k+1}])$ . Yet, the optimal action when the

sender's signal is in  $[\theta_k, \theta_{k+1}]$  is  $A_R(\theta_k - \epsilon)$  and the optimal action of S when his type is  $\theta_k$  is  $A_S(\theta_k - \epsilon)$ . Because of the misalignment of playser (assumption 2), it can not be that  $A_R(\theta_k - \epsilon) > A_S(\theta_k - \epsilon)$ .

The evaluation of actions by R changes. Take  $B = [\theta_1, \theta_2] \subset \mathcal{C}$ , if R learns that  $s \in B$  it implies that  $\theta_0 \in [\theta_1 - \epsilon, \theta_2 + \epsilon]$ . As a result, given that the sender's type is in  $B$ , R evaluates action  $a$  through:

$$\begin{aligned} V_R^{MEU}(a, B) &= \min_{\theta \in [\theta_1 - \epsilon, \theta_2 + \epsilon]} \mathbb{E}_\theta(u_R(a, \omega)) \\ &= \mathbb{E}_{\theta_2 + \epsilon}(u_R(a, \omega)) \end{aligned}$$

Thus, R's evaluation of actions, for a given interval of parameters, still depends only on the upper bound of that interval. As a result, Theorem 2 still holds as well.

However, the characterisation in the linear quadratic will differ. The arbitrage condition of proposition 4 gives that :

$$\theta_{k+1} = \theta_k + 2b + \epsilon$$

Thus, it is as if the bias of the sender was  $b + \frac{\epsilon}{2}$ . The cells' length will change to a size of  $2b + \epsilon$ . This will have an effect on ex-ante evaluation of welfare, as the noise and the sender's ambiguity aversion decreases the precision of communication.

## 1.8.2. Proofs of the results in the main text

**Proof of lemma 1:**

$$\frac{\partial^2 \mathbb{E}_\theta(u_i(a, \omega))}{\partial \theta \partial a} = \frac{\partial u_i(a, 1)}{\partial a} - \frac{\partial u_i(a, 0)}{\partial a}$$

Assumption 1 gives that the latter is strictly positive.

□

### Proof of Proposition 1:

In order to prove our result we need to study the variations of  $\mathbb{E}_\theta(u_R(a, \omega))$  as a function of  $\theta$ . For  $a \in \mathcal{A}$ ,

$$\frac{\partial \mathbb{E}_\theta(u_R(a, \omega))}{\partial \theta} = u_R(a, 1) - u_R(a, 0)$$

Thus, we are interested in the sign of  $u_R(a, 1) - u_R(a, 0)$ . First, we need to prove the following lemma:

**Lemma 3.** *Define  $B \subset \mathcal{C}$  the belief of the receiver with minimal element  $\theta_1$  and maximal element  $\theta_2$ . Given this belief, her optimal action  $A_R(B) \subset [A_R(\theta_1), A_R(\theta_2)]$ .*

### Proof of lemma 3:

We prove this lemma in the more general context of  $\alpha$ -MEU preferences. This criteria coincides with MEU when  $\alpha = 1$ .

First, notice that  $\forall a \in \mathcal{A}$ , there is  $\theta_m(a) \in B$  such that  $\min_{\theta \in B} \mathbb{E}_\theta(u_R(a, \omega)) = \mathbb{E}_{\theta_m(a)}(u_R(a, \omega))$ . Similarly,  $\forall a \in \mathcal{A}$ , there is  $\theta_M(a) \in B$  such that  $\max_{\theta \in B} \mathbb{E}_\theta(u_R(a, \omega)) = \mathbb{E}_{\theta_M(a)}(u_R(a, \omega))$ .

As a result,  $\forall a \in \mathcal{A}$ ,  $\alpha \min_{\theta \in B} \mathbb{E}_\theta(u_R(a, \omega)) + (1 - \alpha) \max_{\theta \in B} \mathbb{E}_\theta(u_R(a, \omega)) =$

$\alpha \mathbb{E}_{\theta_m(a)}(u_R(a, \omega)) + (1 - \alpha) \mathbb{E}_{\theta_M(a)}(u_R(a, \omega)) = \mathbb{E}_{\alpha\theta_m(a) + (1-\alpha)\theta_M(a)}(u_R(a, \omega))$ . As, for all  $a \in \mathcal{A}$ ,  $\theta_1 \leq \alpha\theta_m(a) + (1 - \alpha)\theta_M(a) \leq \theta_2$  and that  $A_R(\theta)$  is a strictly increasing function, it must be that  $A_R(B) \subset [A_R(\theta_1), A_R(\theta_2)]$ .

□

A consequence of the Lemma 3 is that when looking for optimal actions for a given  $B$ , it is sufficient to look for actions in  $[A_R(\theta_1), A_R(\theta_2)]$ . Notice that  $[A_R(\theta_1), A_R(\theta_2)] \subset [a_R(0), a_R(1)]$  and that for all  $a \in [a_R(0), a_R(1)]$  either:

1.  $u_R(a_R(0), 0) < u_R(a_R(0), 1)$ .

For  $a > a_R(0)$ ,  $u_R(a, 0)$  is decreasing and  $u_R(a, 1)$  is increasing, utilities in both states are never equal and  $u_R(a, 0) < u_R(a, 1)$  for all  $a \in \mathcal{A}$ . As in this case  $\tilde{a} = a_R(0)$  and thus  $\tilde{\theta} = 0$ ,  $\mathbb{E}_\theta(u_R(a, \omega))$  is strictly increasing with  $\theta$  for all  $a \in [a_R(0), a_R(1)]$ . As a result,  $A_R(B) = A_R(\theta_1)$ .

2.  $u_R(a_R(0), 0) > u_R(a_R(0), 1)$  and  $u_R(a_R(1), 0) > u_R(a_R(1), 1)$ .

For  $a > a_R(0)$ ,  $u_R(a, 0)$  is decreasing and  $u_R(a, 1)$  is increasing, but as  $u_R(a_R(1), 0) > u_R(a_R(1), 1)$  it must be that utilities in both states are never equal. As a result,  $u_R(a, 0) > u_R(a, 1)$  for all  $a \in \mathcal{A}$ . Thus, in this case  $\tilde{a} = a_R(1)$  and  $\tilde{\theta} = 1$ . It follows that  $\mathbb{E}_\theta(u_R(a, \omega))$  is strictly decreasing with  $\theta$  for all  $a \in [a_R(0), a_R(1)]$ . As a result,  $A_R(B) = A_R(\theta_2)$ .

3.  $u_R(a_R(0), 0) > u_R(a_R(0), 1)$  and  $u_R(a_R(1), 0) \leq u_R(a_R(1), 1)$ .

As for  $a > a_R(0)$ ,  $u_R(a, 0)$  is strictly decreasing and  $u_R(a, 1)$  is strictly increasing. Thus, both utilities are equal for a unique given action and by definition of  $\tilde{a}$  it must be that this point is  $\tilde{a}$ . As a result:

$$\begin{cases} u_R(a, 0) > u_R(a, 1) \text{ for } a < \tilde{a} \\ u_R(a, 0) = u_R(a, 1) \text{ for } a = \tilde{a} \\ u_R(a, 0) < u_R(a, 1) \text{ for } a > \tilde{a} \end{cases}$$

Thus, for  $a \in [A_R(\theta_1), A_R(\theta_2)]$ ,  $\mathbb{E}_\theta(u_R(a, \omega))$  is strictly decreasing with  $\theta$  when  $\theta_2 < \tilde{\theta}$  and strictly increasing with  $\theta$  when  $\theta_1 > \tilde{\theta}$ , which gives the corresponding result. The above system also implies that when  $\tilde{\theta} \in B$ ,  $\mathbb{E}_\theta(u_R(a, \omega))$  is always minimal for  $\theta = \tilde{\theta}$ . As a result, for all  $a \in [A_R(\theta_1), A_R(\theta_2)]$  the minimal pay-off of the receiver as a function of the sender's type is given by:

$$\min_{\theta \in B} \mathbb{E}_\theta(u_R(a, \omega)) = \begin{cases} \mathbb{E}_{\theta_2}(u_R(a, \omega)) & \text{if } a < \tilde{a} \\ \mathbb{E}_{\tilde{\theta}}(u_R(a, \omega)) & \text{if } a = \tilde{a} \\ \mathbb{E}_{\theta_1}(u_R(a, \omega)) & \text{if } a > \tilde{a} \end{cases}$$

The above system implies that when  $\tilde{\theta} \in B$ ,  $\min_{\theta \in B} \mathbb{E}_\theta(u_R(a, \omega))$  is increasing on  $(A_R(\theta_1), \tilde{a})$  (as  $\mathbb{E}_{\theta_2}(u_R(a, \omega))$  is maximal at  $A_R(\theta_2) > \tilde{a}$ ) and decreasing on  $(\tilde{a}, A_R(\theta_2))$  (as  $\mathbb{E}_{\theta_1}(u_R(a, \omega))$  is maximal at  $A_R(\theta_1) < \tilde{a}$ ). As a result, it is always maximal for  $\tilde{a}$ . As a result,  $\min_{\theta \in B} A_R(B) = A_R(\tilde{\theta})$ .

□

### Proof of Proposition 2

The proof is structured as follows. First, I show that the number of outcome actions induced at equilibrium is finite. Then, I prove that the set of types which get the same equilibrium outcome must form an interval. The continuity and the strict monotonicity of the sender's preferences closes the argument.

**Lemma 4.** *There exists  $\epsilon > 0$  such that if  $u$  and  $v$  are actions induced in equilibrium,  $|u - v| \geq \epsilon$ . Further the set of actions induced in equilibrium is finite.*

### Proof of Lemma 4

I say that action  $u$  is induced by an S-type  $\theta$  if it is a best response to a given equilibrium message  $m : u \in \{A_R(\theta) | \theta \in \sigma^{-1}(m)\}$ . Let  $Y$  be the set of all actions

induced by some S-type  $\theta$ . First, notice that if  $\theta$  induces  $\bar{a}$ , it must be that  $V_S^\theta(\bar{a}) = \max_{a \in Y} V_S^\theta(a)$ . Since  $u_S$  is strictly concave,  $V_S^\theta(a)$  can take on a given value for at most two values of  $a$ . Thus,  $\theta$  can induce no more than two actions in equilibrium.

Let  $u$  and  $v$  be two actions induced in equilibrium,  $u < v$ . Define  $\Theta_u$  the set of S types who induce  $u$  and  $\Theta_v$  the set of S types who induce  $v$ . Take  $\theta \in \Theta_u$  and  $\theta' \in \Theta_v$ . By definition,  $\theta$  reveals a weak preference for  $u$  over  $v$  and  $\theta'$  reveals a weak preference for  $v$  over  $u$  that is:

$$\begin{cases} V_S^\theta(u) \geq V_S^\theta(v) \\ V_S^{\theta'}(v) \geq V_S^{\theta'}(u) \end{cases}$$

Thus, by continuity of  $\theta \rightarrow V_S^\theta(u) - V_S^\theta(v)$ , there is  $\hat{\theta} \in [\theta, \theta']$  such that  $V_S^{\hat{\theta}}(u) = V_S^{\hat{\theta}}(v)$ . Since  $u_S$  is strictly concave, we have that:

$$u < A_S(\hat{\theta}) < v$$

Then, notice that since  $\frac{\partial^2 \mathbb{E}_\theta(u_S(a, \omega))}{\partial a \partial \theta} > 0$  (Lemma 1), it must be that all types that induce  $u$  are below  $\hat{\theta}$ . Similarly, it must be that all types that induce  $v$  are above  $\hat{\theta}$ . That is:

$$\begin{aligned} \forall \theta \in \Theta_u, \theta &\leq \hat{\theta} \\ \forall \theta \in \Theta_v, \theta &\geq \hat{\theta} \end{aligned}$$

Thus, when  $R$  is MEU, Lemma 3 implies that the optimal action of the receiver, given that  $\theta \in \Theta_u$  is below the optimal action when the type is  $\hat{\theta}$ . Similarly, the

optimal action of the receiver, given that  $\theta \in \Theta_v$  is above the optimal action when the type is  $\hat{\theta}$ . The same is true when when  $R$  is SEU. That is:

$$\begin{cases} A_R(\Theta_u) \leq A_R(\hat{\theta}) \\ A_R(\Theta_v) \geq A_R(\hat{\theta}) \end{cases} \\ \iff u \leq A_R(\hat{\theta}) \leq v$$

However, as  $A_R(\theta) \neq A_S(\theta)$  for all  $\theta \in \mathcal{C}$ , there is  $\epsilon > 0$  such that  $|A_R(\theta) - A_S(\theta)| \geq \epsilon$  for all  $\theta \in \mathcal{C}$ . It follows that  $|u - v| \geq \epsilon$ .

Lemma 3 implies that for any belief  $B \subset \mathcal{C}$ , the optimal action of the receiver is in  $[A_R(\underline{\theta}), A_R(\bar{\theta})]$ . Thus, the set of actions induced in equilibrium is bounded by  $A_R(\underline{\theta})$  and  $A_R(\bar{\theta})$  and at least  $\epsilon$  away from one another, which completes the proof.

□

**Lemma 5.** *In every equilibrium of the game, if  $a$  is an action induced by type  $\theta$  and by type  $\theta''$  for some  $\theta < \theta''$ , then  $a$  is also induced by  $\theta' \in (\theta, \theta'')$*

**Proof of Lemma 5:**

For the purpose of the proof, we introduce the notation  $W^\theta(a) = \mathbb{E}_\theta(u_S(a, \omega))$ , which is the evaluation of  $a \in \mathcal{A}$  by a sender of type  $\theta$ .

We proceed by contradiction. Suppose  $a_1$  is induced by type  $\theta$  and by type  $\theta''$  and that there is  $\theta' \in (\theta, \theta'')$  such that  $a_1$  is not induced. Then there must be  $a_2 \neq a_1$  that type  $\theta'$  prefers and that  $\theta''$  does not. Formally, this is:

$$\begin{cases} W^\theta(a_2) \leq W^\theta(a_1) \\ W^{\theta'}(a_1) \leq W^{\theta'}(a_2) \\ W^{\theta''}(a_2) \leq W^{\theta''}(a_1) \end{cases} \quad (1.6)$$

Notice that for  $a \in \mathcal{A}$ :

$$\frac{\partial W^\theta(a)}{\partial \theta} = u_S(a, 1) - u_S(a, 0)$$

Similarly to  $S$ , define  $\tilde{a}_S = \operatorname{argmax}_{a \in \mathcal{A}} \min_{\omega \in \Omega} u_S(a, \omega)$ .  $\tilde{a}_S$  is the action that maximises the worst possible expected utility of the sender among the set of models. Two special cases are to be noticed. Either the high state is sufficiently worst than the good one for it to give a lower utility at its optimal point:  $u_S(a_S(1), 1) \leq u_S(a_S(1), 0)$ . Then the hedging action is the optimal action in the high state  $\tilde{a}_S = a_S(1)$ . Either the former is not true ( $u_S(a_S(1), 1) > u_S(a_S(1), 0)$ ) and both states must give the same utility for a given action in  $(a_S(0), a_S(1))$ . In that case  $\tilde{a}_S$  is the action that gives the same utility in both states.

As a result,  $W^\theta(a)$  is strictly decreasing for  $a < \tilde{a}_S$ , constant for  $a = \tilde{a}_S$  and strictly increasing for  $a > \tilde{a}_S$ . Assume that  $a_1 < a_2$ :

- When  $a_1 < \tilde{a}_S$  and  $a_2 \geq \tilde{a}_S$  can cross at most once and system (1.6) is impossible.
- Assume  $\tilde{a}_S \leq a_1 < a_2$ . Then:

$$\frac{\partial(W^\theta(a_1) - W^\theta(a_2))}{\partial \theta} = u_S(a_1, 1) - u_S(a_1, 0) - (u_S(a_2, 1) - u_S(a_2, 0))$$

As, for  $a \geq \tilde{a}_S$ ,  $u_S(a, 1)$  is a strictly increasing function and  $u_S(a, 0)$  a strictly decreasing one, we have that  $a_1 < a_2$  implies that  $u_S(a_1, 1) - u_S(a_1, 0) <$

$u_S(a_2, 1) - u_S(a_2, 0)$ . Thus,  $W^\theta(a_1) - W^\theta(a_2)$  is a strictly decreasing function of  $\theta$  and  $W^\theta(a_2)$  and  $W^\theta(a_1)$  can cross at most once, making system (1.6) impossible.

- Assume  $a_1 < a_2 < \tilde{a}_S$ . Then,  $W^\theta(a_1) - W^\theta(a_2)$  is a strictly increasing function of  $\theta$  and  $W^\theta(a_2)$  and  $W^\theta(a_1)$  can cross at most once, making system (1.6) impossible.

The case  $a_2 > a_1$  is symmetric.

□

By Lemma 4 there is a finite number of outcomes induced in equilibrium. The continuity of  $A_S(\theta)$  gives that there is a type of the sender which is indifferent between any pair of outcomes induced in equilibrium and the monotony of  $A_S(\theta)$  implies there are only a finite number of types which are indifferent between any pair of outcomes. Hence, Lemma 5 implies that there is a partitioning of  $\mathcal{C}$  in a finite number of cells where every cell induces a given action at equilibrium.

□

### Proof of Proposition 3

The outline of the proof is as follows. I start by showing that the cut-off types of any equilibrium must satisfy condition (1.1). Any other equilibrium strategies would be outcome equivalent.

Consider a couple of strategy  $(\sigma_q^*, y_q^*)$  and write  $C_k^q = [\theta_k^q, \theta_{k+1}^q]$ .

- Assume  $y_q^*$  is the equilibrium strategy of R. Given Proposition 2, any type  $\theta \in C_k^q$  induces the same action and prefers it to any other equilibrium action. Thus, for  $\sigma_q^*$  to be an equilibrium strategy, it is without loss of generality to assume that any type  $\theta \in C_k^q$  sends the same message  $m_k$  and prefer it

to any other message<sup>8</sup>. In particular, it must be preferred to message  $m_{k-1}$  which induces the preferred equilibrium action of types in  $C_{k-1}^q$ . For all  $\theta \in C_k^q$ :

$$V_S^\theta(y^*(m_k^q)) \geq V_S^\theta(y^*(m_{k-1}^q))$$

Similarly, any type  $\theta \in C_{k-1}^q$  must prefer sending  $m_{k-1}$  to  $m_k$ . For all  $\theta \in C_{k-1}^q$ :

$$V_S^\theta(y^*(m_k^q)) \leq V_S^\theta(y^*(m_{k-1}^q))$$

Thus, for  $\sigma_q^*$  to be an equilibrium strategy a necessary condition is that:

$$V_S^{\theta_k^q}(y^*(m_{k-1}^q)) = V_S^{\theta_k^q}(y^*(m_k^q))$$

- Assume  $\sigma_q^*$  is the equilibrium strategy of S. Then, for any  $\theta \in \mathcal{C}$ , the best response of R in the MEU case to any equilibrium message  $\sigma_q^*(\theta)$  is:

$$\operatorname{argmax}_{a \in A} V_R^{\text{MEU}}(a, \sigma_q^{*-1}(\sigma_q^*(\theta))) = y_q^*(\sigma_q^*(\theta))$$

Similarly, in the SEU case, the best response of R to any equilibrium message  $\sigma_q^*(\theta)$  is:

$$\operatorname{argmax}_{a \in A} V_R^{\text{SEU}}(a, \sigma_q^{*-1}(\sigma_q^*(\theta))) = y_q^*(\sigma_q^*(\theta))$$

□

---

<sup>8</sup>Any other signaling strategy must induce the same action from R and will thus lead to the same pay-offs for both players, whatever the sender's type.

**Proof of Theorem 1**

Assume there is a  $q + 1$  cut-off equilibrium and that  $\theta_{q-1} < \tilde{\theta} < \theta_q$ . As  $\theta_q > \tilde{\theta}$ , we have that :

$$\begin{cases} y^*(m_{q-1}) = A_R(\sigma^{*-1}(m_{q-1})) = A_R([\theta_{q-1}, \theta_q)) = A_R(\tilde{\theta}) = \tilde{a} \\ y^*(m_q) = A_R(\sigma^{*-1}(m_q)) = A_R([\theta_q, \theta_{q+1})) = A_R(\theta_q) \end{cases}$$

As  $A_R$  is a strictly increasing function and because  $S$  is upwards misaligned, we have that  $y^*(m_{q-1}) < y^*(m_q) < A_S(\theta_q)$ . As by definition,  $a \rightarrow \mathbb{E}_\theta(u_S(a, \omega))$  is strictly increasing on  $[0, A_S(\theta_q)]$ , we have that  $\mathbb{E}_{\theta_q}(u_S(y^*(m_{q-1}), \omega)) < \mathbb{E}_{\theta_q}(u_S(y^*(m_q), \omega)) \iff V_S^{\theta_q}(m_{q-1}) < V_S^{\theta_q}(m_q)$ , which is a contradiction with the assumption that  $\theta_q$  is a cut-off type.

□

**Proof of Proposition 4:**

The structure of the proof is as follows. First, I provide an algorithm that characterises the cut-off types of the equilibrium that has most cut-offs:  $\theta_0 < \dots < \theta_M$  (step 1). Define  $\mathcal{E} = \{(\theta_{1+k}, \dots, \theta_M) | 1 \leq k \leq M\}$ . Then, I show that any non-babbling partitioned strategy of the sender characterised by cut-offs which are elements of  $\mathcal{E}$  is an equilibrium strategy (step 2). I conclude by showing that this describes every equilibrium of the game (step 3).

In the following, I call  $C_q = [\theta_q, \theta_{q+1}]$ , for  $1 \leq q < M - 1$ ,  $C_M = [\theta_M, \bar{\theta}]$  and  $C_0 = [\underline{\theta}, \theta_1]$

**Step 1:**

Assume there is a  $M$  cut-off equilibrium. Then the signalling strategy of the

sender  $\sigma$  must be such that for  $q \in 0, \dots, M, \forall \theta \in C_q, \sigma(\theta) = m_k$

First notice that  $V_R^{MEU}(a, C_0) = \mathbb{E}_{\theta_1}(u_R(a, \omega))$ . For  $\sigma$  to be an equilibrium strategy we need that  $\forall \theta \in C_0$  and  $m \neq m_0$ :

$$V_S^\theta(m_0) \geq V_S^\theta(m)$$

In  $C_0$ , type  $\theta_1$  has the most incentive to deviate from sending  $m_0$  to sending  $m_1$ , which would induce a higher action, as,  $V_R^{MEU}(a, C_1) = \mathbb{E}_{\theta_2}(u_R(a, \omega))$  and  $A_R(\theta)$  is strictly increasing by Lemma 1.

Thus, a necessary condition for all types in  $C_0$  to send  $m_0$  is that:

$$V_S^{\theta_1}(m_0) \geq V_S^{\theta_1}(m_1)$$

Furthermore, it is also necessary that all types in  $C_1$  prefer message  $m_1$ . In particular it must be the case for type  $\theta_1$ , thus:  $V_S^{\theta_1}(m_1) \geq V_S^{\theta_1}(m_0)$ . As a consequence, a necessary condition for  $\sigma$  to be an equilibrium strategy is:

$$V_S^{\theta_1}(m_0) = V_S^{\theta_1}(m_1) \tag{1.7}$$

By repeating the argument for all  $C_q, q \in 1, \dots, M$ , a necessary condition for  $\sigma$  to be an equilibrium strategy is for all  $q \in 1, \dots, M$ :

$$V_S^{\theta_q}(m_{q-1}) = V_S^{\theta_q}(m_q) \quad (1.8)$$

Furthermore, the fact that  $\cup_{k=0}^M C_k = \mathcal{C}$  and the fact that for every pair of consequent cell of the partition the incentive constraints are transitive gives that conditions (1.8) is both necessary and sufficient. As  $A_R(\theta)$  is strictly monotone, it implies that  $A_R(\theta_k) \neq A_R(\theta_{k+1})$ .  $\underline{\theta}$  being known, it is possible to derive  $\theta_1$  directly from (1.7). By repeating the reasoning by induction,  $\theta_{k+1}$  can be derived from  $\theta_k$  for  $k \in 1, \dots, M - 1$  from (1.8) as long as there is  $\theta_M < \tilde{\theta}$ .

### Step 2:

I show that any partitional strategy of the sender characterised by  $\theta_0 < \dots < \theta_q$  is an equilibrium strategy. I proceed by iteration:

- Step 1 proves that  $\theta_0 < \dots < \theta_M$  characterise an equilibrium. Let's show that  $\theta_0, \theta_2 < \dots < \theta_M$  does as well.

Assume  $S$ 's strategy is  $\sigma_{M-1}$  such that for  $2 \leq k \leq M, \forall \theta \in C_k, \sigma_{M-1}(\theta) = m_k$  and  $\forall \theta \in [\underline{\theta}, \theta_2], \sigma_{M-1}(\theta) = m_0$ . Then for  $k \in 1, \dots, n - 1$ , when learning its type  $\theta \in C_k$ , by construction of the previous equilibrium,  $S$ 's preferred message is  $m_k$ . When  $\theta \in [\theta_{M-1}, \theta_M], m_{M-1}$  induces the same outcome as in the  $M$  cut-off equilibrium and is preferred to all other messages.

When  $\theta \in [\underline{\theta}, \theta_2]$  the fact that, for every pair of consequent cell of the partition, the incentive constraints are transitive implies that message  $m_0$  is preferred to any other message.

- Let's assume that for  $q \geq 2$ ,  $\sigma_q$  defined as above is an equilibrium strategy for  $S$ . By the same reasoning as above, it is straightforward to show that  $\sigma_{q-1}$  is one as well. This completes the proof of step 2.

### Step 3:

Assume there is an equilibrium strategy of the sender  $\sigma$  which is not described above. Recall  $A_R(B)$  to be the optimal action of R under the belief that  $\theta_0 \in B$  for  $B \subset \mathcal{C}$  and  $W^\theta(a) = \mathbb{E}_\theta(u_S(a, \omega))$  the evaluation of action  $a \in \mathcal{A}$  by a sender of type  $\theta$ .

- Proposition 2 gives that all equilibria are partitional. First I'll show that any equilibria only characterised by elements of  $\theta_0, \dots, \theta_q$  must be characterised by elements of  $\mathcal{E}$ . It is straightforward to see that any equilibria only characterised by elements of  $\theta_0, \dots, \theta_q$  which is not in  $\mathcal{E}$  can be constructed from an element of  $\mathcal{E}$  by removing at least one element which is not an extrema. To prove our claim, it is thus sufficient to prove that no equilibrium constructed from an element of  $\mathcal{E}$  by removing exactly one element which is not an extrema exists.

For  $1 \leq q \leq M$ , consider a strategy  $\sigma_p$  characterised by cut-offs  $\theta_0, \theta_q, \dots, \theta_{p-1}, \theta_{p+1}, \dots, \theta_M$  for  $q+1 \leq p \leq M$  and assume it is an equilibrium strategy<sup>9</sup>. It must be that that type  $\theta_{p+1}$  prefers outcome  $A_R([\theta_{p-1}, \theta_{p+1}])$  to outcome  $A_R([\theta_{p+1}, \theta_{p+2}])$ . Yet, by construction of the equilibrium of  $q$  cut-offs, types  $\theta_{p+1}$  is exactly indifferent between outcome  $A_R([\theta_p, \theta_{p+1}])$  and outcome  $A_R([\theta_{p+1}, \theta_{p+2}])$ . As  $A_R([\theta_{p-1}, \theta_{p+1}]) < A_R([\theta_p, \theta_{p+1}])$ , the previous implies that type  $\theta_{p+1}$  prefers outcome  $A_R([\theta_{p+1}, \theta_{p+2}])$  to outcome  $A_R([\theta_{p-1}, \theta_{p+1}])$ , which is a contradiction.

- Thus  $\sigma$  must have a cut-off type  $\theta^* \notin \{\theta_1, \dots, \theta_M\}$ . Assume without loss of generality that  $\theta_p < \theta^* < \theta_{p+1}$  for  $p \in 1, \dots, M-1$ . Then we have that:

$$\begin{aligned} W^{\theta^*}([\theta_p, \theta^*]) &= W^{\theta^*}([\theta^*, \theta_{p+1}]) \\ \iff \mathbb{E}_{\theta^*}(u_S(A_R(\theta^*))) &= \mathbb{E}_{\theta^*}(u_S(A_R(\theta_{p+1}))) \end{aligned}$$

Yet, by the construction in step 1, the above implies that  $\theta^* = \theta_{p+1}$ , which is a contradiction.

□

<sup>9</sup>the choice of removing an element  $\theta_p < \tilde{\theta}$  is without loss of generality

**Proof of Theorem 2:**

Assume the equilibrium with most cut-offs has  $M$  elements. For any  $1 \leq q \leq M$  let a  $q$  cut-off equilibrium be characterised by  $S$ 's strategy  $\sigma_q^*$  and elements  $\theta_0, \theta_{M-q}, \dots, \theta_M$ .

First I'll show that  $S$  is interim better-off in the  $q + 1$  cut-off equilibrium than in the  $q$  cut-off equilibrium. Then a simple iteration gives that  $S$  is better-off in the  $M$  cut-off equilibrium than in the  $q$  cut-off equilibrium, for any  $q < M$ .

- Assume  $\theta_0 \in [\theta_q, \bar{\theta}]$ .

Then,  $S$ 's interim utility in the  $q + 1$  cut-off equilibrium and in the  $q$  cut-off equilibrium is  $\mathbb{E}_{\theta_0}(u_S((A_R(\tilde{\theta})))$ . Thus  $S$  is indifferent between both equilibria.

- Assume  $\theta_0 \in [\theta_k, \theta_{k+1}]$ , for  $M - q \leq k \leq M$ .

Then,  $S$ 's interim utility in the  $q + 1$  cut-off equilibrium and in the  $q$  cut-off equilibrium is  $\mathbb{E}_{\theta_0}(u_S((A_R(\theta_{M-q-1})))$ . Thus  $S$  is indifferent between both equilibria.

- Assume  $\theta_0 \in [\underline{\theta}, \theta_{M-q-1}]$ , for  $M - q \leq k \leq M$ .

Then,  $S$ 's interim utility in the  $q + 1$  cut-off equilibrium is  $\mathbb{E}_{\theta_0}(u_S((A_R(\theta_{M-q-1})))$  and  $S$ 's interim utility in the  $q$  cut-off equilibrium is  $\mathbb{E}_{\theta_0}(u_S((A_R(\theta_{M-q})))$ .

Yet, because  $\theta_{M-q-1}$  is a cut-off type in the  $q + 1$  cut-off equilibrium, for any  $\theta \in [\underline{\theta}, \theta_{M-q-1})$ ,

$$\mathbb{E}_{\theta_0}(u_S((A_R(\theta_{M-q-1}))) > \mathbb{E}_{\theta_0}(u_S((A_R(\theta_{M-q})))$$

Thus, any type of sender in  $[\underline{\theta}, \theta_{M-q-1})$  is interim better-off in the  $q + 1$  cut-off equilibrium than in the  $q$  cut-off equilibrium.

□

**Proof of Proposition 5:**

1. Assume R has SEU preferences. Assume there are  $n$  equilibrium cut-offs in  $[0, 1]$ :  $\theta_0, \dots, \theta_n$  and thus  $\theta_0 = 0$   $\theta_n = 1$ . When receiving equilibrium message  $m_k$  sent by types  $\theta \in [\theta_k, \theta_{k+1})$  S evaluates action through:

$$\begin{aligned} V_R(a|m_k) &= \int_{\theta \in [\theta_k, \theta_{k+1}]} (1 - \theta)u_R(a, 0) + \theta u_R(a, 1) d\theta \\ &= (1 - \mathbb{E}(\theta|m_k))u_R(a, 0) + \mathbb{E}(\theta|m_k)u_R(a, 1) \end{aligned}$$

where  $\mathbb{E}(\theta|m_k) = \int_{\theta \in [\theta_k, \theta_{k+1}]} \theta d\theta = \frac{\theta_k + \theta_{k+1}}{2}$ . A first order condition on the above gives that when evaluating actions through  $V_R(a|m_k)$ , the optimal action is  $\mathbb{E}(\theta|m_k)$ . It follows that the equilibrium action of R is  $y^*(m_k) = \frac{\theta_k + \theta_{k+1}}{2}$ . The optimal action in the eyes of S is  $A_S(\theta_0) = \theta_0 + b$ . The arbitrage condition gives that a sender of type  $\theta_k$  must be indifferent between  $m_{k-1}$  and  $m_k$ . That is, for  $k \in 2, \dots, n$ :

$$A_S(\theta_{k+1}) = \frac{y^*(m_k) + y^*(m_{k+1})}{2}$$

Notice that this arbitrage condition translates in the similar condition as in CS's example:

$$\theta_{k+1} - \theta_k = \theta_k - \theta_{k-1} + 4b \quad (1.9)$$

Equation (1.9) further gives that:

$$\theta_k = k(\theta_1 - \theta_0) + \frac{k(k-1)}{2}4b$$

Specifically,  $1 = \mathbb{E}(\theta_n) = n(\theta_1) + \frac{n(n-1)}{2}4b$  which gives  $\theta_1 = \frac{1}{n} - 2(n-1)b$  and:

$$\mathbb{E}(\theta_k) = \theta_k = \frac{k}{n} - 2kb(n-k)$$

It follows that a  $n$  cut-off equilibrium exists if and only if:

$$0 < b < \frac{1}{2n(n-1)}$$

2. Assume  $R$  has MEU preferences and that there is a  $n$ -cut-off equilibrium. When receiving message  $m_k^n$ , for  $k \geq 2$ :

$$V_R(a|m_k) = \min_{\theta \in [\theta_k, \theta_{k+1}]} \mathbb{E}_\theta(u_R(a))$$

Thus, when  $\theta_1 \leq \tilde{\theta}$ ,  $V_R(a|m_0) = \mathbb{E}_{\theta_1}(u_R(a))$  and the arbitrage condition giving the cut-off types gives that  $A_S(\theta_1) = \theta_1 + b$  must thus be at equal distance from  $\theta_1$  and  $\theta_2$ . For this to be possible, it must be that  $b > 0$ . Thus, when there is a  $n$ -cut-off equilibrium, it must be that  $\tilde{\theta} > \theta_n$ . When receiving message  $m_k$ , for  $k \geq 1$ :

$$V_R(a|m_k) = \mathbb{E}_{\theta_{k+1}}(u_R(a))$$

The equilibrium action of  $R$  when receiving the equilibrium message  $[\theta_k, \theta_{k+1}]$  is  $y(m_k^n) = \mathbb{E}(\theta_{k+1})$ . The arbitrage condition giving the cut-off types gives that  $A_S(\theta_{k+1})$  must thus be at equal distance from  $\mathbb{E}(\theta_{k+1})$  and  $\mathbb{E}(\theta_{k+2})$ , giving

$$\begin{aligned} \theta_{k+1} + b &= \frac{\theta_{k+1} + \theta_{k+2}}{2} \\ \iff \theta_{k+2} &= \theta_{k+1} + 2b \end{aligned}$$

When receiving message  $m_n$ , the equilibrium action of  $R$  is  $y(m_n) = \tilde{\theta} = \frac{1}{2}$ . The arbitrage condition when  $S$  is of type  $\theta_{n-1}$  gives that:

$$\begin{aligned}\frac{\tilde{\theta} + \theta_{n-1}}{2} &= \theta_{n-1} + b \\ \Leftrightarrow \theta_{n-1} &= 1 - 2b\end{aligned}$$

Which implies that, for all  $1 \leq k \leq n - 1$ :

$$\theta_k = \theta_k = 1 - 2b(n - k)$$

It follows that a  $n$  cut-off equilibrium exists if and only if:

$$\begin{aligned}\theta_1 &> 0 \\ \Leftrightarrow 1 - 2bn &> 0 \\ \Leftrightarrow 0 < b < \frac{1}{2n}\end{aligned}$$

□

### Proof of Corollary 1:

It is possible to derive from Proposition 5 that in the SEU case:

$$\theta_{k+1} - \theta_k = \theta_k - \theta_{k-1} + 4b$$

It is also possible to derive from Proposition 5 that in the MEU case:

$$\theta_{k+1} - \theta_k = 2b$$

□

**Proof of Proposition 6 :**

In order to prove our result we need to study the variations of  $\mathbb{E}_\theta(u_R(a, \omega))$  as a function of  $\theta$ . For  $a \in \mathcal{A}$ ,

$$\frac{\partial \mathbb{E}_\theta(u_R(a, \omega))}{\partial \theta} = u_R(a, 1) - u_R(a, 0)$$

Thus, we are interested in the sign of  $u_R(a, 1) - u_R(a, 0)$ . We want to prove that :

$$\min_{\theta \in B} A_R(\theta) = \begin{cases} A_R(\theta_2) & \text{if } \theta_2 < \tilde{\theta} \\ A_R(\theta_1) & \text{if } \theta_1 > \tilde{\theta} \\ A_R(\tilde{\theta}) & \text{if } \tilde{\theta} \in B \end{cases}$$

The maximal pay-off of the receiver as a function of the sender's type is given by :

$$\max_{\theta \in B} A_R(\theta) = \begin{cases} A_R(\theta_1) & \text{if } \theta_2 < \tilde{\theta} \\ A_R(\theta_2) & \text{if } \theta_1 > \tilde{\theta} \\ A_R(\theta_M) & \text{if } \tilde{\theta} \in B \end{cases}$$

where  $\theta_M = \operatorname{argmax}_{\theta \in \{\theta_1, \theta_2\}} \mathbb{E}_\theta(u_R(a, \omega))$ . A consequence of Lemma 3 is that when looking for optimal actions for a given  $B$ , it is sufficient to look for actions in  $[A_R(\theta_1), A_R(\theta_2)]$ . Notice that  $[A_R(\theta_1), A_R(\theta_2)] \subset [a_R(0), a_R(1)]$  and that for all

$a \in [a_R(0), a_R(1)]$  either :

1.  $u_R(a_R(0), 0) < u_R(a_R(0), 1)$ .

For  $a > a_R(0)$ ,  $u_R(a, 0)$  is decreasing and  $u_R(a, 1)$  is increasing, utilities in both states are never equal and  $u_R(a, 0) < u_R(a, 1)$  for all  $a \in \mathcal{A}$ . As in this case  $\tilde{a} = a_R(0)$  and thus  $\tilde{\theta} = 0$ ,  $\mathbb{E}_\theta(u_R(a, \omega))$  is strictly increasing with  $\theta$  for all  $a \in [a_R(0), a_R(1)]$ . As a result,  $\min_{\theta \in B} A_R(B) = A_R(\theta_1)$  and  $\max_{\theta \in B} A_R(B) = A_R(\theta_2)$

2.  $u_R(a_R(0), 0) > u_R(a_R(0), 1)$  and  $u_R(a_R(1), 0) > u_R(a_R(1), 1)$ .

For  $a > a_R(0)$ ,  $u_R(a, 0)$  is decreasing and  $u_R(a, 1)$  is increasing, but as  $u_R(a_R(1), 0) > u_R(a_R(1), 1)$  it must be that utilities in both states are never equal. As a result,  $u_R(a, 0) > u_R(a, 1)$  for all  $a \in \mathcal{A}$ . Thus, in this case  $\tilde{a} = a_R(1)$  and  $\tilde{\theta} = 1$ . It follows that  $\mathbb{E}_\theta(u_R(a, \omega))$  is strictly decreasing with  $\theta$  for all  $a \in [a_R(0), a_R(1)]$ . As a result,  $\min_{\theta \in B} A_R(B) = A_R(\theta_2)$  and  $\max_{\theta \in B} A_R(B) = A_R(\theta_1)$

3.  $u_R(a_R(0), 0) > u_R(a_R(0), 1)$  and  $u_R(a_R(1), 0) \leq u_R(a_R(1), 1)$ .

As for  $a > a_R(0)$ ,  $u_R(a, 0)$  is strictly decreasing and  $u_R(a, 1)$  is strictly increasing. Thus, both utilities are equal for a unique given action and by definition of  $\tilde{a}$  it must be that this point is  $\tilde{a}$ . As a result:

$$\begin{cases} u_R(a, 0) > u_R(a, 1) \text{ for } a < \tilde{a} \\ u_R(a, 0) = u_R(a, 1) \text{ for } a = \tilde{a} \\ u_R(a, 0) < u_R(a, 1) \text{ for } a > \tilde{a} \end{cases} \quad (1.10)$$

It follows from system (1.10) that, for all  $a \in [A_R(\theta_1), A_R(\theta_2)]$  the minimal pay-off of the receiver as a function of the sender's type is given by:

$$\min_{\theta \in B} \mathbb{E}_{\theta}(u_R(a, \omega)) = \begin{cases} \mathbb{E}_{\theta_2}(u_R(a, \omega)) & \text{if } a < \tilde{a} \\ \mathbb{E}_{\tilde{\theta}}(u_R(a, \omega)) & \text{if } a = \tilde{a} \\ \mathbb{E}_{\theta_1}(u_R(a, \omega)) & \text{if } a > \tilde{a} \end{cases}$$

The above system implies that when  $\tilde{\theta} \in B$ ,  $\min_{\theta \in B} \mathbb{E}_{\theta}(u_R(a, \omega))$  is increasing on  $(A_R(\theta_1), \tilde{a})$  (as  $\mathbb{E}_{\theta_2}(u_R(a, \omega))$  is maximal at  $A_R(\theta_2) > \tilde{a}$ ) and decreasing on  $(\tilde{a}, A_R(\theta_2))$  (as  $\mathbb{E}_{\theta_1}(u_R(a, \omega))$  is maximal at  $A_R(\theta_1) < \tilde{a}$ ). As a result, it is always maximal for  $\tilde{a}$ . As a result,  $\min_{\theta \in B} A_R(B) = A_R(\tilde{\theta})$ .

It also follows from system (1.10) that, for all  $a \in [A_R(\theta_1), A_R(\theta_2)]$  the maximal pay-off of the receiver as a function of the sender's type is given by:

$$\max_{\theta \in B} \mathbb{E}_{\theta}(u_R(a, \omega)) = \begin{cases} \mathbb{E}_{\theta_1}(u_R(a, \omega)) & \text{if } a < \tilde{a} \\ \mathbb{E}_{\theta_M}(u_R(a, \omega)) & \text{if } a = \tilde{a} \\ \mathbb{E}_{\theta_2}(u_R(a, \omega)) & \text{if } a > \tilde{a} \end{cases}$$

where  $\theta_M = \operatorname{argmax}_{\theta \in \{\theta_1, \theta_2\}} \mathbb{E}_{\theta}(u_R(a, \omega))$ . The above system implies that when  $\tilde{\theta} \in B$ ,  $\max_{\theta \in B} \mathbb{E}_{\theta}(u_R(a, \omega))$  is decreasing on  $(A_R(\theta_1), \tilde{a})$  (as  $\mathbb{E}_{\theta_1}(u_R(a, \omega))$  is maximal at  $A_R(\theta_1)$ ) and increasing on  $(\tilde{a}, A_R(\theta_2))$  (as  $\mathbb{E}_{\theta_2}(u_R(a, \omega))$  is maximal at  $A_R(\theta_2)$ ). As a result, it is maximal at either  $A_R(\theta_1)$  or  $A_R(\theta_2)$ . As a result,  $\max_{\theta \in B} A_R(B) = A_R(\theta_M)$ .

Notice that when utilities are quadratic, a simple algebra gives that for  $\theta < \theta'$  :

$$\begin{aligned} \operatorname{argmax}_{a \in \mathcal{A}} [\alpha \mathbb{E}_{\theta}(u_i(a, \omega)) + (1 - \alpha) \mathbb{E}_{\theta'}(u_i(a, \omega))] &= \alpha (\operatorname{argmax}_{a \in \mathcal{A}} \mathbb{E}_{\theta}(u_i(a, \omega))) + (1 - \alpha) (\operatorname{argmax}_{a \in \mathcal{A}} \mathbb{E}_{\theta'}(u_i(a, \omega))) \\ &= A_i(\alpha \theta + (1 - \alpha) \theta') \end{aligned}$$

which implies that :

$$\begin{aligned}
A_R(B) &= \operatorname{argmax}_{a \in \mathcal{A}} \left[ \alpha \min_{\theta \in B} \mathbb{E}_\theta(u_R(a, \omega)) + (1 - \alpha) \max_{\theta \in B} \mathbb{E}_\theta(u_R(a, \omega)) \right] \\
&= \begin{cases} \alpha A_R(\theta_2) + (1 - \alpha) A_R(\theta_1) & \text{if } \theta_2 < \tilde{\theta} \\ \alpha A_R(\tilde{\theta}) + (1 - \alpha) A_R(\theta_M) & \text{if } \tilde{\theta} \in B \\ \alpha A_R(\theta_1) + (1 - \alpha) A_R(\theta_2) & \text{if } \theta_1 > \tilde{\theta} \end{cases}
\end{aligned}$$

□

### Proof of Proposition 7

**Lemma 6.** *There exists  $\epsilon > 0$  such that if  $u$  and  $v$  are actions induced in equilibrium,  $|u - v| \geq \epsilon$ . Further the set of actions induced in equilibrium is finite.*

### Proof of Lemma 6

I say that action  $u$  is induced by an S-type  $\theta$  if it is a best response to a given equilibrium message  $m : u \in \{A_R(\theta) | \theta \in \sigma^{-1}(m)\}$ . Let  $Y$  be the set of all actions induced by some S-type  $\theta$ . First, notice that if  $\theta$  induces  $\bar{a}$ , it must be that  $V_S^\theta(\bar{a}) = \max_{a \in Y} V_S^\theta(a)$ . Since  $u_S$  is strictly concave,  $V_S^\theta(a)$  can take on a given value for at most two values of  $a$ . Thus,  $\theta$  can induce no more than two actions in equilibrium.

Let  $u$  and  $v$  be two actions induced in equilibrium,  $u < v$ . Define  $\Theta_u$  the set of S types who induce  $u$  and  $\Theta_v$  the set of S types who induce  $v$ . Take  $\theta \in \Theta_u$  and  $\theta' \in \Theta_v$ . By definition,  $\theta$  reveals a weak preference for  $u$  over  $v$  and  $\theta'$  reveals a weak preference for  $v$  over  $u$  that is :

$$\begin{cases} V_S^\theta(u) \geq V_S^\theta(v) \\ V_S^{\theta'}(v) \geq V_S^{\theta'}(u) \end{cases}$$

Thus, by continuity of  $\theta \rightarrow V_S^\theta(u) - V_S^\theta(v)$ , there is  $\hat{\theta} \in [\theta, \theta']$  such that  $V_S^{\hat{\theta}}(u) = V_S^{\hat{\theta}}(v)$ . Since  $u_S$  is strictly concave, we have that :

$$u < A_S(\hat{\theta}) < v$$

Then, notice that since  $\frac{\partial^2 \mathbb{E}_\theta(u_S(a, \omega))}{\partial a \partial \theta} > 0$  (Lemma 1), it must be that all types that induce  $u$  are below  $\hat{\theta}$ . Similarly, it must be that all types that induce  $v$  are above  $\hat{\theta}$ . That is :

$$\begin{aligned} \forall \theta \in \Theta_u, \theta &\leq \hat{\theta} \\ \forall \theta \in \Theta_v, \theta &\geq \hat{\theta} \end{aligned}$$

Thus, when  $R$  is  $\alpha$ -MEU, Lemma 3 implies that the optimal action of the receiver, given that  $\theta \in \Theta_u$  is below the optimal action when the type is  $\hat{\theta}$ . Similarly, the optimal action of the receiver, given that  $\theta \in \Theta_v$  is above the optimal action when the type is  $\hat{\theta}$ . That is :

$$\begin{cases} A_R(\Theta_u) \leq A_R(\hat{\theta}) \\ A_R(\Theta_v) \geq A_R(\hat{\theta}) \end{cases} \\ \iff u \leq A_R(\hat{\theta}) \leq v$$

However, as  $A_R(\theta) \neq A_S(\theta)$  for all  $\theta \in \mathcal{C}$ , there is  $\epsilon > 0$  such that  $|A_R(\theta) - A_S(\theta)| \geq \epsilon$  for all  $\theta \in \mathcal{C}$ . It follows that  $|u - v| \geq \epsilon$ .

Lemma 3 implies that for any belief  $B \subset \mathcal{C}$ , the optimal action of the receiver

is in  $[A_R(\underline{\theta}), A_R(\bar{\theta})]$ . Thus, the set of actions induced in equilibrium is bounded by  $A_R(\underline{\theta})$  and  $A_R(\bar{\theta})$  and at least  $\epsilon$  away from one another, which completes the proof.

□

By Lemma 6 there is a finite number of outcomes induced in equilibrium. The continuity of  $A_S(\theta)$  gives that there is a type of the sender which is indifferent between any pair of outcomes induced in equilibrium and the monotony of  $A_S(\theta)$  implies there are only a finite number of types which are indifferent between any pair of outcomes. Hence, Lemma 5 implies that there is a partitioning of  $\mathcal{C}$  in a finite number of cells where every cell induces a given action at equilibrium.

□

### Proof of Proposition 8:

I focus on the case  $c = 1$ . The case  $c = -1$  is symmetric.

Assume there is a  $n > 0$  cut-off equilibrium. It follows from the characterisation of cut-off types in the linear-quadratic example given in the proof of Proposition 5 and the characterisation of optimal actions in the  $\alpha$ -MEU case given in Proposition 6 that, for  $1 \leq k \leq n - 2$  and  $\alpha > 0$ :

$$\begin{aligned} \underline{\theta}_k^n + b &= \frac{\alpha \underline{\theta}_k^n + (1 - \alpha) \underline{\theta}_{k-1}^n + \alpha \underline{\theta}_{k+1}^n + (1 - \alpha) \underline{\theta}_k^n}{2} \\ \iff \theta_{k+1} - \theta_k &= \frac{1 - \alpha}{\alpha} (\underline{\theta}_k^n - \underline{\theta}_{k-1}^n + \frac{2b}{\alpha}) \end{aligned}$$

Set  $V_k = \underline{\theta}_{k+1}^n - \underline{\theta}_k^n$ . It follows from the previous equality that  $(V_k)_k$  is an arithmetico-geometrical sequence. As a result, for  $1 \leq k \leq n - 2$  and  $\alpha \notin \{0, \frac{1}{2}\}$ :

$$V_k = \left(\frac{1-\alpha}{\alpha}\right)^k \left(\underline{\theta}_1^n - \frac{2b}{2\alpha-1}\right) + \frac{2b}{2\alpha-1}$$

By induction, it follows that :

$$\begin{aligned} \underline{\theta}_{k+1}^n &= \sum_{j=1}^k \left[ \left(\frac{1-\alpha}{\alpha}\right)^j \left(\underline{\theta}_1^n - \frac{2b}{2\alpha-1}\right) + \frac{2b}{2\alpha-1} \right] + \underline{\theta}_1 \\ \iff \underline{\theta}_{k+1}^n &= \sum_{j=0}^k V_j \\ \iff \underline{\theta}_k^n &= \left(\underline{\theta}_1 - \frac{2bn}{2\alpha-1}\right) \left(\frac{1 - \left(\frac{1-\alpha}{\alpha}\right)^k}{1 - \left(\frac{1-\alpha}{\alpha}\right)}\right) + \frac{2bk}{2\alpha-1} \end{aligned}$$

In particular, it must be that  $\underline{\theta}_n^n = \frac{1}{2}$  which give that  $\underline{\theta}_1^n = \left(\frac{1}{2} - \frac{2bn}{2\alpha-1}\right) \left(\frac{1 - \left(\frac{1-\alpha}{\alpha}\right)}{1 - \left(\frac{1-\alpha}{\alpha}\right)^n}\right) + \frac{2b}{2\alpha-1}$ . As a result, we get that :

$$\underline{\theta}_k^n = \left(\frac{1}{2} - \frac{2bn}{2\alpha-1}\right) \left(\frac{1 - \left(\frac{1-\alpha}{\alpha}\right)^k}{1 - \left(\frac{1-\alpha}{\alpha}\right)^n}\right) + \frac{2bk}{2\alpha-1}$$

□

### Proof of Proposition 9 :

1. I start by proving that for  $n \geq 2$ ,  $\underline{\theta}_{n-1}^n(\alpha)$  is a strictly increasing function. Define  $f(a) = \frac{1-a^{n-1}}{1-a^n}$ , for  $a \in (0, 1/2)$ . Notice that :

$$\frac{\partial f(a)}{\partial a} = \frac{a^{n-2}(a^n - na + n - 1)}{(1 - a^n)^2}$$

Thus :

$$\begin{aligned} & \frac{\partial f(a)}{\partial a} < 0 \\ \iff & = \frac{a^{n-2}(a^n - na + n - 1)}{(1 - a^n)^2} < 0 \\ \iff & = a^n > n(a - 1) + 1 \end{aligned}$$

Yet,  $a \in (0, \frac{1}{2}) \Rightarrow a^n > 0$  and  $n(a - 1) + 1 < 0 \iff a < 1 - \frac{1}{n}$  which is true because  $a \in (0, 1/2)$  and  $n \geq 2$ . As a result,  $\frac{\partial f(a)}{\partial a} < 0$  and  $f$  is a decreasing function. Yet :

$$\underline{\theta}_{n-1}^n(\alpha) = \frac{1}{2}f\left(\frac{1-\alpha}{\alpha}\right) + \frac{2bn}{2\alpha-1}\left(1 - f\left(\frac{1-\alpha}{\alpha}\right)\right) - \frac{2b}{2\alpha-1}$$

$\frac{1-\alpha}{\alpha} \in (0, 1/2)$  for  $\alpha \in (\frac{1}{2}, 1)$  and is decreasing in  $\alpha$ . As a result  $f(\frac{1-\alpha}{\alpha})$  is increasing in  $\alpha$  and  $\underline{\theta}_{n-1}^n(\alpha)$  as well as a sum and product of increasing functions of  $\alpha$ . In addition, we have that :

$$\begin{aligned} & \frac{\partial \underline{\theta}_{n-1}^n(\alpha)}{\partial b} < 0 \\ \iff & = \frac{2n}{(2\alpha-1)}\left(1 - f\left(\frac{1-\alpha}{\alpha}\right)\right) - \frac{2}{(2\alpha-1)} < 0 \\ \iff & = f\left(\frac{1-\alpha}{\alpha}\right) > 0 \end{aligned}$$

which is true. By a symmetrical process, one can prove that  $\bar{\theta}_{n-1}^n(\alpha)$  is a strictly decreasing function and that  $\frac{\partial \bar{\theta}_{n-1}^n(\alpha)}{\partial b} < 0$ . Yet :

$$\lim_{\alpha \rightarrow 1} \bar{\theta}_{n-1}^n(\alpha) = -2b < \frac{1}{2}$$

Thus, as  $\bar{\theta}_{n-1}^n(\alpha)$  is strictly decreasing and continuous, there is  $\alpha(b) \in (1/2, 1)$  such that  $\bar{\theta}_{n-1}^n(\alpha) = \frac{1}{2}$ . As  $\bar{\theta}_{n-1}^n(\alpha)$  is strictly decreasing, for  $\alpha \geq \alpha(b)$ , no

information transmission is possible in  $\bar{C}$ . In addition, because  $\frac{\partial \bar{\theta}_{n-1}^n(\alpha)}{\partial b} < 0$ , it follows that  $\alpha(b)$  is a decreasing function.

2. I start by proving that for  $n \geq 2$ ,  $\underline{\theta}_1^n(\alpha)$  is a strictly increasing function. Define  $f(a) = \frac{1-a}{1-a^n}$ , for  $a \in (0, 1/2)$ . Notice that :

$$\frac{\partial f(a)}{\partial a} = \frac{n(1-a)a^{n-1}}{(1-a^n)^2} - \frac{1}{1-a^n}$$

Thus :

$$\begin{aligned} \frac{\partial f(a)}{\partial a} &< 0 \\ \iff &= n - (n-1)a < \frac{1}{a^{n-1}} \end{aligned}$$

Yet,  $a \in (0, 1/2) \Rightarrow \frac{1}{a^{n-1}} > 2^{n-1}$  and  $a \in (0, 1/2) \Rightarrow n - (n-1)a < n$ . As a result, for  $n \geq 2$ ,  $n - (n-1)a < n \leq 2^{n-1} < \frac{1}{a^{n-1}}$  which implies that  $\frac{\partial f(a)}{\partial a} < 0$  and  $f$  is a decreasing function. Yet :

$$\underline{\theta}_{n-1}^n(\alpha) = \frac{1}{2}f\left(\frac{1-\alpha}{\alpha}\right) + \frac{2bn}{2\alpha-1}\left(1 - f\left(\frac{1-\alpha}{\alpha}\right)\right) - \frac{2b}{2\alpha-1}$$

$\frac{1-\alpha}{\alpha} \in (0, 1/2)$  for  $\alpha \in (\frac{1}{2}, 1)$  and is decreasing in  $\alpha$ . As a result  $f(\frac{1-\alpha}{\alpha})$  is increasing in  $\alpha$  and  $\underline{\theta}_1^n(\alpha)$  as well as a sum and product of increasing functions of  $\alpha$ . By a symmetric process, one can prove that  $\bar{\theta}_1^n(\alpha)$  is a decreasing function.

Consider two receivers  $\alpha_1$  and  $\alpha_2$  such that  $\alpha_1 < \alpha_2$ . Assume there is a  $n$  cut-off equilibrium between S and  $\alpha_1$ . Then  $\bar{\theta}_1^n(\alpha_1) \in (0, 1)$ . As  $\bar{\theta}_1^n(\alpha)$  is a decreasing function, it must be that  $\bar{\theta}_1^n(\alpha_2) < 1$ . In addition, as  $\bar{\theta}_{n-1}^n(\alpha)$  is an decreasing function, it follows that  $\bar{\theta}_{n-1}^n(\alpha_2) > \lim_{\alpha \rightarrow 1} \bar{\theta}_{n-1}^n(\alpha) = \frac{1}{2} - 2b > 0$  for  $b < \frac{1}{4}$ , which is the existence condition of the considered equilibrium. As a result, there is a  $n$  cut-off equilibrium between S and  $\alpha_2$

□

**Proof of Proposition 10 :**

We are in the case where  $c = 0$ . Recall that in this case  $\tilde{\theta} = \frac{1}{2}$ .

- Consider the case of an  $n$  cut-off equilibrium that has  $n \leq N - 3$  cut-off types below  $\frac{1}{2}$ .

I call  $\theta_k^N(\alpha)$  the  $k$ -th cut-off of our equilibrium. By Proposition 6, we get that for  $n + 2 \leq k \leq N - 1$  and  $\frac{1}{2} < \alpha < 1$ :

$$\theta_k^N(\alpha) + b = \frac{(1 - \alpha)\theta_k^N(\alpha) + \alpha\theta_{k-1}^N(\alpha) + (1 - \alpha)\theta_{k+1}^N(\alpha) + \alpha\theta_k^N(\alpha)}{2}$$

as before, we get by induction that:

$$\theta_k^N(\alpha) = \left(\theta_{n+1}^N(\alpha) - \frac{2bN}{2\alpha - 1}\right) \left(\frac{1 - \left(\frac{\alpha}{1-\alpha}\right)^k}{1 - \left(\frac{\alpha}{1-\alpha}\right)}\right) - \frac{2bk}{2\alpha - 1} - \frac{1}{2}$$

In particular, it must be that  $\theta_N^N = 1$  which give that  $\theta_{n+1}^N = \left(1 + \frac{2bN}{2\alpha - 1}\right) \left(\frac{1 - \left(\frac{\alpha}{1-\alpha}\right)^N}{1 - \left(\frac{\alpha}{1-\alpha}\right)}\right) + \frac{2b}{2\alpha - 1}$ . As a result, we get that :

$$\theta_{N-1}^N(\alpha) = \left(1 - \frac{2bN}{2\alpha - 1}\right) \left(\frac{1 - \left(\frac{\alpha}{1-\alpha}\right)^{N-1}}{1 - \left(\frac{\alpha}{1-\alpha}\right)}\right) - \frac{2b(N-1)}{2\alpha - 1} + \frac{1}{2}$$

Yet, reproducing the reasoning in the proof of Proposition 9, one can show that  $\theta_{N-1}^N(\alpha)$  is a strictly decreasing function and that :

$$\lim_{\alpha \rightarrow 1} \theta_{N-1}^N(\alpha) = \frac{1}{2} - 2b < \frac{1}{2}$$

Thus, as  $\theta_{N-1}^N(\alpha)$  is strictly decreasing and continuous, there is  $\alpha(b) \in (1/2, 1)$

such that  $\theta_{N-1}^N(\alpha) = \frac{1}{2}$ . It follows that there is  $\alpha(b) \in (0, 1)$  such that, for  $\alpha \geq \alpha(b)$ , only one action can be induced by types in  $[\frac{1}{2}, 1]$ .

- Second, consider the case where  $n = N - 2$ . Then, there is a single cut-off type in  $(\frac{1}{2}, 1)$ . Call that type  $\theta_{N-1}^N(\alpha) \in (\frac{1}{2}, 1)$ . By Proposition 6 it must be that:

$$\Leftrightarrow \begin{cases} \theta_{N-1}^N(\alpha) + b = \frac{\alpha \frac{1}{2} + (1-\alpha)\theta_n^N + \alpha\theta_{N-1}^N(\alpha) + (1-\alpha)1}{2} \\ \text{or} \\ \theta_{N-1}^N(\alpha) + b = \frac{\alpha \frac{1}{2} + (1-\alpha)\theta_{n+1}^N + \alpha\theta_{N-1}^N(\alpha) + (1-\alpha)1}{2} \\ \text{or} \\ \theta_{N-1}^N(\alpha) = \frac{1}{2} - \frac{1-\alpha}{2-\alpha}\theta_n^N - 2b \\ \text{or} \\ \theta_{N-1}^N(\alpha) = \frac{1}{2} - \frac{1-\alpha}{2-\alpha}\theta_{n+1}^N - 2b \end{cases}$$

In both cases we have that  $\lim_{\alpha \rightarrow 1} \theta_{N-1}^N(\alpha) = \frac{1}{2} - 2b < \frac{1}{2}$ . By the same argument as above, there must be  $\alpha(b) \in (0, 1)$  such that, for  $\alpha \geq \alpha(b)$ , only one action can be induced by types in  $[\frac{1}{2}, 1]$ .

- Finally, consider the case where  $N = 2$ . Then, either  $\theta_{N-1}^N(\alpha) \leq \frac{1}{2}$  for any  $\alpha > \frac{1}{2}$ , either there is  $\alpha > \frac{1}{2}$  such that  $\theta_{N-1}^N(\alpha) > \frac{1}{2}$ . In that second case, following Proposition 6 it must be that:

$$\begin{aligned} \theta_{N-1}^N(\alpha) + b &= \frac{\alpha\tilde{\theta} + (1-\alpha)0 + \alpha\theta_{N-1}^N(\alpha) + (1-\alpha)1}{2} \\ \Leftrightarrow \theta_{N-1}^N(\alpha) &= \frac{2-\alpha-4b}{4-2\alpha} \end{aligned}$$

As  $\lim_{\alpha \rightarrow 1} \frac{2-\alpha-4b}{4-2\alpha} = \frac{1}{2} - 2b < \frac{1}{2}$ . By the same argument as above, there must be  $\alpha(b) \in (0, 1)$  such that, for  $\alpha \geq \alpha(b)$ , only one action can be induced by types in  $[\frac{1}{2}, 1]$ .

In all three cases one can show that  $\frac{\partial \theta_{N-1}^N(\alpha)}{\partial b} < 0$ , which implies that that  $\alpha(b)$  is a decreasing function.

□

## 2. Eliciting Multiple Prior Beliefs <sup>1</sup>

(with Mohammed Abdellaoui and Brian Hill)

### Abstract

Despite the increasing relevance of multiple priors in various domains of economics and the significant theoretical work on them, choice-based incentive compatible multiple-prior elicitation largely remains an open problem. This paper develops a new solution, comprising of a preference-based identification of a subject's probability interval for an event, and two procedures for eliciting it. The method does not rely on specific assumptions about subjects' ambiguity attitude or probabilistic sophistication. To demonstrate its feasibility, we implement it in two incentivized experiments to elicit the multiple-prior equivalent of subjects CDFs over continuous-valued sources of uncertainty. The method finds a predominance of non-degenerate probability intervals among subjects for all explored sources, with intervals being wider for less familiar sources. Finally, we use our method to undertake the first elicitation of the mixture coefficient in the Hurwicz  $\alpha$ -maxmin EU model that fully controls for beliefs.

---

<sup>1</sup>I would like to thank Maxim Forlov and Thibaud Choppin de Janvry for very precious research assistance.

## 2.1. Introduction

The standard Bayesian model of decision under uncertainty in economics stipulates that a decision maker's beliefs are fully captured by a single probability measure over the states of the world (Savage, 1954; Anscombe and Aumann, 1963). However, in the face of contrary empirical evidence, starting with Ellsberg (1961b)'s famous examples, more general theories have weakened the standard assumption of probabilistic beliefs. One of the most popular models involves a 'belief component' consisting of a set of priors over the states of the world, rather than imposing a single prior from the outset (Gilboa and Schmeidler, 1989b). Multiple prior models have found a growing number of applications in macro-economics (Ilut and Schneider, 2014), finance (Garlappi et al., 2007; Epstein and Schneider, 2010), mechanism design (Bose and Renou, 2014), econometrics (Manski, 2003, 2013), health economics (Giustinelli et al., 2021), but also beyond economics, in climate science (Kriegler et al., 2009), risk analysis (Cox, 2012) and uncertainty communication (Dieckmann et al., 2010), including by central banks (Carney et al., 2019). Despite obvious theoretical appeal, empirical applications of multiple prior models still have to muddle along in the absence of appropriate choice-based procedures for eliciting their 'belief component'. To date, almost all attempts to operationalize multiple prior elicitation have focused on subjects' *stated* probability intervals for individual events (Giustinelli et al., 2021; Kriegler et al., 2009), and hence involve procedures that are neither choice based nor incentive compatible. This paper proposes, and implements in a laboratory experiment, a choice-based incentive-compatible elicitation method for probability intervals.

Our elicitation is theoretically robust, in a double sense. On the one hand, we work with a decision framework that does not postulate a specific ambiguity attitude from the outset, e.g. universal ambiguity aversion (Gilboa and Schmeidler, 1989b). Specifically, we assume a Hurwicz  $\alpha$ -maxmin expected utility decision setup, which is consistent with recent experimental findings suggesting that people could possibly exhibit ambiguity aversion in some choices and be ambiguity

seeking in others (Kocher et al., 2018). On the other hand, we purposefully eschew any assumption that preferences are generated by precise probabilistic beliefs, i.e. that they are probabilistically sophisticated (Machina and Schmeidler, 1992). Multiple priors clearly come to the fore precisely in situations where such assumptions are unwarranted. To our knowledge, we develop the first choice-based incentive-compatible implementable approach to multiple prior elicitation that is theoretically robust in this double sense (Section 2.5).

Elicitation under Hurwicz  $\alpha$ -maxmin expected utility in the absence of probabilistic sophistication faces well-known theoretical difficulties with the identification of this model (Ghirardato et al., 2004a; Siniscalchi, 2006; Eichberger et al., 2011). Our first contribution is a new resolution of them. Our proposal is inspired by the matching probability (MP) method for determining the subjective probability of an uncertain event (Anscombe and Aumann, 1963). Under Subjective Expected Utility (SEU), the subjective probability of a target event  $E$  coincides with its MP, which can be inferred from preferences between a bet on  $E$  and bets on events generated by extraneous random devices with known probability, e.g. the color of a randomly drawn ball from an urn of known composition. Specifically, for urns containing only red or blue balls, the MP is given by the proportion  $r$  of red balls such that the subject is indifferent between the gamble that pays out a monetary prize  $z$  if  $E$  occurs, and nothing otherwise, and the gamble on the urn with proportion  $r$  of red balls, that pays  $z$  if the next ball drawn from the urn is red (Abdellaoui et al., 2005; Dimmock et al., 2015).

Our insight for eliciting probability intervals is to use extraneous random devices with interval-valued rather than precise probabilities. To illustrate, consider a partially known urn containing only red or blue balls, where all that is known is that at least proportion  $r$  of balls in the urn are red, and at least proportion  $b$  are blue (with  $r + b \leq 1$ ). Here, the probabilities of getting red or blue on the next draw from the urn are summarized by the intervals  $[r, 1 - b]$  and  $[b, 1 - r]$ , respectively. Whilst a single indifference between the gamble on the target event  $E$  and bets on such urns does not suffice to identify the subject's probability interval for  $E$ , we show that the latter can be identified from a *pair* of correctly-chosen

indifferences involving bets concerning  $E$  and such ‘interval-valued’ urns. As explained in detail in Section 2.2, we thus propose a preference-based association of an ‘interval-valued’ urn to each event, which can be thought of as the equivalent of matching probabilities for probability intervals. This *matching probability interval* notion resolves the problem of choice-based incentive-compatible probability-interval elicitation in theory.

Our second contribution is to develop and apply in a laboratory setting two methods for eliciting matching probability intervals: a two-dimensional extension of well-known (one-dimensional) choice lists, and an adaptive binary-choice procedure that can be thought of as an interval analogue of the bisection or staircase method for eliciting matching probabilities. Whilst most elicitation exercises in behavioral economics have focused on beliefs about single events (e.g. it will rain tomorrow), many elicitation applications in economics and beyond require subjects’ probability distributions or CDFs over a continuous variable of interest (e.g. US inflation in 2023, Eurozone GDP in 2022, average global temperature in 2030). Motivated by this observation, we implement our methods on two pairs of sources of uncertainty of the latter sort, to elicit the interval-valued CDFs generated by subjects’ multiple priors. Interval-valued CDFs are commonly used in applications to go beyond the assumption of precise subjective probabilities (Karanki et al., 2009). Application to elicit CDFs provides a test of concept, showing that our approach can operate in such contexts.

Our central findings attest to the feasibility of our approach. Our method yields generally consistent results, eliciting, for the vast majority of subjects, non-degenerate interval-valued CDFs. Our elicitation suggests that imprecise beliefs—i.e. non-singleton intervals for some events—are widespread, with only a handful of subjects having fully precise probabilities for all elicited events. This finding, which is consistent with elicitations using stated probability intervals (Giustinelli et al., 2021), attests to the relevance of multiple-prior belief elicitation. Moreover, by eliciting subjects’ beliefs for two similar sources which intuitively have different degrees of familiarity or predictability (e.g. the temperature in Paris and in Sydney for subjects in Paris), our elicitations provide insight into the relationship

between intuitive familiarity or predictability and probability intervals. Specifically, we observe that the width of the subjective probability intervals is typically larger for intuitively less familiar or less predictable sources. Again, the reasonableness of this correlation provides corroboration of the solidity of our method.

Finally, we connect our elicited beliefs with the Hurwicz  $\alpha$ -maxmin EU model, and perform what to our knowledge is the first elicitation of the mixture coefficient  $\alpha$  in that model that fully controls for beliefs without making strong assumptions about their form, such as probabilistic sophistication.

The paper is structured as follows. Section 2.2 sets out the theoretical background and presents the central planks of our methods (the ‘matching probability interval’ notion, the two-dimensional choice lists and the binary-choice procedure), with the relevant theoretical results. Section 2.3 sets out our experimental implementations, in the form of two studies. Section 2.4 contains our results and supporting analyses, whereas in Section 2.5 we discuss connected issues, next steps and related literature. Proofs, data analyses and experimental details are contained in the Appendices.

## 2.2. Theoretical Background

### 2.2.1. Preliminaries

We consider decision making situations where the objects of choice are two-outcome prospects that pay a fixed monetary outcome  $z$  if an event occurs, and nothing otherwise. Prospects with general winning event  $E$  and winning amount  $z$  are denoted  $(z, E, 0)$  and called *bets*. The *complementary* bet, which pays out when the event  $E$  does not occur, is denoted  $(0, E, z)$ .

As mentioned previously, we use extraneous interval-valued random devices realised by urns containing red and blue balls with partial information about the

composition. For instance, consider the urn where subjects are only told that at least a proportion  $r$  of its balls are red, at least a proportion  $b$  are blue (with  $r + b \leq 1$ ) but receive no information about the colour composition of the remaining balls (except that each is either red or blue). For such an urn, the information only allows assignment of the interval  $[r, 1 - b]$  for the probability of the next ball drawn from the urn being red; similarly, there is the interval  $[b, 1 - r]$  for the next ball being blue. Using these probability intervals for parametrization, we denote the urn with at least proportion  $r$  of red balls and at least proportion  $b$  of blue balls by  $[r, 1 - b]$ . We denote the set of such *interval-valued urns* by  $\mathcal{I}$ .<sup>1</sup>

Each urn  $[r, 1 - b]$  in  $\mathcal{I}$  can be related to two (sorts of) prospects. One is the prospect which pays  $z$  if the next ball drawn from the urn is red, and nothing otherwise. For such a prospect, the probability of winning is characterized by the interval  $[r, 1 - b]$ ; we denote this prospect by  $(z, [r, 1 - b], 0)$ . The other prospect involves the complementary bet on this urn—that is, the bet on the next ball drawn from it being blue. We denote this prospect by  $(0, [r, 1 - b], z)$ . Note that the probability of *losing* here is characterised by the interval  $[r, 1 - b]$ , so the probability of winning is given by  $[b, 1 - r]$ ; this prospect is thus essentially equivalent to  $(z, [b, 1 - r], 0)$ . Since these prospects all involve objectively given information about the probability of winning, albeit in interval rather precise probability form, we call them *interval lotteries* (IL).<sup>2</sup> Standard lotteries correspond to the special case where the composition of the urn is fully known—i.e.  $r = 1 - b$ .

The set  $\mathcal{I}$  of interval-valued urns can be visually represented by the black triangle in Figure 2.1. Each point  $(x, y)$  represents the urn  $[x, y]$ —i.e. with at least proportion  $x$  of red balls and at least proportion  $1 - y$  of blue ones. As such, it represents two interval lotteries: the bet on red,  $(z, [x, y], 0)$ , where all that is known is that the winning probability is in the interval  $[x, y]$ , and the bet on blue  $(0, [x, y], z)$ , where all that is known is that the losing probability is in this

<sup>1</sup>Formally:  $\mathcal{I} = \{[x, y] : (x, y) \in \mathbb{R}^2, 0 \leq x \leq y \leq 1\}$ .

<sup>2</sup>Our notion of interval lottery is distinct from that used by Gul and Pesendorfer (2014). They use ‘interval lottery’ to denote (precise) probability measures over the set of intervals of (monetary) prizes; here, ‘interval lottery’ denotes assignments of probability intervals to (fully determined, precise) outcomes. In particular, the interval lotteries  $(z, [r, 1 - b], 0)$  used here clearly do not belong to the concept used by Gul & Pesendorfer’s (zero probability is assigned to each outcome between 0 and  $z$ ).

interval. Standard lotteries, and urns with fully known composition correspond to the points on the diagonal ( $x = y$ ).

### 2.2.2. Upper and lower probabilities and CDFs

The sources of uncertainty considered here are real-valued variables, e.g. the daily minimum temperature in Paris between November and March. In the precise probability case, elicitation aims at revealing the subjective probability over the variable, which can be represented as a subjective cumulative distribution function (CDF). One common way of doing so is by eliciting subjective probabilities of events corresponding to the variable lying below certain fixed values, i.e. if the variable takes values in the real interval  $T$ , the events considered are of the form  $E_t = \{t' \in T : t' \leq t\}$ . For future reference, we call these *cumulative events*. We now set out the aim of the corresponding exercise for multiple priors.

Multiple prior belief representations involve a convex, closed set  $\mathcal{C}$  of probability measures: measures over the values of variable of interest, in our case.<sup>3</sup> For each event  $E_t$ , the set of priors generates a *probability interval*  $\{p(E_t) : p \in \mathcal{C}\} = [\underline{p}(E_t), \bar{p}(E_t)]$ , where  $\underline{p}(E_t) = \min \{p(E_t) : p \in \mathcal{C}\}$  and  $\bar{p}(E_t) = \max \{p(E_t) : p \in \mathcal{C}\}$  are the *lower* and *upper probabilities* for  $E_t$  respectively. As is well-known, a set of priors contains more information than the collection of upper and lower probabilities for all events generated from it, but the latter (or sometimes less) is often sufficient for applications, and sometimes preferable, insofar as it is easier to communicate.

For continuous-valued variables, CDFs are often used. Recall that for a probability measure  $p \in \Delta(T)$ , the CDF is defined as  $F_p(t) = p(\{t' \in T : t' \leq t\}) = p(E_t)$ ; when the probability measure is a subjective probability, this is the corresponding subjective CDF. In this context, a set of priors  $\mathcal{C}$  generates the *interval-valued CDF*  $F_{\mathcal{C}}(t) = \{p(E_t) : p \in \mathcal{C}\}$ , which takes the probability interval corresponding to  $E_t$  as value, for each  $t$ . This can be visually represented in terms of two (real-

<sup>3</sup>Technically,  $\mathcal{C} \subseteq \Delta(T)$ , the set of probability measures over  $T$ .

valued) functions: the *lower CDF*,  $\underline{F}_{\mathcal{C}}(t) = \min \{p(E_t) : p \in \mathcal{C}\} = \underline{p}(E_t)$ , and the *upper CDF*,  $\overline{F}_{\mathcal{C}}(t) = \max \{p(E_t) : p \in \mathcal{C}\} = \overline{p}(E_t)$ . These are widely used for representing, communicating and studying sets of priors over continuous variables, where they often go under the name of distribution bands or p-boxes (Berger et al., 2000; Karanki et al., 2009), though, like upper and lower probabilities, they involve an information loss as compared to sets of priors. In the implementation of our elicitation procedure conducted here, our aim is to elicit subjective upper and lower CDFs for the variables considered.

### 2.2.3. Decision model

For the purposes of presentation, we will focus on one of the most popular and general models of decision under uncertainty involving sets of priors, the Hurwicz  $\alpha$ -maxmin EU model. (In Section 2.5 and Appendix 2.7.3, we discuss how our proposals extend to generalisations.) Under the  $\alpha$ -maxmin model, a bet  $(z, E, 0)$  is evaluated according to:

$$\alpha \underline{p}(E).u(z) + (1 - \alpha) \overline{p}(E).u(z) \quad (2.1)$$

where  $\overline{p}(E)$  and  $\underline{p}(E)$  are the upper and lower probabilities of  $E$  generated by the subjects' set of priors, as defined above, and  $u$  is a utility function normalized so that  $u(0) = 0$ . The coefficient  $\alpha$  is often associated with ambiguity attitude in this model, with  $\alpha > \frac{1}{2}$  considered as reflecting typical ambiguity aversion and  $\alpha < \frac{1}{2}$  as typical ambiguity seeking. For illustration, the standard behavior in the Ellsberg two-urn example can be accommodated by  $\alpha > \frac{1}{2}$  but not by  $\alpha < \frac{1}{2}$ . This model coincides with the Gilboa-Schmeidler maxmin-EU model when  $\alpha = 1$ ; whenever  $\alpha \neq 1$ , the model does not satisfy the Gilboa-Schmeidler uncertainty aversion axiom—which can be thought of as characterizing universal ambiguity aversion—and hence can accommodate ambiguity seeking behavior in certain choices (even for  $1 > \alpha > \frac{1}{2}$ ). Since typical findings suggest some ambiguity seeking behavior, but not in situations that give reason to believe that  $\alpha < \frac{1}{2}$ , we take  $\alpha > \frac{1}{2}$  to be typical and assume that preferences are represented according

to (2.1) with  $\alpha > \frac{1}{2}$  in the sequel (except where specified). As discussed in Section 2.7.3, the full strength of this assumption is not required for the central elements of the elicitation method. The aim is to elicit  $\underline{p}(E)$  and  $\bar{p}(E)$ .

We also assume the same representation for interval lotteries: i.e. preferences concerning them are represented by an evaluation of  $(z, [r, 1 - b], 0)$  by:

$$\alpha r u(z) + (1 - \alpha)(1 - b)u(z) \quad (2.2)$$

#### 2.2.4. Matching Probability Benchmark

As discussed in the Introduction, our proposal can be thought of as an interval analogue of the matching probability method. The latter is based on the insight that, under SEU, the decision maker's subjective probability of an event  $E$ ,  $p(E)$ , is equal to the *matching probability* (MP) for  $E$ , defined as the probability  $p$  such that  $(z, [p, p], 0) \sim (z, E, 0)$ . So eliciting the MP is tantamount to eliciting the subject's subjective probability. In this way, the concept of MP resolves the problem of incentive-compatible probability elicitation in theory: it suffices to elicit preferences between the bet  $(z, E, 0)$  and the lottery  $(z, [r, r], 0)$  for values of  $r$  ranging from 0 to 1. Due to stochastic dominance, the bet is expected to be preferred for all values below some threshold  $r^*$ , with the lottery being preferred above. In practice, the subject may be asked to fill in a choice list with each line offering the choice between the bet  $(z, E, 0)$  and a lottery  $(z, [r, r], 0)$ , for equally spaced values of  $r$  ranging from 0 to 1. Then, the shifting value  $r^*$ , i.e. the MP of event  $E$ , is determined by the maximum value of  $r$  for which the subject prefers the bet. To incentivize the procedure, a lottery—or equivalently, a row in the choice list—is selected at random and the stated choice between the bet and that lottery is played 'for real'. This mechanism is *incentive compatible*: reporting one's real preferences is weakly dominant, for if one does not do so, there is a non-zero probability of playing a choice 'for real' in which one will obtain one's less preferred option.

Another elicitation procedure, sometimes called the ‘bisection’ procedure, uses a chained sequence of binary choice questions to ‘hone in’ on the matching probability. Starting with a binary choice between  $(z, [\frac{1}{2}, \frac{1}{2}], 0)$  and  $(z, E, 0)$ , it then asks a binary choice with the midpoint of the lower (respectively upper) interval  $[0, \frac{1}{2}]$  (resp.  $[\frac{1}{2}, 1]$ ) whenever the subject chooses the former (resp. latter) option, and so on. While this procedure optimizes the time needed to elicit MPs as compared to choice lists, the application of the random incentive mechanism to it is sometimes criticized given that the binary choice offered at each iteration depends on the subject’s choices at previous iterations. Hence there is a possibility of strategic considerations in responses; accordingly, this mechanism is incentive compatible only under the assumption that subjects treat each binary choice in isolation from the others.

A hybrid elicitation procedure—proposed and implemented by [Abdellaoui et al. \(2019\)](#)—combines these two procedures. The subject first undertakes the bisection procedure, to aid her in filling in a choice list. At the end of the bisection procedure, she is presented with a choice list filled in according to her replies on the bisection procedure (i.e. for all entries with probability above the MP found by the bisection procedure, the lottery is ‘pre-filled’ as chosen, and for all entries below, the bet is). At this stage, the subject must confirm the choice-list replies deduced from her bisection choices or correct the responses accordingly, and then confirm. The prospect played ‘for real’ is determined by her reply in this choice list task, according to the associated incentivisation mechanism specified above. As argued in the cited paper, this combination of the two methods retains the advantages of both: the bisection procedure—consisting of easier tasks—aids the subject to fill in the choice list—which has more robust incentive compatibility properties.

In this paper, we provide analogues of each of these three elements for multiple priors. First, we propose an analogue of MPs, and show that eliciting all preferences between bets for and against an event  $E$  and interval lotteries (ILs) is sufficient to yield the subject’s probability interval for  $E$ , in theory. Turning to implementation, we then develop an extended notion of choice list under which

a single ‘switching point’ question reveals the subject’s probability interval for the event in question. We also develop a binary-choice procedure, reminiscent of the bisection procedure. In our experiment we implement both, with the binary-choice procedure helping to ‘fill in’ the extended choice list, as under the hybrid approach. We now present these elements in turn.

### 2.2.5. Theory: Matching Probability Intervals

Our approach is based on the following notion. The *matching probability interval (MPI)* of an event  $E$  is an  $[r, 1 - b] \in \mathcal{I}$  such that:

$$(z, [r, 1 - b], 0) \sim (z, E, 0) \quad (2.3)$$

$$(0, [r, 1 - b], z) \sim (0, E, z) \quad (2.4)$$

Plugging these indifferences into (2.1) and (2.2) yields the following equations:

$$\begin{aligned} \alpha r + (1 - \alpha)(1 - b) &= \alpha \underline{p}(E) + (1 - \alpha)\bar{p}(E), \\ \alpha(1 - (1 - b)) + (1 - \alpha)(1 - r) &= \alpha(1 - \bar{p}(E)) + (1 - \alpha)(1 - \underline{p}(E)). \end{aligned} \quad (2.5)$$

Clearly, these equations are satisfied by  $r = \underline{p}(E)$ ,  $1 - b = \bar{p}(E)$ . Moreover, whenever there is a unique pair  $r, 1 - b$  satisfying them, then there is a unique matching probability interval, which indicates precisely the subjective probability interval for  $E$ : i.e.  $[\underline{p}(E), \bar{p}(E)] = [r, 1 - b]$ . Under the  $\alpha$ -maxmin EU model with  $\alpha \neq \frac{1}{2}$ , the MPI is unique (Proposition 3, Appendix 2.7.1). So to elicit the subjects’ probability interval for the event  $E$ , it suffices to find the MPI of  $E$ .

The MPI can be illustrated in Figure 2.1. The red hatched area represents the upper contour set of the bet  $(z, E, 0)$  in the space of interval lotteries corresponding to bets on red: that is, the set of  $(x, y)$  such that  $(z, [x, y], 0) \succeq (z, E, 0)$ . The blue hatched area is the upper contour set of the complementary bet  $(0, E, z)$  in

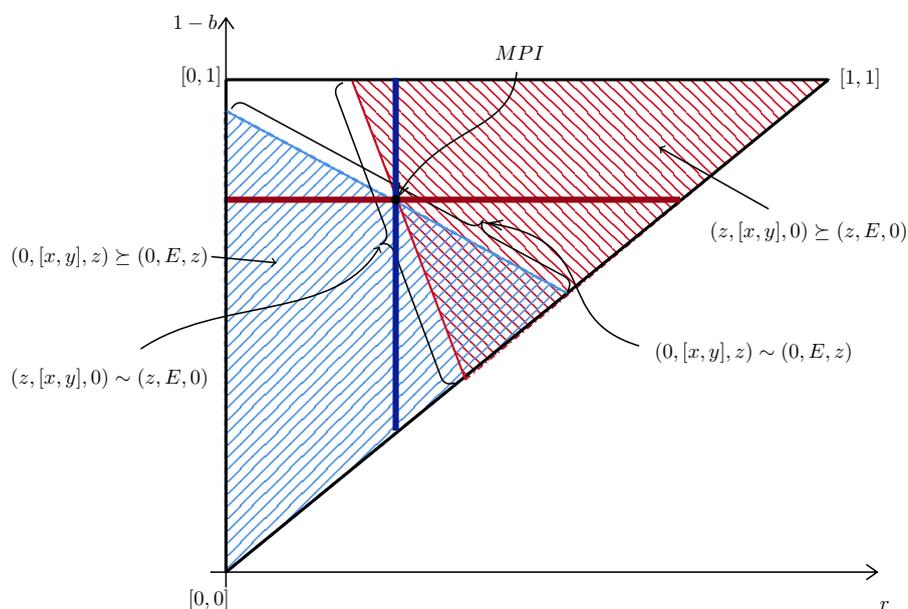


Figure 2.1: Matching Probability Interval in space  $\mathcal{I}$  of interval-valued urns, for an event  $E$ .

the space of complementary ILs (corresponding to bets on blue): that, it is the set of  $(x, y)$  such that  $(0, [x, y], z) \succeq (0, E, z)$ . The boundaries of these sets (the diagonal red and blue lines respectively) represent the indifference curves of  $(z, E, 0)$  (resp.  $(0, E, z)$ ), in the space of ‘red’ (resp. ‘blue’) ILs. The matching probability interval corresponds to the black point at the intersection of these two lines.

This Figure also brings out the contribution of ILs as compared to standard lotteries and the long-standing identification problem for the  $\alpha$ -maxmin EU model (Section 2.5). The MP of the bet  $(z, E, 0)$  is given by the point where the red indifference curve meets the diagonal; clearly eliciting it is insufficient to pin down the subject’s probability interval for  $E$ . Similarly, the MP of the complementary bet  $(0, E, z)$  is given by the point where the blue indifference curve meets the diagonal. Eliciting both of these MPs is sufficient to pin down the subject’s probability interval (as the intersection of the indifference curves) *only if the slope of the indifference curves is known: but this is determined by the  $\alpha$  mixture coefficient in (2.1), which also needs to be elicited*. The use of ILs, and the notion of MPI built upon it, allows elicitation of the subjective probability interval without requiring elicitation of the  $\alpha$  mixture coefficient. Indeed, we shall use our probability interval elicitation in tandem with MPs to estimate subjects’  $\alpha$  (Section 2.4.5).

Just like the notion of MP for subjective probabilities under EU, the notion of MPI resolves the challenge of incentive-compatible probability-interval elicitation *in theory*. Obtaining the subject's preferences between each pair consisting of a bet (for or against  $E$ ) and an IL provides the MPI—and hence the subject's probability interval for  $E$ —as the point satisfying (2.3) and (2.4). It is well known that there are fully incentive-compatible mechanisms for eliciting such preferences. For instance: the subject states her preference between each pair consisting of a bet (for or against  $E$ ) and an IL; a random bet (for or against  $E$ ) and IL are then chosen and she is remunerated according to the prospect between the two which she stated as more preferred. By the standard argument, it is in the subject's best interest to report preferences truthfully, for if not there is a chance of receiving her less preferred prospect in the choice which is 'played for real'. These elicited preferences provide, *inter alia*, the MPI and hence the subject's probability interval for  $E$ .

Of course, implementation typically requires a method involving fewer preference questions. This is especially challenging for probability intervals, since the target is a point in a two-dimensional space, whereas elicitation of the precise probability of an event only needs to search a one-dimensional space. As we shall see, there is a trade-off in this context between parsimony and the strength of the incentive compatibility. For the purposes of implementation, we develop two more parsimonious methods for eliciting MPIs, whilst making no claim to have exhausted all possibilities.

### 2.2.6. Implementation 1: 2D Choice lists

Consider any MPI  $[r, 1 - b]$  of an event  $E$ , so that the indifference (2.3) is satisfied. Since  $\alpha > 0$  in the representation (2.1), it follows that:

$$\begin{aligned} (z, [q, 1 - b], 0) &\succ (z, E, 0) && \text{for all } q > r && (2.6) \\ (z, [q, 1 - b], 0) &\prec (z, E, 0) && \text{for all } q < r && \end{aligned}$$

On Figure 2.1, this determines the preferences on the ‘red’ ILs corresponding to the bold red (horizontal) line. To the left of the MPI, the bet on  $E$  is preferred to the IL corresponding to the bet on red from the urn  $[q, 1 - b]$  (i.e. with probability  $[q, 1 - b]$  of winning); to the right of the MPI, the IL is preferred to the bet; and at the MPI, the two are indifferent.

Similar reasoning applies to complementary bets. By the indifference (2.4) and the fact that  $\alpha > 0$ , for an MPI  $[r, 1 - b]$  of  $E$ , it follows that:

$$\begin{aligned} (0, [r, q], z) &\prec (0, E, z) && \text{for all } q > 1 - b \\ (0, [r, q], z) &\succ (0, E, z) && \text{for all } q < 1 - b \end{aligned} \quad (2.7)$$

On Figure 2.1, this determines the preferences on the (complementary) ‘blue’ ILs corresponding to the bold blue (vertical) line. Above the MPI, the bet against  $E$  is preferred to the IL corresponding to the bet on blue from urn  $[r, q]$  (i.e. with probability  $[r, q]$  of losing); below the MPI, the IL is preferred to the bet; and at the MPI, the two are indifferent.

*Remark 1.* Note the property behind preferences (2.6) and (2.7):

*Definition 1 (Lower Stochastic Dominance).* For every  $r, r', b \in [0, 1]$  with  $1 - b \geq r, r'$ ,  $(z, [r, 1 - b], 0) \prec (z, [r', 1 - b], 0)$  whenever  $r < r'$ .

Between ILs  $(z, [r, 1 - b], 0)$  and  $(z, [r', 1 - b], 0)$  corresponding to bets on red from urns with the same minimum proportion of blue balls, the decision maker prefers the prospect where the minimum proportion of red balls is higher. This can be thought of as an analogue of the standard stochastic dominance property for lotteries. Given the weakness of this assumption, this part of the proposal extends beyond the  $\alpha$ -maxmin EU model to models satisfying Lower Stochastic Dominance (see Section 2.5 and Appendix 2.7.3).

It follows from the previous observations that the only points supporting the specified preference patterns on the corresponding horizontal and vertical lines are MPIs.

**Proposition 1.** For any event  $E$ , let  $[r, 1 - b] \in \mathcal{I}$  be such that:

$$\begin{aligned} (z, [q, 1 - b], 0) \succ (z, E, 0) & \quad \text{for all } q > r \\ (z, [q, 1 - b], 0) \prec (z, E, 0) & \quad \text{for all } q < r \\ (0, [r, q], z) \prec (0, E, z) & \quad \text{for all } q > 1 - b \\ (0, [r, q], z) \succ (0, E, z) & \quad \text{for all } q < 1 - b \end{aligned}$$

Then  $[r, 1 - b]$  is a matching probability interval of  $E$ .

The red (horizontal) and blue (vertical) bold lines can thus be thought of as a pair of choice lists, and the MPI is the switching point on each of them. We henceforth refer to the combination of the two as a *2D choice list*. Inspired by this observation, consider the following mechanism for eliciting a subject's MPI for an event  $E$ . A subject reports an interval-valued urn  $[r, 1 - b]$  for  $E$ . She is then remunerated as follows. First, an urn  $[x, y]$  is chosen at random from the 2D choice list.<sup>4</sup> Then she 'receives' or 'plays' a bet or IL according to the following scheme:

- if  $y = 1 - b, x < r$ , then she gets  $(z, E, 0)$  (i.e. she 'plays' the bet on  $E$ )
- if  $y = 1 - b, x \geq r$ , then she gets  $(z, [x, y], 0)$  (i.e. she 'plays' this IL)
- if  $x = r, y < 1 - b$ , then she gets  $(0, [x, y], z)$  (i.e. she 'plays' this IL)
- if  $x = r, y \geq 1 - b$ , then she gets  $(0, E, z)$  (i.e. she 'plays' the bet against  $E$ )

It follows from the previous Proposition that this mechanism is incentive compatible in the following sense: on each choice list, reporting the urn reflecting one's true upper or lower probability is in one's best interest—it weakly dominates any other report in the respective choice list. Hence asking a subject for an urn  $[r, 1 - b]$  such that, in each of the branches on the 2D choice list, she prefers the option she would receive under the mechanism, incentivises reporting of a MPI. Note that,

<sup>4</sup>I.e. the interval is chosen at random from  $\{[x, y] : (x, y) \in \mathcal{I}, y = 1 - b\} \cup \{[x, y] : (x, y) \in \mathcal{I}, x = r\}$ , the union of the horizontal and vertical lines going through  $(r, 1 - b)$  in the Figure 2.1.

since precise probabilities (and SEU) are a special case of multiple priors (respectively,  $\alpha$ -maxmin EU), this mechanism functions equally for Bayesian decision makers, who are incentivised to report their precise probabilities.

Despite the considerably higher complexity involved in eliciting probability intervals as opposed to precise probability values, this incentive mechanism is as parsimonious as standard choice lists for MPs. Like them, it only asks for one point. Like them, this point can be used to ‘fill in’ the other preferences in the 2D choice list using the Lower Stochastic Dominance property (Definition 1). However here, this does not suffice to determine all preferences between bets and ILs. This is due to the space of ILs being a dimension larger than that of lotteries: whereas stochastic dominance allows filling in all bet-lottery preferences in a standard MP choice list, no stochastic-dominance property can do the same for all bet-IL preferences.<sup>5</sup> Because of this, the branches of the 2D choice list are not fixed, but depend on the subject’s reports, e.g. the choice list for the lower probabilities (red horizontal line in Figure 2.1) depends on the reported upper probability. It is thus possible for a subject to choose strategically on the basis of the bet-IL choices she would like to face, thus reporting preferences that do not satisfy the conditions in Propositions 1 and do not yield the MPI.

It is well-known that the incentive compatibility of an elicitation mechanism typically depends on some form of *isolation assumption*. For instance, existing preference elicitation methods under ambiguity concerning several events are unbiased only if subjects treat different preference tasks in isolation from others (Bade, 2015): a property that Azrieli et al. (2018) analyze under the name of state-wise monotonicity. The proposed mechanism is incentive compatible under an isolation assumption of this sort: namely, that subjects treat the branches in the 2D choice list in isolation from each other. Experimental design can favor such isolation. As set out in Section 2.3.3 (see also Appendix 2.9), 2D choice lists are realized in our implementation by a single scrollbar with two cursors: visually very different from that in Figure 2.1, this presentation promotes considering each

---

<sup>5</sup>This is also true for stronger stochastic dominance properties than Lower Stochastic Dominance.

branch individually and is less suggestive of strategic opportunities of changing the choice lists by ‘moving around the 2D space’. That said, the extent to which isolation holds in an experiment is ultimately an empirical question, and we treat it as such. On this front, an advantage of the proposed elicitation method is that strategic reasoning leads to easily recognizable choice patterns. As discussed in Appendix 2.7.3, for any subject represented by (2.1) with  $\alpha \in (0, 1)$  (whatever her set of priors  $\mathcal{C}$ ), her optimal response to the 2D choice list task when reasoning strategically is one of the points  $[0, 0]$ ,  $[0, 1]$ ,  $[1, 1]$ : i.e., one of the vertices of the space of interval-valued urns in Figure 2.1 (see Appendix 2.7.3 for details). Moreover, the only case in which a response in the interior of the space is not sub-optimal is when  $\alpha = 1$  and the subject assigns a precise probability of 0.5 to the event under consideration. Checking for concentration of responses at the vertices will thus provide insight into the extent of strategic reasoning in our subject pool (Section 2.4.1).

### 2.2.7. Implementation 2: Binary-choice procedure

In our implementation, the 2D choice list task is preceded by a ‘bisection-style’ binary-choice procedure for identifying the MPI. Here we set out its general principles; full details are provided in Appendix 2.7.2. The logic can again be illustrated on Figure 2.1. The space of interval-valued urns (or probability intervals) is divided into four preference-defined areas, summarised in Table 2.1. The procedure is based on the following observation.

**Proposition 2.** *Suppose preferences are represented according to (2.1) with  $\alpha > \frac{1}{2}$ , and let  $E$  be an event.*

- a. *For any point  $[x, y]$  in the R-B region (i.e. such that the corresponding preferences in Table 2.1 hold, for  $E$ ),  $\underline{p}(E) \leq x$  and  $\bar{p}(E) \geq y$ . Moreover, for any point  $[x, y]$  in the W region,  $\underline{p}(E) \geq x$  and  $\bar{p}(E) \leq y$ .*

Name	Preferences	Colour (in Figure 2.1)
R-B	$(z, [x, y], 0) \succeq (z, E, 0)$ $(0, [x, y], z) \succeq (0, E, z)$	Red & Blue
W	$(z, [x, y], 0) \preceq (z, E, 0)$ $(0, [x, y], z) \preceq (0, E, z)$	White (neither Red nor Blue)
R	$(z, [x, y], 0) \succeq (z, E, 0)$ $(0, [x, y], z) \preceq (0, E, z)$	Red
B	$(z, [x, y], 0) \preceq (z, E, 0)$ $(0, [x, y], z) \succeq (0, E, z)$	Blue

Table 2.1: Preference-based division of  $\mathcal{I}$ 

- b. For any point  $[x, y]$  in the R region (i.e. such that the corresponding preferences in Table 2.1 hold, for  $E$ ), every  $[x', y']$  with  $x' \geq x$  and  $y' \geq y$  is also in R. Moreover, for any point  $[x, y]$  in the B region, every  $[x', y']$  with  $x' \leq x$  and  $y' \leq y$  is also in B.

It follows from part a. that if the experimenter has found a point (interval-valued urn)  $[x_{RB}, y_{RB}]$  in the R-B region (i.e. such that there is the preference pattern in Table 2.1, row 1), and a point  $[x_W, y_W]$  in the W region, then the MPI is contained in the ‘box generated’ by these points, i.e. it is in the set  $\{[x, y] : x_W \leq x \leq x_{RB}, y_{RB} \leq y \leq y_W\}$ . The procedure works by searching the smallest such generated box for further points in R-B or W, in order to ‘reduce’ the size of the boxes and hence ‘home into’ the MPI. In this sense, it is analogous to the bisection procedure for MPs, where preferences indicate that the MP is in a particular interval, and the procedure searches to reduce the width of that interval.

Note that a similar result to Proposition 2 a. for the R and B regions does not hold.<sup>6</sup> However, by part b. it can be concluded, for any point (interval-valued urn)  $[x, y]$  in R, that every point North-East of  $[x, y]$  is also in R, and similarly for a point in B. So, if the experimenter has just discovered a point in R (i.e. the elicited preferences for the relevant urn are as set out in Table 2.1), then, to seek a point in R-B or W, she needs to look South-West of this point; and analogously

<sup>6</sup>This can be seen by considering the horizontal and vertical bold lines in Figure 2.1 to define four quadrants, and by noting that there are both red and blue areas in the upper left-hand and lower right-hand quadrants.

for points in B. The procedure works, when at a point in R and B, by performing a bisection along one-dimensional cuts of the space  $\mathcal{I}$  guided by this observation, until a point in R-B or W is found. Then the procedure between the closest R-B and W points re-applies, in an attempt to generate a ‘smaller’ box. Details are available in Appendix 2.7.2.

One final important property of the procedure used is an in-built ‘precision bias’. Whenever no point in the areas R-B or W has been found, and hence at a stage when it is searching for points in these areas, the procedure deliberately moves closer to the space of precise urns (the  $45^\circ$  line in Figure 2.1); again, see Appendix 2.7.2 for details. In this way, if there is any misclassification of subjects, the tendency would be for the procedure to represent them as more precise than they actually are.

### 2.3. Experimental Methods

We carried out two experiments in which we used our method to elicit upper and lower CDFs for various sources of uncertainty. Both experiments involved two comparable yet different sources of uncertainty (Table 2.2). EXP 1 implemented a faster elicitation, eliciting probability intervals for fewer points per source. This left time for standard matching probability elicitation for the same events, which yield insight into the  $\alpha$ -maxmin EU model (see Section 2.4.5). By contrast, EXP 2 implemented a slower upper and lower CDF elicitation, eliciting more points per source. It also involved an omnibus confirmation screen, allowing subjects to confirm (or revise) all choices concerning events in a source, after elicitation. In EXP 2, no MPs were elicited.

	Type of Source of uncertainty	Treatments	# Events	Elicited	# subjects
EXP 1	Minimum winter temperature (°C)	Where: Paris; Sydney	4	MPIs, MPs (Paris)	80
EXP 2	Entrance exam grade ( /20)	Exam: Maths, Contraction	5	MPIs	52

Table 2.2: Summary of studies

### 2.3.1. Subjects

132 subjects (undergraduate students) were recruited in total from two French academic institutions: 80 from university of Paris 1 for EXP 1 and 52 from HEC Paris Business School for EXP 2 (Table 2.2). EXP 1 took place in June 2018; EXP 2 took place in September-October 2019. Subjects' choices were collected through computer-based individual interviews that lasted about one hour in each of the two studies. Each individual interview started with a video presentation of the experimental instructions, followed by comprehension questions and one training MPI elicitation task (on an event not involved in the ensuing experiment).<sup>7</sup> Appendix 2.9 reports the typical screenshots faced by subjects for the tasks. In both experiments, subjects were told that there were no wrong or right answers, and that they could ask any question regarding the experiment. Any differences in experimental instructions between the experiments are explained in the sequel.

### 2.3.2. Stimuli and treatments

#### Sources of uncertainty

Each experiment involved two comparable sources of uncertainty, with one treatment for each source (Table 2.2). The type of source in EXP 1 was the minimum daily temperature over the previous November–March period; the sources dif-

<sup>7</sup>The video presentations are available upon request.

ferred on the city whose temperature was of interest—Paris, where the experiment was carried out, and Sydney. EXP 2 involved two of the previous year’s entrance exams for admission at undergraduate level through the ‘ECS’ and ‘ECE’ entrance streams to a prominent French business school, HEC Paris. The subjects in the experiment—students admitted at this level at the school—had sat the exams either the same year or the previous year. The sources differed in the exam considered: the probability and statistics exam (officially called ‘Mathématiques II’), which is generally considered to be ‘objectively marked’, and the ‘Contraction’ exam—a summary of a philosophical or literary text—whose marking is considered more ‘subjective’, ‘random’ and ‘unpredictable’ among candidates and students. Indeed, the marks in the latter exam have higher variance.<sup>8</sup>

Each source of uncertainty involves a variable (temperature in °C, mark out of 20), the aim was to elicit subjects’ multiple prior beliefs—in the form of the generated upper and lower CDFs (Section 2.2.2)—over the variable. In each experiment, the subject chose a number at the beginning of the experiment<sup>9</sup> which identified, according to a spreadsheet to which the subject would only have access at the end of the experiment, a random day  $D$  between 01/11/2017 and 31/03/2018 (in EXP 1), and a random candidate  $C$  for entry to HEC Paris in 2019 (in EXP 2). We estimated upper and lower CDFs by eliciting subjects’ upper and lower probabilities for cumulative events, i.e. events  $E_{t_i}$  of the form: “the minimum temperature on day  $D$  in Paris (resp. Sydney) was less than or equal to  $t_i$ ”, or “candidate  $C$  obtained a mark less than or equal to  $t_i$  in the Maths (resp. Contraction) exam” (Section 2.2.2), for various fixed values of  $t_i$  given in Table 2.3.<sup>10</sup>

Note that the events used pertained to time periods several months before subjects participated in the experiment (2017-2018 Winter temperature, for sub-

<sup>8</sup>The variance of marks for Maths is 3.77, where it is 9.92 for Contraction.

<sup>9</sup>They chose a number between 1 and 150 in EXP1 (the number of days in the period considered), 1 and 456 (the number of candidates) in EXP 2.

<sup>10</sup>For each source in EXP 1, we chose temperature values close to the 10%, 33%, 66% and 90% percentiles of the true distribution. For EXP 2, we used the same values for both sources (Maths and Contraction), picked so they would seem to reasonably scan the range and correspond to comparable points in the true distribution over Contraction scores, where they were at the 3% 15% 33% 68% and 86% percentiles. They were at the 0%, 0%, 2%, 21% and 60% of the true distribution of Maths scores.

	Treatment	Events $E_{t_i} = \{t' \in T : t' \leq t_i\}$ for $t_i$ :
EXP 1	Paris	-2, 2, 5, 8
	Sydney	15, 18, 20, 22
EXP 2	Maths	7, 10, 12, 15, 17
	Contraction	7, 10, 12, 15, 17

Table 2.3: Sources of uncertainty and events

jects taking part in the experiment in Spring 2018; exams sat in Spring 2019, for subjects taking part in the experiment in Autumn 2019). Moreover, there is a natural difference in the familiarity with or predictability of the sources involved in each experiment—with Paris’s weather being more familiar to Paris subjects than Sydney’s, and Maths considered a more predictable exam than Contraction. Finally, we had access to the real data for all the sources, which were used for incentivisation (Section 2.3.4).<sup>11</sup>

## Organisation of EXP 1

EXP 1 consisted of three blocks of tasks. Each of the first two blocks concerned a single source (Paris or Sydney), and involved the elicitation of the upper and lower probabilities for each of the events in the source (Table 2.3). The order of these two blocks was randomized. In each block, the subject first declared, in a non-incentivized manner and using a scrollbar, her estimated maximum and minimum values for the minimum temperature on the unidentified day selected. This is standard procedure in expert elicitation for unbounded sources, aimed at combatting anchoring bias (Morgan, 2014), and played no role in our elicitation. Then the elicitation procedure set out in Section 2.2 and implemented as described in Section 2.3.3 was applied for each event in the source. Within each block, the two extreme events (i.e. lowest and highest temperature points) were asked first, in a random order, followed by the other two events, in a random order.

<sup>11</sup>For the weather, the data source was Météo France (Paris Orly [meteofrance.fr](http://meteofrance.fr)) and the Australian Bureau of Metereology (Sydney Observatory Hill [bom.gov.au](http://bom.gov.au)); for the marks, they were provided by HEC admission services.

The final block involved the elicitation of MPs for the events in Paris treatment. MPs were elicited for each event  $E_{t_i}$  in this source and its complement  $E_{t_i}^c$  (Table 2.3). The order of elicitations was randomized in this block.

## Organisation of EXP 2

EXP 2 consisted of two blocks of tasks, corresponding to the first two blocks of EXP 1. Each of the blocks concerned a single source (Maths or Contraction), and involved the elicitation of the upper and lower probabilities for each of the events in the source (Table 2.3). The order of the blocks was randomized. In each block, the elicitation procedure set out in Section 2.2 and implemented as described in Section 2.3.3 was applied for each event in the source. The order of events in each block was randomized. Each block ended with a omnibus confirmation screen, in which the confirmation 2D choice list scrollbars (Section 2.3.3) for each of the events in the source were displayed and graphed, and the subject was given the opportunity to go back and modify any of her responses for the events in the source. This screen, the sources and the larger number of events elicited per source were the central differences with respect to EXP 1.

### 2.3.3. Elicitation techniques

#### Elicitation of upper and lower probabilities

The elicitation procedure followed the steps whose underlying theory was set out in Section 2.2. For each event  $E_{t_i}$  (Table 2.3), we first applied the chained binary-choice procedure set out in Section 2.2.7 and Appendix 2.7.2. Each step of the procedure involved an event  $E_{t_i}$ , and a 100-ball urn with a specified minimum number of blue and red balls, where nothing was known about the colour of the remaining balls. At each step, two choices were elicited from subjects: their choice in the decision between the bet on the event  $E_{t_i}$  and the bet on the next ball

drawn from the urn being red, and their choice in the decision between the bet on  $E_{t_i}^c$  (or against  $E_{t_i}$ ) and the bet on the next ball drawn from the same urn being blue. (Details on the display are provided in Appendix 2.9.) The urn proposed in the next step depended on the preferences elicited in the previous step according to the procedure (Section 2.2.7 and Appendix 2.2.7). The subjective probability interval for  $E_{t_i}$  elicited at the end of the procedure is deduced from the preferences over such bets, as specified in the cited sections. The procedure continued until the interval was estimated to a precision of 0.15 if it was not degenerate, 0.05 if it was degenerate (i.e. corresponding to a precise probability), or up to 12 steps, whichever came first.

At the end of this procedure, the ‘confirmation’ 2D-choice list was displayed for verification. Although the ‘space of choices’ to be confirmed is the two-dimensional ‘cross’ in Figure 2.1 (Section 2.2.6), we implemented it via a one-dimensional scrollbar based display with two cursors (see Appendix 2.9). The cursors specified the minimum number of red and blue balls respectively, and hence together determined an interval-valued urn. They were initially set at the values determined by the binary-choice procedure. To confirm the whole 2D-choice list, the subject had to scan all the associated choices. When moving the red cursor, the blue cursor remained fixed at the pre-specified value. This accentuates the separate nature of the cursors, which cannot be moved in tandem. It thus promotes isolation of the branches of the 2D choice list, which is the condition for incentive compatibility (Section 2.2.6). By moving the red cursor, the subject scanned all the urns with the same minimum number of blue balls but differing minimum numbers of red balls, i.e. the choices represented by the bold red horizontal line in Figure 2.1. During this scan the corresponding choices between the bets on the event  $E_{t_i}$  and the bets on red from the urn were displayed, with the ‘chosen bet’, specified as in Section 2.2.6, being indicated (i.e. the bet on the urn when there are more red balls than the provisionally elicited point; the bet on the event otherwise). The subject also had to scan the choices associated with moving the blue cursor—there, the red cursor (and hence minimum number of red balls) was held fixed. When moving the blue cursor, the choices between the bets on  $E_{t_i}^c$  and the bets on blue from the urn were displayed (with the corresponding choice,

following the logic in Section 2.2.6). This corresponds to scanning the choices represented by the vertical bold blue line in Figure 2.1. By clicking on the appropriate bet (on the event or the urn) in any of the displayed choices, subjects could revise their reported preferences, hence modifying the specified position of the fixed cursors (and the associated provisionally elicited point). After such modifications, subjects had to reconfirm all of the associated choices, by moving one and then the other cursor, before moving on to the next stage of the experiment. The precision of the scrollbar, and hence subject responses, was to the nearest 0.01 (to the precise minimum number of red and blue balls respectively out of 100).

### Omnibus confirmation screen (EXP 2)

In EXP 2, after the procedure described above was completed for all the events in the source, the subject was asked to confirm all the elicited values, and given the opportunity to modify responses. The confirmation screen displayed the interval-valued urns elicited for the five events in the source. Moreover, this information was summarized in a graph displaying the minimum number of red and blue balls for each event (see Appendix 2.9). Hovering the mouse over the points on the graph caused the associated interval-valued urn to be highlighted. By clicking on the point on the graph or the urn, the subject could access the corresponding two-cursor scrollbar confirmation screen at the end of the binary-choice procedure for that event, where she could change her choices in exactly the same way as set out above.

### Elicitation of MPs (EXP 1)

The MP of the bet on a given event was elicited through a fairly standard two-step procedure, from which our multiple prior elicitation procedure was inspired (Section 2.2.4). First, a candidate MP was determined through a bisection process (Abdellaoui et al., 2008) that consisted in a series of single pairwise choices between the bet on the event and an urn whose composition was fully known. The

latter realises a standard lottery, i.e. a prospect for which the probability of winning is perfectly known (Section 2.2.1). The displays used were similar to those described above. Then the complete confirmation (one-dimensional, single cursor) scrollbar-based choice list, filled in according to the prior bisection choices, was displayed for verification. The precision of the elicited MP was to the nearest 0.05.

#### 2.3.4. Incentivizing subjects

Participants in all studies were offered a flat payment of  $\text{€}10$ . Additionally, a random incentive system was implemented, which was entirely analogous to those standardly used to implement elicitation of matching probabilities. As noted above, after the presentation of the instructions and before the beginning of the experiment, the subject chose a number from a given range, which identified the individual case of the variable of interest (the day, if the source was minimum temperature; the candidate, if the source was the mark). The exact case identified was specified according to a spreadsheet that would only be revealed at the end of the experiment. This is in concordance with the approach set out by [Johnson et al. \(2021\)](#), who argue that it reduces hedging motivations, given the well-known fact that ambiguity models are indifferent to ex ante hedging. At the end of the experiment, a choice list (a 2D-choice list or MP-choice list in EXP 1; a 2D-choice list in EXP 2) and choice on it were chosen at random by the computer. The subject was then paid according to the decision she had made on that choice. If she had chosen, say, the bet on the event that the minimum temperature in Paris is less than or equal to  $2^{\circ}\text{C}$ , then the day which she chose was revealed, as well as daily temperature data for the November–March period, and she won if the minimum temperature on that day was indeed  $2^{\circ}\text{C}$  or less; if not, she lost. If she had chosen the urn, then she composed the appropriate urn—she counted the specified minimum numbers of red and blue balls, with the remaining balls coming from pre-constructed Ellsberg urns (of unknown composition). Then a ball was drawn from the constructed urn, and she was paid according to whether she bet on the color of that ball or not. All bets yielded  $20\text{€}$  if won, and nothing otherwise.

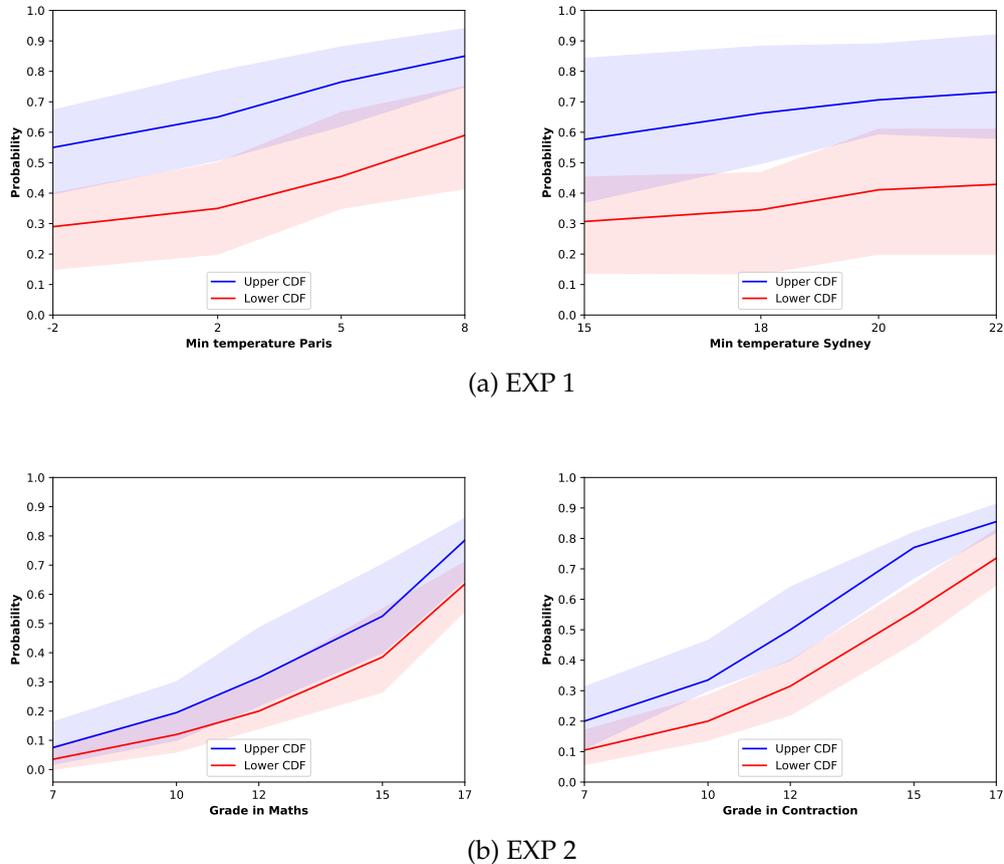


Figure 2.2: Median, 25% and 75% quantile ranges of upper and lower CDFs

## 2.4. Results

### 2.4.1. Descriptive Statistics and Performance

Figure 2.2 plots the median, 25% and 75% quantile upper and lower CDFs for all elicited events and both experiments (see Tables 2.5–2.6 in Appendix 2.8.1 for basic descriptive statistics). This Figure already give some early indications about our results, and the performance of our elicitation method.

First of all, it reports ‘well-behaved’ upper and lower CDFs with probabilities differing across subjects and events—thus suggesting the consistency of the method. The elicited points for both upper and lower CDFs were also consistent

across the successive steps of the elicitation procedure: the binary-choice procedure and the confirmation 2D choice list. This convergence accords with an absence of strategic reasoning. Further, as noted in Section 2.2.6 (see also Appendix 2.7.3), subjects who engage in strategic reasoning when responding to the 2D choice list will be characterised by a large number of responses at the vertices of the space of interval-valued urns, i.e. at the points  $[0, 0]$ ,  $[0, 1]$ ,  $[1, 1]$ . In EXP 1, only one subject (out of 80) reported points in  $[0, 0]$ ,  $[0, 1]$ ,  $[1, 1]$  for more than half of the elicited events<sup>12</sup>, whereas for no subject in EXP 2 were more than half of the elicited points among these extremes (Table 2.9, Appendix 2.8.1).

Moreover, Figure 2.2 suggests that, on aggregate, upper and lower CDFs are increasing, as they should be. Figure 2.3 plots descriptive statistics of the empirical distribution of individual Kendall  $\tau_b$  rank correlation coefficients between the size of events and the upper (resp. lower) probabilities or MPs elicited for each source (see also Table 2.10, Appendix 2.8.1). As is clear from the Figure, the median Kendall  $\tau_b$  is far greater than 0 for all sources—suggesting that the general direction of the upper and lower CDFs is increasing—with a notable difference between EXP 1 and EXP 2. In EXP 2, where subjects were given the opportunity to confirm all their replies on all the 2D choice lists for a source (Section 2.3.3), CDFs were strictly increasing for the vast majority of subjects; in EXP 1, where 2D choice lists were confirmed after consideration of the event and there was no opportunity to reconfirm later, there were more violations of monotonicity. Comparison of the (cognitively less demanding) MPs with upper and lower CDFs in the Paris treatment, whose median Kendall ranks are similar (Figure 2.3a), suggests that such violations were not unique to the elicitation method proposed here. As could have been expected, the frequency of monotonicity violations appears to increase with the difficulty of the choice task, with the MP task being arguably easier than that for probability-interval elicitation, and the task for Paris, the more familiar source for our subjects, being easier than that for Sydney.

<sup>12</sup>He / she reported 3 points out of 4 as  $[0, 1]$ , for both sources.

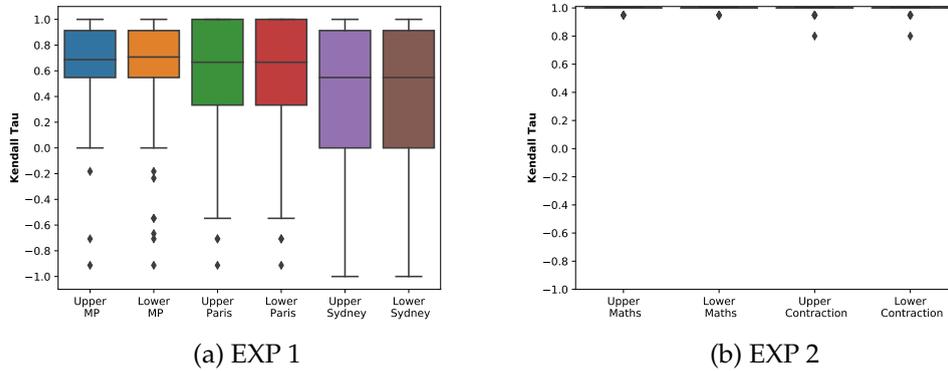


Figure 2.3: Individual-level Kendall  $\tau_b$  for upper and lower CDFs for the sources for each experiment, and for the MPs in EXP 1, Paris treatment. In the case of MPs, the Lower MP is calculated using the MPs of the events  $E_t$  (and should be increasing in  $t$ ), whereas the Kendall  $\tau_b$  for upper MPs are calculated using one minus the MPs of complementary events  $E_t^c$  (which should be increasing with  $t$ ).

*Note:* The Kendall  $\tau_b$  is an indication of ordinal association: the value 1 indicates that the CDFs or MPs are strictly increasing; 0 suggests that there is no association between the elicited probability and the size of the event;  $-1$  indicates strictly decreasing CDFs.

## 2.4.2. Bayesian analysis

We also adopt a standard Bayesian approach, estimating hyperparameters for upper and lower CDFs using a MCMC procedure. We run estimations for each source under the assumption that upper and lower CDFs follow a (truncated) normal distribution, and under a Beta distribution (Table 2.15, Appendix 2.8.2). As shown in Table 2.4, the Beta distribution has the best goodness of fit under both the AIC and BIC criteria for the sources in EXP 1, whereas the truncated Normal distribution performs better according to both criteria for the sources in EXP 2. Henceforth, we present the results under these distributions (the analyses under the other distributions are given in Appendix 2.8.2).

Figure 2.4 plots 1000 MCMC samples for each of the upper and lower distributions, for each source. (Statistics on the distributions of parameters are given in Tables 2.16–2.23, Appendix 2.8.2.) They suggest that the proposed elicitation technique supports a clean Bayesian estimation of subjective probability intervals in the population, insofar as they chime with expectations given the nature of the

		Paris	Sydney	Mathematics	Contraction
AIC	normal	706.65	700.79	411.22	385.52
	beta	648.26	684.36	416.18	390.64
BIC	normal	711.42	705.56	415.12	389.42
	beta	653.02	689.12	420.08	394.54

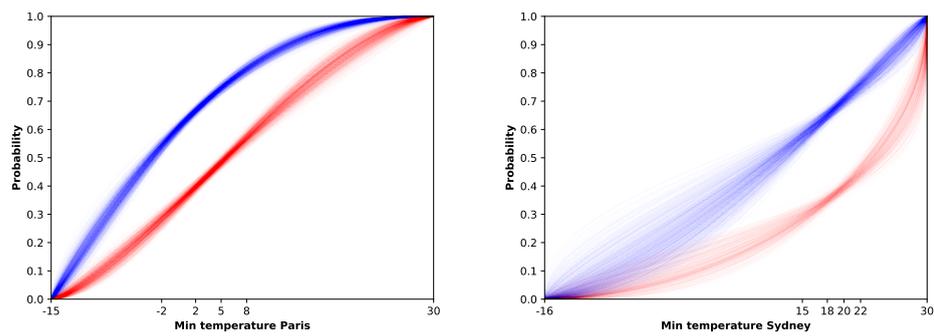
Table 2.4: AIC and BIC under (truncated) normal and Beta specifications for CDFs (Table 2.15).

events. For instance, they suggest that the dispersion of subjective upper and lower probabilities is larger for the temperature source (EXP 1) than the grade source (EXP 2), which could be related to the fact that all subjects in EXP 2 had sat both exams, and were very interested in the marking, several months before. Also, within EXP 1, there is more dispersion in the estimated distributions for Sydney than for Paris, as would be expected given the less familiar nature of the former source, for Paris subjects.<sup>13</sup>

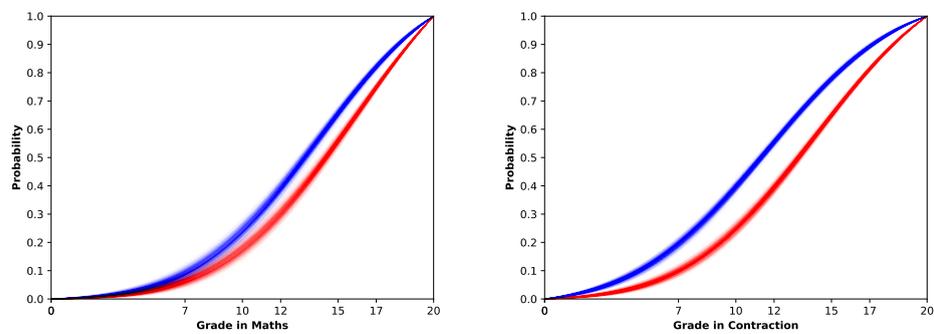
### 2.4.3. Imprecision

Both the graphs of raw data (Figure 2.2) and those emerging from the Bayesian analysis (Figure 2.4) suggest that subjects' beliefs are often *imprecise*: i.e. there is a gap between their upper and lower probabilities. Indeed, two-sided Kolmogorov-Smirnoff tests of the hypothesis that the median upper and lower CDFs are drawn from the same distribution reject the hypothesis for each source ( $p < 0.0001$  in all cases), suggesting a gap between upper and lower CDFs. For further analysis, we define the following index. For an event  $E$  from a given source (e.g. minimum temperature in Paris), we say that a subject's *imprecision concerning  $E$*  is  $\bar{p}(E) - \underline{p}(E)$ , i.e. the width of her (elicited) probability interval for  $E$ . A subject's *Imprecision Index* for a source is defined to be her average imprecision across all

<sup>13</sup>More precisely, it is clear from Tables 2.17 and 2.19 that the standard deviations of the parameters for the Paris source are lower than for Sydney.



(a) EXP 1



(b) EXP 2

Figure 2.4: Bayesian estimation of lower and upper CDFs: plots of 1000 samples from MCMC.

(Beta distribution for EXP 1; Truncated Normal distribution for EXP 2)

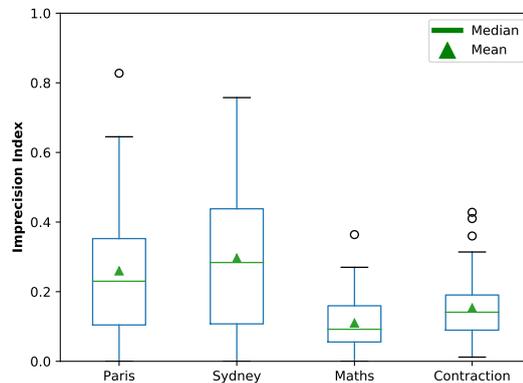


Figure 2.5: Imprecision Index (Eq. (2.8)) across sources in EXP 1 and EXP 2

elicited events in the source:

$$II = \frac{1}{n} \sum_{i=1}^n (\bar{p}(E_n) - \underline{p}(E_n)) \quad (2.8)$$

This clearly gives an indication of how imprecise the subject's beliefs are, on average, for events in the source. Naturally, an SEU decision maker will assign precise probabilities to all events, and hence have an imprecision index of 0 (for all sources).

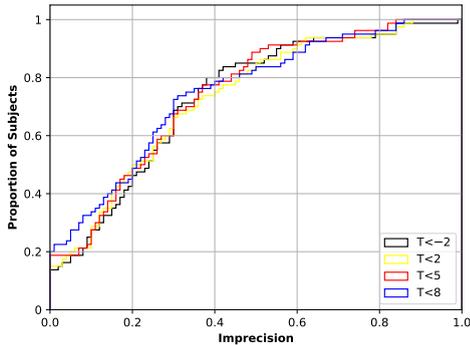
Figure 2.5 displays the mean, median, 25% and 75% quantile, and max and min Imprecision Indices across all sources in both experiments (see also Table 2.11, Appendix 2.8.1). It clearly suggests a tendency towards imprecision, with mean and median Imprecision Indices greater than 0.1 for all sources. Two-sided binomial tests reject the hypothesis of equal probability for the Imprecision Index to be equal to vs. greater than 0 for each source ( $p < 0.0001$  in all cases), with a clear majority of subjects—74 out of 80 in EXP 1, and 49 out of 52 in EXP 2—having strictly positive Imprecision Indices.

The general message of widespread imprecision is confirmed by data on the number of precise events—events for which the subject's elicited upper and lower probabilities coincide (Table 2.12, Appendix 2.8.1). Not more than around 5% of subjects gave precise probabilities for all events in a single source. Only 2 subjects

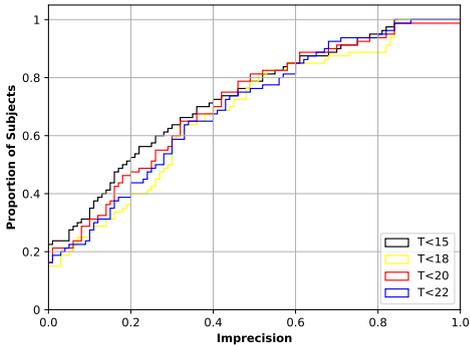
(out of the 132 participating in both experiments) gave precise probabilities for all events elicited. The data in Table 2.12 also allows a check on the extent to which this imprecision could be driven by the binary-choice procedure, insofar as it gives the number of precise events after the binary-choice procedure and before the confirmation 2D choice list, as well as after confirmation. The general finding of few fully precise subjects holds both before and after the confirmation stage. Moreover, relatively few subjects change to fully precise probabilities for all events of the source (at most 3 out of 80, for Sydney in EXP 1), with several fully precise subjects *introducing* imprecision during the confirmation stage, especially in EXP 2.

Delving further, we also investigate imprecision at the event level within sources. Figure 2.6 presents CDFs of the imprecision for each elicited event in each of the experiments and sources, across subjects. One-way ANOVAs of the imprecision (dependent variable) against the event (factor) reject the null hypothesis of identical imprecision across all events for the sources in EXP 2 ( $p < 0.001$  for Maths;  $p = 0.003$  for Contraction), whilst failing to reject it for the sources in EXP 1 (Table 2.13, Appendix 2.8.1). This suggests not only that imprecision is widespread, but that imprecision may be event dependent within sources, as one would expect if some events are intuitively more uncertain than others. For instance, the least imprecise event in EXP 2 involves, for both sources, the lowest grade, where many subjects are presumably more sure of their judgements.

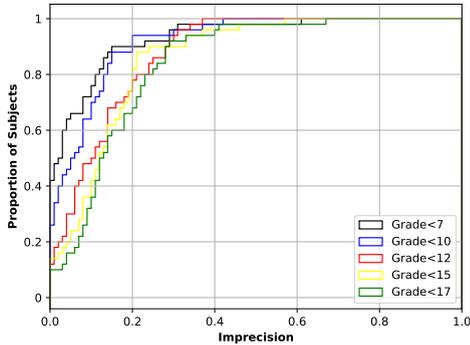
In summary, the development of a method for eliciting multiple priors does not emerge from this analysis as devoid of relevance: rather, it reveals that, when given the possibility to ‘express’ the imprecision implied by non-degenerate probability intervals, many subjects do, at least for the events considered here. Moreover, at least within some sources, the extent of imprecision may depend on the event.



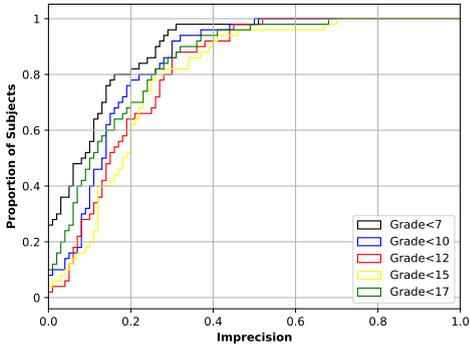
(a) Paris source, EXP 1



(b) Sydney source, EXP 1



(c) Maths source, EXP 2



(d) Contraction source, EXP 2

Figure 2.6: CDFs of Imprecision across subjects, for each elicited event

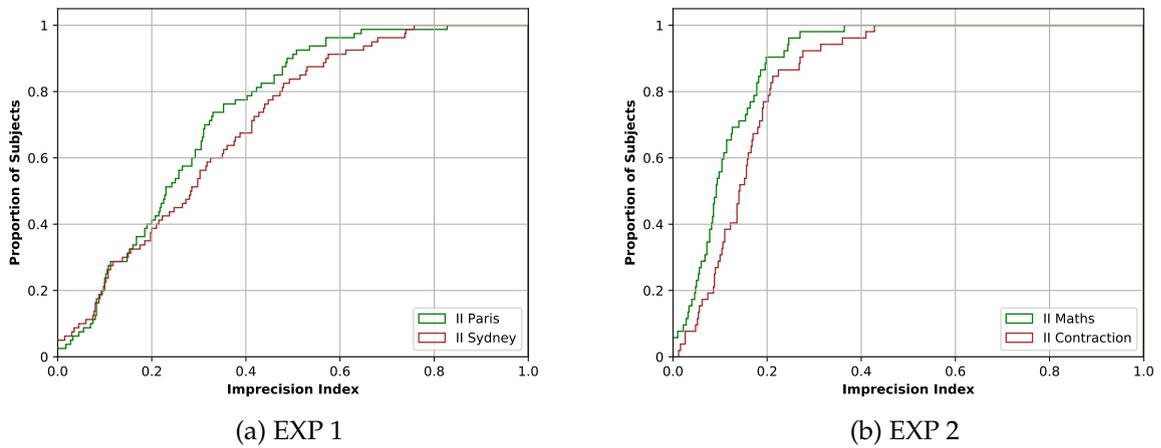


Figure 2.7: CDFs of Imprecision Index (Eq. (2.8))

#### 2.4.4. Imprecision and familiarity

One reasonable hypothesis is that *ceteris paribus* subjects' beliefs are more imprecise concerning events of which they are less familiar, or about which they feel as if they have less knowledge. In terms of multiple priors models, this corresponds to the probability intervals for the events being wider. Since, as explained in Section 2.3, each of our experiments features two sources with which our subjects will typically have different levels of familiarity, or which they naturally consider as having different levels of predictability, a natural conjecture would be that imprecision would be larger for Sydney than Paris, and for the Contraction grade than the Maths one. After all, Paris subjects are less familiar with the weather in Sydney than that in Paris; and the Contraction exam is generally considered to be 'less predictable' than the Maths one (Section 2.3.2).

Figure 2.7 plots the CDFs of the Imprecision Index defined above (Eq. (2.8)) across subjects, for the pair of sources in each experiment. A two-sided paired  $t$ -test barely fails to reject the null hypothesis of identical Imprecision Indices across the sources in EXP 1 ( $p = 0.0895$ ), whilst it rejects it for EXP 2 ( $p = 0.0016$ ). A two-sided Binomial test with null hypothesis that an equal number of subjects have larger Imprecision Index under one source than the other fails to reject the null hypothesis for EXP 1 ( $p = 0.576$ ), but rejects it for EXP 2 ( $p = 0.017$ ). Rerunning

the latter test using the results of the Bayesian analysis—i.e. for each experiment and source, using the posterior distribution over parameters obtained from the Bayesian estimation to sample 1000 tuples of parameters determining the upper and lower distributions, and computing, for each tuple, the Imprecision Index as defined in Eq. (2.8)—yields a rejection of the null hypothesis across the sources for both EXP 1 and EXP 2 ( $p < 0.001$  in both cases). These findings confirm the expected relationship between imprecision and predictability in EXP 2: indeed, in Figure 2.7, the CDF for Contraction—known as the less predictable exam—is entirely to the right of that for Math, indicating a larger Imprecision Index. They also point to a similar relationship between imprecision and familiarity in EXP 1: again, with Figure 2.7 suggesting that CDFs for Paris are generally more precise.

That an expected relationship between imprecision and familiarity or predictability emerges clearly can also be seen as providing further indirect evidence as to the solidity of the proposed elicitation method.

#### 2.4.5. Matching probabilities and the $\alpha$ -maxmin EU mixture coefficient

Recall that EXP 1 contained a supplementary treatment in which the MPs were elicited for the Paris events  $E_{t_i}$  and  $E_{t_i}^c$  for which probability intervals had been elicited (Table 2.2). Henceforth, we denote the MP of an event  $E$  by  $MP(E)$ . Under SEU,  $MP(E_{t_i}) = 1 - MP(E_{t_i}^c) = p(E_{t_i})$ , the subjective probability of  $E_{t_i}$ , for all  $E_{t_i}$ . So, as is well-known, comparing  $MP(E_{t_i})$  and  $1 - MP(E_{t_i}^c)$  provides an indication into the violation of SEU. Under the  $\alpha$ -maxmin EU model (2.1), we have the following equations:

$$MP(E_{t_i}) = \alpha \underline{p}(E_{t_i}) + (1 - \alpha) \bar{p}(E_{t_i}) \quad (2.9)$$

$$1 - MP(E_{t_i}^c) = \alpha \bar{p}(E_{t_i}) + (1 - \alpha) \underline{p}(E_{t_i}) \quad (2.10)$$

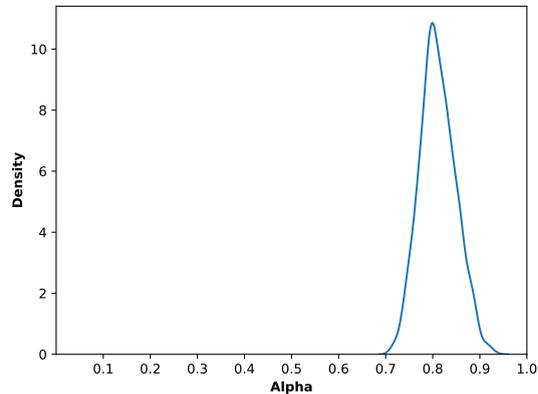


Figure 2.8: PDF of  $\alpha$  from the Bayesian estimation (EXP 1, Paris treatment)

Drawing on the elicited MPs and our elicitation of upper and lower probabilities, the equations (2.9) and (2.10) can be used to elicit the mixture coefficient  $\alpha$  in the Hurwicz  $\alpha$ -maxmin EU model. Under analysis using the raw data, the median  $\alpha$  across subjects is 0.80 (Table 2.25, Appendix 2.8.3). We also perform a Bayesian estimation of the  $\alpha$  in tandem with the lower and upper CDFs, combining equations (2.9) and (2.10) and the MP data with our upper and lower CDF elicitation (see Appendix 2.8.2 and Tables 2.16 and 2.17). Figure 2.8 plots the distribution over  $\alpha$  resulting from this estimation. The Bayesian mean for  $\alpha$  is at 0.81, which is broadly consistent with the finding from the raw data. As discussed at more length in Section 2.5, this is, to our knowledge, the first direct choice-based elicitation of the  $\alpha$  in the  $\alpha$ -maxmin EU model that fully controls for the set of priors by eliciting the relevant information about them without making any assumption about their shape.

## 2.5. Discussion

Our results attest to the feasibility of the proposed elicitation method. It elicits non-degenerate and reasonable upper and lower CDFs. Our elicitation also shows

that imprecision—a gap between upper and lower probabilities—is widespread, with few subjects having precise probabilities for all events. Moreover, they bring out some determinants of imprecision. For some sources, the width of probability intervals may vary according to the event elicited; moreover, average imprecision decreases with the familiarity of the source of uncertainty, as one might expect. Finally, we draw on our probability interval elicitation to elicit the mixture coefficient in the Hurwicz  $\alpha$ -maxmin EU model—the first such elicitation, to our knowledge, to fully control for beliefs.

We now discuss the robustness of our procedure, some related literature, and its potential contributions going forward.

**Robustness** Our approach has been presented in terms of the popular Hurwicz  $\alpha$ -maxmin EU decision model (Section 2.2), which is doubtless the most general decision model in which the ‘belief component’ of the representation is just a set of priors. However, many of the central elements of the approach generalise to extensions building on sets of priors but weakening the linearity of the Hurwicz function form (2.1), to account for probability weighting for instance (see Appendix 2.7.3 for details). First of all, the notion of MPI remains well-defined for all such extensions, and the decision maker’s subjective probability interval is always an MPI. Though MPIs are not guaranteed to be unique for every conceivable extension of this sort, they are essentially unique for a family of reasonable extensions (Appendix 2.7.3). The 2D choice list incentivization mechanism only relies on the weak Lower Stochastic Dominance property of preferences (Definition 1, Section 2.2.6). Apart from the maxmax-EU model (i.e. (2.1) with  $\alpha = 0$ ), which is very rarely found in subjects, this property is satisfied by any reasonable decision model generalising  $\alpha$ -maxmin EU to allow for nonlinear dependence of preferences on upper and lower probabilities (Appendix 2.7.3). In this sense, the 2D choice list incentivization mechanism is widely valid. Finally, whilst the binary-choice procedure relies on the strongest assumption made in Section 2.2—that  $\alpha > \frac{1}{2}$ —there is independent evidence that this holds for most of our subjects (Appendix 2.7.3). As noted in Section 2.2, it is the 2D choice list confirmation

task that counts for incentivizing subjects' choices, the binary-choice procedure playing the role of an aid to completing it.

**Related literature** Our elicitation method relates to existing experimental and theoretical literature on multiple prior models, and the  $\alpha$ -maxmin EU model in particular. Part of this literature is concerned with testing such models, or comparing them to others (e.g. [Hey et al., 2010](#); [Baillon and Bleichrodt, 2015](#)); by contrast, the aim here is to elicit probability intervals in the context of a fairly general multiple prior model. Similarly, there is a literature studying matching probabilities or certainty equivalents of bets on objectively-given probability intervals based on interval-valued urns (e.g. [Baillon et al. \(2012\)](#); [Chew et al. \(2017\)](#)); here, by contrast, we use such urns as elicitation devices for subjective probability intervals.

On the theory side, the challenge of incentive-compatible elicitation of multiple prior beliefs under  $\alpha$ -maxmin EU is related to identification issues with this model, arising from the fact that different pairs of mixture coefficient  $\alpha$  and sets of priors can represent the same preferences over (Savage or Anscombe-Aumann) acts. Proposed approaches to this challenge include pinning down the set of priors using 'unambiguous preferences' ([Ghirardato et al., 2004a](#)), though this has problems in finite state spaces ([Eichberger et al., 2011](#)), or enriching the state space to include an infinite product structure and invoking symmetry axioms ([Klibanoff et al., 2021](#)). Another line of attack concentrates on special cases of the  $\alpha$ -maxmin EU model, notably involving some form of probabilistic sophistication, i.e. the assumption that there are precise probabilistic beliefs which completely determine the contributions of events to preferences ([Machina and Schmeidler, 1992](#); [Chew and Sagi, 2006](#)). Working with a rich state space à la Savage, [Gul and Pesendorfer \(2014, 2015\)](#) obtain a unique identification of  $\alpha$  and the set of priors whenever the latter is generated as the set of extensions of a precise probability measure on a subalgebra of events. ([Grant et al. \(2019\)](#) have extended this approach beyond the assumption of linearity in upper and lower probabilities built into  $\alpha$ -maxmin EU; see Appendix 2.7.) [Chateauneuf et al. \(2007\)](#) obtain

a unique identification of  $\alpha$  and the set of priors whenever the latter is generated from a precise probability measure via  $\varepsilon$ -contamination, i.e. mixture with the set of all probability measures. As stated in the Introduction, we specifically avoid the sort of probabilistic sophistication assumption behind these approaches, motivated by the observation that such assumptions are inadmissible precisely in those situations where multiple prior beliefs are most relevant. Indeed, our procedure takes a different approach, following the theoretical contribution of [Hill \(2019\)](#), who shows that invoking (the equivalent of this paper's) interval lotteries allows full identification of  $\alpha$ -maxmin EU with no need for specific richness assumptions on the state space, probabilistic sophistication, or any other (non-standard) assumptions on the set of priors. As we shall see below, our method provides data to evaluate the sort of probabilistic sophistication assumption behind the aforementioned approaches.

On the experimental front, there is a small literature dealing with incentive-compatible elicitation of multiple priors. One family of approaches purport to elicit multiple priors as the support of second-order beliefs, represented as a measure over the space of probability measures. Beyond the assumption of second-order beliefs, which is foreign to the original multiple prior models ([Gilboa and Schmeidler, 1989b](#); [Bewley, 2002](#); [Ghirardato et al., 2004a](#)), these often make further assumptions about the role of these second-order beliefs in choice. For instance, [Qiu and Weitzel \(2016\)](#) elicit subjects' distributions over the matching probabilities of other participants in the experiment, and purport to deduce subjects' own second-order beliefs from these, relying on the assumption that a subject's opinions about others' matching probabilities coincides with the uncertainty surrounding her own assessment. In a theoretical paper, [Karni \(2020\)](#) develops an ingenious incentive-compatible mechanism for eliciting second-order beliefs and the associated set of priors (as the support), relying on a three-period setup. The mechanism assumes that the subject's second-order beliefs coincide with her beliefs about what she will believe in the interim period. As made clear above, our method relies on no assumptions beyond the  $\alpha$ -maxmin EU model (or appropriate weakenings thereof), and in particular there is no role for second-order beliefs or assumptions on how they relate to other beliefs.

Another family of approaches draws on the theoretical literature discussed above, and in particular on the probabilistically sophisticated special case studied by [Chateauneuf et al. \(2007\)](#), where the subject's set of priors is generated as the  $\varepsilon$ -contamination of a single probability measure with the space of all priors.<sup>14</sup> [Dimmock et al. \(2015\)](#); [Baillon et al. \(2018b,a\)](#) use elicitation of standard MPs (in the case of the last paper, certainty equivalents) to estimate 'ambiguity indices', from which one can back out the mixture coefficient  $\alpha$  and the parameters of the  $\varepsilon$ -contaminated set of priors. As noted previously, multiple prior decision models come to the fore in situations where preferences cannot be reasonably assumed to be generated from precise probabilities, and our elicitation technique was specifically designed to be independent of the assumption of probabilistic sophistication for this reason. Moreover, our data provides empirical insight into the aforementioned probabilistic sophistication assumption. In particular, it implies that, except for events with very high probabilities, the imprecision (in the sense of Section 2.4.1) is the same for all events.<sup>15</sup> As noted in Section 2.4.3 (see also Table 2.13, Appendix 2.8.1), our observations reject this equality for the sources in EXP 2, though not for the sources in EXP 1.<sup>16</sup> This suggests that there are sources for which their method's underlying assumption is violated. That said, it can be viable on some sources; indeed, our data indicate that the Paris source in EXP 1 may be one such source. And in fact, we can estimate the ambiguity indices used in the aforementioned papers on the basis of the data from our study (EXP 1, Paris treatment) under their assumption about the set of priors,<sup>17</sup> and find, for instance, that they yield the value 0.82 for the mixture coefficient  $\alpha$ , which, reassuringly, is

<sup>14</sup>Formally, the assumption is that the set of priors  $\mathcal{C} = \{(1 - \varepsilon)p + \varepsilon\Delta\}$ , where  $\Delta$  is the space of all probability measures,  $p$  is an element of  $\Delta$  and  $\varepsilon \in [0, 1]$ .

<sup>15</sup>If the set of priors is as defined in footnote 14, then, for any  $E$  such that  $(1 - \varepsilon)p(E) \in [0, 1 - \varepsilon]$ , the probability interval for event  $E$  is  $[(1 - \varepsilon)p(E), (1 - \varepsilon)p(E) + \varepsilon]$ , and hence the event has imprecision  $\varepsilon$ .

<sup>16</sup>The  $\varepsilon$ -contamination assumption is consistent with the imprecision being smaller for high-probability events, which in our setup means  $E_t$  for large  $t$ . It is thus inconsistent with the fact, clear in Figure 2.6, that the imprecision is generally smallest for  $E_7$  (the lowest-probability event in EXP 2), and in the case of the Maths source, highest for  $E_{17}$  (the highest probability event).

<sup>17</sup>Specifically, [Baillon et al. \(2018b\)](#) propose the average of  $1 - MP(E) - MP(E^c)$  over a selection of events as their measure of the 'ambiguity aversion index'  $b$ . The average for the events elicited here can be deduced directly from Table 2.26 (Appendix 2.8.4), as around 0.16. On the other hand, under (2.1) with the specified form for the set of priors (see footnote 14), their 'a-insensitivity index'  $a = \varepsilon$ . Under such sets of priors, as noted in footnote 15, every  $E$  with  $(1 - \varepsilon)p(E) \in [0, 1 - \varepsilon]$  has imprecision  $\varepsilon$ . The average imprecision, as measured by the Imprecision Index (Table 2.11), thus gives an estimate of their  $a$ : it is around 0.25. The mixture coefficient  $\alpha$  is related to these indices by  $\alpha = \frac{1}{2} \left( \frac{b}{a} + 1 \right)$  ([Baillon et al., 2018b,a](#)), yielding the value in the text.

close to the Bayesian and raw estimates reported in Section 2.4.5. So not only is our elicitation method more robust, insofar as it applies in situations where the assumptions underlying their approach do not hold, it can evaluate precisely in which cases they do hold; in those cases, their approach, implemented on our data, gives the same result as our ‘direct’ elicitation.

Going beyond the lab, there is a large and growing literature on elicitation of multiple priors or imprecise probabilities in a range of disciplines, from economics to climate science. All such elicitation exercises of which we are aware use stated probability intervals, and as such are not incentive compatible. For instance, [Giustinelli et al. \(2021\)](#) elicit beliefs on dementia and long-term care decisions in a large-scale representative survey (over 1000 subjects), allowing stated probabilities to be interval-valued. Consistently with our results (Section 2.4.3), they find widespread imprecision. They argue forcefully for the importance of probability-interval elicitation for reducing survey bias and understanding attitudes to and behavior in the face of high-uncertainty events, such as whether one will develop dementia and whether one should insure against it. In another approach, in another domain, [Kriegler et al. \(2009\)](#) elicit beliefs of selected scientists (around 50 subjects) concerning climate tipping points, allowing participants to state probability intervals for these (notoriously uncertain) events. Such expert elicitations, which involve often time-consuming and individualised sessions with selected experts, have emerged as a central tool for managing complex uncertainties ([Morgan, 2014](#)). Though they have traditionally aimed at eliciting precise probabilities, [Kriegler et al. \(2009\)](#) shows that imprecision is widespread for some events, and hence once again argue for the relevance of probability-interval elicitation.

**Future Directions** Two leitmotifs emerge from the literature review. On the one hand, our results are consistent with existing studies suggesting that imprecision is widespread, especially for certain events. However, based as they are on an incentive-compatible, choice-based and theoretically robust elicitation method, our results are less open to criticisms of existing studies pointing to a lack

of incentive compatibility or the reliance on a specific model. On the other hand, as we saw on the  $\varepsilon$ -contamination example, our method can be used to evaluate the assumptions behind—and hence the effectiveness of—existing methods.

The latter point suggests one direction for future research. As noted, stated probability intervals are typically used in large-scale surveys (such as [Giustinelli et al. 2021](#)), but how close, or far, are subjects' stated probability intervals from their actual multiple prior beliefs? This is a question that our method can be used to answer, by eliciting upper and lower probabilities with it on some subjects, and comparing these with their stated intervals. As such, our method can be used to corroborate, refine, correct and choose between existing, less well-founded approaches.

Moreover, the main aim of this paper is to demonstrate the possibility of choice-based incentive-compatible probability-interval elicitation with the general methodology proposed. Another future direction is to operationalize simpler, parametrized versions of the method, involving fewer choice questions. These could prove more suitable for elicitation in large field studies.

Finally, at the other end of the spectrum are expert elicitation exercises of the sort cited above. In general, fewer subjects are involved, with each spending more time; the flip side is that more precision is desired of the elicitation at the individual level. Here our longer EXP 2 suggests that our method can provide the appropriate individual-level probability-interval elicitation, whilst being incentive compatible and theoretically well-founded. Note that expert elicitation exercises often deal with continuous variables, and hence aim to elicit CDFs ([Colson and Cooke, 2018](#)); our implementation on continuous sources provides a proof of concept of the method on precisely such cases.

## 2.6. Conclusion

This paper proposes and implements a solution to the open problem of choice-based incentive-compatible elicitation of multiple prior beliefs. It comprises of a new preference-based notion—Matching Probability Intervals—and probability-interval analogues of standard choice lists and bisection elicitation procedures. Theoretically, it operates in the context of the Hurwicz  $\alpha$ -maxmin EU model and in the absence of strong assumptions about subjects' sets of priors, most notably any form of probabilistic sophistication.

Our implementation of the elicitation method, in two experiments to elicit subjective upper and lower CDFs over continuous-valued sources of uncertainty, testifies to its feasibility. It finds a predominance of imprecision—a gap between upper and lower probabilities—across our subjects, for all explored sources, showing it to be related to familiarity or predictability. It also allows us to perform what, to our knowledge, is the first elicitation of the mixture coefficient in the  $\alpha$ -maxmin EU model that fully controls for beliefs.

## 2.7. Theoretical Appendix

In order to bring out the robustness of our proposal, and the assumptions underlying it, we shall at times work with a more general decision model than Hurwicz  $\alpha$ -maxmin EU (2.1). Consider the representation where a bet  $(z, E, 0)$  is evaluated according to:

$$W(\underline{p}(E), \bar{p}(E)) \cdot u(z) \quad (2.11)$$

where  $\underline{p}(E), \bar{p}(E)$  and  $u$  are as in Section 2.2.3, and  $W$  is an ‘aggregation function’, which is continuous in both coordinates and normalised— $W(x, x) = x$  for all  $x$ .  $\alpha$ -maxmin EU is the special case where  $W$  is linear:  $W(x, y) = \alpha x + (1 - \alpha)y$ . See Grant et al. (2019) for an axiomatisation of a special case of (2.11) where the set of priors is generated by a probability measure on a subalgebra, and a thorough discussion of its potential. As in Section 2.2, we assume the same representation for imprecise risky prospects.

Note that, unlike  $\alpha$ -maxmin EU, the general form (2.11) can accommodate non-linear, Prospect-Theory-style weighting of the lower and upper probabilities, for instance taking  $W(x, y) = \alpha w(x) + (1 - \alpha)w(y)$ , where  $w$  is a Prospect-Theory-style weighting function.

### 2.7.1. Proofs

We prove Proposition 1 under representation (2.11) with  $W$  continuous, normalised and strictly increasing in the first coordinate. As noted above,  $\alpha$ -maxmin EU model with  $\alpha > 0$  is a special case.

**Proposition 3.** *For any decision maker represented according to (2.1) with  $\alpha \neq \frac{1}{2}$ , and for any event  $E$ , there is a unique MPI for  $E$ .*

*Proof.* Existence is immediate from Eqs. (2.3) and (2.4). Uniqueness is immediate from the linearity of the indifference curves in  $\mathcal{I}$ -space (see Figure 2.1).  $\square$

*Proof of Proposition 1.* Under (2.11), it follows from the first preference pattern in Proposition 1 that  $W(q, 1 - b) > W(\underline{p}(E), \bar{p}(E))$  for all  $q > r$ , and similarly for the others. By the continuity of  $W$ , it thus follows from the first two preferences that  $W(r, 1 - b) = W(\underline{p}(E), \bar{p}(E))$ , and from the second pair of preferences that  $W(b, 1 - r) = W(1 - \bar{p}(E), 1 - \underline{p}(E))$ . It thus follows that  $(z, [r, 1 - b], 0) \sim (z, E, 0)$  and  $(0, [r, 1 - b], z) \sim (0, E, z)$ , so  $[r, 1 - b]$  is a MPI for  $E$ , as required.  $\square$

*Proof of Proposition 2.* Part a. Plugging in the representations (2.1) and (2.2), any  $[x, y]$  in the R-B region satisfies:

$$\begin{aligned} \alpha x + (1 - \alpha)y &\geq \alpha \underline{p}(E) + (1 - \alpha)\bar{p}(E) \\ \alpha(1 - y) + (1 - \alpha)(1 - x) &\geq \alpha(1 - \bar{p}(E)) + (1 - \alpha)(1 - \underline{p}(E)) \end{aligned}$$

By basic algebra (add  $\alpha$  times the first inequality to  $(1 - \alpha)$  times the second), one obtains  $(\alpha^2 - (1 - \alpha)^2)x \geq (\alpha^2 - (1 - \alpha)^2)\underline{p}(E)$ , whence it follows, since  $\alpha > \frac{1}{2}$ , that  $x \geq \underline{p}(E)$ . Similarly, one obtains  $((1 - \alpha)^2 - \alpha^2)y \geq ((1 - \alpha)^2 - \alpha^2)\bar{p}(E)$ , whence, since  $\alpha > \frac{1}{2}$ ,  $y \leq \bar{p}(E)$ . A similar argument establishes the result for points in W.

Part b. follows directly from the fact that, under 2.2, whenever  $x \leq x'$  and  $y \leq y'$ , then  $(z, [x, y], 0) \preceq (z, [x', y'], 0)$  and  $(0, [x, y], z) \succeq (0, [x', y'], z)$ .  $\square$

We state for completeness the result on the uniqueness of the MPI.

## 2.7.2. Binary-choice procedure

### Introduction and setup

Our binary-choice procedure is fully described in Figures 2.10–2.13. Figure 2.10 sets out the general structure (and stopping rules). At each step of the procedure,

preferences are elicited for a single probability interval  $[\underline{p}_i, \bar{p}_i]$ : i.e. preferences between the bet on the event and the IL  $(z, [\underline{p}_i, \bar{p}_i], 0)$ , and between the bet on the complement event and the complementary IL  $(0, [\underline{p}_i, \bar{p}_i], z)$ . The heart of the procedure, detailed in Figures 2.11–2.13, involves specification of the next probability interval proposed for elicitation on the basis of the preferences concerning the previous intervals. We first set out the notation used in the presentation of these parts of the procedure, before explaining informally its main steps.

The procedure draws on two formal elements. The first is the assignment of partially known urns—or equivalently probability intervals—to preference-defined regions, discussed in Section 2.2.7. Recall from Section 2.2.1 that a partially known urn  $[p, q]$ , i.e. with a minimum proportion  $p$  of red balls and a minimum proportion  $1 - q$  of blue balls, corresponds to a probability interval; we shall present the procedure in terms of the latter here. For every event  $E_i$  and urn  $[p, q]$ , the preferences in the choices between the bet on  $E_i$  and that on a red ball being drawn from the urn, and between the bet on  $E_i^c$  and that on blue from the urn suffice to situate  $[p, q]$  in one of the four regions,  $R - B, W, R, B$  defined in Table 2.1 (Section 2.2.7). For instance, in Figure 2.9, which we shall use to illustrate the procedure, the probability intervals already elicited are the dots coloured white, red, blue and red-blue according to the (preference-based) region they belong to.

The second element is a ‘polar’-style coordinate system for the set of probability intervals  $\mathcal{I}$ , under which, informally,  $(m, \alpha) \in [0, 0.5] \times [0, 1]$  is the probability interval that is  $\alpha$  along the piecewise-linear line that goes through the probability intervals  $[0, 0]$ ,  $[1, 1]$ , and  $[m, 1 - m]$  (corresponding to the urn with at least proportion  $m$  of red balls and at least proportion  $m$  of blue balls). The thick grey line in Figure 2.9 is one such line. Formally,  $\sigma : \mathcal{I} \rightarrow [0, 0.5] \times [0, 1]$  is defined by:

$$\sigma([p, q]) = \begin{cases} \left( \frac{p}{p+q}, \frac{p+q}{2} \right) & p \leq 1 - q, p + q \in (0, 2) \\ \left( \frac{1-q}{2-p-q}, \frac{p+q}{2} \right) & p > 1 - q, p + q \in (0, 2) \\ (0, 0) & p = q = 0 \\ (0, 1) & p = q = 1 \end{cases} \quad (2.12)$$

It is straightforward to check that  $\sigma$  is a well-defined function on  $\mathcal{I}$ . Every point except for  $[0, 0], [1, 1]$  corresponds to a unique line (parametrised by  $m$ ) and ‘distance’ along that line (parametrised by  $\alpha$ ).  $[0, 0]$  (respectively  $[1, 1]$ ) corresponds to a single  $\alpha$ , namely 0 (resp. 1), though it lies on all such lines; we set the corresponding  $m = 0$  by convention. For information, the inverse map is given by:

$$\sigma^{-1}(m, \alpha) = \begin{cases} [2\alpha m, 2\alpha(1 - m)] & \alpha \leq \frac{1}{2} \\ [(2 - 2\alpha)m + (2\alpha - 1), (2 - 2\alpha)(1 - m) + (2\alpha - 1)] & \alpha > \frac{1}{2} \end{cases} \quad (2.13)$$

We write  $\sigma_1([p, q])$  (respectively  $\sigma_2([p, q])$ ) for the first (resp. second coordinate) of  $\sigma([p, q])$ . Since this is a simple change of coordinates, we shall write  $(m, \alpha) \in B$  as short for  $\sigma^{-1}(m, \alpha) \in B$ , and similarly for other cases.

## Presentation of main steps

As discussed in Section 2.2.7 (Proposition 2), elicited points in the R-B and W regions determine an area in  $\mathcal{I}$  ‘between the R-B and the W points’ to which the MPI must belong. The general aim of the procedure is thus to find progressively ‘closer’ points in R-B and W, hence reducing the size of this area. This motivates the two main steps in the determination of the next probability interval to be presented for elicitation,  $[p_{i+1}, q_{i+1}]$ , on the basis of the previously elicited point  $[p_i, q_i]$ .

On the one hand, if  $[p_i, q_i]$  is in the R-B region (respectively, the W region), then by Proposition 2 a. (Section 2.2.7), the MPI will be North-West of  $[p_i, q_i]$  (resp. South-East of  $[p_i, q_i]$ ) in Figure 2.1—i.e.  $\underline{p} \leq p_i$  and  $\bar{p} \geq q_i$  (resp.  $\underline{p} \geq p_i$  and  $\bar{p} \leq q_i$ ), where the MPI is  $[\underline{p}, \bar{p}]$ . In such cases, the procedure proposes a  $[p_{i+1}, q_{i+1}]$  North-West (resp. South-East) of  $[p_i, q_i]$ . This exemplified by the  $[p_{i+1}, q_{i+1}]$  proposed for point X in Figure 2.9. The precise proposal for  $[p_{i+1}, q_{i+1}]$  depends on whether there is a point in W (resp. R-B); technicalities aside, this is the general strategy of

the cases in lines 20-23 and 36-39 of the procedure (Figures 2.12-2.13). If the point  $[p_{i+1}, q_{i+1}]$  turns out to be in R-B or W, this will further restrict the area where the MPI can lie.

On the other hand, if  $[p_i, q_i]$  is in the R or W regions, then Proposition 2 a. does not apply; as discussed in Section 2.2.7, the aim in such cases is to find a point in the R-B or W regions, to continue reducing the area containing the MPI. The procedure draws on two observations; first, as mentioned above, any point  $[p_i, q_i]$  can be equivalently written in another coordinate system, specifying the line it sits on—parametrised by  $m = \sigma_1([p_i, q_i])$ —and how ‘far’ along the line it is—parametrised by  $\alpha = \sigma_2([p_i, q_i])$ . Second, for  $[p_i, q_i]$  in R (respectively W), by Proposition 2 b., all points North-East (resp. South-West) of  $[p_i, q_i]$  are also in R (resp. W). So the only points in R-B and W on the line  $m = \sigma_1([p_i, q_i])$  corresponding to the point  $[p_i, q_i]$  must be South-West of  $[p_i, q_i]$ , i.e. with lower  $\alpha$  (resp. North-West, i.e. with higher  $\alpha$ ). Accordingly, the procedure proposes a point  $[p_{i+1}, q_{i+1}]$  on the line  $m = \sigma_1([p_i, q_i])$  but shifted in the appropriate direction, as illustrated by the  $[p_{i+1}, q_{i+1}]$  proposed for point Y (lying in the R region) in Figure 2.9. Technicalities aside, this is general strategy for Case 1 (lines 1-17) and the cases in lines 24-34 and lines 40-44 of the procedure (Figures 2.11–2.13). Among these cases, all retain the same  $m$  (grey line in Figure 2.9) except those considered in lines 12-17. These treat cases where no point in R-B or W has yet been found; the procedure in these cases increases  $m$  during the search, hence looking closer to the diagonal (ie. the line of  $[p, q]$  with  $p = q$ ). We used a procedure with this in-built precision bias to favour Bayesian replies (i.e. precise probabilities); in the light of it, our finding of widespread imprecision (Section 2.4.3) is all the more remarkable.

## Convergence

Except for extreme cases, the procedure tends to the MPI.<sup>18</sup>

**Proposition 4.** *Suppose preferences are represented according to (2.1) with  $1 > \alpha > \frac{1}{2}$ ,*

<sup>18</sup>In this result, we adopt the Euclidean topology on  $\mathcal{I} \subseteq \mathbb{R}^2$ .

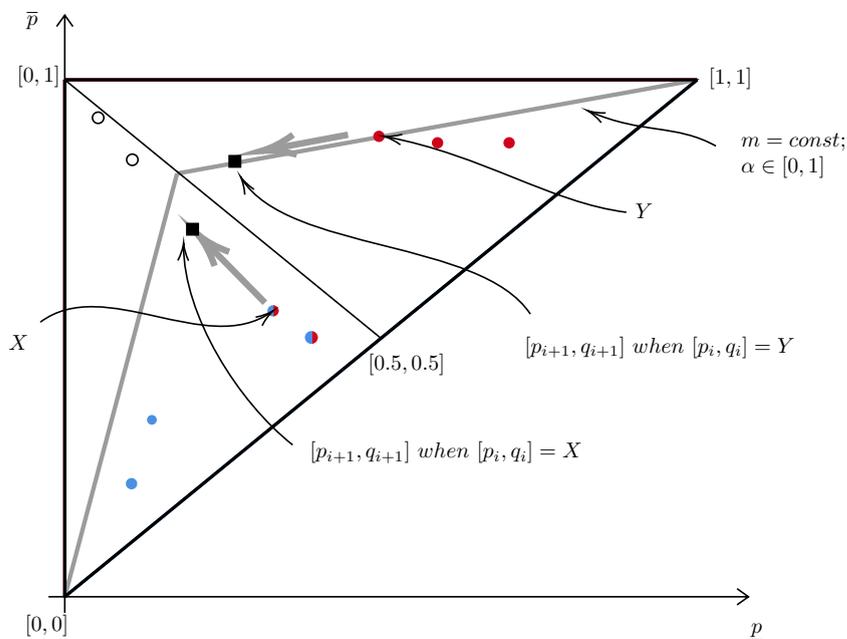


Figure 2.9: Binary Choice Procedure.

let  $E$  be an event, and let  $[\underline{p}_n, \bar{p}_n]$  be the result of the procedure in Figures 2.11–2.13 (with initial values set as in Figure 2.10) applied for  $n$  steps. Then  $[\underline{p}_n, \bar{p}_n] \rightarrow [\underline{p}(E), \bar{p}(E)]$  as  $n \rightarrow \infty$ . Moreover, the procedure also converges in this sense when preferences are represented according to (2.1) with  $\alpha = 1$ ,  $\underline{p}(E) \neq 0$  and  $\bar{p}(E) \neq 1$ .

*Proof.* We provide the main steps of the proof here; they rely on technical Lemmas 1–4, which are detailed in Appendix 2.10. We adopt the notation and initial values from Figure 2.10; in particular, let  $El_n$  be the set of elicited points after  $n$  steps. As discussed in Section 2.2.5, the MPI is  $[\underline{p}(E), \bar{p}(E)]$ . Moreover, by Proposition 2, at stage  $n$ , the MPI is contained in

$$\Phi_n = \left\{ [p, q] \in \mathcal{I} : \begin{array}{l} \max \{p' : [p', q'] \in El_n \cap W\} \leq p \leq \min \{p'' : [p'', q''] \in El_n \cap R - B\}, \\ \max \{q'' : [p'', q''] \in El_n \cap R - B\} \leq q \leq \min \{q' : [p', q'] \in El_n \cap W\} \end{array} \right\} \quad (2.14)$$

where the maximum of an empty set is taken to be 0 and the minimum 1.

We reason referring to the cases in the procedure (Figures 2.11–2.13). At the beginning of the procedure, it is in Case 1 ( $El_0 \cap W = El_0 \cap R - B = \emptyset$ ). By lines 13-16, if no point in  $W$  or  $R-B$  is found, the points elicited by the procedure will reach the space of precise probabilities (i.e. points  $[p, q]$  with  $p = q$ ), where

it will follow a standard bisection procedure. All such points have  $\sigma_1$ -value of 0.5. It follows from Lemma 1 that if the MPI is not precise, then a point will be found in R-B, so the procedure moves to Case 2. On the other hand, if the MPI is precise, then, by Lemma 1 and the bisection character of the procedure on the space of precise probabilities, the points elicited in the procedure will converge to it as required.

Now consider cases where the procedure arrives to Case 2 or 3, i.e. it finds a point in R-B or W. By Lemma 3,  $\sigma_1([p_n, q_n]) \rightarrow \sigma_1([\underline{p}(E), \bar{p}(E)])$  as  $n \rightarrow \infty$ . We distinguish three cases.

- $\sigma_1([\underline{p}(E), \bar{p}(E)]) > 0$  and  $\sigma_1([p_n, q_n]) \neq \sigma_1([\underline{p}(E), \bar{p}(E)])$  for all  $n$ . By Proposition 2 and the definition of  $\sigma$  (and in particular the slopes of the lines  $\sigma_1([p, q]) = m$  for  $m > 0$ ), it follows straightforwardly that  $\min_{[p, q] \in El_n} d([\underline{p}(E), \bar{p}(E)], [p, q]) \rightarrow 0$  as  $n \rightarrow \infty$ , where  $d$  is the Euclidean distance on  $\mathcal{I} \subseteq \mathbb{R}^2$ , whence  $[p_n, q_n] \rightarrow [\underline{p}(E), \bar{p}(E)]$  as required.
- $\sigma_1([\underline{p}(E), \bar{p}(E)]) > 0$  and  $\sigma_1([p_i, q_i]) = \sigma_1([\underline{p}(E), \bar{p}(E)])$  for some  $i$ . By Lemma 1 and Case 2 (lines 24-33) and Case 3 (lines 40-43), the procedure will, from  $i$  onwards, only pass through points with same  $\sigma_1$ -value  $\sigma_1([\underline{p}(E), \bar{p}(E)])$ , where it will only find points in R and B. Moreover, it follows a bisection-style procedure on the line  $\sigma_1([p, q]) = \sigma_1([\underline{p}(E), \bar{p}(E)])$ . It follows from standard arguments, Lemma 1 and representation (2.1) that this procedure converges to  $[\underline{p}(E), \bar{p}(E)]$  as required.
- $\sigma_1([\underline{p}(E), \bar{p}(E)]) = 0$  and  $\alpha < 1$  in the representation (2.1). Suppose  $\underline{p}(E) = 0$ ; the other case ( $\underline{p}(E) \neq 0$  and so  $\bar{p}(E) = 1$ ) is treated similarly. By Lemma 1,  $[p_n, q_n]$  contains a subsequence of points in R-B, with  $\sigma_1$ -value tending to 0. Since  $\alpha < 1$ , by representation (2.1), for every  $q < \bar{p}(E)$ , there exists  $p > 0$  such that  $(z, [p, q], 0) \prec (z, E, 0)$ , and hence such that  $[p, q]$  is not in R-B. Moreover, by the representation and Lower Stochastic Dominance, for every  $q > \bar{p}(E)$  and  $p$ ,  $(0, [p, q], z) \prec (0, [\underline{p}(E), \bar{p}(E)], z) \sim (0, E, z)$ , so such  $[p, q]$  are not in R-B. It follows that the subsequence of  $[p_n, q_n]$  consisting of points in R-B converges to  $[\underline{p}(E), \bar{p}(E)]$ , so  $[p_n, q_n] \rightarrow [\underline{p}(E), \bar{p}(E)]$  as required.

□

### 2.7.3. Robustness of the elicitation method

As stated in Section 2.2, the proposed elicitation method has three (novel) elements. The first is the notion of MPI, and the observation that they yield the probability intervals generated by the subjects' set of priors. The second is the incentivisation mechanism, based on the 2D choice list set out in Section 2.2.6. As for elicitation of subjective probabilities (e.g. choice-list methods for eliciting MPs), this is already sufficient to provide an elicitation mechanism for subjects' sets of priors. However, the proposal also includes a chained binary-choice procedure, in the style of the 'bisection' or 'staircase' method for MPs or certainty equivalents, to aid the subject find the MPI. We now discuss to what extent the proposed elements apply beyond the typical  $\alpha$ -maxmin EU representation with  $\alpha > \frac{1}{2}$  on which we have focused in Section 2.2. We also analyse the consequences of violation of the isolation assumption (Section 2.2.6) for the 2D choice list incentivisation mechanism.

### Matching Probability Intervals

Under the general preferences of the form (2.11), the equations (2.5) for the MPI can be rewritten in the obvious way.<sup>19</sup> Clearly, the notion of MPI is well defined, and the subjective probability interval is an MPI. The form of  $W$  can however affect the uniqueness of the MPI. More precisely, it is guaranteed to be unique whenever there is a unique solution to the equations, and this only occurs if  $W$  satisfies the following 'single-crossing property': every pair of red-and-blue in-

---

<sup>19</sup>Explicitly:

$$W(\underline{p}, \bar{p}) = W(\underline{p}(E), \bar{p}(E)) \quad (2.15)$$

$$W(1 - \bar{p}, 1 - \underline{p}) = W(1 - \bar{p}(E), 1 - \underline{p}(E)) \quad (2.16)$$

---

**Procedure** Binary Choice Procedure: structure and stopping rules

---

```

1 Set  $[p_1, q_1] = [0.3, 0.7]$ ,  $i = 1$  and  $El = \emptyset$ ; /*  $[p_i, q_i]$  is the last
   interval for which preferences were elicited, and  $El \not\ni [p_i, q_i]$ 
   is the set of intervals for which preferences have been
   elicited previously. So  $El \cup \{[p_i, q_i]\}$  is the set of all
   elicited intervals, including the one under consideration.
   */
2 Set  $rb_{El} = \arg \max_{[p,q] \in El \cap RB} q$ ,  $w_{El} = \arg \min_{[p,q] \in El \cap W} q$ ,
    $r_{El} = \arg \min_{[p,q] \in El \cap R} q$ ,  $b_{El} = \arg \max_{[p,q] \in El \cap B} q$ ; /*  $rb_{El}$  is the
   highest elicited point in R-B,  $w_{El}$  is the lowest elicited
   point in W, etc. */
3 while  $|El| < 12$  do
4   while  $d(rb_{El}, w_{El}) \geq 0.15$  do
5     Elicit preferences for  $[p_i, q_i]$ ;
6     Execute algorithm 1;
7     Add  $[p_i, q_i]$  to  $El$ ;
8     if  $d(rb_{El}, w_{El}) < 0.15$  then
9       return  $[p_i, q_i]$ ;
10    Stop
11 if  $El \cap W = \emptyset$  and  $El \cap R - B = \emptyset$  then
12   return  $\frac{b_{El} + r_{El}}{2}$ 
13 if  $El \cap R - B \neq \emptyset$  and  $El \cap W = \emptyset$  then
14   return  $rb_{El}$ 
15 if  $El \cap W \neq \emptyset$  and  $El \cap R - B = \emptyset$  then
16   return  $w_{El}$ 
17 if  $El \cap R - B \neq \emptyset$  and  $El \cap W \neq \emptyset$  then
18   return  $\frac{rb_{El} + w_{El}}{2}$ 

```

---

Figure 2.10: Binary choice procedure: structure

Notation:  $d(\bullet, \bullet)$  is the Euclidean distance in  $\mathcal{I} \subset \mathbb{R}^2$ .

---

**Algorithm 1: Determination of Next Binary Choice**

---

```

1 Case 1
2 if  $(El \cup \{[p_i, q_i]\}) \cap R - B = (El \cup \{[p_i, q_i]\}) \cap W = \emptyset$  then
3   if there is no  $[p, q] \in El$  with  $\sigma_1([p, q]) = m$ , where  $m = \sigma_1([p_i, q_i])$  then
4     if  $[p_i, q_i] \in R$  then
5        $[p_{i+1}, q_{i+1}] = \sigma^{-1}(m, \max\{\alpha_0 | (m_0, \alpha_0) \in El \cap B\});$ 
6     if  $[p_i, q_i] \in B$  then
7        $[p_{i+1}, q_{i+1}] = \sigma^{-1}(m, \min\{\alpha_0 | (m_0, \alpha_0) \in El \cap R\});$ 
8   if there exists  $[p, q] \in El$  with  $\sigma_1([p, q]) = m$ , where  $m = \sigma_1([p_i, q_i])$ , but there
     exists no  $[p, q] \in El$  with  $[p, q] \in B$  and  $\sigma_1([p, q]) = m$  then
9      $[p_{i+1}, q_{i+1}] = \sigma^{-1}(m, \frac{\sigma_2([p_i, q_i])}{2});$ 
10  if there exists  $[p, q] \in El$  with  $\sigma_1([p, q]) = m$ , where  $m = \sigma_1([p_i, q_i])$ , but there
     exists no  $[p, q] \in El$  with  $[p, q] \in R$  and  $\sigma_1([p, q]) = m$  then
11   $[p_{i+1}, q_{i+1}] = \sigma^{-1}(m, \frac{1+\sigma_2([p_i, q_i])}{2});$ 
12  if there exists  $[p, q], [p', q'] \in El$  with  $[p, q] \in B, [p', q'] \in R$  and
      $\sigma_1([p, q]) = \sigma_1([p', q']) = \sigma_1([p_i, q_i]) = m$  then
13    if  $[p_i, q_i] \in R$  then
14      
$$[p_{i+1}, q_{i+1}] = \begin{cases} ((m + 0.5)/2, (\sigma_2([p_i, q_i] + \alpha')/2) & m \leq 0.44 \\ (0.5, (\sigma_2([p_i, q_i] + \alpha')/2) & m > 0.44 \end{cases}$$

      where  $\alpha' = \max\{\alpha_0 | (m, \alpha_0) \in El \cap B\};$ 
15    if  $[p_i, q_i] \in B$  then
16      
$$[p_{i+1}, q_{i+1}] = \begin{cases} ((m + 0.5)/2, (\sigma_2([p_i, q_i] + \alpha'')/2) & m \leq 0.44 \\ (0.5, (\sigma_2([p_i, q_i] + \alpha'')/2) & m > 0.44 \end{cases}$$

      where  $\alpha'' = \min\{\alpha_0 | (m, \alpha_0) \in R\};$ 
17  end

```

---

Figure 2.11: Determination of Next Binary Choice: Part 1  
Notation:  $\sigma$  defined in (2.12) and (2.13).

18 **Case 2**19 **if either  $(El \cup \{[p_i, q_i]\}) \cap R - B \neq \emptyset$  or  $El \cup \{[p_i, q_i]\} \cap W \neq \emptyset$ , but not both then**20 **if  $[p_i, q_i] \in R - B$  then**

21

$$[p_{i+1}, q_{i+1}] = \begin{cases} \left[ \frac{p_i + (p_i + q_i - 1)}{2}, \frac{q_i + 1}{2} \right] & p_i + q_i > 1 \\ \left[ \frac{p_i}{2}, \frac{q_i + (p_i + q_i)}{2} \right] & p_i + q_i \leq 1 \end{cases}$$

22 **if  $[p_i, q_i] \in W$  then**

23

$$[p_{i+1}, q_{i+1}] = \left[ \frac{2p_i + q_i}{2}, \frac{p_i + 2q_i}{2} \right]$$

24 **if  $[p_i, q_i] \in R$  then**

25

**if  $El \cap R - B \neq \emptyset$  then**

26

$$[p_{i+1}, q_{i+1}] = \sigma^{-1} \left( \sigma_1([p_i, q_i]), \frac{\sigma_2([p_i, q_i]) + \alpha'}{2} \right)$$

where

$$\alpha' = \begin{cases} \frac{q'}{2(1-\sigma_1([p_i, q_i]))} & q' \leq 1 - \sigma_1([p_i, q_i]) \\ \frac{q'-1}{2\sigma_1([p_i, q_i])} + 1 & q' > 1 - \sigma_1([p_i, q_i]) \end{cases}$$

with  $[p', q'] = \arg \max_{[p, q] \in El \cap R - B} q$ .

27

**if  $El \cap W \neq \emptyset$  then**

28

$$[p_{i+1}, q_{i+1}] = \sigma^{-1} \left( \sigma_1([p_i, q_i]), \frac{\sigma_2([p_i, q_i]) + \alpha''}{2} \right)$$

where

$$\alpha'' = \begin{cases} \frac{p''}{2\sigma_1([p_i, q_i])} & p'' \leq \sigma_1([p_i, q_i]) \\ \frac{p'' + 1 - 2\sigma_1([p_i, q_i])}{2(1-\sigma_1([p_i, q_i]))} & p'' > \sigma_1([p_i, q_i]) \end{cases}$$

with  $[p'', q''] = \arg \max_{[p, q] \in El \cap W} p$ .

29

**if  $[p_i, q_i] \in B$  then**

30

**if  $El \cap R - B \neq \emptyset$  then**

31

$$[p_{i+1}, q_{i+1}] = \sigma^{-1} \left( \sigma_1([p_i, q_i]), \frac{\sigma_2([p_i, q_i]) + \alpha''}{2} \right)$$

where

$$\alpha'' = \begin{cases} \frac{p''}{2\sigma_1([p_i, q_i])} & p'' \leq \sigma_1([p_i, q_i]) \\ \frac{p'' + 1 - 2\sigma_1([p_i, q_i])}{2(1-\sigma_1([p_i, q_i]))} & p'' > \sigma_1([p_i, q_i]) \end{cases}$$

with  $[p'', q''] = \arg \min_{[p, q] \in El \cap R - B} p$ .

32

**if  $El \cap W \neq \emptyset$  then**

33

$$[p_{i+1}, q_{i+1}] = \sigma^{-1} \left( \sigma_1([p_i, q_i]), \frac{\sigma_2([p_i, q_i]) + \alpha'}{2} \right)$$

where

$$\alpha' = \begin{cases} \frac{q'}{2(1-\sigma_1([p_i, q_i]))} & q' \leq 1 - \sigma_1([p_i, q_i]) \\ \frac{q'-1}{2\sigma_1([p_i, q_i])} + 1 & q' > 1 - \sigma_1([p_i, q_i]) \end{cases}$$

with  $[p', q'] = \arg \min_{[p, q] \in El \cap W} q$ .

---



---

```

34 Case 3
35 if  $(EI \cup \{[p_i, q_i]\}) \cap R - B \neq \emptyset$  and  $(EI \cup \{[p_i, q_i]\}) \cap W \neq \emptyset$  then
36   if  $[p_i, q_i] \in R - B$  then
37     
$$[p_{i+1}, q_{i+1}] = \left[ \frac{p_i + p''}{2}, \frac{q_i + q''}{2} \right]$$

     with  $[p'', q''] = \arg \max_{[p, q] \in EI \cap W} p$ .
38   if  $[p_i, q_i] \in W$  then
39     
$$[p_{i+1}, q_{i+1}] = \left[ \frac{p_i + p'}{2}, \frac{q_i + q'}{2} \right]$$

     with  $[p', q'] = \arg \min_{[p, q] \in EI \cap RB} p$ .
40   if  $[p_i, q_i] \in B$  then
41     
$$[p_{i+1}, q_{i+1}] = \sigma^{-1} \left( \sigma_1([p_i, q_i]), \frac{\sigma_2([p_i, q_i]) + \max(\alpha', \alpha'')}{2} \right)$$

     where
     
$$\alpha' = \begin{cases} \frac{q'}{2(1-\sigma_1([p_i, q_i]))} & q' \leq 1 - \sigma_1([p_i, q_i]) \\ \frac{q'-1}{2\sigma_1([p_i, q_i])} + 1 & q' > 1 - \sigma_1([p_i, q_i]) \end{cases}$$

     with  $[p', q'] = \arg \max_{[p, q] \in EI \cap R - B} q$  and
     
$$\alpha'' = \begin{cases} \frac{p''}{2\sigma_1([p_i, q_i])} & p'' \leq \sigma_1([p_i, q_i]) \\ \frac{p''+1-2\sigma_1([p_i, q_i])}{2(1-\sigma_1([p_i, q_i]))} & p'' > \sigma_1([p_i, q_i]) \end{cases}$$

     with  $[p'', q''] = \arg \min_{[p, q] \in EI \cap W} q$ .
42   if  $[p_i, q_i] \in R$  then
43     
$$[p_{i+1}, q_{i+1}] = \sigma^{-1} \left( \sigma_1([p_i, q_i]), \frac{\sigma_2([p_i, q_i]) + \min(\alpha', \alpha'')}{2} \right)$$

     where
     
$$\alpha' = \begin{cases} \frac{q''}{2(1-\sigma_1([p_i, q_i]))} & q'' \leq 1 - \sigma_1([p_i, q_i]) \\ \frac{q''-1}{2\sigma_1([p_i, q_i])} + 1 & q'' > 1 - \sigma_1([p_i, q_i]) \end{cases}$$

     with  $[p'', q''] = \arg \min_{[p, q] \in EI \cap W} q$  and
     
$$\alpha'' = \begin{cases} \frac{p'}{2\sigma_1([p_i, q_i])} & p' \leq \sigma_1([p_i, q_i]) \\ \frac{p'+1-2\sigma_1([p_i, q_i])}{2(1-\sigma_1([p_i, q_i]))} & p' > \sigma_1([p_i, q_i]) \end{cases}$$

     with  $[p', q'] = \arg \max_{[p, q] \in EI \cap R - B} q$ .
44 end

```

---

Figure 2.13: Determination of Next Binary Choice: Part 3

difference curves in Figure 2.1 cross at most once.<sup>20</sup> Whether this is the case, and how often it is not, will depend on the functional form of  $W$ . We thus consider what form of uniqueness holds for reasonable  $W$ .

For instance, the MPI is clearly unique when  $W$  is linear and non-symmetric<sup>21</sup>—and hence for  $\alpha$ -maxmin EU whenever  $\alpha \neq \frac{1}{2}$ . A more general interesting case is when  $W$  incorporates probability weighting, e.g. is of the form  $W(x, y) = \alpha w(x) + (1 - \alpha)w(y)$  for a weighting function  $w$ . Note that this form can incorporate findings on probability weighting for (two-outcome) lotteries, via the  $w$ . For such  $W$ , if  $w$  takes the quasi-linear form often used in literature (Chateauneuf et al., 2007; Wakker, 2010), then MPIs can be shown to remain unique (by a similar reasoning to that for the non-weighted case). Moreover, even for non-linear weighting functions, calculation of relevant cases suggests that MPIs are typically unique. As an example, Figure 2.14 plots red and blue indifference curves for the specified form of  $W$  with  $w$  being the popular Prelec weighting function with the parameters found by Abdellaoui et al. (2011) for a Paris temperature source—i.e. one that is similar to the source we used in EXP 1—and an  $\alpha$  of 0.8 (close to the value we found for  $\alpha$ ; Section 2.4.5). Clearly, red and blue indifference curves typically only cross (at most) once, as required for uniqueness of MPI. Even in the cases where there are multiple MPIs, there will be at most two, with one close to the horizontal or vertical boundary.

In summary, even for reasonable extensions beyond  $\alpha$ -maxmin EU, MPIs are well-defined, and the subject's probability interval is always a MPI. Moreover, there is reason to believe that uniqueness continues to hold largely, and where it does not, there is at most one other possible candidate MPI. Note that even in cases of non-uniqueness, the analysis of the 2D choice list incentivisation mechanism is unaffected, and every MPI remains a weakly dominant strategy. So it will yield a candidate probability interval.

<sup>20</sup>Technically, for every  $A, B \in \mathbb{R}$ ,  $|\{[x, y] \in \mathcal{I} : W(x, y) = A, W(1 - y, 1 - x) = B\}| \leq 1$ .

<sup>21</sup>I.e. it is not the case that  $W(x, y) = W(y, x)$  for all  $x, y$ .

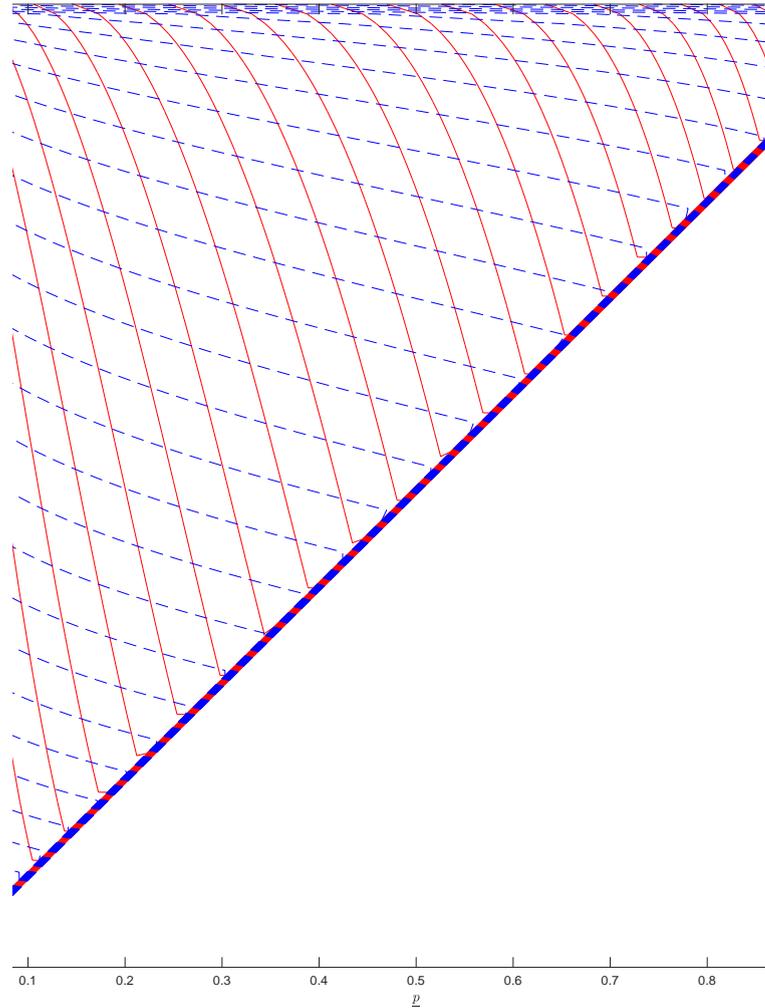


Figure 2.14: Indifference curves in probability interval space  $\mathcal{I}$  under (2.11) with  $W(x, y) = \alpha w(x) + (1 - \alpha)w(y)$ .

Red lines: indifference curves for IL  $(z, [p, q], 0)$ : i.e. curves of the form  $\alpha w(x) + (1 - \alpha)w(y) = C$ .

Blue lines: indifference curves for IL  $(0, [p, q], z)$ : i.e. curves of the form  $\alpha w(1 - y) + (1 - \alpha)w(1 - x) = D$ .

Parametrisation: Prelec weighting function  $w(x) = (e^{-(-\ln(x))^\alpha})^\beta$  with  $\alpha = 0.54$  and  $\beta = 0.85$  (Abdellaoui et al., 2011);  $\alpha = 0.8$ .

## 2D Choice List

As discussed in Section 2.2.6, the 2D choice list is incentive compatible under an isolation assumption—that subjects treat the two branches of the choice list separately—and under the  $\alpha$ -maxmin EU model (2.1). We first consider the consequences of violation of the isolation assumption, before turning to robustness to generalisation of the decision model.

**Robustness to violations of isolation** Suppose that the isolation assumption in Section 2.2.6 does not hold, and the subject reasons strategically across the two branches of the 2D choice list. Then the choice of MPI is conceptualised as the choice of a (second-order) lottery assigning a probability to playing a bet for or against  $E$  or to playing specific ILs according to the mechanism. Assuming the  $\alpha$ -maximin EU model (2.1) at both levels, the subject evaluates each such second-order lottery using the expectation over the values of the bets and ILs. Let  $[\underline{p}(E), \bar{p}(E)] = [\underline{p}, \bar{p}]$ . For any reported point  $[\underline{q}, \bar{q}]$  in this task, by the incentive mechanism defined in Section 2.2.6:

- the probability of receiving the bet on  $E$  is  $\frac{q}{\bar{q}+1-q}$
- the probability of receiving the IL on red is  $\frac{\bar{q}-q}{\bar{q}+1-q}$
- the probability of receiving the bet on  $E^c$  is  $\frac{1-\bar{q}}{\bar{q}+1-q}$
- the probability of receiving the IL on red is  $\frac{\bar{q}-q}{\bar{q}+1-q}$

Using these, and the evaluations of the bets and the ILs according to (2.1) and (2.2) (with  $[\underline{p}, \bar{p}]$ ), one obtains the following form for the utility of reporting  $[\underline{q}, \bar{q}]$  when the true beliefs are  $[\underline{p}, \bar{p}]$ :

$$\begin{aligned} & \frac{(1 - \bar{q}) (\alpha (1 - \bar{p}) - (\alpha - 1) (1 - \underline{p}))}{\bar{q} - \underline{q} + 1} + \frac{(\bar{q} - \underline{q}) \left( \frac{\underline{q}}{2} + \frac{\alpha(1-\bar{q})}{2} - \frac{(\alpha-1)(1-\underline{q})}{2} - \frac{1}{2} \right)}{\bar{q} - \underline{q} + 1} \\ & + \frac{\underline{q} (\alpha \underline{p} - \bar{p} (\alpha - 1))}{\bar{q} - \underline{q} + 1} + \frac{(\bar{q} - \underline{q}) \left( \frac{\bar{q}}{2} + \frac{\alpha \underline{q}}{2} - \frac{\bar{q}(\alpha-1)}{2} \right)}{\bar{q} - \underline{q} + 1} \end{aligned}$$

Finding the optimum numerically for a grid of values of  $\underline{p}, \bar{p}, \alpha \in [0, 1]$  using Matlab, we find that, for every  $(\underline{p}, \bar{p}, \alpha)$  (with  $\bar{p} \geq \underline{p}$ ) except for  $\underline{p} = 0, \bar{p} = 1, \alpha = 0$ , and those with  $\underline{p} = 0.5, \alpha = 1$  or  $\bar{p} = 0.5, \alpha = 1$ , the maximum is one or several of the ‘vertices’ of the triangle in Figure 2.1, i.e.  $[0, 0], [0, 1], [1, 1]$ . For  $\underline{p} = 0, \bar{p} = 1, \alpha = 0$  and  $\underline{p} = 0.5, \alpha = 1$  or  $\bar{p} = 0.5, \alpha = 1$  with  $\underline{p} \neq \bar{p}$ , the maxima are all points on one of the boundaries of the triangle, i.e.  $\{[0, y] : y \in [0, 1]\}, \{[x, 1] : x \in [0, 1]\}, \{[x, y] : x \in [0, 1], y = x\}$ . When  $\underline{p} = 0.5, \bar{p} = 0.5, \alpha = 1$ , the utility above is constant, so all points maximise it.

Clearly, in our experiment involving nested events, a subject who gives precise probability of 0.5 to all elicited events would have a very particular (bimodal) distribution across the variable of interest (temperature, marks). Given the implausibility of such beliefs, we conclude from this analysis that, for a subject violating the isolation assumption discussed in Section 2.2.6 and responding to the choice list strategically, every response except perhaps one or two will be at a vertex of the space  $\mathcal{I}$  (i.e. the triangle in Figure 2.1). Certainly, every response except at most one will be on the boundary of this set.

**Robustness to generalizations of the decision model** As suggested in Section 2.2.6, the incentivization mechanism implemented by the 2D choice list relies solely on the weak Lower Stochastic Dominance property (Definition 1). Formulated in terms of Eq. (2.11), this is just the assumption that  $W$  is strictly increasing in the first coordinate—or, in terms of preferences, decision makers are sensitive to the lower winning probability. The only reasonable model in the family of form (2.11) violating this property is the maxmax EU— $\alpha$ -maxmin EU with  $\alpha = 0$ . Since there is basically no evidence for a significant number of subjects with such

preferences, the incentive compatibility of the 2D choice list discussed in Section 2.2.6 generalizes widely.

## Binary-choice procedure

The binary-choice procedure is based on the division of  $\mathcal{I}$  into regions, displayed in Table 2.1, and Proposition 2, in particular part a. dictating ‘where’ the MPI is relative to points in two of the regions (the W and R-B regions). For decision makers represented according to the  $\alpha$ -maxmin EU model (2.1), Proposition 2 a. only holds if  $\alpha > \frac{1}{2}$ .<sup>22</sup> When  $\alpha < \frac{1}{2}$ , the opposite of the statement in the Proposition holds: the MPI is ‘North-West’ of the elicited point (on Figure 2.1) not when the latter is in R-B, but when it is in W (and similarly for South-East). So the algorithm applied to such decision makers would ‘move’ in the wrong direction: instead of looking ‘South-East’ for the MPI after finding a point in R-B, it would look ‘North-West’, for instance. Note that, even if the algorithm does not work properly for such decision makers, the 2D choice list incentivisation mechanism is still valid, and hence they would, in principle, correct any issues at the 2D choice list confirmation stage. To gain some insight into the extent of procedure malfunction due to  $\alpha < \frac{1}{2}$ , we can look at the evidence on the value of  $\alpha$  for our subjects, as well as some statistics on the functioning of the procedure.

We find little evidence for widespread  $\alpha < \frac{1}{2}$  among our subjects. First of all, the elicitation of  $\alpha$  reported in Section 2.4.5 finds median and 25 percentile values way over  $\frac{1}{2}$  (Table 2.25), indicating that less than 25% of subjects have  $\alpha < \frac{1}{2}$ . Moreover, under the  $\alpha$ -maxmin EU model, the sum of the MP of an event and that of its complement is less than (respectively, greater than) one precisely when  $\alpha > \frac{1}{2}$  (resp.  $\alpha < \frac{1}{2}$ ; see Appendix 2.8.4), indicating that we can use our matching probability data to check for the sign of  $\alpha - \frac{1}{2}$ . Table 2.26 (Appendix 2.8.4) displays the descriptive statistics on this sum for the Paris treatment where MPs were elicited, confirming again that  $\alpha > \frac{1}{2}$  for over 75% of subjects.

<sup>22</sup>It also holds under the probability weighting specification of (2.11) mentioned in Section 2.7.3, with  $\alpha > \frac{1}{2}$ .

As concerns its functioning, since the procedure ‘moves’ in the wrong direction for subjects with  $\alpha < \frac{1}{2}$ , no such subjects will pass through both points in W and points in R-B during the procedure. However, 383 applications of the procedure out of 704 in EXP 1 passed through points in W and R-B (300 out of 606 in EXP 2). Whilst there were nevertheless applications of the procedure which passed through points in R-B but not W (152 in EXP 1, 77 in EXP 2) and in W but not R-B (114 in EXP 1, 105 in EXP 2), these would be expected if the procedure functioned correctly and the probability intervals were large (respectively small). The evidence thus does not support a hypothesis involving malfunctioning of the procedure over explanations, such as this, relating to proper functioning and the character of the elicited intervals.

## 2.8. Supplementary Statistics

### 2.8.1. Descriptive Statistics

Tables 2.5–2.8 report the basic descriptive statistics on the upper and lower elicited probabilities after the ‘confirmation’ 2D choice list, and before the confirmation screen but after the binary-choice procedure, respectively.

**Elicited points on a vertex** Table 2.9 reports counts of the number of subjects with a given number of elicited points at the vertex of the space  $\mathcal{I}$  of partially known urns (and corresponding probability intervals) in Figure 2.1.

**Monotonicity** Tables 2.10a and 2.10b report the descriptive statistics for the individual-level Kendall  $\tau_b$ , calculated over the events in each source.<sup>23</sup>

<sup>23</sup>Note that  $\tau_b$  for not defined for some subjects in EXP 1 (because of too many ties), and they were dropped. Hence the count for EXP 1 may vary across sources.

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
-2	80.0	0.29	0.17	-0.01	0.15	0.29	0.40	0.7
2	80.0	0.38	0.22	0.00	0.20	0.35	0.50	1.0
5	80.0	0.48	0.23	0.00	0.35	0.46	0.66	1.0
8	80.0	0.57	0.24	0.05	0.42	0.59	0.75	1.0

(a) Lower probabilities Paris

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
-2	80.0	0.55	0.21	0.09	0.40	0.55	0.67	1.0
2	80.0	0.65	0.19	0.23	0.51	0.65	0.80	1.0
5	80.0	0.74	0.17	0.25	0.62	0.76	0.88	1.0
8	80.0	0.82	0.14	0.50	0.75	0.85	0.94	1.0

(b) Upper probabilities Paris

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
15	80.0	0.31	0.22	0.00	0.14	0.26	0.45	0.95
18	80.0	0.35	0.26	0.00	0.14	0.32	0.47	1.00
20	80.0	0.41	0.27	-0.01	0.20	0.40	0.61	1.00
22	80.0	0.43	0.26	-0.01	0.20	0.39	0.61	1.00

(c) Lower probabilities Sydney

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
15	80.0	0.58	0.27	0.01	0.37	0.56	0.84	0.99
18	80.0	0.66	0.24	0.03	0.50	0.69	0.88	1.00
20	80.0	0.71	0.23	0.01	0.60	0.76	0.89	1.00
22	80.0	0.73	0.23	0.00	0.58	0.80	0.92	1.00

(d) Upper probabilities Sydney

Table 2.5: Descriptive Statistics: Elicited lower and upper probabilities after 2D choice list, EXP 1

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.05	0.07	-0.01	0.00	0.04	0.08	0.35
10	52.0	0.15	0.12	0.00	0.06	0.12	0.19	0.50
12	52.0	0.24	0.15	0.00	0.14	0.20	0.31	0.63
15	52.0	0.40	0.19	0.08	0.26	0.38	0.55	0.73
17	52.0	0.60	0.16	0.18	0.54	0.64	0.71	0.86

(a) Lower probabilities Maths

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.12	0.14	0.00	0.02	0.08	0.16	0.60
10	52.0	0.23	0.17	0.00	0.10	0.20	0.30	0.65
12	52.0	0.35	0.18	0.04	0.22	0.32	0.48	0.72
15	52.0	0.54	0.20	0.08	0.40	0.52	0.70	0.87
17	52.0	0.75	0.15	0.22	0.65	0.78	0.86	1.01

(b) Upper probabilities Maths

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.12	0.08	0.01	0.06	0.11	0.17	0.36
10	52.0	0.22	0.12	0.02	0.14	0.20	0.29	0.50
12	52.0	0.33	0.13	0.14	0.22	0.32	0.40	0.60
15	52.0	0.54	0.14	0.19	0.46	0.56	0.65	0.83
17	52.0	0.71	0.13	0.25	0.65	0.74	0.83	0.90

(c) Lower probabilities Contraction

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.22	0.14	0.02	0.11	0.20	0.31	0.56
10	52.0	0.37	0.14	0.06	0.30	0.34	0.46	0.65
12	52.0	0.51	0.14	0.20	0.40	0.50	0.64	0.74
15	52.0	0.74	0.11	0.40	0.67	0.77	0.82	0.90
17	52.0	0.86	0.07	0.60	0.82	0.86	0.91	1.00

(d) Upper probabilities Contraction

Table 2.6: Descriptive Statistics: Elicited lower and upper probabilities after 2D choice list, EXP 2

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
-2	80.0	0.28	0.15	0.0	0.15	0.30	0.39	0.64
2	80.0	0.36	0.21	0.0	0.20	0.33	0.45	1.00
5	80.0	0.48	0.23	0.0	0.35	0.46	0.67	1.00
8	80.0	0.54	0.25	0.05	0.35	0.55	0.74	1.00

(a) Lower probabilities Paris

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
-2	80.0	0.55	0.21	0.0	0.39	0.56	0.68	0.99
2	80.0	0.64	0.18	0.23	0.52	0.67	0.78	1.00
5	80.0	0.75	0.17	0.25	0.61	0.79	0.89	1.00
8	80.0	0.81	0.14	0.50	0.73	0.85	0.92	1.00

(b) Upper probabilities Paris

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
15	80.0	0.28	0.21	0.0	0.12	0.26	0.38	0.95
18	80.0	0.33	0.27	0.0	0.11	0.30	0.45	1.00
20	80.0	0.43	0.28	0.0	0.19	0.42	0.61	1.00
22	80.0	0.42	0.26	0.0	0.21	0.35	0.61	1.00

(c) Lower probabilities Sydney

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
15	80.0	0.60	0.27	0.00	0.42	0.61	0.83	0.99
18	80.0	0.67	0.25	0.03	0.50	0.66	0.89	1.00
20	80.0	0.73	0.23	0.00	0.60	0.79	0.89	1.00
22	80.0	0.74	0.22	0.00	0.62	0.80	0.91	1.00

(d) Upper probabilities Sydney

Table 2.7: Descriptive Statistics: Elicited lower and upper probabilities after binary-choice procedure and before 2D choice list, EXP 1

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.11	0.16	0.00	0.01	0.06	0.14	1.00
10	52.0	0.18	0.14	0.00	0.08	0.14	0.30	0.50
12	52.0	0.25	0.15	0.00	0.15	0.22	0.32	0.60
15	52.0	0.38	0.22	0.04	0.19	0.33	0.53	1.00
17	52.0	0.54	0.18	0.07	0.43	0.60	0.66	0.86

(a) Lower probabilities Maths

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.18	0.23	0.00	0.03	0.08	0.28	1.00
10	52.0	0.27	0.22	0.00	0.10	0.23	0.44	0.70
12	52.0	0.40	0.23	0.00	0.22	0.38	0.61	0.78
15	52.0	0.56	0.21	0.08	0.39	0.58	0.71	1.00
17	52.0	0.73	0.16	0.22	0.62	0.76	0.85	0.99

(b) Upper probabilities Maths

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.15	0.09	0.01	0.08	0.15	0.20	0.40
10	52.0	0.26	0.12	0.00	0.18	0.23	0.35	0.48
12	52.0	0.32	0.16	0.04	0.21	0.30	0.40	0.71
15	52.0	0.53	0.19	0.16	0.40	0.55	0.66	1.00
17	52.0	0.67	0.20	0.07	0.62	0.68	0.83	1.00

(c) Lower probabilities Contraction

Event $E_t$ for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.30	0.23	0.03	0.12	0.24	0.41	0.85
10	52.0	0.41	0.18	0.00	0.31	0.35	0.56	0.86
12	52.0	0.54	0.16	0.20	0.44	0.55	0.65	0.87
15	52.0	0.74	0.14	0.37	0.66	0.74	0.84	1.00
17	52.0	0.85	0.09	0.50	0.82	0.86	0.91	1.00

(d) Upper probabilities Contraction

Table 2.8: Descriptive Statistics: Elicited lower and upper probabilities after binary-choice procedure and before 2D choice list, EXP 2

Point	Paris # subjects					Sydney # subjects				
	0	1	2	3	4	0	1	2	3	4
[0, 0]	80	0	0	0	0	80	0	0	0	0
[0, 1]	80	0	0	0	0	80	0	0	0	0
[1, 1]	79	0	0	1	0	78	1	0	1	0
[0, 0], [0, 1] or [1, 1]	79	0	0	1	0	78	1	0	1	0

(a) EXP 1

Point	Maths # subjects						Contraction # subjects					
	0	1	2	3	4	5	0	1	2	3	4	5
[0, 0]	46	5	1	0	0	0	52	0	0	0	0	0
[0, 1]	52	0	0	0	0	0	52	0	0	0	0	0
[1, 1]	52	0	0	0	0	0	52	0	0	0	0	0
[0, 0], [0, 1] or [1, 1]	46	5	1	0	0	0	52	0	0	0	0	0

(b) EXP 2

Table 2.9: For each type of point, the table indicates the number of subjects with the specified number of elicited points being of this type.

	MP Lower	MP Upper	Paris Lower	Paris Upper	Sydney Lower	Sydney Upper
count	74.00	78	79	78	78	78
mean	0.62	0.66	0.56	0.56	0.27	0.41
std	0.46	0.38	0.45	0.47	0.59	0.50
min	-0.91	-0.91	-0.91	-0.91	-1.00	-1.00
25%	0.55	0.55	0.33	0.33	-0.14	0.00
50%	0.71	0.69	0.67	0.67	0.33	0.55
75%	0.91	0.91	1.00	1.00	0.67	0.91
max	1.00	1.00	1.00	1.00	1.00	1.00

(a) EXP 1

	Contraction Lower	Contraction Upper	Maths Lower	Maths Upper
count	52	52	52	52
mean	0.99	0.99	0.98	1.00
std	0.02	0.03	0.07	0.01
min	0.95	0.80	0.53	0.95
25%	1.00	1.00	1.00	1.00
50%	1.00	1.00	1.00	1.00
75%	1.00	1.00	1.00	1.00
max	1.00	1.00	1.00	1.00

(b) EXP 2

Table 2.10: Individual-level Kendall  $\tau_b$  descriptive statistics

	EXP 1		EXP 2	
	Paris	Sydney	Maths	Contraction
count	80	80	52	52
mean	0.25	0.29	0.13	0.19
std	0.17	0.20	0.09	0.11
min	0	0	0	0.01
25%	0.10	0.10	0.06	0.11
50%	0.23	0.28	0.11	0.17
75%	0.35	0.43	0.19	0.23
max	0.82	0.75	0.45	0.53

Table 2.11: Imprecision Index (Eq. (2.8)) descriptive statistics; EXP 1 and EXP 2.

# Precise events	EXP 1				EXP 2			
	Paris		Sydney		Maths		Contraction	
	2D C.L.	B-C Proc	2D C.L.	B-C Proc	2D C.L.	B-C Proc	2D C.L.	B-C Proc
0	51	48	48	44	20	12	31	19
1	14	14	18	23	14	14	12	20
2	7	11	8	8	12	10	5	3
3	6	0	2	4	3	10	3	4
4	2	1	4	1	0	2	1	7
5	-	-	-	-	3	4	0	2
Total	80	80	80	80	52	52	52	52

Table 2.12: Number of subjects with given number of precise events, per source. Data given after the 2D choice list confirmation screen (2D C.L.) and after the binary-choice procedure but before the confirmation screen (B-C Proc).

## Imprecision

Table 2.11 presents the descriptive statistics for the Imprecision Index, whereas Table 2.12 displays counts of the number of subjects with various numbers of precise elicited points, as well as differences before the 2D choice list confirmation stage of the experiment as opposed to after. Table 2.13 presents the results of ANOVAs of the imprecision concerning an event against the event, for each source, where the null hypothesis is that imprecision is invariant across events.

	Source	$F$	p-value
EXP 1	Paris	0.1048	0.957
	Sydney	0.4769	0.698
EXP 2	Contraction	4.0352	0.003
	Maths	5.863	0.00015

Table 2.13: One-sided ANOVAs of the imprecision related to an event (dependent variable) on the event (factor), for each source. ( $H_0$ : the imprecision is identical across all events in the source.)

## Binary choice procedure

Table 2.14 reports data on the number of steps, and mean and median widths of elicited probability intervals, over the implemented binary-choice procedure in the two experiments.

### 2.8.2. Bayesian estimation

#### Statistical approach

**Estimation of upper and lower CDFs in EXP 1 and EXP 2** Recall that  $T$  denotes the space of possible values of the variables of interest (minimum temperatures in EXP 1, grades in EXP 2). For each source, we estimate general models of the form:

$$\begin{cases} \underline{p}(E) = \underline{f}(E) + \underline{\epsilon} \\ \bar{p}(E) = \bar{f}(E) + \bar{\epsilon} \end{cases} \quad (2.17)$$

where  $\underline{p}(E)$  (resp.  $\bar{p}(E)$ ) are the elicited lower (resp. upper) probabilities of the cumulative event  $E$  (Section 2.2.2),  $\underline{f}$  and  $\bar{f}$  are CDFs over  $T$  from specified two-parameter families, with parameters  $\underline{a}, \underline{b}$  (resp.  $\bar{a}, \bar{b}$ ), and  $\underline{\epsilon}$  and  $\bar{\epsilon}$  are zero-mean normal distributions with variance  $\underline{\sigma}^2$  and  $\bar{\sigma}^2$  respectively.

# steps	Occurrences		Interval size	
	#	%	Mean	Median
4	43.0	6.10	0.45	0.30
5	88.0	12.5	0.29	0.29
6	45.0	6.39	0.21	0.15
7	42.0	5.96	0.24	0.19
8	85.0	12.07	0.33	0.33
9	54.0	7.67	0.31	0.25
10	61.0	8.66	0.19	0.14
11	54.0	7.67	0.31	0.33
12	50.0	7.10	0.18	0.17
13	61.0	8.66	0.27	0.23
14	44.0	6.25	0.34	0.30
15	55.0	7.81	0.34	0.39
16	17.0	2.41	0.32	0.27
17	2.0	0.28	0.45	0.45
18	3.0	0.42	0.43	0.49

(a) EXP 1

# steps	Occurrences		Interval size	
	#	%	Mean	Median
4	30.0	4.95	0.28	0.15
5	58.0	9.57	0.17	0.14
6	35.0	5.77	0.12	0.14
7	31.0	5.11	0.21	0.15
8	59.0	9.73	0.23	0.18
9	80.0	13.20	0.10	0.0
10	72.0	11.88	0.7	0.0
11	42.0	6.93	0.19	0.12
12	45.0	7.42	0.17	0.15
13	33.0	5.44	0.15	0.5
14	49.0	8.08	0.23	0.20
15	58.0	9.57	0.13	0.10
16	11.0	1.81	0.24	0.16
17	2.0	0.33	0.15	0.15
19	1.0	0.16	0.06	0.06

(b) EXP 2

Table 2.14: Bisection-style choice procedure: number of steps, mean and median width of elicited interval

For each equation, the parameter space is  $\Theta \subseteq \mathbb{R}^3$ , with a typical point  $(\underline{a}, \underline{b}, \underline{\sigma})$  (resp.  $(\bar{a}, \bar{b}, \bar{\sigma})$ ) specifying an  $\underline{f}$  (resp.  $\bar{f}$ ) and the variance of the relevant error term. We specify the following priors over the hyperparameters :  $a, b, \sigma$  are realisations from  $A \sim N(\mu_a, \sigma_a^2)$ ,  $B \sim N(\mu_b, \sigma_b^2)$  and  $\Sigma = \sigma_\sigma | Y |$  with  $Y \sim N(0, 1)$ .

We use a MCMC-like approach to estimate the posterior distributions of these distributions through the use of the Python package PyMC3, and more specifically, the No-U-Turn Sampler algorithm (NUTS) (Hoffman and Gelman, 2014).

The likelihood of observations  $x_1, \dots, x_n$  pertaining to  $t_1, \dots, t_n$  (e.g. elicited lower probabilities for cumulative events  $E_{t_i} = \{t \in T : t \leq t_i\}$ ) given the point  $(a, b, \sigma) \in \Theta$  is:

$$L(a, b, \sigma | x_1, \dots, x_n) = \prod_{i \in \{1, \dots, n\}} \varphi \left( \frac{x_i - f_{(a,b)}(\{t \leq t_i\})}{\sigma} \right)$$

where  $f_{(a,b)}$  is the CDF with parameters  $a, b$  and  $\varphi$  is the density of the normal distribution. Hence the likelihood of hyperparameters  $\mu_a, \sigma_a^2, \mu_b, \sigma_b^2, \mu_\sigma, \sigma_\sigma^2$  given observations  $x_1 \dots x_n$  is :

$$\begin{aligned} & L(\mu_a, \sigma_a^2, \mu_b, \sigma_b^2, \mu_\sigma, \sigma_\sigma^2 | x_1, \dots, x_n) \\ &= \int_{(a,b,\sigma) \in \Theta} L(a, b, \sigma | x_1, \dots, x_n) dp(a, b, \sigma | \mu_a, \sigma_a^2, \mu_b, \sigma_b^2, \mu_\sigma, \sigma_\sigma^2) \end{aligned}$$

$L(\mu_{\underline{a}}, \sigma_{\underline{a}}^2, \mu_{\underline{b}}, \sigma_{\underline{b}}^2, \mu_{\underline{\sigma}}, \sigma_{\underline{\sigma}}^2 | \underline{x}_1, \dots, \underline{x}_n)$  and  $L(\mu_{\bar{a}}, \sigma_{\bar{a}}^2, \mu_{\bar{b}}, \sigma_{\bar{b}}^2, \mu_{\bar{\sigma}}, \sigma_{\bar{\sigma}}^2 | \bar{x}_1, \dots, \bar{x}_n)$  are used by the NUTS algorithm to estimate the posterior distributions of  $A, B$  and  $\Sigma$ , where  $\underline{x}_1, \dots, \underline{x}_n, \bar{x}_1, \dots, \bar{x}_n$  are the elicited lower and upper probabilities respectively, under the parametric families for  $f$  given in Table 2.15.

**Likelihood estimation of alpha in EXP 1 (Paris treatment)** For the Bayesian estimation of the mixture coefficient  $\alpha$  in the  $\alpha$ -maxmin EU model, we supplement

	Temperature (EXP 1)	Grade (EXP 2)
Family 1	Truncated Normal $\mathcal{N}(a, b)$	Truncated Normal $\mathcal{N}(a, b)$
Family 2	Beta $B(a, b)$	Beta $B(a, b)$
Support	[min of min stated temperature, max of max stated temperature]	[0,20]

Table 2.15: Families of distributions over  $T$  (temperature; grade)

the general model (2.17) with the following equations

$$\begin{cases} MP(E) = \alpha \underline{p}(E) + (1 - \alpha) \bar{p}(E) + \epsilon_{\alpha} \\ 1 - MP(E^c) = \alpha \bar{p}(E) + (1 - \alpha) \underline{p}(E) + \epsilon_{\bar{\alpha}} \end{cases} \quad (2.18)$$

which are discussed in Section 2.4.5. We assume that  $\alpha$  follows a beta distribution  $B(a_{\alpha}, b_{\alpha})$ , and the  $\epsilon_{\alpha}$  are zero-mean normal distributions, with the hyperparameters independent and normally distributed, as above.

The MPs have been elicited for the Paris treatment in EXP 1. The hyperparameters concerning the upper and lower CDFs discussed above and those for  $\alpha$  were estimated under the model composed of (2.17) and (2.18) using the NUTS algorithm, with the procedure set out above.

## Analysis

Figure 2.15 displays the upper and lower distributions under the parametric families not shown in Figure 2.4. Tables 2.16-2.23 give statistics on the distribution over parameters under the estimated hyperparameters.

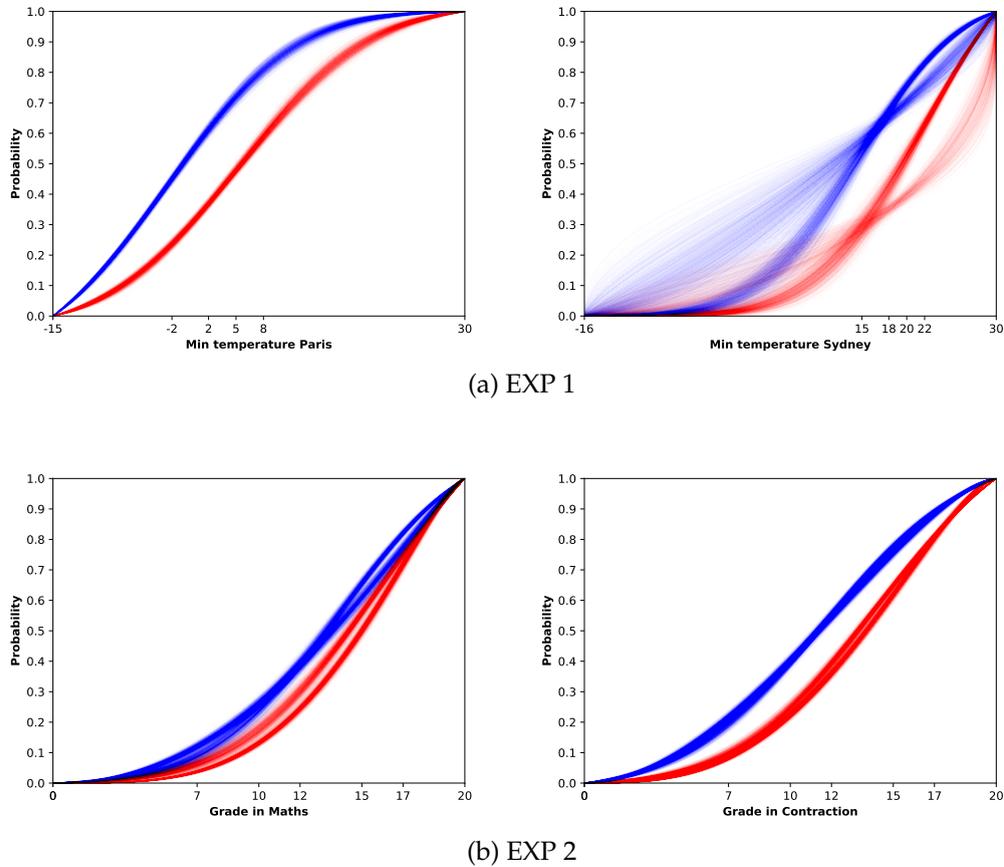


Figure 2.15: Bayesian estimation of lower and upper CDFs: plots of 1000 samples from MCMC (Truncated Normal distribution for EXP 1; Beta distribution for EXP 2)

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
$\underline{a}$	5.25	0.31	0.01	4.65	5.87	1009.11	1.0
$\bar{a}$	-2.57	0.37	0.01	-3.27	-1.87	542.99	1.0
$a_\alpha$	3.42	1.62	0.06	0.53	6.38	630.90	1.0
$b_\alpha$	1.80	1.04	0.05	0.10	3.80	429.61	1.0
$\underline{b}$	11.35	0.75	0.02	9.71	12.66	1214.33	1.0
$\bar{b}$	11.00	0.64	0.03	9.78	12.17	614.95	1.0
$\underline{s}$	0.22	0.01	0.00	0.20	0.23	1043.67	1.0
$\bar{s}$	0.18	0.01	0.00	0.17	0.19	1055.83	1.0
$s_{\bar{\alpha}}$	0.21	0.01	0.00	0.20	0.23	930.43	1.0
$s_{\underline{\alpha}}$	0.19	0.01	0.00	0.17	0.20	909.78	1.0
$\alpha$	0.81	0.04	0.00	0.74	0.88	754.54	1.0

Table 2.16: Statistics for parameters under Bayesian estimation; Paris (EXP 1); Normal parametrisation

Note mc\_error: Monte Carlo procedure standard error; hpd\_2.5 / hpd\_97.5: Highest posterior density 2.5 and 97.5 percentiles; n\_eff : count of iteration in the MCMC procedure.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
$\underline{a}$	1.43	0.23	0.01	1.04	1.91	569.30	1.0
$\underline{b}$	1.64	0.29	0.01	1.13	2.24	571.99	1.0
$\bar{a}$	1.07	0.16	0.01	0.73	1.39	541.73	1.0
$\bar{b}$	2.46	0.35	0.01	1.76	3.17	522.49	1.0
$a_\alpha$	4.32	1.76	0.06	1.06	7.87	812.98	1.0
$b_\alpha$	1.90	1.08	0.04	0.18	3.92	606.36	1.0
$\underline{s}$	0.22	0.01	0.00	0.20	0.23	1163.74	1.0
$\bar{s}$	0.18	0.01	0.00	0.17	0.19	1239.11	1.0
$s_{\bar{\alpha}}$	0.21	0.01	0.00	0.20	0.23	1134.15	1.0
$s_{\underline{\alpha}}$	0.19	0.01	0.00	0.17	0.20	1409.46	1.0
$\alpha$	0.81	0.04	0.00	0.74	0.88	1079.99	1.0

Table 2.17: Statistics for parameters under Bayesian estimation; Paris (EXP 1); Beta parametrisation

Note mc\_error: Monte Carlo procedure standard error; hpd\_2.5 / hpd\_97.5: Highest posterior density 2.5 and 97.5 percentiles; n\_eff : count of iteration in the MCMC procedure.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
$\underline{a}$	22.03	0.45	0.01	21.11	22.80	1130.92	1.0
$\bar{a}$	14.66	0.46	0.01	13.78	15.48	876.71	1.0
$\underline{b}$	9.62	0.95	0.03	7.88	11.67	1018.57	1.0
$\bar{b}$	9.04	0.85	0.02	7.42	10.78	882.98	1.0
$\underline{s}$	0.26	0.01	0.00	0.23	0.28	933.24	1.0
$\bar{s}$	0.25	0.01	0.00	0.23	0.27	831.75	1.0

Table 2.18: Statistics for parameters under Bayesian estimation; Sydney (EXP 1); Normal parametrisation

Note mc\_error: Monte Carlo procedure standard error; hpd\_2.5 / hpd\_97.5: Highest posterior density 2.5 and 97.5 percentiles; n\_eff : count of iteration in the MCMC procedure.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
$\underline{a}$	1.12	0.55	0.04	0.31	2.21	1.91	1.57
$\bar{a}$	0.14	0.32	0.03	-0.27	0.67	1.07	4.19
$\underline{b}$	1.32	0.48	0.02	0.49	2.24	320.38	1.00
$\bar{b}$	0.94	0.24	0.01	0.50	1.37	321.66	1.00
$\underline{s}$	0.25	0.01	0.00	0.23	0.27	522.03	1.00
$\bar{s}$	0.24	0.01	0.00	0.22	0.26	671.56	1.00

Table 2.19: Statistics for parameters under Bayesian estimation; Sydney (EXP 1); Beta parametrisation

Note mc\_error: Monte Carlo procedure standard error; hpd\_2.5 / hpd\_97.5: Highest posterior density 2.5 and 97.5 percentiles; n\_eff : count of iteration in the MCMC procedure.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
$\underline{a}$	15.88	0.17	0.01	15.57	16.22	788.64	1.0
$\bar{a}$	5.40	0.28	0.01	4.86	5.95	955.04	1.0
$\underline{b}$	13.97	0.17	0.00	13.65	14.33	1218.24	1.0
$\bar{b}$	5.03	0.25	0.01	4.58	5.53	928.02	1.0
$\underline{s}$	0.14	0.01	0.00	0.13	0.16	1158.52	1.0
$\bar{s}$	0.17	0.01	0.00	0.16	0.19	958.54	1.0

Table 2.20: Statistics for parameters under Bayesian estimation; Maths (EXP 2); Normal parametrisation

Note mc\_error: Monte Carlo procedure standard error; hpd\_2.5 / hpd\_97.5: Highest posterior density 2.5 and 97.5 percentiles; n\_eff : count of iteration in the MCMC procedure.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
$\underline{a}$	3.76	0.09	0.0	3.58	3.93	978.44	1.0
$\bar{a}$	1.46	0.05	0.0	1.36	1.55	900.67	1.0
$\underline{b}$	2.27	0.09	0.0	2.10	2.43	1049.31	1.0
$\bar{b}$	1.22	0.05	0.0	1.12	1.32	1031.36	1.0
$\underline{s}$	0.16	0.01	0.0	0.14	0.17	1311.22	1.0
$\bar{s}$	0.19	0.01	0.0	0.17	0.20	1247.27	1.0

Table 2.21: Statistics for parameters under Bayesian estimation; Maths (EXP 2); Beta parametrisation

Note mc\_error: Monte Carlo procedure standard error; hpd\_2.5 / hpd\_97.5: Highest posterior density 2.5 and 97.5 percentiles; n\_eff : count of iteration in the MCMC procedure.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
$\underline{a}$	14.17	0.13	0.00	13.93	14.42	1090.28	1.0
$\bar{a}$	5.32	0.22	0.01	4.90	5.78	1068.01	1.0
$\underline{b}$	11.61	0.13	0.00	11.37	11.87	1167.95	1.0
$\bar{b}$	5.39	0.21	0.01	4.99	5.78	1209.38	1.0
$\underline{s}$	0.12	0.01	0.00	0.11	0.13	1501.55	1.0
$\bar{s}$	0.12	0.01	0.00	0.11	0.13	1144.85	1.0

Table 2.22: Statistics for parameters under Bayesian estimation; Contraction (EXP 2); Normal parametrisation

Note mc\_error: Monte Carlo procedure standard error; hpd\_2.5 / hpd\_97.5: Highest posterior density 2.5 and 97.5 percentiles; n\_eff : count of iteration in the MCMC procedure.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
$\underline{a}$	3.06	0.09	0.0	2.90	3.23	888.13	1.00
$\bar{a}$	1.57	0.05	0.0	1.48	1.67	937.24	1.00
$\underline{b}$	1.96	0.07	0.0	1.82	2.11	478.84	1.01
$\bar{b}$	1.56	0.06	0.0	1.45	1.66	460.19	1.00
$\underline{s}$	0.14	0.01	0.0	0.13	0.15	966.76	1.00
$\bar{s}$	0.13	0.01	0.0	0.12	0.15	894.44	1.00

Table 2.23: Statistics for parameters under Bayesian estimation; Contraction (EXP 2); Beta parametrisation

Note mc\_error: Monte Carlo procedure standard error; hdp\_2.5 / hdp\_97.5: Highest posterior density 2.5 and 97.5 percentiles; n\_eff : count of iteration in the MCMC procedure.

### 2.8.3. Matching Probability data and analysis of $\alpha$

Table 2.24 provides descriptive statistics on the elicited MPs. Table 2.25 provide descriptive statistics on the  $\alpha$  estimated from the raw data (from equations (2.9) and (2.10)). These equations cannot be applied to estimate  $\alpha$  whenever the upper and lower probabilities of an event coincide, i.e.  $\underline{p}(E) = \bar{p}(E)$ ; Table 2.25 performs the estimates using all events for which the equations can be applied—and hence only removes the two subjects for which the upper and lower probabilities coincide for all events (Table 2.12).

### 2.8.4. Elicitation-free check of $\alpha > \frac{1}{2}$

Under the  $\alpha$ -maxmin EU model (2.1), it follows from Eqs. 2.9 and 2.10 that

$$MP(E) + MP(E^c) = 1 + (\bar{p}(E) - \underline{p}(E)) \cdot (1 - 2\alpha)$$

Since  $\bar{p}(E) - \underline{p}(E) \geq 0$  by definition, it follows that, whenever there is imprecision,  $MP(E) + MP(E^c) < 1$  if and only if  $\alpha > \frac{1}{2}$ .

$MP(E_t)$	count	mean	std	min	25%	50%	75%	max
Event $t =$								
-2	80	0.35	0.21	0.02	0.17	0.37	0.47	1.00
2	80	0.44	0.20	0.02	0.27	0.47	0.57	0.97
5	80	0.54	0.23	0.02	0.37	0.55	0.68	0.97
8	80	0.60	0.21	0.17	0.47	0.57	0.76	0.97

$1 - MP(E_t^c)$	count	mean	std	min	25%	50%	75%	max
Event $t =$								
-2	80	0.50	0.19	0.03	0.38	0.48	0.63	0.98
2	80	0.59	0.19	0.23	0.48	0.57	0.74	0.98
5	80	0.71	0.20	0.23	0.53	0.73	0.92	0.98
8	80	0.77	0.17	0.43	0.63	0.80	0.93	0.98

Table 2.24: Descriptive statistics for  $MP(E_i)$  and  $1 - MP(E_i^c)$  in Paris treatment, EXP 1

$\alpha$	
count	78
mean	0.97
std	0.66
min	-0.32
25%	0.62
50%	0.80
75%	1.17
max	3.84

Table 2.25: Descriptive statistics for  $\alpha$ , estimated from raw data according to Eqns (2.9) and (2.10).

Estimation conducted across all subjects such that, for any least one event  $E$ ,  $\underline{p}(E) \neq \bar{p}(E)$ .

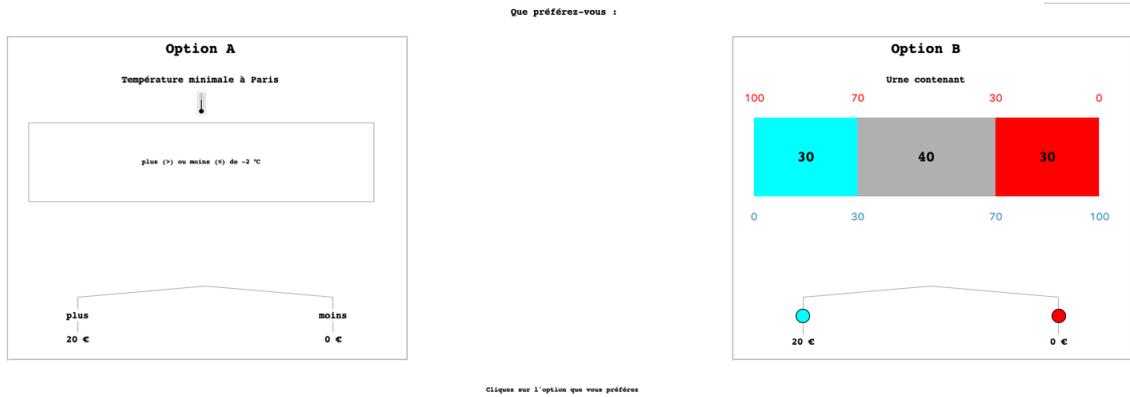
$MP(E_t) +$ $MP(E_t^c)$ Event $t =$	count	mean	std	min	25%	50%	75%	max
-2	80	0.84	0.20	0.29	0.71	0.89	0.98	1.31
2	80	0.85	0.20	0.29	0.73	0.89	0.99	1.34
5	80	0.83	0.22	0.24	0.69	0.89	0.99	1.29
8	80	0.83	0.18	0.39	0.69	0.89	0.99	1.26

Table 2.26:  $MP(E) + MP(E^c)$  descriptive statistics for all events for which MPs were elicited (those concerning Paris temperature in EXP1).

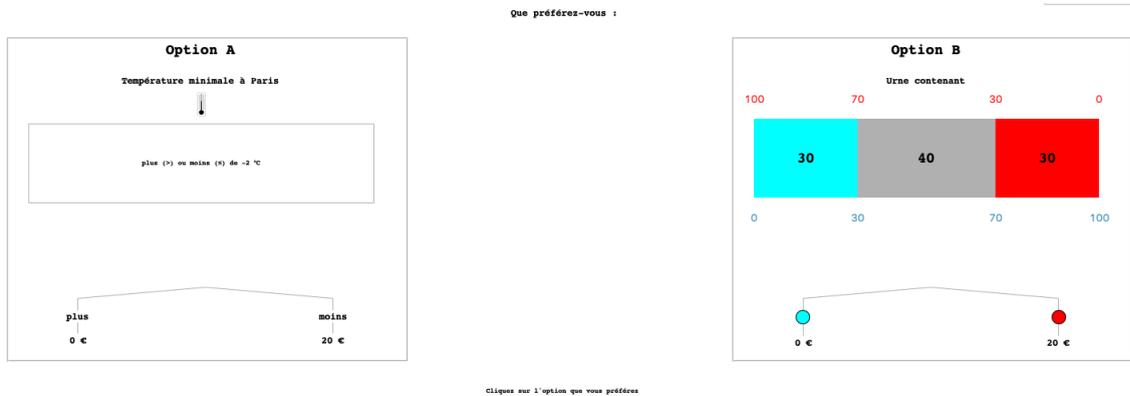
Table 2.26 displays the descriptive statistics for the sum  $MP(E) + MP(E^c)$  for the Paris source in EXP1. It is clear that the vast majority of subjects have a sum of MPs less than 100 1 indicating an  $\alpha$  greater than 0.5. Indeed, over 80% of subjects have sum of MPs less than or equal to 1.

## 2.9. Experimental design and displays

### 2.9.1. Probability interval elicitation: displays



(a) Bet on the complement of the event  $E_{-2}$



(b) Bet on the event  $E_{-2}$

Figure 2.16: Displays for a step in the binary-choice procedure

Figure 2.16 shows the display in a typical step of the binary-choice procedure. Specifically, the two figures show the two choice questions making up the step, involving bets on complementary events (temperature below vs above; bet on red vs blue).

At the end of the binary choice procedure, the two-cursor scrollbar, realising the 2D-choice list described in Section 2.2.6, is displayed, and the subject is invited to verify all choices, and correct them if required, prior to confirmation. The top



(a) First display



(b) Moving the blue cursor to confirm choices

Figure 2.17: Two-cursor scrollbar confirmation screen, implementing the 2D choice list

pane of Figure 2.17 shows a typical confirmation screen that appears at the end of the binary-choice procedure, where the retained values for red and blue balls are 66 and 29@ respectively. The red lines below then above the bar indicate that, for an urn with at least 29 blue balls and a minimum number of red balls greater than 66, option B (the bet on red from the urn) is preferred over A (the bet on the temperature being less than  $-2^{\circ}\text{C}$ ), whereas when there are at least 29 blue urns and the minimum number of red balls is less than 66, option A is preferred over B. The blue lines indicate the same for how preferences vary over urns with at least 66 red balls, for different minimum numbers of blue balls.

The bottom pane of Figure 2.17 illustrates a situation where the subject is verifying preferences as the number of blue balls vary, by moving the blue cursor (which is thus highlighted). The red cursor is kept fixed at its provisional value,<sup>24</sup> and, for each position of the blue cursor, the choice between the bet on the event (temperature greater than  $2^{\circ}$ ) and the bet on the urn with the specified minimum number of blue balls and at least 66 red balls is presented. The chosen option (as per Section 2.2.6) is indicated. In this case, the subject prefers the bet on the event over that on the urn, when only 3 balls are guaranteed to be blue. He may change this choice by clicking on the other option, or on the cursor. To confirm his response for the event, the subject has to scroll the blue cursor across the entire confirmation line, scanning all the choices, and likewise for the red cursor.

In EXP 2, there was a final confirmation screen after the elicitation for all events in a given source, presented in Figure 2.18. All interval-valued urns corresponding to the choices made and confirmed by the subject for the source are presented on the left. They are graphically depicted on the right: the red line shows the minimum number of red balls for each event (mark, in the case of this source), whereas the blue line plots 100 minus the minimum number of blue balls. To change a choice, a subject can either click on the choice on the right hand plot or on the corresponding urn in the sidebar on the left. By doing so, she returns to the corresponding two-cursor scrollbar confirmation screen, as in Figure 2.17. She may modify her choices on this screen as described above, and must reconfirm

<sup>24</sup>If the subject tries to move the red cursor, the blue cursor returns to its provisional value, and remains fixed there whilst the red cursor is being moved.

before proceeding.

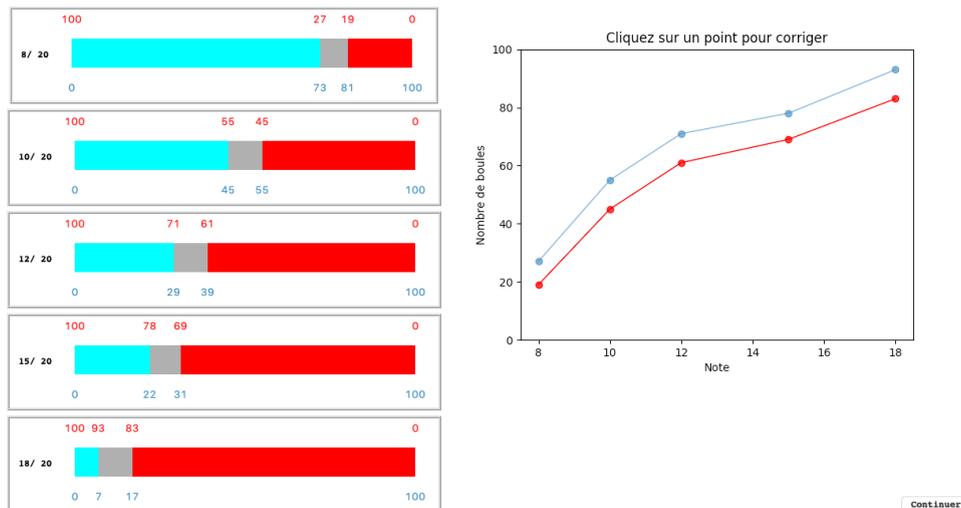
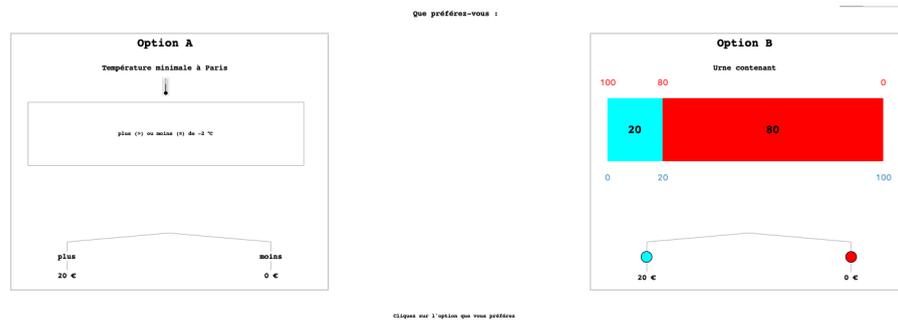


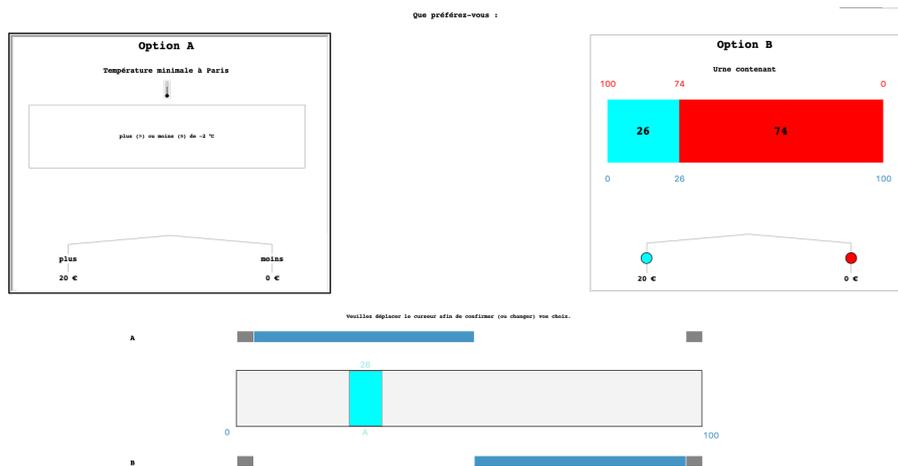
Figure 2.18: Omnibus confirmation screen

### 2.9.2. MP elicitation: displays

Figure 2.19 shows the displays for a typical choice in the MP elicitation (top pane) and the confirmation screen (bottom pane). These are comparable to the displays for probability interval elicitation, with the exception that a standard single-cursor scrollbar is used for confirmation. For the latter, as for the probability interval confirmation screen, the subject may use the cursor to scan choices and may click on the relevant option to modify her choice. She must scan all choices before confirming.



(a) Binary-choice procedure



(b) Confirmation scrollbar

Figure 2.19: MP displays

## 2.10. Lemmas for the proof of Proposition 4

In the following Lemmas, we suppose that preferences are represented according to (2.1) with  $\alpha > \frac{1}{2}$ , with  $E$  the event of interest with the subjective probability interval  $[\underline{p}(E), \bar{p}(E)]$ .

**Lemma 1.** *For every  $m \in [0, 0.5]$ :*

- *If  $\sigma_1([\underline{p}(E), \bar{p}(E)]) < m$ , there exists  $[p, q] \in R - B$  with  $\sigma_1([p, q]) = m$ , but no  $[p, q] \in W$  with  $\sigma_1([p, q]) = m$ ;*
- *If  $\sigma_1([\underline{p}(E), \bar{p}(E)]) > m$ , there exists  $[p, q] \in W$  with  $\sigma_1([p, q]) = m$ , but no  $[p, q] \in R - B$  with  $\sigma_1([p, q]) = m$ ;*
- *If  $\sigma_1([\underline{p}(E), \bar{p}(E)]) = m$ , each  $[p, q]$  with  $\sigma_1([p, q]) = m$  and  $[p, q] \neq [\underline{p}(E), \bar{p}(E)]$  belongs to either  $R$  or  $B$ .*

*Proof.* Straightforward to check from the representation (2.1) and the definition of  $\sigma$  (2.12). (See also Figures 2.1 and 2.9.) □

**Lemma 2.** *If  $\underline{p}(E) = \bar{p}(E)$  and Case 1 arrives at a point  $[p_i, q_i]$  with  $p_i = q_i$ , then the procedure remains in Case 1, and  $[\underline{p}_n, \bar{p}_n] \rightarrow [\underline{p}(E), \bar{p}(E)]$  as  $n \rightarrow \infty$ .*

*Proof.* Once the procedure reaches the subspace of precise probabilities, it executes a standard bisection procedure (lines 13–16, Figure 2.11). □

**Lemma 3.** *Suppose that the procedure reaches a point  $[p_i, q_i]$  in  $R$ - $B$  or  $W$ . Then the sequence  $\sigma_1([p_n, q_n]) \rightarrow \sigma_1([\underline{p}(E), \bar{p}(E)])$  as  $n \rightarrow \infty$ .*

*Proof.* Consider a stage  $i$  in the procedure where a point has just been found in

R-B or W. So the area containing the MPI is  $\Phi_i$  (Eq. (2.14)). Let

$$\begin{aligned} m_i^W &= \min \sigma_1(\Phi_i) \\ &= \sigma_1([\max \{p' : [p', q'] \in El_n \cap W\}, \min \{q' : [p', q'] \in El_n \cap W\}]) \\ m_i^{RB} &= \max \sigma_1(\Phi_i) \\ &= \begin{cases} \sigma_1([\min \{p'' : [p'', q''] \in El_n \cap R - B\}, \max \{q'' : [p'', q''] \in El_n \cap R - B\}]) & El_n \cap R - B \\ 0.5 & otherwise \end{cases} \end{aligned}$$

and  $|\Phi_i| = m_i^{R-B} - m_i^W$ . The latter is the maximum difference in  $\sigma_1$  values across all pairs of points in  $\Phi_i$ . In the first two subcases of Case 3 (lines 35-39), the next probability interval elicited is

$$\begin{aligned} [p_{i+1}, q_{i+1}] &= \frac{1}{2}[\max \{p' : [p', q'] \in El_n \cap W\}, \min \{q' : [p', q'] \in El_n \cap W\}] \\ &\quad + \frac{1}{2}[\min \{p'' : [p'', q''] \in El_n \cap R - B\}, \max \{q'' : [p'', q''] \in El_n \cap R - B\}] \end{aligned}$$

In the second subcase of Case 2 (lines 22-23), where a point in W has been found, but no point in R-B, the next probability interval elicited is

$$\begin{aligned} [p_{i+1}, q_{i+1}] &= \frac{1}{2} \left[ \frac{1}{2} \left( \begin{array}{c} \min \{p'' : [p'', q''] \in El_n \cap W\} + \\ \max \{q'' : [p'', q''] \in El_n \cap W\} \end{array} \right), \frac{1}{2} \left( \begin{array}{c} \min \{p'' : [p'', q''] \in El_n \cap W\} + \\ \max \{q'' : [p'', q''] \in El_n \cap W\} \end{array} \right) \right] \\ &\quad + \frac{1}{2}[\min \{p'' : [p'', q''] \in El_n \cap W\}, \max \{q'' : [p'', q''] \in El_n \cap W\}] \end{aligned}$$

where  $\frac{1}{2}(\min \{p'' : [p'', q''] \in El_n \cap R - B\} + \max \{q'' : [p'', q''] \in El_n \cap R - B\})$ ,  $\frac{1}{2}(\min \{p'' : [p'', q''] \in El_n \cap W\} + \max \{q'' : [p'', q''] \in El_n \cap W\})$  is the point on the diagonal of precise probabilities (i.e. degenerate probability intervals) that is closest to  $[\min \{p'' : [p'', q''] \in El_n \cap W\}, \max \{q'' : [p'', q''] \in El_n \cap W\}]$  (it is on the downwards sloping  $45^\circ$  line from  $[\min \{p'' : [p'', q''] \in El_n \cap W\}, \max \{q'' : [p'', q''] \in El_n \cap W\}]$ ). So this point has  $\sigma_1$ -value 0.5.

In first subcase of Case 2 (lines 20-21), where a point in R-B has been found, but no point in W, the next probability interval elicited,  $[p_{i+1}, q_{i+1}]$ , is a  $\frac{1}{2} - \frac{1}{2}$  mix of  $[\min \{p'' : [p'', q''] \in El_n \cap R - B\}, \max \{q'' : [p'', q''] \in El_n \cap R - B\}]$  with

$$\left\{ \begin{array}{l} \left[ \begin{array}{l} \min \{p'' : [p'', q''] \in El_n \cap R - B\} + \\ \max \{q'' : [p'', q''] \in El_n \cap R - B\} - 1 \end{array} , 1 \right] \\ \left[ 0, \begin{array}{l} \min \{p'' : [p'', q''] \in El_n \cap R - B\} + \\ \max \{q'' : [p'', q''] \in El_n \cap R - B\} \end{array} \right] \end{array} \right. \begin{array}{l} \min \{p'' : [p'', q''] \in El_n \cap R - B\} \\ + \max \{q'' : [p'', q''] \in El_n \cap R - B\} \\ \min \{p'' : [p'', q''] \in El_n \cap R - B\} \\ + \max \{q'' : [p'', q''] \in El_n \cap R - B\} \end{array} \begin{array}{l} > 1 \\ \leq 1 \end{array}$$

which is the point on the upper boundary (with either lower bound for the probability interval 0 or upper bound 1) that is on the downwards sloping 45° line from  $[\min \{p'' : [p'', q''] \in El_n \cap R - B\}, \max \{q'' : [p'', q''] \in El_n \cap R - B\}]$ . This point has  $\sigma_1$ -value 0.

Clearly, in all cases,  $m_i^W < \sigma_1([p_{i+1}, q_{i+1}]) < m_i^{RB}$ . Moreover, by the rest of the subcases in Cases 2 & 3, if this point is not in R-B or W, all the subsequent points elicited will have the same  $\sigma_1$ -value as  $[p_{i+1}, q_{i+1}]$ . And whenever a point in R-B is found, the next area containing the MPI,  $\Phi_{i+1}$ , will have the same minimum  $\sigma_1$ -value  $m_i^W$ , but its maximum value will be replaced by  $\sigma_1([p_{i+1}, q_{i+1}])$ . By Lemma 4, it follows that

$$\begin{aligned} |\Phi_i| \cdot \frac{m_i^W}{m_i^{RB} + m_i^W} &\leq |\Phi_{i+1}| \\ &\leq |\Phi_i| \cdot \frac{1 - m_i^W}{(1 - m_i^{RB}) + (1 - m_i^W)} \end{aligned}$$

Similarly, whenever a point in W is found, the next area containing the MPI,  $\Phi_{i+1}$ , will have the same maximum  $\sigma_1$  value  $m_i^{RB}$ , but its minimum value will be replaced by  $\sigma_1([p_{i+1}, q_{i+1}])$ , whence

$$\begin{aligned} |\Phi_i| \cdot \frac{1 - m_i^{RB}}{(1 - m_i^{RB}) + (1 - m_i^W)} &\leq |\Phi_{i+1}| \\ &\leq |\Phi_i| \cdot \frac{m_i^{RB}}{m_i^{RB} + m_i^W} \end{aligned}$$

Since, for any  $j > i$ ,  $m_j^{RB} \leq m_i^{RB}$  and  $m_j^W \geq m_i^W$ , for any such  $j$ ,  $\frac{1 - m_j^W}{(1 - m_j^{RB}) + (1 - m_j^W)} \leq \frac{1 - m_i^W}{(1 - m_i^{RB}) + (1 - m_i^W)}$  and  $\frac{m_j^{RB}}{m_j^{RB} + m_j^W} \leq \frac{m_i^{RB}}{m_i^{RB} + m_i^W}$ . So, for any  $j = i + k$  with  $k \in \mathbb{N}$ ,  $k \geq 1$ ,  $|\Phi_j| \leq \left( \max \left\{ \frac{1 - m_i^W}{(1 - m_i^{RB}) + (1 - m_i^W)}, \frac{m_i^{RB}}{m_i^{RB} + m_i^W} \right\} \right)^k \cdot |\Phi_i|$ . So the sequence  $[m_n^W, m_n^{RB}]$  is a bisection-like sequence of decreasing intervals (in the sense of containment), each of which contains  $\sigma_1([p(E), \bar{p}(E)])$ . Moreover, by the previous observation, whenever a point  $[p, q]$  is found in  $W$  with  $\sigma_1([p, q]) > 0$ , then the sequence  $|\Phi_n| = m_n^{RB} - m_n^W \rightarrow 0$  as  $n \rightarrow \infty$ , so  $\sigma_1([p_n, q_n]) \rightarrow \sigma_1([p(E), \bar{p}(E)])$  as required. (Recall that  $0.5 \geq m_n^{RB} \geq m_n^W \geq 0$  for all  $n$ .)

We now separate two cases, according to whether  $\sigma_1([p(E), \bar{p}(E)]) = 0$  or not. Suppose first that  $\sigma_1([p(E), \bar{p}(E)]) = \delta > 0$ . We show that the procedure will either arrive at a point with  $\sigma_1$ -value  $\delta$ , or a point in  $W$ . At a stage  $i$  in the procedure where no points in  $W$  have been found, but a point in R-B has,  $m_i^W = 0$  and  $0.5 \geq m_i^{RB} > 0$ . At each subsequent stage, by Lemma 1, either i. no point is found in  $W$  or R-B; ii. a point is found in  $W$  or R-B, and the next such point is in  $W$ ; iii. a point is found in  $W$  or R-B, and the next such point is in R-B. In case ii., the claim is established; in case i., by Lemma 1, the procedure is examining points with  $\sigma_1$ -value  $\delta$ , and the claim is established. Assume for reductio that at all such stages, the  $\sigma_1$ -value of the explored points is not  $\delta$ , but no point in  $W$  is found—i.e. we are always in case iii. Then, by the previous observations, for every  $j = i + k$  with  $k \in \mathbb{N}$ ,  $k \geq 1$ ,  $|\Phi_j| \leq \left( \frac{1 - m_i^W}{(1 - m_i^{RB}) + (1 - m_i^W)} \right)^k \cdot |\Phi_i| = \left( \frac{1}{2 - m_i^{RB}} \right)^k \cdot m_i^{RB}$ . Hence  $|\Phi_j| = m_j^{RB} \rightarrow 0$ , contradicting the fact that there are no points with  $\sigma_1$ -value less than  $\delta$  in R-B. Hence the procedure eventually finds a point in  $W$ . By the previous observation it follows that  $\sigma_1([p_n, q_n]) \rightarrow \sigma_1([p(E), \bar{p}(E)])$  as required.

Now consider the case where  $\sigma_1([p(E), \bar{p}(E)]) = 0$ . By Lemma 1, whenever the procedure searches for a point on a line  $\sigma_1([p, q]) = m > 0$ , it will find a point

in R-B. Hence, by the previous argument, it produces a sequence of points  $[p_n, q_n]$  in R-B, defining  $\Phi_n$  and associated  $[m_n^W, m_n^{RB}]$ , with  $m_n^W = 0$  and  $m_n^{RB} \rightarrow 0$ , as required.  $\square$

**Lemma 4.** *Let  $[p_W, q_W]$  be a point in  $W$ , with  $\sigma_1([p_W, q_W]) = m_W$  and suppose that the line  $\sigma_1([p, q]) = m_{R-B}$  contains a point in R-B but not in  $W$ . Then, for any point  $[p_{R-B}, q_{R-B}] \in R - B$  with  $\sigma_1([p_{R-B}, q_{R-B}]) = m$*

$$\sigma_1\left(\left[\frac{p_W + p_{R-B}}{2}, \frac{q_W + q_{R-B}}{2}\right]\right) \in \left[\frac{2m_W \cdot m_{R-B}}{m_W + m_{R-B}}, \frac{m_W(1 - m_{R-B}) + m_{R-B}(1 - m_w)}{(1 - m_{R-B}) + (1 - m_w)}\right]$$

Moreover, the same holds for a given point  $[p_{R-B}, q_{R-B}] \in R - B$  and any point  $[p_W, q_W] \in W$  on the line  $\sigma_1([p, q]) = m_W$ .

*Proof.* We first restrict attention to points  $[p, q]$  with  $p < 1 - q$  (or, in the polar-style coordinate system,  $\alpha < \frac{1}{2}$ ). For any points  $[p_1, q_1]$  and  $[p_2, q_2]$ , written in polar-style coordinate system as  $(m_1, \alpha_1)$  and  $(m_2, \alpha_2)$ , by (2.12) and (2.13), the midpoint (in Cartesian coordinates),  $\frac{1}{2}[p_1, q_1] + \frac{1}{2}[p_2, q_2]$  is  $\left(\frac{\alpha_1 m_1 + \alpha_2 m_2}{\alpha_1 + \alpha_2}, \frac{\alpha_1 + \alpha_2}{2}\right)$  in the polar system. Written in the polar coordinate system, let  $[p_W, q_W]$  be  $(m_W, \alpha_W)$ ; the points on the line  $\sigma_1([p, q]) = m_{R-B}$  are  $(m_{R-B}, \alpha)$ , for varying  $\alpha$ . Note that, by Proposition 2,  $m_{R-B} > m_W$ . It follows from representation 2.1 that  $(z, [p', q'], 0) \prec (z, [p, q], 0)$  whenever  $q' < q$  and  $p' < p$ , whence, since  $[p_W, q_W] \in W$ , we have that  $(z, [p', q'], 0) \prec (z, E, 0)$  for all  $q' < q_W$  and  $p' < p_W$ , so such points are not in R-B. So any point  $[p, q]$  on  $\sigma_1([p, q]) = m_{R-B}$  which is in R-B is such that  $p \geq p_W$ . By a similar argument (using the fact that  $(0, [p', q'], z) \prec (0, E, z)$  for all  $q' > q_W$  and  $p' > p_W$ ), any point  $[p, q]$  on  $\sigma_1([p, q]) = m_{R-B}$  which is in R-B is such that  $q \geq q_W$ . So any point  $[p, q]$  on  $\sigma_1([p, q]) = m_{R-B}$  which is in R-B has  $\alpha > \frac{\alpha_W m_W}{m_{R-B}}$  (where, by 2.13, this is in the  $\alpha$  of the point on  $\sigma_1([p, q]) = m_{R-B}$  with  $p = p_W = 2\alpha_W m_W$ ); similarly, any such point has  $\alpha < \frac{\alpha_W(1 - m_W)}{(1 - m_{RB})}$ . Plugging these bounds into the expression for the midpoint yields the result. Similar calculations yield the same result for the cases of  $p > 1 - q$  for some or all of the point considered. Finally, analogous arguments establish the conclusion for  $[p_{R-B}, q_{R-B}] \in R - B$  fixed and  $[p_W, q_W] \in W$  on the line  $\sigma_1([p, q]) = m_W$ .  $\square$

### 3. Communicating over climate risk<sup>1</sup>

#### Abstract

It is often argued that insufficient climate change awareness is responsible for its limited mitigation. This paper discusses that claim in the context of a game of contribution to a public bad, green house gases (GHG) emissions, under uncertainty. Contributors receive information regarding potential climate damages from a benevolent expert. Because of climate science complexity, I assume that expert claims are non-verifiable and that communication is cheap-talk. I show that no information transmission can happen at equilibrium. As a result, the contributors' expected social welfare is always lower than under perfect information. This is because, in this game, communication can only happen if the equilibrium over-provision of GHG emissions is not too high. This result highlights the necessity of commitment power for the informed party when it comes to communicating over climate risk.

---

<sup>1</sup>I would like to thank Jean-Marc Tallon, Stéphane Zuber, Frédéric Koessler, Valentina Bossetti, Marco Ottaviani, H el ene Ollivier, Loic Berger, Francis Bloch, Larry Samuelson and Guillaume Pomme y for helpful discussions. I also thank seminar and conference participants at PSE (TOM, SRE), Bocconi and FUR 2018 and Bocconi University for its hospitality. Financial support through ANR CHOp (ANR-17-CE26-0003), ANR ADE (ANR-18-ORAR-0005-01), ANR INDUCED (ANR-17-CE03-0008) and EUR PGSE is gratefully acknowledged.

### 3.1. Introduction

For decades, there has been scientific consensus on human responsibility regarding global warming. Considerable effort and international coordination has been devoted to raising awareness on the negative consequences of green house gases (GHG) emissions on the environment, a task most famously embodied by the Intergovernmental Panel on Climate Change's (IPCC) creation in 1988. Climate awareness has generally been considered as an obvious factor in favour of public willingness to act on the matter (see for instance [Halady and Rao, 2010](#); [Lee et al., 2015](#)) and often raised as top policy recommendation (see [Douenne and Fabre, 2020](#), for a recent example). Yet, interestingly, the connection between information provision to the general public regarding climate risk and the reduction of GHG emissions has rarely been discussed from a theoretical perspective—a question this paper aims to address.

Global warming is undoubtedly one of the most dramatic example of strategic behaviours leading to overall inefficiency. Because individuals fail to internalise the negative consequences of their GHG emissions on others, the non-cooperative equilibrium level of emission is higher than the socially optimal one. Although it is known since [Blackwell \(1950\)](#) that information has always positive value in Bayesian decision problems, the value of information in the context of strategic interactions is a more subtle question. In games, information can have a negative effect on social welfare.<sup>1</sup> In a seminal paper, [Angeletos and Pavan \(2007\)](#) shows that for a large class of games, information can decrease social welfare. Although the game I consider in this paper does not fall within this class, it is also a case where information does not always have positive value for social welfare.<sup>2</sup>

In this paper, I consider a game of contribution to a public bad where there is uncertainty over climate damages resulting from global warming. Contribution to GHG emissions is made through positive energy consumption. Contributors benefit from energy consumption to the extent that consumption goods depend

---

<sup>1</sup>For an early example of this see [Hirshleifer \(1978\)](#).

<sup>2</sup>An illustration of this is given in section 3.4.

on carbon-intensive energies to be produced, but expose themselves to potential climate damages. Before choosing their energy consumption levels, contributors receive expert advice regarding climate risk. The expert is known to be benevolent and aim for socially optimal energy consumption. Yet, because of the complexity of science, the information he provides is not verifiable by contributors. More precisely, his channel of communication is cheap-talk. The main result of this paper is that under this channel of communication, no information transmission can happen from the sender to the receivers. As a result, this situation is welfare decreasing compared to perfect information. This is because in this game, communication can only happen if the equilibrium over-provision of GHG emissions is not too high. Yet, this is only the case in the degenerate situation where receivers' optimal emission levels are null.

The key assumption underlying this result is that when experts communicate over climate science to the general public, their claims are non-certifiable. Although scientific results regarding climate change are well established for scientists, for non-scientists the supporting evidence is simply too complex to be acquired. One leading example is the *black box* prospective computer simulations over which the estimation of the effects of GHG on global temperature heavily rely. Not only is the process through which these simulations provide predictions obscure: as pointed out by [Pidgeon and Fischhoff \(2011\)](#), black box simulations are hardly considered as convincing evidence, even for scientists whose disciplines use observational methods. Surveying a large sample of citizens in Europe and the USA, [Lorenzoni and Pidgeon \(2006\)](#) shows that most people have limited knowledge over climate change and mostly relate to it through trust in experts.

But when non-experts are involved in a game of contribution to a public bad, there are good reasons to doubt that they blindly trust experts. This is because in that context, individual contributors interests are not aligned with those of a benevolent expert. Said otherwise, there is a normative difference between the informed party, who aims for social welfare, and the non-informed one, who aims for his own private good. In recent work, [Gabel et al. \(2021\)](#) provides evidence in line with that concern. Their paper shows that American conservatives display less confidence in science than liberals because scientists are believed to prioritise

regulation over individual freedom—a normative choice over which they differ. [Ehret et al. \(2018\)](#) and [Van Boven et al. \(2018\)](#) also provide evidence on how distrust for climate science can emerge from political partisanship and perceived normative misalignment. When information is non-certifiable and the informed party has a difference of interest with the non-informed one, full revelation becomes doubtful.

The assumption that climate science is non-certifiable information does not necessarily lead to cheap-talk communication. One could for instance assume that scientists could commit to some signal structure in the manner of [Kamenica and Gentzkow \(2011\)](#)'s information design approach. Yet, cheap-talk communication is generally considered as the *weakest* form of communication for the expert because no commitment is possible for him.<sup>3</sup> As information can have negative value in the game I consider, it is natural to ask if, under the weakest form of non-certifiable communication, information can be provided in a welfare improving way.

This paper relates to several strands of literature. First, it relates to the literature on international environmental agreements (IEA), where emission policies regarding climate change are negotiated between countries. They can formally be seen as a game of contribution to a public bad. Yet, in these models, players can agree to form a coalition to mitigate GHG emissions. Various concepts of coalition are used to model welfare improving counterparts to the simple Nash equilibrium of the game (see [Finus, 2008](#), for a review). In this paper, I choose to focus on the latter, as it serves as reference point to all the coalitional achievable agreements. As a consequence, the modelling of the emission game is similar to the canonical model of [Bergstrom et al. \(1986\)](#). My model also integrates the uncertainty surrounding climate change, which is essential in that context.<sup>4</sup> In the IEA literature, [Na and Shin \(1998\)](#) were the first to take this route. [Finus and Pintassilgo \(2013\)](#) added risk behaviour to the picture. [Asheim \(2010\)](#) introduces

---

<sup>3</sup>[Min \(2017\)](#) provides a general argument in favour of that claim. It shows that if one assumes that the sender could commit to a signal structure, even with some limited probability, then he can design information in order to obtain a higher expected pay-off than under pure cheap-talk communication. As in my model the sender is benevolent, this result implies that more commitment power for the sender necessarily implies a higher expected social welfare compared with cheap-talk.

<sup>4</sup>On that topic and for a review of its economic modelling, see [Pindyck \(2007\)](#).

a benevolent environmental agency which can strategically reveal certifiable information over climate risk. My paper is the first to introduce non-certifiable information to the picture.

Second, this paper relates to the literature on strategic communication, starting with [Crawford and Sobel \(1982\)](#) cheap-talk model with one sender and one receiver. [Goltsman and Pavlov \(2008\)](#) study cheap-talk with public and private communication with multiple receivers. Yet, unlike in my paper, these actions of these receivers do not affect each other. [Galeotti et al. \(2013\)](#) studies cheap-talk communication in the context of networks, where multiple imperfectly informed senders are also decision-makers and can influence each other both by their actions and messages.

Section [3.2](#) introduces the base model. Section [3.3](#) solves the game of contribution to a public bad for any given message of the informed party. Section [3.4](#) discusses the welfare implications of uncertainty and in particular shows that information doesn't always have positive value. Section [3.5](#) establishes that no information can be communicated by the sender at equilibrium. The appendix contains all the proofs.

## 3.2. Setup

I consider a game between a scientific authority acting as a sender  $S$  (he) and  $N$  receivers (she)  $R_i$  having to choose a level of energy consumption  $e_i \geq 0$ . When consuming  $e_i$  units of energy, receiver  $i$  emits  $\alpha e_i$  units of GHG emissions, where  $\alpha > 0$  is called the level of carbon dependence. This is because the higher  $\alpha$ , the higher the level of GHG emissions for a unit of energy. The set of possible actions for each receiver is thus  $\mathcal{A} = \mathbb{R}^+$ . For a given level of total GHG emissions, there is uncertainty on the severity of damages suffered by receivers. Let  $\Omega = [0, 1]$  be the possible global warming scenarios, where climate damages are increasing with  $\omega$ . I will refer to them as the states. The scientific authority  $S$  learns the state from Nature, but cannot certify his information. He can send a costless message

$m \in [0, 1]$  to the receivers indicating a subset of states where the true state might be. The timing of the game is as follows:

1. Nature draws the state  $\omega_0$  according to a uniform distribution over  $\Omega$  of density  $g$ .
2.  $S$  is privately informed of  $\omega_0$ .
3. **Communication stage:** the sender sends a message to the receivers regarding its type.
4. **Emission stage:** the receivers choose simultaneously a level of energy consumption  $e_i$

Receiver  $i$ 's utility function will be:

$$u_i(e_i, e_{-i}, \omega) = e_i - \frac{1}{2} \left( \sum_{i=1}^N \alpha e_i + \omega \right)^2$$

Let us decompose the components of this utility function. Energy consumption positively impacts receivers' utility through:

$$b(e_i) = e_i$$

$b$  represents the fact that energy consumption correlates with economic consumption which increases receivers' utility. Yet, it is also responsible for GHG emission and potential damages due to global warming. This is captured through:

$$d(e_i, e_{-i}, \omega) = \frac{1}{2} \left( \sum_{i=1}^N \alpha e_i + \omega \right)^2$$

$d$  represents the cost of energy consumption. It is increasing with the total

level of emissions. The higher  $\omega$ , the more severe the consequences of global warming and the greater the cost of emissions. Finally, notice that  $\alpha$  is an exogenous variable over which contributors have no control. Yet, its value will be useful to our analysis, for the purpose of comparative statics.

Thus, overall:

$$u_i(e_i, e_{-i}, \omega) = b(e_i) - d(e_i, e_{-i}, \omega)$$

Receiver  $i$  does not take into account the impact of her emissions on other receivers. As a result, emissions are a public bad and its collective consequences are not internalised by players. The sender seeks to maximise social welfare. That is:

$$\begin{aligned} u_S(e_1, \dots, e_N, \omega) &= \sum_{i=1}^N u_i(e_i, e_{-i}, \omega) \\ &= \sum_{i=1}^N e_i - \frac{N}{2} \left( \sum_{i=1}^N \alpha e_i + \omega \right)^2 \end{aligned}$$

The choice of a receiver's energy consumption is the result of a trade-off between economic growth and potential damages created by global warming. They are determined by the state and the level of carbon dependence. The more severe climate damages (the higher the state), the lower the optimal level of energy consumption in order to mitigate the consequences of climate damages. Notice that utility functions are single-peaked. As a result, for a given state, there is a single optimal energy consumption level  $e_i(\omega)$ . A higher energy consumption level  $e > e_i(\omega)$  is not optimal for  $i$  because it might create too much climate damage. A lower energy consumption level  $e < e_i(\omega)$  is neither optimal for  $i$  as it implies to reduce economic growth by too much.

A strategy for  $S$  is  $\sigma : \Omega \rightarrow [0, 1]$  which consists in transmitting a message  $m$  to the receivers regarding its type. A strategy for a receiver consists in choosing an energy consumption level as a function of message  $m$ . An equilibrium consists in a signalling strategy  $\sigma(\omega)$  and an action rule for each receiver  $y_1(m), \dots, y_N(m)$  such that :

1.  $S$  chooses a strategy  $\sigma$  such that for all  $m \in [0, 1]$ :

$$u_S(y_1(\sigma(\omega)), \dots, y_N(\sigma(\omega)), \omega) \geq u_S(y_1(m), \dots, y_N(m), \omega)$$

2. Having received an equilibrium public message  $m \in \text{supp}(\sigma)$ ,  $R_i$  updates her prior using Bayes' rule such that:

$$g(\omega|m) = \begin{cases} \frac{g(\omega)}{\int_{\sigma^{-1}(m)} g(\omega) d\omega} & \text{if } \omega \in \sigma^{-1}(m) \\ 0 & \text{if not} \end{cases}$$

and chooses action  $y_i(m)$ , such that for all  $e \in \mathcal{A}$  :

$$\mathbb{E}(u_i(y_i(m), y_{-i}(m), \omega)|m) \geq \mathbb{E}(u_i(e, y_{-i}(m), \omega)|m)$$

where  $\mathbb{E}(u_i(e_i, e_{-i}, \omega)|m) = \int_{\omega \in \Omega} g(\omega|m) u_i(e_i, e_{-i}, \omega) d\omega$ . Any message  $m$  such that  $m \notin \text{supp}(\sigma)$  is interpreted as some equilibrium message  $m_* \in \text{supp}(\sigma)$ .

### 3.3. Emission stage

I start by focusing on the emission game. Consider any message  $m \in [0, 1]$  and set  $e_i(m)$ , for any  $i \in 1, \dots, N$ , the solution to the maximisation problem:  $\max_{e_i \in \mathbb{R}^+} E(u_i(e_i, e_{-i}(m)|m))$ .  $e_i(m)$  is the equilibrium level of energy consumption of receiver  $i$  having received message  $m$ .

First order condition gives that the equilibrium levels of energy consumption of the receivers must be such that:

$$\sum_{i=1}^N e_i(m) = \max\left(\frac{1}{\alpha^2} - \frac{\mathbb{E}(\omega|m)}{\alpha}, 0\right)$$

Having received message  $m$ , there is a continuum of equilibria in the emission game which all give the same total level of energy consumption. As my focus is on the latter quantity, and for the sake of simplicity, in the following I will restrict attention to symmetric individual equilibrium energy consumption levels. Thus, for any given equilibrium message  $m$ , the equilibrium symmetric energy consumption of a receiver is:

$$e(m) = \max\left(\frac{1}{N\alpha^2} - \frac{\mathbb{E}(\omega|m)}{N\alpha}, 0\right) \quad (3.1)$$

Similarly, set  $e_i^W(m)$ , for any  $i \in 1, \dots, N$ , the solution to the maximisation problem:  $\max_{e_i^W \in \mathbb{R}^+} \sum_{i=1}^N \mathbb{E}(u_i(e_i^W, e_{-i}^W(m), \omega)|m)$ .  $e_i^W(m)$  is the socially optimal level of energy consumption of receiver  $i$  having received message  $m$ .

First order condition gives that the socially optimal levels of energy consumption of the receivers must be such that:

$$\sum_{i=1}^N e_i^W(m) = \max\left(\frac{1}{N\alpha^2} - \frac{\mathbb{E}(\omega|m)}{\alpha}, 0\right)$$

As before, I will restrict attention to symmetric individual socially optimal energy consumption levels. For any given equilibrium message  $m$ , the symmetric

socially optimal energy consumption of a receiver is:

$$e^W(m) = \max\left(\frac{1}{(N\alpha)^2} - \frac{\mathbb{E}(\omega|m)}{N\alpha}, 0\right)$$

### 3.4. Welfare analysis

Before turning to the resolution of the communication stage, let's examine the two sources of loss of welfare in our model: over-provision of the public bad (energy consumption) and uncertainty regarding climate damages. In a first-best world, levels of energy consumption would be efficient. This could be because receivers could commit to adopt the socially optimal level of energy consumption through some mechanism. In addition, the sender would be able to certify his private information, something that would then be a dominant strategy<sup>5</sup>.

I will draw my attention to the expected welfare of a given receiver. Under the assumptions we have made, the social welfare can be directly derived by multiplying the expected welfare of a given receiver by  $N$ . For  $\alpha < 1$  we have that for any  $m \in [0, 1]$ ,  $e(m) > 0$ <sup>6</sup> and that for  $\alpha < \frac{1}{N}$  we have that for any  $m \in [0, 1]$ ,  $e^W(m) > 0$ . In the present section I'll focus on the case  $\alpha < \frac{1}{N}$ . In the first-best world, a receiver's energy consumption is the socially optimal one  $e^W(\omega) = \frac{1}{(N\alpha)^2} - \frac{\omega}{N\alpha}$  and the expected welfare of a given receiver is:

<sup>5</sup>In fact, in this case, even if information is non-certifiable, because there is no difference of interest between the sender and the outcome of the receiver's game, full revelation is a dominant equilibrium for the sender.

<sup>6</sup>Write  $e(\omega) = \frac{1}{N\alpha^2} - \frac{\omega}{N\alpha}$ . It must be that for any  $m \in [0, 1]$ , there is  $\tilde{\omega} \in \Omega$  such that  $e(m) = e(\tilde{\omega})$ . Yet,  $e(\omega) > 0$  for any  $\omega \in \Omega$  when  $\alpha < 1$ . The same reasoning applies for  $e^W(m) > 0$  resulting in condition  $\alpha < \frac{1}{N}$ .

$$\begin{aligned}
W_1(N, \alpha) &= \int_0^1 u_i(e^W(\omega), e^W(\omega), \omega) d\omega \\
&= \frac{1}{N\alpha} \left( \frac{1}{N\alpha} - \frac{1}{2} \right) - \frac{1}{2} \left( \frac{1}{N\alpha} \right)^2 \\
&= \frac{1}{2N\alpha} \left( \frac{1}{N\alpha} - 1 \right)
\end{aligned}$$

In the first-best world, the state is known. The socially-optimal action is implemented in every state, making the level of climate damage  $d$  state-independent. As a result, the expectation over state is only taken over the benefit term  $b$ .

Now consider a second-best world where receivers can no longer commit to adopt socially optimal levels of energy consumption but are constrained to non-cooperative levels. The sender certifies his private information and we are under perfect information. When  $\alpha < \frac{1}{N}$ , the resulting equilibrium energy consumption is  $e(\omega) = \frac{1}{N\alpha^2} - \frac{\omega}{N\alpha}$ . The expected welfare of a given receiver then is:

$$\begin{aligned}
W_2(N, \alpha) &= \int_0^1 u_i(e(\omega), e(\omega), \omega) d\omega \\
&= \frac{1}{N\alpha} \left( \frac{1}{\alpha} - \frac{1}{2} \right) - \frac{1}{2} \left( \frac{1}{\alpha} \right)^2 \\
&= W_1(N, \alpha) + \frac{N-1}{2} b(\alpha, N)
\end{aligned}$$

First, notice that as  $b(\alpha, N) < 0$ , non-cooperative energy consumption choices result in a decrease in expected welfare for  $N > 1$ . This is because, as energy is a public bad in our game, non-cooperative energy consumption is too high compared to social welfare. As a result, receivers' welfare is lower in the second best-world than in the first-best one. As before, the level of climate damage  $d$  is state-independent and the expectation over state is only taken over the benefit term  $b$ .

Let's step in a third-best world where receivers choose non-cooperative levels of energy consumption and, in addition, climate damage is uncertain. Here, receivers act entirely based on their prior. When  $\alpha < \frac{1}{N}$ , the expected welfare of a receiver then is:

$$\begin{aligned} W_3(N, \alpha) &= \frac{1}{N\alpha} \left( \frac{1}{\alpha} - \int_0^1 \omega d\omega \right) - \frac{1}{2} \int_0^1 \left( \frac{1}{\alpha} - \frac{1}{2} + \omega \right)^2 d\omega \\ &= \frac{1}{N\alpha} \left( \frac{1}{\alpha} - \frac{1}{2} \right) - \frac{1}{6} \left[ \left( \frac{1}{\alpha} + \frac{1}{2} \right)^3 - \left( \frac{1}{\alpha} - \frac{1}{2} \right)^3 \right] \\ &= W_2(N, \alpha) - \frac{1}{24} \end{aligned}$$

Notice that whatever the number of contributors  $N$  and the level of carbon dependence  $\alpha$ , uncertainty reduces the welfare compared to the second-best world. Full uncertainty is harmful for a receiver's expected welfare. As for the second-best world, the expected welfare of a contributor is decreasing with the number of total contributors  $N$ . Because of uncertainty, receivers are not able to make climate damages state-independent.

Now, notice that in our game, the value of information is not always positive. To see this assume that the sender could certify  $\omega_0 \in [\frac{1}{2}, 1]$  and that  $N = 2$ . When  $\alpha < \frac{1}{2}$ , the expected welfare of a receiver would be:

$$\begin{aligned} W_C(2, \alpha) &= \frac{1}{2\alpha} \left( \frac{1}{\alpha} - \int_{\frac{1}{2}}^1 \omega d\omega \right) - \frac{1}{2} \int_{\frac{1}{2}}^1 \left( \frac{1}{\alpha} - \frac{1}{2} + \omega \right)^2 d\omega \\ &= \frac{1}{2\alpha} \left( \frac{1}{\alpha} - \frac{3}{4} \right) - \frac{1}{6} \left[ \left( \frac{1}{\alpha} + \frac{1}{2} \right)^3 - \left( \frac{1}{\alpha} \right)^3 \right] \\ &= W_2(2, \alpha) - \frac{\alpha^2 + 12\alpha - 12}{48\alpha^2} \end{aligned}$$

For  $0 < \alpha < \frac{1}{2}$ , we have that  $W_C(2, \alpha) > W_2(2, \alpha)$  and a receiver is better-off

knowing only that  $\omega_0 \in [\frac{1}{2}, 1]$  than under perfect information.<sup>7</sup> Could a benevolent sender take advantage of the fact that information can have negative value to increase social welfare compared to the second-best world through cheap-talk communication ?

To answer this question, it is time to turn to the communication stage and see how the sender's influence places the receiver's welfare with respect to the first, second and third-best worlds.

## 3.5. Communication stage

### 3.5.1. Strictly positive energy consumption

I first consider the case where, whatever the state, the equilibrium and socially optimal energy consumption levels are strictly positive. When  $\alpha < \frac{1}{N}$ , energy consumption levels are strictly positive, and we have that:

$$e^W(m) = e(m) - \frac{N-1}{(N\alpha)^2}$$

Notice that when  $N = 1$  sender and receiver have exactly the same utility function and aim for the same total level of energy consumption. Yet when  $N > 1$ , energy consumption is always higher in the non-cooperative equilibrium than what would be socially optimal:  $e^W(m) < e(m)$ .

I now turn to the communication stage. I prove that the equilibrium messages and the resulting total level of energy consumption in the game I study can be derived through the analysis of a parallel one sender one receiver game.

<sup>7</sup>The exact constraint is  $\alpha < 4\sqrt{3} - 6$  but  $\frac{1}{2} < 4\sqrt{3} - 6$ .

**Proposition 11.** *Consider the parallel one sender, one receiver communication game where the set of types is  $\Omega = [0, 1]$  the action variable is  $a \in \mathbb{R}$ , the receiver has uniform prior over  $\Omega$ , the utility of the sender is:*

$$U_S(a, \omega) = -\left(a - \left(\frac{1}{N\alpha^2} - \frac{\omega}{\alpha}\right)\right)^2$$

*and the utility of the receiver is:*

$$U_R(a, \omega) = -\left(a - \left(\frac{1}{\alpha^2} - \frac{\omega}{\alpha}\right)\right)^2$$

When  $\alpha < \frac{1}{N}$ , a strategy of the sender is an equilibrium strategy if and only if it is an equilibrium strategy of the sender of the parallel game.

Let me introduce the definition of a partional equilibrium adapted from [Crawford and Sobel \(1982\)](#) Lemma 1:

**Definition 5.** *Define  $\{\omega_0, \dots, \omega_q\} \subseteq [0, 1]$  such that:*

- $0 = \omega_0 < \dots < \omega_q = 1$  where  $\omega_k$ , for  $0 \leq k \leq q$ , is called the  $k$ -th cut-off.
- $\cup_{k=1}^q [\omega_{k-1}, \omega_k] = [0, 1]$ , where  $[\omega_{k-1}, \omega_k)$ , for  $1 \leq k < q - 1$ , is called the  $k$ -th cell and  $[\omega_{q-1}, 1]$  the  $q$ -th cell.

A  $q$ -cut-off partition equilibrium is an equilibrium of the game where the signaling strategy of  $S$  is uniform on every cell. That is, for  $\omega \in [\omega_{k-1}, \omega_k)$ ,  $\sigma^*(\omega) = m_k$ , for  $1 \leq k \leq q - 1$  and for  $\omega \in [\omega_{q-1}, 1]$ ,  $\sigma^*(\omega) = m_{q-1}$ .

The proof of Proposition 11 works in two steps. I first show that all equilibria of the original game are partional. An important assumption for this to be

the case is that the sender of the original game cares only about the total energy consumption of the receivers, making the outcome variable unidimensional. As a result, the logic of the proof of Crawford and Sobel (1982) can be reproduced.

Then, I show that all equilibria of the original game and the parallel game coincide. Given that both sets of equilibria are partitional, one needs to show that for a given signalling strategy, any type of sender in the original game which is indifferent between two messages is also indifferent between the same two messages in the parallel game if he follows the same signalling strategy, and conversely. This is true because for a given type, the senders in the original and parallel games have the same optimal action and because their utility functions are both symmetric. As a result, for a given signalling strategy, any pair of messages that induce the same utility to the parallel sender also provide the same utility to the original sender. Thus, cut-off types in both games are the same, which suffices to identify partition equilibria.

Given that the equilibrium messages in the parallel game are the same as in the original one, that utility functions of the parallel game are symmetric around their maximum and that the prior is uniform over  $[0, 1]$ , the same procedure as in CS section 4 can be followed to derive cut-off types. That is, a  $n$ -cut-off equilibrium  $\omega_0^n, \dots, \omega_n^n$  is defined such that, for all  $k \in 1, \dots, n-1$ :

$$e^W(\omega_k^n) = \frac{1}{2} \left( \frac{1}{\omega_k^n - \omega_{k-1}^n} \int_{\omega_{k-1}^n}^{\omega_k^n} e(\omega) d\omega + \frac{1}{\omega_{k+1}^n - \omega_k^n} \int_{\omega_k^n}^{\omega_{k+1}^n} e(\omega) d\omega \right) \quad (3.2)$$

$$\iff \frac{1}{N\alpha^2} - \frac{\omega_k^n}{\alpha} = \frac{1}{2} \left[ \frac{1}{\alpha^2} - \frac{\omega_{k-1}^n + \omega_k^n}{2\alpha} + \frac{1}{\alpha^2} - \frac{\omega_k^n + \omega_{k+1}^n}{2\alpha} \right]$$

Given that  $\omega_0^n = 0$  and  $\omega_n^n = 1$ , the incentive constraints above imply that for  $k \in 1, \dots, n$ :

$$\omega_k^n = \frac{k}{n} + 2k(n-k)b(\alpha, N)$$

where  $b(\alpha, N) = -\frac{N-1}{N\alpha}$  is the difference between the total socially optimal emission level and the total non-cooperative one (note that the bias is strictly negative here). Because cut-offs are exactly as in the linear quadratic case of Crawford and Sobel (1982), their existence depends entirely on  $b(\alpha, N)$ . Therefore, I call  $b(\alpha, N)$  the communication bias. Notice that the communication bias is exactly the individual over-provision in public bad (emissions) compared to socially optimal level.

It follows that there a  $n$ -cut-off equilibrium exists if and only if all cut-offs are in  $[0, 1]$ , which is equivalent to:

$$\begin{aligned} & \omega_1^n > 0 \\ \Leftrightarrow & b(\alpha, N) > -\frac{1}{2n(n-1)} \\ \Leftrightarrow & \alpha > \frac{2n(n-1)(N-1)}{N} \end{aligned} \tag{3.3}$$

This constraint states that if receivers' carbon dependence is low enough, the sender cannot convey information through a  $n$ -cut-off equilibrium. In particular, if  $\alpha \leq \frac{4(N-1)}{N}$ , no information can be conveyed at all. In other words, the only equilibrium is the babbling one where all types send the same message.

The fact that the level of carbon dependence  $\alpha$  has to be high for information transmission to be possible can seem surprising at first sight. It comes from the fact that reducing carbon dependence decreases climate damages at a quadratic rate, whereas benefits from energy consumption only vary linearly. When  $\alpha$  is high, receivers reduce drastically their energy consumption because the damage they face is potentially very high. But when  $\alpha$  is low, receivers face little potential damage, neglecting emission reduction in favour of energy consumption. As a result, because they fail to internalise the consequences of energy consumption on the other receivers, they increase the inefficiency of the emission stage. Consequently, the equilibrium energy consumption gets further apart from the

socially optimal level, increasing the communication misalignment. As the communication stage is essentially equivalent to a special case of Crawford and Sobel (1982)'s linear quadratic example, the maximum number of cut-offs decreases with this misalignment. The assumption that damages are more convex in energy consumption than benefits is thus essential for the result we get.

Notice that interval of carbon dependence  $\alpha$  compatible with the existence of a  $n$ -cut-off equilibrium is smaller when the number of receivers  $N > 2$  increases. For a given symmetric level of individual energy consumption, more contributors who fail to internalise the consequences of energy consumption means more over-provision of the public bad. As a result, the misalignment between the non-cooperative equilibrium and the socially optimal level of energy consumption increases. In other words, increasing the number of receivers only makes communication more difficult for the sender.

Yet, recall that in order to focus on strictly positive socially optimal energy consumption levels, we have assumed that  $\alpha < \frac{1}{N}$ . It follows that a  $n$ -cut-off equilibrium exists if and only if:

$$\frac{2n(n-1)(N-1)}{N} < \alpha < \frac{1}{N}$$

For  $N > 1$  and  $n > 1$ , the interval of  $\alpha$  which verifies this inequality is empty. In other words, the only equilibrium is the babbling one.

To understand this result, consider the case of a 1-cut-off equilibrium. For it to exist, there must be  $\omega_1$  such that a sender of that type is indifferent between informing receivers that his type is below and above  $\omega_1$ . This is the incentive constraint of the sender given in (3.2) which states that the induced actions by those messages have to be at equal distance of the optimal action of the sender. Yet, for this to be possible, it must be that the communication bias  $b(\alpha, N)$ , that is the

individual over provision in public bad compared to the socially optimal level, is not too high. That bias is increasing with the number of receivers and decreasing with the level of carbon dependence. But even when  $N = 2$ , (3.3) gives that for a 1-cut-off equilibrium to exist it must be that  $\alpha > 2$ , a value for which socially optimal energy consumption is null. In other words, in the game I consider, when contributions must be strictly positive, the equilibrium over provision in public bad is always too important for a benevolent sender to be able to convey some information regarding climate damages.

Thus, the results of this section imply that under cheap-talk communication and strictly positive energy consumption levels, because no information transmission can occur, receiver's welfare is the same as in the third-best world. In other words, the absence of certification power for the sender, makes him unable to affect the receivers' welfare.

Notice that if the damage function was more convex than what I have assumed, the marginal damage of a unit of energy would be higher. Consider the following damage function:  $d(e_i, e_{-i}, \omega) = \frac{1}{\beta} (\sum_{i=1}^N \alpha e_i + \omega)^\beta$ . For  $\beta > 2$ , the inefficiency would increase and the communication bias as well. But the condition for the level of energy consumption to be strictly positive would stay the same, whatever  $\beta > 1$ , making the existence of a 1-cut-off equilibrium even harder. To the contrary, when the damage function is less convex, for  $1 < \beta < 2$ , the inefficiency of the emission stage decreases and the incentive constraint of the sender is easier to meet. Yet, this gain is still insufficient to overcome the constraint on positive energy consumption.

### 3.5.2. Null energy consumption

Now I consider the case of a 1-cut-off equilibrium when equilibrium and socially optimal energy consumption can be equal to zero. For it to exist, there must be  $\omega_1$  such that a sender of that type is indifferent between informing receivers that his type is below and above  $\omega_1$ . When  $\omega \geq \frac{1}{N\alpha}$ , the socially optimal energy con-

sumption is equal to zero. Given that the sender's utility function is symmetric and single-peaked, for  $\omega_1$  to exist, the induced actions must be strictly lower and higher than  $e^W(\omega_1)$ , the optimal action for a sender of that type. This is impossible when  $\omega_1 \geq \frac{1}{N\alpha}$  because no action can be strictly negative. Thus, we want to know if there is  $\omega_1 < \frac{1}{N\alpha}$  such that strategy  $\sigma_1 : \omega \rightarrow \mathbb{1}(\omega \geq \omega_1)$  is an equilibrium strategy<sup>8</sup>. For this to be the case it must be that:

$$\begin{aligned} u_S(e(0), \omega_1) &= u_S(e(1), \omega_1) \\ \iff e^W(\omega_1) &= \frac{1}{2}(e(0) + e(1)) \end{aligned}$$

The equilibrium energy consumption levels are strictly positive only for states below  $\bar{\omega} = \frac{1}{\alpha}$ . Therefore, the above is equivalent to:

$$\frac{1}{(N\alpha)^2} - \frac{\omega_1}{N\alpha} = \frac{1}{2} \left( \frac{1}{N\alpha^2} - \frac{\omega_1}{2N\alpha} + \frac{1}{(\bar{\omega} - \omega_1)} \left( \frac{1}{N\alpha^2} - \frac{1}{N\alpha} \int_{\omega_1}^{\bar{\omega}} \omega d\omega \right) \right)$$

computation gives that:

$$\omega_1 = \frac{\sqrt{(16a + 9)N^2 - 24N + 16} + N + 4}{4aN}$$

and that there is no  $\alpha > 0$  such that  $0 < \omega_1 < \frac{1}{N\alpha}$ . In other words, a 1-cut-off equilibrium is never possible, even allowing for null energy consumption levels. This result shows that in the game I consider, even when contributions can be null, a benevolent sender is unable to convey any information regarding climate damages. Although we have not computed welfare allowing for null emission

<sup>8</sup>Testing any other strategy giving a 1-cut-off separation would be equivalent

levels, this result also suggests that even then the absence of certification power for the sender, makes him unable to affect the receivers' welfare.

### 3.6. Conclusion

In this paper, I considered a game of contribution to a global temperature-increasing public bad, where there is uncertainty over potential climate damages. Contributors receive expert advice regarding climate risk from a benevolent expert through cheap-talk communication. I showed that this channel of communication is welfare decreasing compared to perfect information because the equilibrium over-provision of GHG emissions is always too high for information transmission to be possible.

This result highlights the necessity of stronger forms of communication when it comes to climate risk. Although the value of information can be negative in the game I consider, under cheap-talk only the lack of information transmission is always harmful for social welfare. The fact that no communication is possible under cheap-talk derives from the assumption that public bad contribution is always positive. Yet, allowing for negative contributions, which could represent investment in carbon capture technologies, could potentially change that result. This possibility is left for future research. Forms of communication with partial commitment, such as information design, could take advantage of the negative value of information aspect of the game. Addressing that question is also an important avenue for future research.

### 3.7. Appendix

#### Proof of Proposition 11:

Notice that  $U_i(a, \omega)$ ,  $i \in \{S, R\}$  is twice continuously differentiable, concave, admits a unique maximum in their action variable and that :

$$\frac{\partial^2 U_i(a, \omega)}{\partial a \partial \omega} > 0$$

Thus,  $U_i(t, \omega)$  verifies all requirements of utility functions in Crawford and Sobel (1982) such that all equilibria of the parallel game are partitional. Notice that the same requirements are met by  $u_i(e_i, e_{-i}, \omega)$ ,  $i \in \{1, \dots, N\}$  in their action variable  $e_i$ . The sender cares only about the total energy consumption, I identify his utility function to  $u_S(t, \omega) = Ct - \frac{N}{2}(\alpha t + \omega)^2$  where  $t$  is the total energy consumption level of receivers and let  $t_S(\omega)$  be the corresponding optimal action.  $u_S(t, \omega)$  meets the requirements of Crawford and Sobel (1982) with respect to  $t$ . Call  $t(\omega) = \sum_{i=1}^N e_i(\omega)$  the total equilibrium energy consumption level in state  $\omega$ . We have that  $t(\omega) \neq t_S(\omega)$  for all  $\omega \in \Omega$  so by directly adapting Crawford and Sobel (1982)'s lemma 1 one can show that all equilibrium of the game are partitional. I'll do it as follows: first, I show that the number of aggregate energy consumption levels of the receivers induced at equilibrium is finite (lemma 7). Then, I prove that the set of types which get the same equilibrium outcome must form an interval. The continuity and the strict monotonicity of the sender's preferences closes the argument.

**Lemma 7.** *There exists  $\epsilon > 0$  such that if  $u$  and  $v$  are actions induced in equilibrium,  $|u - v| \geq \epsilon$ . Further the set of aggregate energy consumption levels induced in equilibrium is finite.*

#### Proof of Lemma 7

I say that action  $u$  is induced by an S-type  $\omega$  if it is a best response to a given

equilibrium message  $m : u \in \{\sum_{i=1}^N e_i(\omega) | \omega \in \sigma^{-1}(m)\}$ . Let  $Y$  be the set of all actions induced by some S-type  $\omega$ . First, notice that if  $\omega$  induces  $\bar{t}$ , it must be that  $u_S(\bar{t}, \omega) = \max_{t \in Y} u_S(t, \omega)$ . Since  $u_S$  is strictly concave, it can take on a given value for at most two values of  $t$ . Thus,  $\omega$  can induce no more than two levels of aggregate energy consumption of the receivers in equilibrium.

Let  $u$  and  $v$  be two levels of aggregate energy consumption induced in equilibrium,  $u < v$ . Define  $\Theta_u$  the set of S types who induce  $u$  and  $\Theta_v$  the set of S types who induce  $v$ . Take  $\omega \in \Theta_u$  and  $\omega' \in \Theta_v$ . By definition,  $\omega$  reveals a weak preference for  $u$  over  $v$  and  $\omega'$  reveals a weak preference for  $v$  over  $u$  that is:

$$\begin{cases} u_S(u, \omega) \geq u_S(v, \omega) \\ u_S(v, \omega') \geq u_S(u, \omega') \end{cases}$$

Thus, by continuity of  $\omega \rightarrow u_S(u, \omega) - u_S(v, \omega)$ , there is  $\hat{\omega} \in [\omega, \omega']$  such that  $u_S(u, \hat{\omega}) = u_S(v, \hat{\omega})$ . Since  $u_S$  is strictly concave, we have that:

$$u < t_S(\hat{\omega}) < v$$

Then, notice that since  $\frac{\partial^2 u_S(t, \omega)}{\partial t \partial \omega} > 0$ , it must be that all types that induce  $u$  are below  $\hat{\omega}$ . Similarly, it must be that all types that induce  $v$  are above  $\hat{\omega}$ . That is:

$$\begin{aligned} \forall \omega \in \Theta_u, \omega &\leq \hat{\omega} \\ \forall \omega \in \Theta_v, \omega &\geq \hat{\omega} \end{aligned}$$

Given that  $u_1$  and  $u_2$  verify the assumptions of [Crawford and Sobel \(1982\)](#), the

sum of optimal action of the receivers, given that  $\omega \in \Theta_u$  is below the optimal action when the type is  $\hat{\omega}$ . Similarly, the sum of optimal actions of the receivers, given that  $\omega \in \Theta_v$  is above the optimal action when the type is  $\hat{\omega}$ . That is:

$$\begin{cases} \sum_{i=1}^N e_i(\Theta_u) \leq \sum_{i=1}^N e_i(\hat{\omega}) \\ \sum_{i=1}^N e_i(\Theta_v) \geq \sum_{i=1}^N e_i(\hat{\omega}) \end{cases} \\ \iff u \leq \sum_{i=1}^N e_i(\hat{\omega}) \leq v$$

However, as  $\sum_{i=1}^N e_i(\omega) \neq t_S(\omega)$  for all  $\omega \in \Omega$ , there is  $\epsilon > 0$  such that  $|\sum_{i=1}^N e_i(\omega) - t_S(\omega)| \geq \epsilon$  for all  $\omega \in \Omega$ . It follows that  $|u - v| \geq \epsilon$ .

For any belief  $B \subset \Omega$ , the sum of optimal action of the receivers is in  $[0, 1]$ . Thus, the set of actions induced in equilibrium is bounded and at least  $\epsilon$  away from one another, which completes the proof. □

Notice also that because  $u_S$  verifies all the requirement of [Crawford and Sobel \(1982\)](#), in every equilibrium of the game, if  $t$  is a level of aggregate energy consumption induced by type  $\omega$  and by type  $\omega''$  for some  $\omega < \omega''$ , then  $t$  is also induced by  $\omega' \in (\omega, \omega'')$ .

By [Lemma 7](#) there is a finite number of outcomes induced in equilibrium. The continuity of  $t_S(\omega)$  gives that there is a type of the sender which is indifferent between any pair of outcomes induced in equilibrium and the monotony of  $t_S(\omega)$  implies there are only a finite number of types which are indifferent between any pair of outcomes. Hence, the point made just above implies that there is a partitioning of  $\Omega$  in a finite number of cells where every cell induces a given level of aggregate energy consumption at equilibrium. This implies that any equilibrium is outcome equivalent to a partition equilibrium.

Call  $(\sigma^P, y^P)$  an equilibrium of the parallel game. To conclude the proof, I want to show that  $(\sigma, e_1, \dots, e_N)$  is an equilibrium of the game if and only if  $(\sigma^P, y^P)$  is an equilibrium of the parallel game where  $\sigma^P = \sigma$  and  $y^P = e_1 + \dots + e_N$ .

First, I show that if  $(\sigma^P, y^P)$  is an equilibrium of the parallel game, then  $\sigma = \sigma^P$  and  $e_1, \dots, e_N$  such that  $e_1 + \dots + e_N = y^P$  is an equilibrium of the game.

1. Assume that  $R_1, \dots, R_N$  play the equilibrium strategies  $e_i$  defined in section 2. Take an equilibrium of the parallel game  $(\sigma^P, y^P)$ . It follows that for any  $\omega \in \Omega$ ,  $\sum_{i=1}^N e_i(\sigma^P(\omega)) = \frac{C}{\alpha^2} - \frac{\mathbb{E}(\omega|\sigma^P(\omega))}{\alpha} = y^P(\sigma^P(\omega))$ . Thus, by playing strategy  $\sigma^P$ , the sender induces exactly the same aggregate energy consumption from the receivers as the action from the parallel receiver. In addition, for any  $\omega \in \Omega$ , the parallel sender has the same optimal action as the sum of socially optimal actions for the sender: for any  $\omega \in \Omega$ ,  $\operatorname{argmax}_{a \geq 0} U_S(a, \omega) = \sum_{i=1}^N e_i^W(\omega)$ .  $u_S$  and  $U_S$  are also both single-peaked and quadratic. As a result, for any parallel sender to be indifferent between two actions, it must be that they are at equal distance of his optimal action. Similarly for the sender to be indifferent between two aggregate energy consumption levels, it must be that they are at equal distance of his optimal action. To see this, assume that the parallel sender is indifferent between two actions  $a_1$  and  $a_2$ . Then it must be that:

$$\begin{aligned} U_S(a_1, \omega) &= U_S(a_2, \omega) \\ \iff -\left(a_1 - \left(\frac{C}{N\alpha^2} - \frac{\omega}{\alpha}\right)\right)^2 &= -\left(a_2 - \left(\frac{C}{N\alpha^2} - \frac{\omega}{\alpha}\right)\right)^2 \\ \iff \frac{C}{N\alpha^2} - \frac{\omega}{\alpha} &= \frac{a_1 + a_2}{2} \end{aligned}$$

Similarly, assume the sender is indifferent between  $\sum_{i=1}^N e_i$  and  $\sum_{i=1}^N e'_i$  such that  $\sum_{i=1}^N e_i = a_1$  and  $\sum_{i=1}^N e'_i = a_2$ . Then it must be that:

$$\begin{aligned}
& u_S\left(\sum_{i=1}^N e_i, \omega\right) = u_S\left(\sum_{i=1}^N e'_i, \omega\right) \\
\iff & C \sum_{i=1}^N e_i - \frac{N}{2}(\alpha \sum_{i=1}^N e_i + \omega)^2 = C \sum_{i=1}^N e'_i - \frac{N}{2}(\alpha \sum_{i=1}^N e'_i + \omega)^2 \\
& \iff \frac{C}{N\alpha^2} - \frac{\omega}{\alpha} = \frac{\sum_{i=1}^N (e_i + e'_i)}{2} = \frac{a_1 + a_2}{2}
\end{aligned}$$

As these optimal actions are the same, if the parallel sender is indifferent between two actions  $a_1$  and  $a_2$ , the sender is also indifferent between  $\sum_{i=1}^N e_i$  and  $\sum_{i=1}^N e'_i$  such that  $\sum_{i=1}^N e_i = a_1$  and  $\sum_{i=1}^N e'_i = a_2$ . As a result, as  $\sigma^P$  is an equilibrium strategy for the parallel sender,  $\sigma = \sigma^P$  is an equilibrium strategy for the sender.

2. Take an equilibrium of the parallel game  $(\sigma^P, y^P)$ . Assume that  $S$  and  $R_{-i}$  play equilibrium strategies  $\sigma^P$  and  $e_{-i}$  defined in section 2. Then, because the sender's equilibrium message is public,  $R_i$  and  $R_{-i}$  have the same belief and it follows from  $R_i$ 's first order condition that  $e_i = y^P - \sum_{j \neq i} e_j$  is an equilibrium strategy.

Now I show that if  $(\sigma, e_1, \dots, e_N)$  is an equilibrium of the game, then  $\sigma^P = \sigma$  and  $y^P$  such that  $y^P = e_1 + \dots + e_N$  is an equilibrium of the parallel game.

1. Take an equilibrium of the game  $(\sigma, e_1, \dots, e_N)$ . Assume that  $R$  plays the equilibrium strategies  $y^P = \sum_{i=1}^N e_i$ . Following lemma 7,  $\sigma$  must be a partitioned signalling strategy. Then, following a similar argument as above, because the parallel sender has the same optimal action as the sum of symmetric socially optimal actions for the sender, and that  $u_S$  and  $U_S$  are both single-peaked and quadratic, every equilibrium message instructed by  $\sigma$  induces exactly the same welfare to the sender and the parallel sender. As a result,  $\sigma^P = \sigma$  must be an equilibrium strategy for the parallel sender.
2. Take an equilibrium of the game  $(\sigma, e_1, \dots, e_N)$ . Assume that the parallel sender plays the equilibrium strategies  $\sigma^P = \sigma$ . Then, because for any given

belief the parallel receiver picks the same optimal action as the sum of optimal actions of  $R_i$ , for  $i \in 1, \dots, N$ , and because  $e_1, \dots, e_N$  is an equilibrium strategy, it must be that  $y^P = \sum_{i=1}^N e_i$  is an equilibrium strategy.

□

## References

- ABDELLAOUI, M., A. BAILLON, L. PLACIDO, AND P. P. WAKKER (2011): "The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation," *The American Economic Review*, 101, 695–723.
- ABDELLAOUI, M., H. BLEICHRODT, AND O. L'HARIDON (2008): "A Tractable Method to Measure Utility and Loss Aversion under Prospect Theory," *Journal of Risk and Uncertainty*, 36, 245–266.
- ABDELLAOUI, M., E. KEMEL, A. PANIN, AND F. M. VIEIDER (2019): "Measuring Time and Risk Preferences in an Integrated Framework," *Games and Economic Behavior*, 115, 459–469.
- ABDELLAOUI, M., F. VOSSMANN, AND M. WEBER (2005): "Choice-Based Elicitation and Decomposition of Decision Weights for Gains and Losses under Uncertainty," *Management Science*, 51, 1384–1399.
- ANDERSON, R. M., H. HEESTERBEEK, D. KLINKENBERG, AND T. D. HOLLINGSWORTH (2020): "How will country-based mitigation measures influence the course of the COVID-19 epidemic?" *The Lancet*, 395, 931–934.
- ANGELETOS, G.-M. AND A. PAVAN (2007): "Efficient use of information and social value of information," *Econometrica*, 75, 1103–1142.
- ANSCOMBE, F. J. AND R. J. AUMANN (1963): "A Definition of Subjective Probability," *The Annals of Mathematical Statistics*, 34, 199–205.
- ASHEIM, G. B. (2010): "Strategic use of environmental information," *Environmental and Resource Economics*, 46, 207–216.

- AZRIELI, Y., C. P. CHAMBERS, AND P. J. HEALY (2018): "Incentives in Experiments: A Theoretical Analysis," *Journal of Political Economy*, 126, 1472–1503.
- BADE, S. (2015): "Randomization Devices and the Elicitation of Ambiguity-Averse Preferences," *Journal of Economic Theory*, 159, 221–235.
- BAILLON, A. AND H. BLEICHRODT (2015): "Testing Ambiguity Models through the Measurement of Probabilities for Gains and Losses," *American Economic Journal: Microeconomics*, 7, 77–100.
- BAILLON, A., H. BLEICHRODT, U. KESKIN, O. L'HARIDON, AND C. LI (2018a): "The Effect of Learning on Ambiguity Attitudes," *Management Science*, 64, 2181–2198.
- BAILLON, A., L. CABANTOUS, AND P. P. WAKKER (2012): "Aggregating Imprecise or Conflicting Beliefs: An Experimental Investigation Using Modern Ambiguity Theories," *Journal of Risk and Uncertainty*, 44, 115–147.
- BAILLON, A., Z. HUANG, A. SELIM, AND P. P. WAKKER (2018b): "Measuring Ambiguity Attitudes for All (Natural) Events," *Econometrica*, 86, 1839–1858.
- BARBEROUSSE, A., D. BONNAY, AND M. COZIC (2011): "Précis de philosophie des sciences," .
- BARRETT, S. (1994): "Self-enforcing international environmental agreements," *Oxford Economic Papers*, 878–894.
- BATTAGLINI, M. (2002): "Multiple referrals and multidimensional cheap talk," *Econometrica*, 70, 1379–1401.
- BERGER, J. O., D. R. INSUA, AND F. RUGGERI (2000): "Bayesian Robustness," *Robust Bayesian Analysis*, 1–32.
- BERGER, L., N. BERGER, V. BOSETTI, I. GILBOA, L. P. HANSEN, C. JARVIS, M. MARINACCI, AND R. SMITH (2020): "Uncertainty and decision-making during a crisis: How to make policy decisions in the COVID-19 context?" *University of Chicago, Becker Friedman Institute for Economics Working Paper*.

- BERGER, L., J. EMMERLING, AND M. TAVONI (2016): "Managing catastrophic climate risks under model uncertainty aversion," *Management Science*, 63, 749–765.
- BERGSTROM, T., L. BLUME, AND H. VARIAN (1986): "On the private provision of public goods," *Journal of public economics*, 29, 25–49.
- BEWLEY, T. F. (2002): "Knightian Decision Theory. Part I," *Decisions in Economics and Finance*, 25, 79–110.
- BLACKWELL, D. (1950): "Comparison of experiments," Tech. rep., Howard University Washington United States.
- BLOOR, D. (1976): *Knowledge and social imagery*, Londres: Routledge.
- BOSE, S. AND L. RENO (2014): "Mechanism Design with Ambiguous Communication Devices," *Econometrica*, 82, 1853–1872.
- BURGE, T. (1993): "Content preservation," *The Philosophical Review*, 102, 457–488.
- CARNEY, M., J. CUNLIFFE, D. RAMSDEN, A. HALDANE, J. HASKEL, M. SAUNDERS, S. TENREYRO, AND G. Vlieghe (2019): "Inflation Report," *Bank of England*, <https://www.bankofengland.co.uk/-/media/boe/files/inflation-report/2019/may/inflation-report-may-2019.pdf>.
- CHATEAUNEUF, A., J. EICHBERGER, AND S. GRANT (2007): "Choice under Uncertainty with the Best and Worst in Mind: Neo-Additive Capacities," *Journal of Economic Theory*, 137, 538–567.
- CHEN, Y., N. KARTIK, AND J. SOBEL (2008): "Selecting cheap-talk equilibria," *Econometrica*, 76, 117–136.
- CHEW AND J. S. SAGI (2006): "Event Exchangeability: Probabilistic Sophistication without Continuity or Monotonicity," *Econometrica*, 74, 771–786.
- CHEW, S. H., B. MIAO, AND S. ZHONG (2017): "Partial Ambiguity," *Econometrica*, 85, 1239–1260.
- CHISHOLM, R. M., R. M. CHISHOLM, R. M. CHISHOLM, AND R. M. CHISHOLM (1989): *Theory of knowledge*, vol. 3, Prentice-Hall Englewood Cliffs, NJ.

- COADY, C. A. J. (1992): "Testimony: A philosophical study," .
- COLLINS, H. M. (1981): "Stages in the empirical programme of relativism," .
- COLSON, A. R. AND R. M. COOKE (2018): "Expert Elicitation: Using the Classical Model to Validate Experts' Judgments," *Review of Environmental Economics and Policy*, 12, 113–132.
- COOKE, R. M. (1991): "Experts in uncertainty: opinion and subjective probability in science," 125–127.
- COX, L. A. T. (2012): "Confronting Deep Uncertainties in Risk Analysis," *Risk Analysis*, 32, 1607–1629, <http://onlinelibrary.wiley.com/doi/10.1111/j.1539-6924.2012.01792.x/full>.
- CRAWFORD, V. P. AND J. SOBEL (1982): "Strategic information transmission," *Econometrica*, 1431–1451.
- DESSEIN, W. (2002): "Authority and communication in organizations," *The Review of Economic Studies*, 69, 811–838.
- DIECKMANN, N. F., R. MAURO, AND P. SLOVIC (2010): "The Effects of Presenting Imprecise Probabilities in Intelligence Forecasts," *Risk Analysis: An International Journal*, 30, 987–1001.
- DIMMOCK, S. G., R. KOUWENBERG, AND P. P. WAKKER (2015): "Ambiguity Attitudes in a Large Representative Sample," *Management Science*, <http://pubsonline.informs.org/doi/abs/10.1287/mnsc.2015.2198>.
- DOUENNE, T. AND A. FABRE (2020): "French attitudes on climate change, carbon taxation and other climate policies," *Ecological Economics*, 169, 106496.
- EHRET, P. J., L. VAN BOVEN, AND D. K. SHERMAN (2018): "Partisan barriers to bipartisanship: Understanding climate policy polarization," *Social Psychological and Personality Science*, 9, 308–318.
- EICHBERGER, J., S. GRANT, D. KELSEY, AND G. A. KOSHEVOY (2011): "The  $\alpha$ -MEU Model: A Comment," *Journal of Economic Theory*, 146, 1684–1698.

- ELLSBERG, D. (1961a): "Risk, ambiguity, and the Savage axioms," *The Quarterly Journal of Economics*, 643–669.
- (1961b): "Risk, Ambiguity, and the Savage Axioms," *Quart. J. Econ.*, 75, 643–669.
- EPSTEIN, L. G. AND M. SCHNEIDER (2010): "Ambiguity and Asset Markets," *Annual Review of Financial Economics*, 2, 315–346.
- FANTL, J. AND M. MCGRATH (2009): *Knowledge in an uncertain world*, Oxford University Press.
- FERGUSON, N. M., M. J. KEELING, W. J. EDMUNDS, R. GANI, B. T. GRENFELL, R. M. ANDERSON, AND S. LEACH (2003): "Planning for smallpox outbreaks," *Nature*, 425, 681–685.
- FINUS, M. (2008): "Game theoretic research on the design of international environmental agreements: insights, critical remarks, and future challenges," *International Review of Environmental and Resource Economics*, 2, 29–67.
- FINUS, M. AND P. PINTASSILGO (2013): "The role of uncertainty and learning for the success of international climate agreements," *Journal of Public Economics*, 103, 29–43.
- GABEL, M., J. GOOBLAR, C. M. ROE, AND J. C. MORRIS (2021): "The ideological divide in confidence in science and participation in medical research," *Scientific reports*, 11, 1–9.
- GAJDOS, T., T. HAYASHI, J.-M. TALLON, AND J.-C. VERGNAUD (2008): "Attitude toward imprecise information," *Journal of Economic Theory*, 140, 27–65.
- GALEOTTI, A., C. GHIGLINO, AND F. SQUINTANI (2013): "Strategic information transmission networks," *Journal of Economic Theory*, 148, 1751–1769.
- GARLAPPI, L., R. UPPAL, AND T. WANG (2007): "Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach," *Review of Financial Studies*, 20, 41–81, <http://rfs.oxfordjournals.org/content/20/1/41.short>.

- GEOFFARD, P.-Y. AND T. PHILIPSON (1997): "Disease eradication: private versus public vaccination," *The American Economic Review*, 87, 222–230.
- GHIRARDATO, P., F. MACCHERONI, AND M. MARINACCI (2004a): "Differentiating Ambiguity and Ambiguity Attitude," *J. Econ. Theory*, 118, 133–173.
- GHIRARDATO, P., F. MACCHERONI, M. MARINACCI, ET AL. (2004b): "Differentiating ambiguity and ambiguity attitude," *Journal of Economic Theory*, 118, 133–173.
- GHIRARDATO, P. AND M. MARINACCI (2002): "Ambiguity made precise: A comparative foundation," *Journal of Economic Theory*, 102, 251–289.
- GILBOA, I. AND D. SCHMEIDLER (1989a): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18, 141–153.
- (1989b): "Maxmin Expected Utility with Non-Unique Prior," *J. Math. Econ.*, 18, 141–153.
- GIUSTINELLI, P., C. F. MANSKI, AND F. MOLINARI (2021): "Precise or Imprecise Probabilities? Evidence from Survey Response on Late-Onset Dementia," *Journal of the European Economic Association*.
- GOLDMAN, A. I. (2001): "Experts: Which ones should you trust?" *Philosophy and Phenomenological Research*, 63, 85–110.
- GOLTSMAN, M. AND G. PAVLOV (2008): "How to talk to multiple audiences," .
- GRANT, S., P. RICH, AND J. D. STECHER (2019): "An Ordinal Theory of Worst- and Best-Case Expected Utility," SSRN Scholarly Paper ID 3369078, Social Science Research Network, Rochester, NY, <https://papers.ssrn.com/abstract=3369078>.
- GUL, F. AND W. PESENDORFER (2014): "Expected Uncertain Utility Theory," *Econometrica*, 82, 1–39.
- (2015): "Hurwicz Expected Utility and Subjective Sources," *Journal of Economic Theory*, 159, 465–488.

- HALADY, I. R. AND P. H. RAO (2010): "Does awareness to climate change lead to behavioral change?" *International Journal of Climate Change Strategies and Management*.
- HANANY, E., P. KLIBANOFF, AND S. MUKERJI (2020): "Incomplete information games with ambiguity averse players," *American Economic Journal: Microeconomics*, 12, 135–87.
- HANSEN, L. AND T. J. SARGENT (2001): "Robust control and model uncertainty," *American Economic Review*, 91, 60–66.
- HANSEN, L. P., T. J. SARGENT, G. TURMUHAMBETOVA, AND N. WILLIAMS (2006): "Robust control and model misspecification," *Journal of Economic Theory*, 128, 45–90.
- HARDWIG, J. (1985): "Epistemic dependence," *The Journal of Philosophy*, 82, 335–349.
- (1991): "The role of trust in knowledge," *The Journal of Philosophy*, 88, 693–708.
- HEAL, G. AND A. MILLNER (2014): "Reflections: Uncertainty and decision making in climate change economics," *Review of Environmental Economics and Policy*, 8, 120–137.
- HELLEWELL, J., S. ABBOTT, A. GIMMA, N. I. BOSSE, C. I. JARVIS, T. W. RUSSELL, J. D. MUNDAY, A. J. KUCHARSKI, W. J. EDMUNDS, F. SUN, ET AL. (2020): "Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts," *The Lancet Global Health*.
- HEY, J., G. LOTITO, AND A. MAFFIOLETTI (2010): "The Descriptive and Predictive Adequacy of Theories of Decision Making under Uncertainty / Ambiguity," *Journal of Risk and Uncertainty*, 41, 81–111, <http://dx.doi.org/10.1007/s11166-010-9102-0>.
- HILL, B. (2019): "Beyond Uncertainty Aversion," *HEC Paris Research Paper, Forthcoming*.

- HIRSHLEIFER, J. (1978): "The private and social value of information and the reward to inventive activity," in *Uncertainty in economics*, Elsevier, 541–556.
- HOFFMAN, M. D. AND A. GELMAN (2014): "The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo." *J. Mach. Learn. Res.*, 15, 1593–1623.
- HOLLINGSWORTH, T. D., D. KLINKENBERG, H. HEESTERBEEK, AND R. M. ANDERSON (2011): "Mitigation strategies for pandemic influenza A: balancing conflicting policy objectives," *PLoS Comput Biol*, 7, e1001076.
- HUME, D. (1740): "A treatise of human nature," .
- (1748): *Enquête sur l'entendement humain*, Paris, Librairie générale.
- ILUT, C. L. AND M. SCHNEIDER (2014): "Ambiguous Business Cycles," *American Economic Review*, 104, 2368–99.
- JOHNSON, C., A. BAILLON, H. BLEICHRODT, Z. LI, D. VAN DOLDER, AND P. WAKKER (2021): "Prince: An Improved Method For Measuring Incentivized Preferences," *Journal of Risk and Uncertainty*.
- KAMENICA, E. AND M. GENTZKOW (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615.
- KANT, E. (1787): *Critique de la raison pure*, Editions Flammarion.
- KARANKI, D. R., H. S. KUSHWAHA, A. K. VERMA, AND S. AJIT (2009): "Uncertainty Analysis Based on Probability Bounds (P-Box) Approach in Probabilistic Safety Assessment," *Risk Analysis*, 29, 662–675.
- KARNI, E. (2020): "A Mechanism for the Elicitation of Second-Order Belief and Subjective Information Structure," *Economic Theory*, 69, 217–232.
- KELLNER, C. AND M. T. LE QUEMENT (2017): "Modes of ambiguous communication," *Games and Economic Behavior*, 104, 271–292.
- (2018): "Endogenous ambiguity in cheap talk," *Journal of Economic Theory*, 173, 1–17.

- KLIBANOFF, P., S. MUKERJI, K. SEO, AND L. STANCA (2021): "Foundations of Ambiguity Models under Symmetry:  $\alpha$ -MEU and Smooth Ambiguity," *Journal of Economic Theory*, 105202.
- KOCHER, M. G., A. M. LAHNO, AND S. T. TRAUTMANN (2018): "Ambiguity Aversion Is Not Universal," *European Economic Review*, 101, 268–283.
- KRIEGLER, E., J. W. HALL, H. HELD, R. DAWSON, AND H. J. SCHELLNHUBER (2009): "Imprecise Probability Assessment of Tipping Points in the Climate System," *Proceedings of the national Academy of Sciences*, 106, 5041–5046.
- KRISHNA, V. AND J. MORGAN (2001): "A model of expertise," *The Quarterly Journal of Economics*, 116, 747–775.
- LEE, T. M., E. M. MARKOWITZ, P. D. HOWE, C.-Y. KO, AND A. A. LEISEROWITZ (2015): "Predictors of public climate change awareness and risk perception around the world," *Nature climate change*, 5, 1014–1020.
- LENTON, T. M., H. HELD, E. KRIEGLER, J. W. HALL, W. LUCHT, S. RAHMSTORF, AND H. J. SCHELLNHUBER (2008): "Tipping elements in the Earth's climate system," *Proceedings of the national Academy of Sciences*, 105, 1786–1793.
- LORENZONI, I. AND N. F. PIDGEON (2006): "Public views on climate change: European and USA perspectives," *Climatic change*, 77, 73–95.
- MACHINA, M. J. AND D. SCHMEIDLER (1992): "A More Robust Definition of Subjective Probability," *Econometrica*, 60, 745–780.
- MANSKI, C. F. (2003): *Partial Identification of Probability Distributions*, Springer Science & Business Media.
- (2013): *Public Policy in an Uncertain World : Analysis and Decisions*, Cambridge, Mass: Harvard University Press.
- MARINACCI, M. (2015): "Model uncertainty," *Journal of the European Economic Association*, 13, 1022–1100.
- MILLNER, A., S. DIETZ, AND G. HEAL (2013): "Scientific ambiguity and climate policy," *Environmental and Resource Economics*, 1–26.

- MIN, D. (2017): "Bayesian persuasion under partial commitment," *Work. Pap., Univ. Ariz., Tucson*.
- MORGAN, M. G. (2014): "Use (and Abuse) of Expert Elicitation in Support of Decision Making for Public Policy," *Proceedings of the National Academy of Sciences*, 111, 7176–7184, <http://www.pnas.org/content/111/20/7176.short>.
- MORGAN, M. G., M. HENRION, AND M. SMALL (1992): *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis*, Cambridge university press.
- NA, S.-L. AND H. S. SHIN (1998): "International environmental agreements under uncertainty," *Oxford Economic Papers*, 50, 173–185.
- PIDGEON, N. AND B. FISCHHOFF (2011): "The role of social and decision sciences in communicating uncertain climate risks," *Nature climate change*, 1, 35–41.
- PINDYCK, R. S. (2007): "Uncertainty in environmental economics," *Review of environmental economics and policy*, 1, 45–65.
- QIU, J. AND U. WEITZEL (2016): "Experimental Evidence on Valuation with Multiple Priors," *Journal of risk and uncertainty*, 53, 55–74.
- SAVAGE, L. J. (1954): *The Foundations of Statistics*, New York: Dover.
- (1972): *The Foundations of Statistics*, Courier Corporation.
- SINISCALCHI, M. (2006): "A Behavioral Characterization of Plausible Priors," *Journal of Economic Theory*, 128, 1–17.
- VAN BOVEN, L., P. J. EHRET, AND D. K. SHERMAN (2018): "Psychological barriers to bipartisan public support for climate policy," *Perspectives on Psychological Science*, 13, 492–507.
- WAKKER, P. P. (2010): *Prospect Theory For Risk and Ambiguity*, Cambridge: Cambridge University Press.
- WALD, A. (1949): "Statistical decision functions," *The Annals of Mathematical Statistics*, 165–205.

WEATHERSON, B. (2012): "Knowledge, bets, and interests," *Knowledge ascriptions*, 75–103.