



Analyse expérimentale du biais de négligence de la corrélation dans un contexte de sondage pré-électoral

Mémoire

Nathan Roger Lea Jombi

Maîtrise en économie - avec mémoire
Maître ès arts (M.A.)

Québec, Canada

Analyse expérimentale du biais de négligence de la corrélacion dans un contexte de sondage pré-électoral

Mémoire

Nathan Roger Lea Jombi

Sous la direction de:

Sabine Kroger, directrice de recherche

Résumé

L'objectif de ce mémoire est de mieux comprendre un des biais cognitifs susceptibles d'affecter le pouvoir prédictif des sondages : le biais de négligence de la corrélation. Intuitivement, le biais de négligence de la corrélation est une simplification du processus de décision chez les votants qui les amène à ignorer la corrélation entre leurs préférences électorales et les coûts de participation au vote. Afin d'analyser ce biais, nous avons conduit une expérience consistant en une série de questions-réponses sous la forme de sondages. Les participants se réfèrent à un ensemble d'informations sur les gains et les coûts rattachés à leurs choix de vote. Les résultats expérimentaux montrent que le biais de négligence envers la corrélation affecte significativement les réponses des participants aux questions de sondage. De plus, certaines caractéristiques socioéconomiques des participants semblent augmenter la probabilité de souffrir du biais. Enfin, l'analyse des données recueillies lors de l'expérience montre que la mesure proposée du biais n'était pas calculable pour certaines observations car les participants avaient effectué des choix certains.

Table des matières

| | |
|--|-----------|
| Résumé | iv |
| Table des matières | iv |
| Liste des tableaux | vi |
| Liste des figures | vii |
| Remerciements | ix |
| Introduction | 1 |
| 1 Revue de littérature | 4 |
| 1.1 Le biais de négligence envers la corrélation | 4 |
| 1.2 Les mécanismes d'élicitation et d'agrégation des préférences | 6 |
| 1.3 Choix du format de questions | 7 |
| 2 L'expérience | 9 |
| 2.1 Plan Expérimental | 10 |
| 2.2 Types de questions de sondage dans l'expérience | 14 |
| 2.3 Présentation de la mesure empirique | 15 |
| 2.4 Traitement de l'expérience | 16 |
| 3 Résultats | 18 |
| 3.1 Statistiques descriptives | 18 |
| 3.2 Analyse descriptive des <i>coefficients $\hat{\Phi}$ calculables</i> | 23 |
| 3.3 Modèle empirique | 25 |
| 3.4 Résultats | 28 |
| 4 Analyse des coefficients Φ indéterminables | 38 |
| 4.1 Analyse des coefficients Φ indéterminables pour les prédictions agrégées | 38 |
| 4.2 Analyse des coefficients Φ indéterminables pour les prédictions individuelles | 43 |
| Conclusion | 49 |
| Bibliographie | 51 |
| A Distribution des Φ déterminables | 53 |

| | | |
|----------|---|-----------|
| B | Aperçu de l'interface présentée aux participants lors d'un tour entier | 54 |
| C | Liste complète des questions posées aux participants | 64 |

Liste des tableaux

| | | |
|----|--|----|
| 1 | Ensemble des informations présentées aux participants | 10 |
| 2 | Table de contingence | 15 |
| 3 | Randomisation des participants selon le type et le traitement | 16 |
| 4 | Paramètres de l'expérience | 17 |
| 5 | Statistiques descriptives | 19 |
| 6 | Distribution des erreurs commises par les participants | 20 |
| 7 | Exemple de table de contingence pour un coefficient indéterminable | 22 |
| 8 | Coefficient du modèle présenté | 29 |
| 9 | Effets marginaux : corrélation inverse | 31 |
| 10 | Effets marginaux : sous-estimation ou négligence de la corrélation | 33 |
| 11 | Effets marginaux : corrélation exacte | 34 |
| 12 | Effets marginaux : surestimation de la corrélation | 36 |
| 13 | Distribution des coefficients Φ individuels indéterminables selon les traitements | 38 |
| 14 | prédictions théoriques pour choix rationnels agrégés | 39 |
| 15 | Paiements avec et sans frais associés au traitement 1 | 40 |
| 16 | Choix optimal du participant rationnel pour le traitement 1 | 40 |
| 17 | Probabilités et prédictions théoriques des choix rationnels agrégés pour le traitement 1 | 41 |
| 18 | Distribution des Φ non calculable selon le type | 44 |
| 19 | Distribution des coefficients phi individuels indéterminables selon les traitements | 44 |
| 20 | Prédictions théoriques pour choix individuels du type vert | 44 |
| 21 | Fréquence des probabilités assignées aux choix A et B pour les types bleus et jaunes dans les observations avec Φ individuels indéterminables | 47 |
| 22 | Statistiques descriptives | 53 |

Liste des figures

| | | |
|---|--|----|
| 1 | Exemple d'écran présenté aux participants | 11 |
| 2 | Distribution de la moyenne des gains théoriques et empiriques à travers les 8 traitements | 20 |
| 2 | Distribution des $\hat{\Phi}$ empiriques agrégés autour du Φ théorique (ligne verte) . . . | 24 |
| 3 | Distribution des $\hat{\Phi}$ empiriques individuels autour du Φ théorique (ligne verte) . | 24 |
| 5 | Distribution des Φ indéterminables selon les probabilités théoriques de voter associées à chaque traitement. | 43 |
| 6 | Distribution des Φ indéterminables selon les traitements. | 47 |

"We tend to assume that if something works out, we're genius decision-makers. But the outcome of a situation is often down to luck — or information we don't have. All we can do is make the best choice possible given what we know, be prepared to be wrong, and move forward : The more you can get comfortable with uncertainty, the better off you are."

Annie Duke

Remerciements

La réalisation de ce mémoire a été possible grâce au concours de plusieurs personnes à qui je voudrais témoigner ma gratitude.

Je voudrais tout d'abord adresser toute ma reconnaissance à la directrice de ce mémoire, Madame Sabine Kroger, pour son soutien financier, sa patience, sa disponibilité et surtout ses précieux conseils qui m'ont permis d'achever ce mémoire.

Je désire aussi remercier les professeurs de l'université Laval, qui m'ont fourni les outils nécessaires à la réussite de mes études.

Je tiens à remercier spécialement le professeur Sylvain Dessy non seulement pour son soutien moral mais aussi nos longs échanges qui m'ont inspiré tout le long de ce mémoire.

Je voudrais exprimer ma reconnaissance envers mes amis et collègues pour les conseils et accompagnements lorsque je rencontrais des difficultés avec mon code et Latex. Je pense notamment à Antoine Dedewanou, Thuo Kossa et Valère Nketcha.

Je tiens à témoigner toute ma gratitude à Kouamé Marius Sossou pour son aide précieuse durant les séances d'expérience.

Enfin, un grand merci à mes parents et frères qui m'ont appuyé tout le long de mon parcours universitaire. Merci Maman.

Introduction

Les élections jouent un rôle important pour les agents économiques dans les pays à démocratie avancée. Dans un sens, ils révèlent au gouvernement les préférences des ménages et du secteur des affaires. Dans l'autre sens, ils informent les ménages et le secteur des affaires sur les impacts découlant des orientations du gouvernement élu. Pour ces raisons, les agents économiques ont toujours ressenti le besoin de prédire les résultats d'une élection afin d'avoir un coup d'avance et d'ajuster leurs comportements face à d'éventuels changements.

À cet effet, des études se sont penchées au fil du temps sur la question du choix de la meilleure méthode d'élicitation et d'agrégation des préférences électorales. Les méthodes les plus connues d'entre elles sont les sondages et les marchés de prédiction. Les sondages permettent de révéler les anticipations des individus sur leurs propres intentions de vote (Manski (2006), Erikson et Wlezien (2008) et Erikson et Wlezien (2012)). Quant aux marchés de prédiction, ils sont conçus pour éliciter les anticipations de l'ensemble des électeurs sur le résultat des élections (Forsythe (1992), Kou et Sobel (2004), Wolfers et Zitzewitz (2004), Wolfers et Zitzewitz (2006), Arrow *et al.* (2008) et Snowberg *et al.* (2013)).

Dans ce mémoire, nous nous intéressons au pouvoir prédictif des questions de sondages pré-électorales. Ces derniers sont un outil de prédiction sur lequel les agents économiques comptent largement pour anticiper le résultat d'une élection et ajuster leur comportement à l'avance. L'avalanche de sondages auxquels sont confrontés les Canadiens durant les campagnes électorales provinciales et fédérales montre bien la place importante qu'elle occupe chez les individus.

Malgré la popularité des sondages et leur utilisation de plus en plus grandissante, il est vrai que le pouvoir prédictif n'est pas sans faille. Le référendum sur l'appartenance du Royaume-Uni à l'Union européenne est un bon exemple. Ce référendum s'est tenu en juin 2016 et opposait les partisans du *Bremain* - c'est-à-dire ceux qui faisaient campagne pour le maintien du Royaume-Uni dans l'Union européenne - et les partisans du *Brexit* - c'est-à-dire ceux qui faisaient campagne pour la sortie du Royaume-Uni de l'Union européenne. Selon les prédictions des agences de sondage, la tendance était bonne pour les partisans du *Bremain* : des 168 sondages depuis l'annonce de la tenue du référendum, plus des 2/3 (113 sondages plus précisément) d'entre eux prédisaient un maintien du Royaume-Uni dans l'Union européenne. Cependant, grande fut la surprise des observateurs dans le monde au soir des résultats. Les résultats finaux

ont annoncé 51.9% pour un *Brexit* contre 48.1% pour un *Bremain*.

Un autre évènement marquant de l'histoire récente durant lequel les sondages pré-électoraux ont failli à leur mission est la victoire de Donald Trump aux élections présidentielles américaines de 2016. En effet, la candidate démocrate Hillary Clinton menait la course à la présidence selon la majorité des agences de sondage. D'après la majorité des sondages pré-électoraux, les chances de victoire d'Hillary Clinton étaient évaluées à 70% ; certains sondages parlant même de 99% de chance de victoire pour la candidate démocrate. À la grande surprise des observateurs, les prédictions des agences de sondage se sont avérées erronées. Bien que Hillary Clinton ait devancé Donald Trump de 2.1 points de pourcentage sur le plan national, le candidat républicain est sorti victorieux de la course à la présidence grâce au vote des grands électeurs en obtenant 306 d'entre eux contre 232 pour Hillary Clinton.

La littérature antérieure nous renseigne sur les facteurs potentiellement responsables des erreurs de prédiction dans les résultats de sondages. Selon Erikson *et al.* (2002) les erreurs de prédiction des sondages peuvent s'expliquer par l'influence des résultats des sondages sur le comportement des électeurs et des candidats. Forsythe *et al.* (1996) attribuent ces erreurs de prédiction au fait que les votants ne révèlent pas leurs véritables intentions de vote lors des sondages. Enfin Manski (1990) nous apprend que le choix du format de question utilisé pour réaliser le sondage peut également influencer les résultats.

Pour réaliser les sondages pré-électoraux, les sondeurs utilisent deux types de formats de questions : **1)** les questions binaires qui demandent aux votants de déclarer le candidat pour lequel il a l'intention de voter et **2)** les questions probabilistes qui invitent les électeurs à attribuer une probabilité de vote à chacun des candidats.

Au moment de répondre à un sondage pré-électoral, les individus peuvent être influencés par de biais cognitifs. L'un des biais comportementaux potentiellement néfastes est le biais de négligence, encore appelé biais d'inattention à de la corrélation. Intuitivement, le biais de négligence de la corrélation est une simplification du processus de décision chez les individus qui les amènent à ignorer la corrélation entre les facteurs pertinents en contexte de vote. Par conséquent, l'individu peut mal rapporter ses intentions de vote si elle ignore cette corrélation. Morton (2006) a notamment mis en évidence l'idée selon laquelle les institutions politiques peuvent réduire le coût de participation des partisans de certains partis politiques en facilitant leur inscription sur les listes électorales. En pratique comprendre ce biais et connaître les individus à risque pourraient permettre l'adoption de comportements optimaux et plus adéquats au moment de faire le sondage.

La littérature sur le biais de négligence de la corrélation est relativement récente. Elle met en lumière la variété de champs dans lesquels le biais de négligence peut être observé. Dans la finance il amène les investisseurs à ignorer la corrélation entre les différents actifs financiers ; ce qui a pour conséquence que ces derniers détiennent des portefeuilles mal diversifiés et

potentiellement trop risqués. D'autres études sur ce biais portent sur la corrélation entre différentes sources d'information sur la formation d'anticipation des individus.

L'objectif de ce mémoire est de mieux comprendre comment le biais de négligence de la corrélation est susceptible d'affecter les réponses des individus à des questions de sondages pré-électorales. Plus précisément, une expérience en laboratoire nous permettra de recueillir des données que nous utiliserons afin de développer notre mesure du biais de négligence de la corrélation. Cette mesure capturera la corrélation entre la probabilité de voter pour l'un des choix et la probabilité de participer au vote. Ainsi nous pourrons étudier les déterminants dudit biais.

Les résultats de notre recherche contribueront à alimenter la littérature encore récente sur le biais de négligence de la corrélation dans le contexte de sondages pré-électorales. En effet, peu d'études ont été menées sur les implications de la corrélation entre les préférences électorales des électeurs et les coûts de participation au vote. Les résultats de nos analyses révèlent que les participants souffrent d'un biais de sous-estimation de la corrélation. Ce biais diffère selon le sexe, le niveau d'étude et l'âge du participant.

Ce mémoire est divisé en cinq chapitres. Le premier chapitre sera composé d'une revue de la littérature sur le biais de négligence de la corrélation, sur les méthodes d'élicitation et d'agrégation des préférences électorales et sur les types de formats de questions de sondage. Le deuxième chapitre décrit l'expérience telle qu'elle a été présentée aux participants et la mesure empirique développée pour mesurer le biais de négligence de la corrélation. Les troisième et quatrième chapitres discutent des données recueillies durant l'expérience, du modèle empirique et des résultats. Enfin, dans le cinquième chapitre nous concluons.

Chapitre 1

Revue de littérature

1.1 Le biais de négligence envers la corrélation

Les études antérieures sur le biais de négligence envers la corrélation sont relativement nouvelles et elles reposent principalement sur des expériences en laboratoire comparant les choix effectués par les individus aux choix rationnels (attendus). Les études se différencient les unes des autres notamment par leurs dispositifs expérimentaux.

Kroll *et al.* (1988) ont conduit une expérience dont le but était de tester les hypothèses de base sur lesquelles reposent le théorème de séparation et le modèle d'évaluation des actifs financiers ("Capital Asset Pricing Model" ou CAPM en abrégé). Durant vingt mises en situation, les sujets avaient pour tâche de répartir leurs dotations dans trois actifs avec des rendements normalement distribués. Leurs résultats suggèrent que les individus ne semblent pas avoir pris la corrélation entre les actifs en compte : leurs décisions d'allocation n'étaient pas affectées par les variations de la corrélation entre les actifs et ce, même après 20 mises en situation.

Kallir et Sonsino (2009) ont mis en place une expérience afin d'examiner les effets d'une variation de corrélation sur les décisions d'allocation des dotations des individus. L'expérience était menée sur plus de 140 sujets avec des connaissances avancées en finance. En effet, il s'agissait principalement d'étudiants de premier et deuxième cycle des écoles de commerce et gestion. Le dispositif expérimental était composé de cinq problèmes de prédiction et d'allocation des dotations. Dans chaque problème les sujets doivent prédire 4 fois les rendements futurs d'un actif A(B) en fonction de celui de l'actif B(A) ; sachant qu'au préalable, il leur est montré la distribution conjointe des 12 précédentes paires d'observations ayant les mêmes distributions que celles des problèmes auxquelles ils sont confrontés. Par la suite les sujets doivent répartir leurs dotations initiales entre les actifs. Les rendements des deux actifs ne peuvent prendre que deux valeurs : *high* pour une valeur élevée et *low* pour une faible valeur. De plus, les valeurs assignées aux rendements sont telles que l'actif A dominait stochastiquement l'actif B. Ainsi, un individu rationnel n'allouera sa dotation dans l'actif B que lorsqu'il y a plus de chance que

ce dernier ait un rendement plus élevé que l'actif A. Enfin, à chaque problème de prédiction et d'allocation, la corrélation des rendements des actifs prennent les valeurs $+2/3$, $+1/3$, 0 , $-1/3$ et $-2/3$. Les résultats montrent que la variation de la corrélation entre le rendement des actifs n'affectait pas, ou que très peu les décisions d'allocation des participants, bien qu'ils percevaient cette variation dans la corrélation.

Eyster et Weizsacker (2010) ont également mené une expérience afin d'examiner les effets d'une variation de corrélation sur les décisions d'allocation des dotations des individus. Les sujets doivent également allouer leurs dotations entre deux actifs. Contrairement à Kallir et Sonsino (2009), le dispositif expérimental comporte huit paires de problèmes de choix de portefeuille. De plus les rendements futurs sont les mêmes entre l'actif A et l'actif B et les rendements historiques des actifs peuvent être corrélés ou non selon le problème. Les paires de problèmes sont conçues de sorte qu'il soit possible de déterminer une allocation équivalente (fournissant le même niveau d'utilité) pour chaque allocation faisable choisie par l'individu dans l'un des problèmes. Les résultats montrent que les sujets ignorent la corrélation en traitant les actifs corrélés comme indépendants. De plus, les sujets tendent à suivre un simple $1/n$ heuristique, c'est-à-dire qu'ils investissent des parts égales de leurs dotations entre les différents actifs.

Le biais de négligence envers la corrélation ne s'observe pas seulement dans le cadre financier. D'ailleurs Enke et Zimmermann (2017) le définissent comme un biais cognitif résultant de l'incapacité des individus à comprendre les implications de la corrélation entre différentes variables sur leurs anticipations subjectives (forming beliefs). Il ne s'agit pas d'un arbitrage rationnel de l'individu entre l'élaboration des anticipations justes et l'économie de ressources cognitives rares. C'est dans ce sens que Enke et Zimmermann (2017) ont mené une expérience afin de mesurer l'effet de la corrélation entre différentes sources d'informations sur les anticipations des individus et sur le résultat d'un marché financier. En effet, plusieurs médias comme les agences de presse partagent une même nouvelle provenant d'une même source. Dès lors, les informations diffusées par ces différents médias (papiers journaux, journal télévisé, site web etc.) tendent à être corrélées. Le dispositif expérimental mis en place par Enke et Zimmermann demande aux sujets de faire une estimation ex ante d'une variable à partir de différentes sources d'information. Il s'agit d'un between-subject design avec deux traitements dans lequel un groupe de participants reçoit des sources d'information corrélées tandis que l'autre source de participant reçoit de l'information non corrélée. Ainsi si les participants du traitement dont l'information est corrélée prend en compte ladite corrélation alors les prédictions des deux types devraient être identiques. Cependant, les résultats suggèrent qu'une bonne partie des participants considèrent les sources d'informations comme indépendantes et par conséquent s'exposent au biais de négligence de la corrélation.

Enfin Levy et Razin (2015) se sont intéressés à l'effet du biais de négligence envers la corrélation dans le cadre d'un modèle théorique. Le biais est présenté ici comme la corrélation entre les différentes sources d'informations auxquelles les électeurs sont exposés et qui causerait des

anticipations erronées de la part des électeurs. Plus précisément, les auteurs ont tenté dans cet article théorique de décrire les caractéristiques de ces anticipations erronées causées par l'inattention à la corrélation. Pour ce faire, les auteurs ont utilisé un modèle de vote avec des électeurs hétérogènes devant choisir entre deux politiques : une de gauche et une de droite. Le choix durant l'élection dépendait des préférences de l'électeur, mais aussi de l'état de monde dans lequel il se trouve ; chaque électeur recevant des signaux concernant l'état de monde. Les auteurs supposaient que les électeurs non rationnels ignoreraient la corrélation entre les signaux concernant l'état de monde contrairement aux électeurs rationnels. Les résultats suggeraient que le biais de négligence polarise les opinions mais ne conduit pas nécessairement à un résultat inefficace. En effet, même si les électeurs ne votent pas individuellement de façon optimale, l'ensemble de l'électorat peut atteindre un niveau de réponse meilleur que celui atteint par un électorat rationnel.

La littérature antérieure nous a renseigné sur l'effet de la négligence envers la corrélation sur les individus dans différents contextes. Dans ce mémoire, nous tentons de mettre en évidence l'impact de la négligence envers la corrélation sur les décisions de votes des individus, dans un contexte de sondages pré-électorales. Plus précisément nous tentons d'évaluer l'impact de la corrélation entre la probabilité de voter pour un candidat et la probabilité de participer au vote. Morton (2006) a notamment mis en évidence l'idée selon laquelle les institutions politiques peuvent réduire le coût de participation des partisans de certains partis politiques en facilitant leur inscription sur les listes électorales. En effet, l'intuition en arrière serait que ces deux probabilités sont corrélées en fonction des gains et coûts associés aux décisions et à la participation de vote. Par coûts, on peut penser par exemple à la difficulté d'accéder à un bureau de vote qui découragerait les individus à participer au vote, ou encore la peur de voter pour le mauvais candidat. Par conséquent plus ces coûts augmentent, moins la probabilité de participer au vote est élevée.

1.2 Les mécanismes d'élicitation et d'agrégation des préférences

Au fil du temps, des études se sont penchées sur la question du choix de la meilleure méthode d'élicitation et d'agrégation des préférences électorales. Les plus connues d'entre elles sont les marchés de prédiction et les sondages. Les marchés de prédiction sont conçus pour éliciter les anticipations des individus sur le résultat des élections. Les sondages quant à eux permettent de révéler les anticipations des individus sur leurs propres intentions de vote. Ces deux méthodes ne conduisent pas généralement aux mêmes résultats. Dès lors, des débats entre les défenseurs des marchés de prédiction (Forsythe (1992), Kou et Sobel (2004), Wolfers et Zitzewitz (2004), Wolfers et Zitzewitz (2006), Arrow *et al.* (2008) et Snowberg *et al.* (2013)) et ceux des sondages (Manski (2006), Erikson et Wlezien (2008) et Erikson et Wlezien (2012)) ont alimenté la

littérature.

Les supporteurs des marchés de prédiction interprètent le prix du marché comme une prédiction du résultat de l'élection. Selon eux, le prix d'équilibre sur le marché de prédiction agrège de façon efficace l'ensemble de l'information que possèdent les agents économiques. L'argument théorique est que les bénéfices potentiels en cas de prédiction exacte encouragent les individus détenant la bonne information à prendre part au marché. Le coût d'achat d'une prédiction quant à lui décourage ceux qui ne détiennent pas l'information pertinente. De plus, des évidences empiriques ont été émises afin de défendre les marchés de prédiction. [Wolfers et Zitzewitz \(2004\)](#) ont comparé les prédictions des marchés et des sondages au cours de la dernière semaine de vote pour 4 campagnes présidentielles américaines. Les résultats ont montré que les marchés avaient prédit les pourcentages de vote des démocrates et des républicains, mieux que les sondages de Gallup : la moyenne d'erreur absolue était de 1.5% pour les marchés et 2.1% pour les sondages de Gallup. Les marchés avaient également une plus faible moyenne d'erreur pour les prédictions datant de 150 jours avant les élections.

Les défenseurs des sondages mettent en garde contre les comparaisons évoquées dans les études qui soutiennent les marchés de prédiction. Selon eux, les deux méthodes ne traduisent pas des résultats comparables. En effet, le prix d'équilibre observé dans le marché prédit le dénouement des élections, alors que les sondages agrègent les préférences individuelles observées le jour du sondage. Dans ce sens, [Erikson et Wlezien \(2008\)](#) ont montré qu'en transformant l'information donnée par les sondages en projection pour le jour de l'élection, les prédictions obtenues étaient plus précises que celles offertes par les marchés ; ceci pour un ratio de trois campagnes présidentielles sur cinq. De plus, selon [Snowberg *et al.* \(2013\)](#), le prix d'équilibre observé sur le marché a tendance à sous-estimer la vraie probabilité des événements très probables et à surestimer la probabilité des événements peu probables.

La littérature antérieure met en relief les avantages et les inconvénients des marchés de prédiction et des sondages. Afin de pallier aux lacunes de ces deux méthodes, une approche alternative serait de combiner les deux méthodes. Cette combinaison consisterait à demander aux individus de donner leurs anticipations sur le comportement agrégé de la population dans le cadre d'un sondage.

1.3 Choix du format de questions

Nous venons de passer en revue les principales méthodes d'élicitation et d'agrégation des préférences électorales. Quelle que soit la méthode retenue, la précision des prédictions du résultat des élections peut aussi être influencée par le choix du format de question. En effet, la distribution des réponses peut varier selon le format de question utilisé. Pour cela, il est important d'utiliser le format de question représentant fiablement les préférences des individus.

Quelle que soit la méthode d'élicitation et d'agrégation des préférences sélectionnée, deux formats de questions sont utilisés : les questions binaires et les questions probabilistes. À partir des questions binaires, l'individu questionné déclare le candidat pour lequel il a l'intention de voter. Cependant, l'individu questionné peut être indécis face aux choix des candidats. Dans ce cas, les questions binaires avec incertitude lui permettent de pallier à cela. Les questions probabilistes quant à elles invitent l'individu questionné à quantifier l'importance qu'il donne à chacun des candidats.

Des études récentes en économie se sont portées sur le choix du format de questions. Manski (1990) a montré que les réponses aux questions binaires comportaient très peu d'informations sur le vote réel des individus. En effet, elles ne permettaient pas de mesurer la probabilité qu'un individu vote pour un candidat ou non. De plus, Manski (2002) a montré que lorsque certains électeurs sont indécis, il n'était pas possible de comparer les réponses obtenues entre différents répondants, contrairement aux questions probabilistes. Enfin Delavande et Manski (2010) ont comparé le pouvoir prédictif des questions binaires avec incertitudes et des questions probabilistes grâce aux données empiriques des élections présidentielles américaines de 2008. Les résultats ont montré que le pouvoir prédictif des questions probabilistes est supérieur à celui des questions binaires avec incertitudes.

Chapitre 2

L'expérience

Plusieurs facteurs peuvent expliquer la déviation d'une prédiction sur la réalisation d'un événement par rapport à la véritable réalisation de ce même événement. Pour cette raison, il serait difficile de mesurer correctement le biais de négligence de corrélation avec des données empiriques. Dès lors, le dispositif expérimental proposé permet de mesurer toute déviation des intentions/prédictions de vote des participants par rapport au choix de vote rationnel. Nous allons attribuer toute éventuelle déviation à la variation du niveau corrélation entre la probabilité de voter pour un candidat et la probabilité de participer au vote.

L'expérience consiste à une série de questions-réponses sous la forme de sondages. Dans la revue de littérature nous avons vu qu'il existe différentes écoles de pensée sur la meilleure manière de poser des questions dans les sondages pré-électorales, notamment les marchés de prédiction qui consiste à éliciter les anticipations des individus sur le résultat d'une élection et l'approche dite conventionnelle des sondages qui permet d'éliciter les intentions personnelles des individus. Nous avons également discuté du choix du format de questions, notamment les questions binaires et les questions probabilistes. Notre expérience combine toutes ces approches avec quatre traitements. Plus précisément, deux types d'information sont exploités. En effet, non seulement nous posons aux participants des questions binaires et probabilistes, mais nous leur demandons également de prédire le comportement des autres participants et leurs propres intentions de vote. De plus, nous faisons varier le niveau de corrélation entre la probabilité de participation et la probabilité de faire un des choix de vote à travers les traitements. Ainsi, notre expérience nous donne des informations sur l'anticipation des sondés quant au comportement agrégé de vote dans une population se trouvant dans la même situation qu'eux et sur l'anticipation propre aux sondés concernant leurs comportements de vote ou de participation au vote. Avec ces deux informations, nous pourrions observer l'écart entre les prédictions des participants et les résultats réels et déterminer les facteurs responsables de cette déviation.

2.1 Plan Expérimental

Dans l'expérience, les participants doivent choisir entre deux options (A et B), ou ne pas faire de choix. À chacun des choix A et B, des bénéfices et des frais sont associés. Ces frais sont élevés ou faibles, de sorte que le rendement net d'un participant découle de la différence entre la valeur de son choix et les frais impliquant ce choix. Également, chaque frais se réalise dépendamment d'une probabilité connue par le participant. Si le participant décide de ne pas faire de choix, alors il reçoit de façon certaine un montant.

Il y a 4 types de participants présents dans la population. Chaque type est caractérisé par une couleur : le bleu, le vert foncé, le vert pâle et le jaune. Chaque participant appartient à un seul type de couleur. Le nombre de personnes pour chaque type est le même tout le long de l'expérience, soit 25 personnes par type. Ici, les deux choix A et B représentent deux candidats en liste pour les élections, tandis que l'option externe équivaut à l'abstention dans le contexte électoral. Les différents types de participants font référence aux préférences des électeurs.

Les participants se réfèrent à un tableau rassemblant toutes les informations sur les rendements et aux coûts rattachés aux différents choix de vote. Le tableau 1 montre l'ensemble des informations présentées aux participants. Dans cette figure, les paramètres P_i , A_i , B_i , q_i et opt_i représentent respectivement la taille de chaque type, les valeurs associées aux options A et B, les probabilités associées aux coûts faibles et élevés et la valeur associée au fait de ne pas choisir.

| | Type | | | |
|--------------------------|---|--------------|--------------|-----------|
| | Bleu | Vert | | Jaune |
| | | Vert Foncé | Vert Pâle | |
| | P_B | P_{VF} | P_{VP} | P_J |
| | Valeur | | | |
| Option | | | | |
| A | A_B | A_{VF} | A_{VP} | A_J |
| B | B_B | B_{VF} | B_{VP} | B_J |
| | Probabilité d'avoir un coût élevé selon le type d'appartenance | | | |
| Frais associés aux choix | | | | |
| C_H | q_B | q_{VF} | q_{VP} | q_J |
| C_L | $1 - q_B$ | $1 - q_{VF}$ | $1 - q_{VP}$ | $1 - q_J$ |
| | Valeur associée au fait de ne pas choisir | | | |
| <i>Ne Pas choisir</i> | opt_B | opt_{VF} | opt_{VP} | opt_J |

Tableau 1: Ensemble des informations présentées aux participants

La figure 1 est un exemple de ce qu'aperçoivent les participants à l'écran durant l'expérience.

Dans cet exemple les valeurs associées à l'option A valent respectivement 120 pour le type bleu, 97 pour le type vert foncé, 44 pour le type vert pâle et 20 pour le type jaune. Ensuite, les valeurs associées à l'option B valent respectivement 20 pour le type bleu, 43 pour le type vert foncé, 96 pour le type vert pâle et 120 pour le type jaune. De même, à l'option de ne pas faire de choix (c'est-à-dire lorsque ni A ni B ne sont choisis), sont respectivement rattachées les valeurs 20 pour le type bleu, 43 pour le type vert foncé, 44 pour le type vert pâle et 20 pour le type jaune. La taille de chaque type est la même dans tous les traitements et tous les types ; soit 25 personnes par traitement. Enfin, les frais associés aux choix entre les options sont les mêmes dans tous les traitements. La valeur des coûts de participation est toujours de 60 lorsque les coûts sont élevés et de 0 lorsque les coûts sont faibles. Cependant, les probabilités de devoir payer les coûts varient non seulement en fonction du type auquel le participant appartient, mais aussi du traitement. Dans l'exemple présenté dans la figure 1, lorsque les frais sont élevés, les probabilités de devoir payer ces frais élevés sont respectivement de 71% pour le type bleu, 27 % pour le type vert foncé, 1% pour le type vert pâle et 1% pour le type jaune. Les probabilités de ne pas devoir payer ces frais élevés sont quant à eux respectivement de 29% pour le type bleu, 73% pour le type vert foncé, 99% pour le type vert pâle et 99% pour le type jaune.

FIGURE 1 – Exemple d'écran présenté aux participants

| Couleur | Groupe | | | | Information supplémentaire: |
|--|--|------------|-----------|-------|------------------------------------|
| | Bleu | Vert foncé | Vert pâle | Jaune | |
| Taille de groupe | 25 | 25 | 25 | 25 | Tour 1 Nombre maximal de tour 9 |
| Option A | Valeurs des options (en UPE) | | | | |
| Option B | 120 | 97 | 44 | 20 | |
| Frais associés au choix entre les options (en UPE) | Probabilité de devoir payer les frais (en %) | | | | |
| Frais élevé : 60 | 71 | 27 | 1 | 1 | |
| Frais bas : 0 | 29 | 73 | 99 | 99 | |
| "Ne pas faire de choix" | Valeur de "ne pas faire de choix" (en UPE) lorsque ni A ni B ne sont choisis | | | | |
| | 20 | 43 | 44 | 20 | |

Considérez les valeurs des options et les frais associés au choix selon le groupe présentés dans le tableau ci-dessus. Il y a 25 personnes dans chacun des groupes bleu, vert foncé, vert pâle et jaune soit au total 100 personnes.

Quelle est selon vous l'option qui sera choisie le plus souvent ?

L'option A sera choisie le plus souvent.
 L'option B sera choisie le plus souvent.

Instructions Suite

L'expérience se déroule en trois grandes tâches.

1) Durant la première tâche, les participants doivent anticiper les choix de 100 personnes

qui sont répartis également à travers les 4 types d'individus existant dans la population. Cette manière de demander les anticipations est comparable aux marchés de prédiction où les individus doivent former des anticipations sur le comportement des autres avant d'acheter ou vendre les options. La valeur des paramètres de l'expérience les guide dans leurs choix. À cette étape, ils répondent à deux questions de prédiction : la question 1 (ou la question 2) et la question 3a (ou la question 3b) ; l'ordre des questions est aléatoire. Les deux questions poursuivent le même objectif, à savoir anticiper les choix de 100 votants. Elles diffèrent par leur formulation. En effet, la manière dont la question est formulée peut affecter la réponse des participants (Manski (1990)). Les questions proviennent d'un questionnaire qui est présenté en annexe B. La question 1 suivante donne un exemple de question à laquelle les participants étaient confrontés.

Question 1

Considérez les valeurs des options et les frais associés au choix selon le type présentés dans le tableau ci-dessus. Il y a 25 personnes dans chacun des types bleu, vert foncé, vert pâle et jaune, soit au total 100 personnes.

Quelle est votre prédiction du choix agrégé de ces 100 personnes ?

Nombre de personnes qui choisiront l'une des options A ou B (sur un total de 100) ...

Nombre de personnes qui ne choisiront aucune des deux options (sur un total de 100) ...

Parmi ceux qui choisiront l'une des options, selon vous, combien de personnes choisiront chacune des options A et B ?

Svp, cochez si vous répondez en nombres absolus ou en termes de proportion et indiquez votre prévision.

Réponse absolue

- Je pense que ... personnes choisiront l'option A.*
- Je pense que ... personnes choisiront l'option B.*

Réponse relative

- Je pense que ... % des personnes choisiront l'option A.*
- Je pense que ... % des personnes choisiront l'option B.*

2) Deuxièmement, les participants doivent faire part de leurs intentions personnelles de vote. À cette étape, de l'information supplémentaire est rajoutée à l'attention des participants. En effet, l'expérimentateur lève le voile sur l'appartenance des individus de type bleus et jaunes. Cependant, l'incertitude demeure pour ceux appartenant aux types verts. Les participants

peuvent toujours se fier aux valeurs des paramètres pour émettre leurs choix. Ils sont informés sur le niveau de coût qui se réalisera. À cette étape, les participants répondent à la question 4a (ou à la question 4b) et aux questions 5a et 5b (ou seulement à la question 6) qu'on retrouve en annexe B. La question 6 suivante donne un exemple de question à laquelle les participants étaient confrontés.

Question 6

Quelles sont les chances que vous choisissiez chacune des options A et B ?

S'il vous plaît, indiquer la chance de choisir chaque option en termes de probabilités. "0" indique que vous êtes certain de ne pas vouloir choisir l'option concernée et "100" indique que vous êtes certain de vouloir choisir l'option concernée. Si vous choisissez un nombre compris entre "0" et "100", vous indiquez que vous n'êtes pas certain de choisir l'option concernée. Plus le nombre choisi est proche de "100", plus vous êtes certain de choisir l'option concernée. Plus le nombre choisi est proche de "0", plus vous êtes certain de ne pas choisir l'option concernée.

Les chances que je choisisse l'option A sont de : (en %) ...

Les chances que je choisisse l'option B sont de : (en %) ...

Les chances que je ne fasse pas de choix (ne choisisse aucune des deux options) sont de : (en %) ...

3) Enfin, la dernière tâche qui est demandée aux participants était de choisir l'une des options (A ou B) ou non. Durant cette ultime étape, tous les participants sont informés du type auquel ils appartiennent et des frais associés aux différentes options. Pour ce faire ils doivent répondre à cette ultime question :

À présent, vous connaissez les valeurs des options et le coût du choix pour vous. Quelle option choisissez-vous ?

– *Je choisis l'option A.*

– *Je choisis l'option B.*

– *Je ne choisis aucune des deux options : je ne fais pas de choix*

Après avoir effectué toutes ces tâches, le participant observe le gain se réaliser. Les participants répètent l'expérience durant 16 tours. Le gain total de chaque participant correspond au cumul des gains réalisés à chaque tour. L'exemple d'un tour complet est donné en annexe.

2.2 Types de questions de sondage dans l'expérience

Rappelons que ce questionnaire nous permet de ressortir les deux types d'information que nous voulons exploiter, à savoir l'anticipation des sondés quant au comportement agrégé de vote dans une population et les anticipations propres des sondés concernant leurs intentions de vote. Afin de disposer de cette information, les compagnies de sondage utilisent généralement les questions binaires. Cependant nous avons vu dans la revue de littérature que les questions binaires ne sont pas le format idéal pour capter l'incertitude (Manski (1990)). Il est donc proposé d'utiliser des questions probabilistes (Manski (1990)).

Les questions probabilistes

Les questions probabilistes invitent le répondant à quantifier l'importance qu'il donne à chaque choix. Ainsi à travers ce format de question le sondé informe sur la probabilité de voter pour un choix, mais aussi quantifie son incertitude face aux choix. Ce format peut être utilisé autant pour les prédictions agrégées qu'individuelles. Les questions suivantes tirées du questionnaire sont des exemples de questions probabilistes.

Question probabiliste - Intentions personnelles de vote

S'il vous plaît, indiquer la chance de choisir chaque option en termes de probabilités. "0" indique que vous êtes certain de ne pas vouloir choisir l'option concernée et "100" indique que vous êtes certain de vouloir choisir l'option concernée. Si vous choisissez un nombre compris entre "0" et "100", vous indiquez que vous n'êtes pas certain de choisir l'option concernée. Plus le nombre choisi est proche de "100", plus vous êtes certain de choisir l'option concernée. Plus le nombre choisi est proche de "0", plus vous êtes certain de ne pas choisir l'option concernée.

Quelles sont les chances que vous choisissiez chacune des options A et B ?

Les chances que je choisisse l'option A sont de : (en %) ...

Les chances que je choisisse l'option B sont de : (en %) ...

Les chances que je ne fasse pas de choix (ne choisisse aucune des deux options) sont de : (en %) ...

Question probabiliste - Prédiction agrégée

Considérez les valeurs des options et les frais associés au choix selon le type présentés dans le tableau ci-dessus. Il y a 25 personnes dans chacun des types bleu, vert foncé, vert pâle et jaune, soit au total 100 personnes.

Quelle est votre prédiction du choix agrégé de ces 100 personnes ?

Nombre de personnes qui choisiront l'option A (sur un total de 100) : ...

Nombre de personnes qui choisiront l'option B (sur un total de 100) : ...

Nombre de personnes qui ne choisiront aucune des deux options (sur un total de 100) : ...

2.3 Présentation de la mesure empirique

La mesure du biais de négligence de la corrélation proposée rejoint directement la définition énoncée plus tôt (la corrélation entre la probabilité de participation et la probabilité de choisir A(B)). En effet, les croyances des individus concernant les distributions jointes des probabilités dans les différents traitements de l'expérience nous permettent de calculer le coefficient Φ . Ce coefficient est une approche non paramétrique consistant à fournir une mesure du niveau d'association entre deux variables et son interprétation est la même que celle du coefficient de corrélation de Pearson.

Afin de calculer le coefficient de corrélation Φ , nous construisons une matrice 2×2 dans laquelle nous incluons les probabilités émises par les participants quant à leurs décisions de participation et leurs choix A(B). Supposons l'exemple fictif dans le tableau 2. Dans cet exemple, le participant i doit donner pour chaque traitement j ses intentions concernant sa participation et son choix entre A et B.

| | Ne participe pas au vote | Participer au vote | Total |
|-----------|--------------------------|--------------------|-------------|
| Choisit A | a_i^j | b_i^j | e_i^j |
| Choisit B | c_i^j | d_i^j | f_i^j |
| Total | g_i^j | h_i^j | Total = 100 |

Tableau 2: Table de contingence

Dans la matrix, a_i^j , b_i^j , c_i^j et d_i^j représentent les probabilités jointes. Par exemple, a_i^j représente la probabilité de ne pas participer et d'avoir une préférence pour le choix A ; et ainsi de suite. g_i^j et h_i^j représentent les probabilités marginales des décisions de participation, tandis que e_i^j et f_i^j représentent les probabilités marginales de choisir A et B. À travers cette matrice, nous sommes capables de calculer pour chaque participant i dans le traitement j , la valeur du coefficient de corrélation Φ de la manière suivante :

$$\Phi_i^j = \frac{(a_i^j \times d_i^j) - (c_i^j \times b_i^j)}{\sqrt{e_i^j \times f_i^j \times g_i^j \times h_i^j}} \quad (2.1)$$

Le coefficient Φ s'interprète comme le coefficient de corrélation de Pearson : sa valeur est entre -1 et $+1$. Plus il tend vers les extrêmes, plus la corrélation est forte. Plus il tend vers 0 alors la corrélation est nulle. Pour chaque observation, nous comparons la valeur empirique Φ_i^j au coefficient Φ (Φ^j) mesurant la corrélation pour chaque traitement de l'expérience tel que :

$$W_i^j = \widehat{\Phi}_i^j - \Phi^j \quad (2.2)$$

W mesure, pour chaque individu, sa déviation par rapport au choix rationnel. Ainsi W prend des valeurs entre -2 et $+2$, sachant que zéro signifie que le participant a fait le choix rationnel. Ainsi, nous parlerons de sous-estimation ou alors de négligence de la corrélation lorsque les $\widehat{\Phi}$ empiriques sont inférieurs aux Φ théoriques en valeurs absolues ; c'est-à-dire quand $0 < \widehat{\Phi}_i^j < \Phi^j$ ou $0 > \widehat{\Phi}_i^j > \Phi^j$.

2.4 Traitement de l'expérience

Initialement, l'expérience comportait au total 8 traitements. Les traitements se différencient par la corrélation entre les préférences et les coûts de participation induits par les différents paramètres de décisions (A_i, B_i, q_i, opt_i) (voir tableau 1.). Pour chaque traitement, il est possible de mesurer la déviation de la prédiction du participant par rapport aux probabilités justes. En effet, le niveau de corrélation théorique entre les probabilités de vote et les probabilités de participation au vote s'obtient grâce aux paramètres de l'expérience. Dès lors, nous pouvons le comparer au niveau de corrélation empirique découlant des choix réalisés par les participants. Les tableaux 3 et 4 présentent respectivement les résultats de la randomisation des participants et les paramètres de décision choisis pour les huit traitements.

| | Traitement | | | | | | | | |
|-------------------|------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|--------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Total |
| Bleu | 16 | 17 | 15 | 9 | 10 | 22 | 16 | 12 | 117 |
| vert foncé | 18 | 8 | 12 | 21 | 22 | 8 | 18 | 20 | 127 |
| vert pâle | 14 | 19 | 16 | 16 | 12 | 13 | 12 | 7 | 109 |
| Jaune | 10 | 14 | 12 | 12 | 14 | 15 | 12 | 19 | 111 |
| Total | 58 | 58 | 58 | 58 | 58 | 58 | 58 | 58 | 464 |

Tableau 3: Randomisation des participants selon le type et le traitement

| Traitement 1 $\phi(\text{Pop}) = 0.08$ $\phi(\text{Ind}) = 0.26$ | | | | |
|--|------|------------|-----------|-------|
| Groupe | Bleu | Vert Foncé | Vert Pâle | Jaune |
| Option A | 110 | 104 | 70 | 37 |
| Option B | 30 | 36 | 70 | 103 |
| Prob. cout élevé | 0.9 | 0.76 | 0.01 | 0.71 |
| Option externe | 30 | 36 | 70 | 37 |
| Traitement 2 $\phi(\text{Pop}) = -0.47$ $\phi(\text{Ind}) = -0.76$ | | | | |
| Groupe | Bleu | Vert Foncé | Vert Pâle | Jaune |
| Option A | 120 | 85 | 55 | 30 |
| Option B | 20 | 55 | 85 | 110 |
| Prob. cout élevé | 0.9 | 0.75 | 0.01 | 0.95 |
| Option externe | 20 | 55 | 55 | 30 |
| Traitement 3 $\phi(\text{Pop}) = 0.25$ $\phi(\text{Ind}) = 0.37$ | | | | |
| Groupe | Bleu | Vert Foncé | Vert Pâle | Jaune |
| Option A | 120 | 96 | 43 | 20 |
| Option B | 20 | 44 | 97 | 120 |
| Prob. cout élevé | 0.9 | 0.76 | 0.01 | 0.71 |
| Option externe | 20 | 44 | 43 | 20 |
| Traitement 4 $\phi(\text{Pop}) = 0$ $\phi(\text{Ind}) = 0$ | | | | |
| Groupe | Bleu | Vert Foncé | Vert Pâle | Jaune |
| Option A | 117 | 113 | 25 | 20 |
| Option B | 23 | 27 | 115 | 120 |
| Prob. cout élevé | 0.01 | 0.01 | 0.9 | 0.95 |
| Option externe | 23 | 27 | 25 | 20 |
| Traitement 5 $\phi(\text{Pop}) = -0.08$ $\phi(\text{Ind}) = -0.26$ | | | | |
| Groupe | Bleu | Vert Foncé | Vert Pâle | Jaune |
| Option A | 103 | 70 | 36 | 30 |
| Option B | 37 | 70 | 104 | 110 |
| Prob. cout élevé | 0.71 | 0.01 | 0.76 | 0.9 |
| Option externe | 37 | 70 | 36 | 30 |
| Traitement 6 $\phi(\text{Pop}) = 0.47$ $\phi(\text{Ind}) = 0.76$ | | | | |
| Groupe | Bleu | Vert Foncé | Vert Pâle | Jaune |
| Option A | 110 | 85 | 55 | 20 |
| Option B | 30 | 55 | 85 | 120 |
| Prob. cout élevé | 0.95 | 0.01 | 0.75 | 0.9 |
| Option externe | 30 | 55 | 55 | 20 |
| Traitement 7 $\phi(\text{Pop}) = -0.25$ $\phi(\text{Ind}) = -0.37$ | | | | |
| Groupe | Bleu | Vert Foncé | Vert Pâle | Jaune |
| Option A | 120 | 97 | 44 | 20 |
| Option B | 20 | 43 | 96 | 120 |
| Prob. cout élevé | 0.71 | 0.27 | 0.01 | 0.01 |
| Option externe | 20 | 43 | 44 | 20 |
| Traitement 8 $\phi(\text{Pop}) = 0$ $\phi(\text{Ind}) = 0$ | | | | |
| Groupe | Bleu | Vert Foncé | Vert Pâle | Jaune |
| Option A | 120 | 115 | 27 | 23 |
| Option B | 20 | 25 | 113 | 117 |
| Prob. cout élevé | 0.95 | 0.9 | 0.01 | 0.01 |
| Option externe | 20 | 25 | 27 | 23 |

Tableau 4: Paramètres de l'expérience

Chapitre 3

Résultats

L'expérience a été réalisée en quatre sessions. Les quatre sessions se sont tenues durant le mois de novembre 2017 dans le Laboratoire d'Économie Expérimentale de l'Université Laval (LEEL). Le nombre des participants variait entre 2 et 12 par séance. Finalement, ce sont 29 individus qui ont pris part à l'expérience. Ils ont répété la même expérience durant 16 tours. Dès lors, notre base de données est constituée de 464 observations. Les participants répondaient au questionnaire à l'aide d'ordinateurs leur présentant les différents écrans réalisés grâce au logiciel ZTree. La disposition des ordinateurs du LEEL est telle que les participants étaient isolés les uns des autres de sorte qu'ils étaient incapables de communiquer entre eux. Une vidéo introductrice donnait les instructions aux participants. Cette dernière était d'une durée de 6 minutes. Par la suite, les participants devaient parcourir les 16 tours de l'expérience. En moyenne, les participants ont mis 42 minutes pour compléter l'expérience.

Les participants étaient rémunérés selon le choix qu'ils faisaient durant l'expérience. En effet, trois choix s'offraient à eux : l'option A, B ou l'option de ne pas voter. Lorsqu'ils décidaient de ne pas voter, le gain du participant correspondait à la valeur associée à cette option. Par contre, en choisissant l'option A ou B le gain du participant correspondait à la valeur associée à cette option, à laquelle le coût associé à la décision de voter est soustrait.

Enfin, les gains étaient affichés en unités expérimentales (UPE) à la fin de chaque tour. Ce n'est qu'à la fin de l'expérience que la conversion en dollar canadien était effectuée; le taux de change étant de 1CAD dollar pour 50 UPE.

3.1 Statistiques descriptives

3.1.1 caractéristiques socio-économiques

Les données collectées ont donné des informations sur les paramètres de l'expérience, les caractéristiques sociodémographiques des participants et les gains résultants des décisions prises par les participants. Le tableau 5 résume la distribution de ces caractéristiques. Il contient

trois colonnes décrivant les caractéristiques des participants lors des prédictions agrégées, lors des prédictions individuelles et à travers l'ensemble de la population questionnée. La distribution des caractéristiques varie à travers les trois colonnes car elles n'ont pas le même nombre d'observations. Mais nous concentrons d'abord sur les caractéristiques dans l'ensemble de la population. Nous reviendrons en détail sur la singularité des deux premières colonnes plus tard.

Les participants sont âgés entre 21 et 59 ans, avec une moyenne d'âge de 31 ans et 52 % sont des femmes. De plus 45 % des participants ont mentionné qu'ils détiennent un diplôme du premier cycle, 24 % du deuxième cycle, 14 % du troisième cycle et 17 % ont mentionné détenir un autre niveau d'éducation. Enfin, 48 % des participants ont indiqué avoir étudié dans un programme d'étude en sciences sociales ; les 51 % restant ayant étudié dans d'autres programmes excluant ceux des sciences sociales (par exemple médecine, philosophie et les programmes de génies). Les participants recevaient 5\$ pour avoir répondu présents à l'expérience. Le gain moyen perçu à la fin de l'expérience s'est élevé à un peu plus de 12\$.

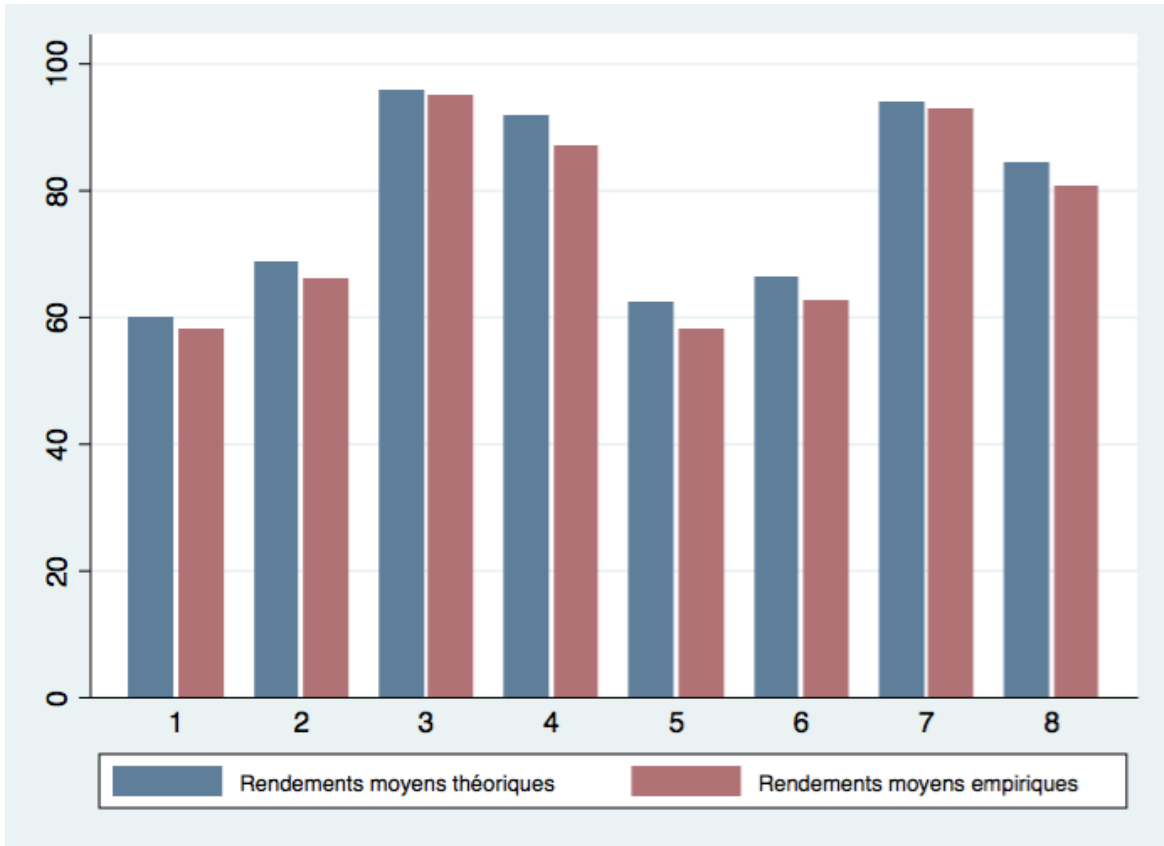
| Variables | Description | Préd. agrégées | Préd. individuelles | Population |
|---|--|----------------|---------------------|------------|
| Distribution selon l'âge | | | | |
| <i>Age</i> ₁ | Catégorie d'âge [21 – 24] | 29.1% | 22.0% | 31.0% |
| <i>Age</i> ₂ | Catégorie d'âge [25 – 34] | 27.2% | 40.3% | 38.0% |
| <i>Age</i> ₃ | Catégorie d'âge [35 – 59] | 33.7% | 37.7% | 31.0% |
| Distribution selon le genre | | | | |
| Femme | = 1 Si femme, = 0 si homme | 50.1% | 58.4% | 51.7% |
| Homme | = 1 Si homme, = 0 si femme | 49.9% | 41.6% | 48.3% |
| Distribution selon le niveau d'éducation | | | | |
| Baccalauréat | = 1, si Baccalauréat, = 0 sinon | 42.3% | 40.3% | 44.8% |
| Maîtrise | = 1, Si Maîtrise, = 0 sinon | 27.1% | 25.8% | 24.1% |
| Doctorat | = 1, Si Doctorat, = 0 sinon | 11.9% | 15.1% | 13.8% |
| Autres | = 1, si autre niveau t, = 0 sinon | 18.7% | 18.9% | 17.3% |
| Distribution selon la faculté | | | | |
| Faculté : sciences sociales | = 1 Si en sciences sociales, = 0 sinon | 47.6% | 51.6% | 51.7% |
| Faculté : autres | = 1 Si pas en sciences sociales, = 0 sinon | 52.4% | 48.4% | 48.3% |
| Distribution selon le traitement | | | | |
| Traitement 1 | = 1 si traitement 1, = 0 sinon | 14.2% | 12.0% | 12.5% |
| Traitement 2 | = 1 si traitement 2, = 0 sinon | 14.4% | 13.8% | 12.5% |
| Traitement 3 | = 1 si traitement 3, = 0 sinon | 11.9% | 15.1% | 12.5% |
| Traitement 4 | = 1 si traitement 4, = 0 sinon | 9.6% | 12.0% | 12.5% |
| Traitement 5 | = 1 si traitement 5, = 0 sinon | 13.7% | 12.0% | 12.5% |
| Traitement 6 | = 1 si traitement 6, = 0 sinon | 14.2% | 10.7% | 12.5% |
| Traitement 7 | = 1 si traitement 7, = 0 sinon | 13.4% | 15.1% | 12.5% |
| Traitement 8 | = 1 si traitement 8, = 0 sinon | 8.6% | 9.4% | 12.5% |
| Nombre total d'observations | | 395 | 159 | 464 |

Tableau 5: Statistiques descriptives

3.1.2 Aperçu de la distribution des gains

Au cours de l'expérience, les participants ont été rémunérés selon les choix de votes qu'ils ont réalisés. Ainsi il est possible de comparer les gains découlant des choix réalisés avec les gains des choix rationnels. La figure 2 compare la distribution moyenne des rendements issus des choix réels et ceux issus des choix rationnels selon les traitements. Dans tous les traitements, les participants ont eu tendance à prendre des décisions menant à des gains inférieurs à ceux résultant du choix rationnel. Le même résultat pouvait être observé en comparant ces deux distributions selon le type auquel ils d'appartiennent.

FIGURE 2 – Distribution de la moyenne des gains théoriques et empiriques à travers les 8 traitements



On a également remarqué que cette déviation par rapport au choix rationnel a été observée 29 fois. En examinant de plus près ces 29 déviations, on a remarqué qu'elles ont été commises par 9 participants ; soit 30% des sujets de notre expérience (Voir tableau 6).

| # d'erreurs commises | # de participants les ayant commises |
|------------------------------------|--------------------------------------|
| 1 | 5 |
| 2 | 1 |
| 6 | 1 |
| 7 | 1 |
| 9 | 1 |
| Total : 29 erreurs commises | Total : 9 participants |

Tableau 6: Distribution des erreurs commises par les participants

Le tableau 6 nous indique qu'un individu a commis à lui seul 6 erreurs, un autre en a commis 7 et un troisième en a commis 9. À eux trois, ces individus auraient commis 76% du nombre total d'erreurs.

En examinant de plus près les 3 individus ayant commis la majorité des erreurs, il a été

possible d'évaluer la perte, c'est-à-dire l'impact de commettre des erreurs. Une analyse de la différence entre le rendement moyen découlant de la décision optimale de l'agent rationnel et le rendement effectivement réalisé par les participants a montré que les conséquences ont été substantielles lorsqu'un individu commettait beaucoup d'erreurs. En effet, si le participant ayant commis 9 erreurs avait fait le bon choix à chacun des 16 tours, il aurait réalisé un gain de 25.9 \$. Les 9 erreurs commises ont réduit ses gains à 21.2 \$. Ensuite, pour celui en ayant 7, le gain potentiel total était de 26.1 \$. Les 7 erreurs commises ont réduit ses gains à 18.4 \$. Enfin, pour celui en ayant commis 6, on est passé d'un gain potentiel total de 23.7 \$ à un gain total réalisé de 16.7 \$ pour les 16 tours. Ces écarts dénotent un manque substantiel à gagner lorsque le participant commettait beaucoup d'erreurs. À titre comparatif, la différence entre les gains potentiellement réalisables et les gains effectivement réalisés durant les 16 tours chez les participants ayant commis 1 seule erreur n'était en moyenne que de 0.3 \$.

Ces observations nous ont amenés à tester si la moyenne des rendements théoriques et empiriques sont statistiquement différentes à travers les traitements. Pour ce faire, nous avons effectué trois tests de différence. Afin de relâcher l'hypothèse de distribution normale, nous avons effectué des tests paramétriques (statistique de Student) et non paramétriques (statistique de Wilcoxon). Le seuil de signification est de 5%.

Premièrement, nous avons effectué un test de différence à travers les traitements sans tenir compte du signe de la corrélation dans les traitements. Dans ce premier test, l'hypothèse nulle selon laquelle la différence entre les gains réels et les gains attendus est la même est rejetée ($t = 5.4$).

Deuxièmement, nous avons testé si ce résultat différait selon que la corrélation soit positive ou négative. Nous avons donc testé l'hypothèse nulle selon laquelle cette différence était la même selon que les traitements soient de corrélation positive ou négative. L'hypothèse nulle n'a pas été rejetée ($t = 1.03$ et $z = 1.04$).

Enfin, nous avons testé si le résultat est le même lorsqu'on ne prenait pas en compte les traitements dont la corrélation était nulle. L'hypothèse nulle n'a pas été rejetée ($t = 0.38$ et $z = -0.7$).

En retirant les trois individus ayant commis la majorité des erreurs, nous obtenons des résultats différents. En effet, lorsqu'on retire les 3 individus, l'hypothèse nulle selon laquelle la différence entre les gains réels et les gains attendus est la même n'est pas rejetée ($t = 1.7$). Ce résultat ne diffère pas selon que les traitements soient de corrélation positive ou négative ($t = -0.3$ et $z = -1.4$). Enfin, en testant si la différence entre les gains réels et les gains attendus est la même lorsqu'on ne tient pas compte des traitements à corrélation nulle, l'hypothèse nulle n'est pas rejetée mais la statistique de test demeure faible. ($t = -1.7$ et $z = 2.08$).

3.1.3 Analyse descriptive du coefficient de corrélation Φ

Dans cette section, nous analysons les choix des participants à travers la distribution de la mesure empirique proposée dans le chapitre 2 : le coefficient de corrélation Φ . Il est important de noter que seules les réponses aux questions probabilistes 1,2,5 et 6 présentées en annexe B ont permis le calcul du coefficient de corrélation Φ . Plus précisément, les questions 1 et 2 demandaient aux individus d'effectuer des prédictions agrégées. Dès lors, à travers les réponses à ces questions, nous avons calculé les coefficients Φ des prédictions agrégées. Quant aux réponses des questions 5 et 6, elles ont donné les coefficients Φ des prédictions individuelles, vu que ces questions demandaient aux individus appartenant aux types verts leurs propres intentions de vote.

Pour 69 observations reliées aux prédictions agrégées et 233 observations reliées aux prédictions individuelles, le dénominateur était nul, car une des quatre probabilités marginales qui le constituent était égal à zéro. Ceci est le cas quand :

- un individu est certain de voter, c'est-à-dire $P(V=1) = 100\%$,
- un individu est certain de ne pas voter, c'est-à-dire $P(V=0) = 100\%$,
- un individu est certain de choisir A, c'est-à-dire $P(A) = 100\%$ et
- un individu est certain de choisir B, c'est-à-dire $P(B) = 100\%$.

Dès lors cela avait des répercussions sur le coefficient de corrélation Φ qui devenait non calculable. Le tableau 7 donne un exemple d'une observation de l'expérience pour laquelle le coefficient Φ était indéterminable.

| | Ne participe pas au vote | Participe au vote | Total |
|-----------|--------------------------|-------------------|-------------|
| Choisit A | 100 | 0 | 100 |
| Choisit B | 0 | 0 | 0 |
| Total | 100 | 0 | Total = 100 |

Tableau 7: Exemple de table de contingence pour un coefficient indéterminable

Dans cet exemple, le participant a clairement manifesté son choix de ne pas participer au vote mais d'avoir une préférence certaine pour le choix A. En intégrant ces réponses dans la formule du coefficient de corrélation on obtient le résultat suivant :

$$\Phi_i^j = \frac{(100 \times 0) - (0 \times 0)}{\sqrt{0 \times 0 \times 100 \times 0}} = \frac{0}{0} \quad (3.1)$$

Le coefficient de corrélation n'est plus déterminable dans ce cas. Cet exemple n'est pas le seul menant à ce résultat. Plusieurs combinaisons peuvent amener à un coefficient de corrélation non déterminable.

Dans la prochaine section nous analyserons d'abord les déterminants socioéconomiques des coefficients $\widehat{\Phi}$ pour les prédictions agrégées et individuelles. Pour cette analyse nous utilisons uniquement les observations avec les $\widehat{\Phi}$ calculables. Ensuite nous examinons $\widehat{\Phi}$ non calculables en profondeur dans le chapitre 4.

3.2 Analyse descriptive des *coefficients $\widehat{\Phi}$ calculables*

À partir des données recueillies au cours de l'expérience nous avons calculé pour chaque observation les coefficients de corrélation $\widehat{\Phi}$ des participants et les avons comparés aux coefficients théoriques issus des paramètres de l'expérience. Tel qu'expliqué dans le chapitre 2, nous définissons le biais comme la différence entre la valeur empirique du coefficient lié au choix du participant et la valeur du coefficient inhérent à chaque traitement de l'expérience

$$W_i^j = \widehat{\Phi}_i^j - \Phi^j \quad (3.2)$$

où :

- W_i^j correspond au biais pour l'observation i dans le traitement j ;
- $\widehat{\Phi}_i^j$ est la valeur empirique du coefficient de corrélation pour l'observation i dans le traitement j ;
- et Φ^j est la valeur théorique du coefficient de corrélation dans le traitement j .

Les graphiques 2 et 3 présentent respectivement pour chaque traitement la distribution des coefficients Φ agrégés et individuels autour du Φ théorique. La valeur des coefficients $\widehat{\Phi}$ empiriques est comprise entre -1 et 1. Les lignes vertes représentent le niveau de corrélation théorique. À travers ces figures, on remarque que l'écart entre les Φ empiriques et les Φ théoriques peut être classé selon 4 cas :

1. Lorsque les Φ empiriques sont inférieurs aux Φ théoriques en valeurs absolues ; c'est-à-dire quand $0 < \widehat{\Phi}_i^j < \Phi^j$ ou $0 > \widehat{\Phi}_i^j > \Phi^j$, nous parlerons de sous-estimation de la corrélation ou de négligence de la corrélation ;
2. Lorsque les valeurs des Φ empiriques et des Φ théoriques sont égales ; c'est-à-dire quand $\widehat{\Phi}_i^j = \Phi^j$, nous dirons que nous avons une estimation exacte de la corrélation ;
3. Lorsque les Φ empiriques sont supérieures aux Φ théoriques en valeurs absolues ; c'est-à-dire quand $0 < \Phi^j < \widehat{\Phi}_i^j$ ou $0 > \Phi^j > \widehat{\Phi}_i^j$, nous parlerons de surestimation de la corrélation ;
4. Lorsque les Φ empiriques et les Φ théoriques se positionnent de part et d'autre de 0 ; c'est-à-dire quand $\Phi^j < 0 < \widehat{\Phi}_i^j$ ou $\Phi^j > 0 > \widehat{\Phi}_i^j$, nous parlerons de perception inverse de la corrélation.

FIGURE 2 – Distribution des $\hat{\Phi}$ empiriques agrégés autour du Φ théorique (ligne verte)

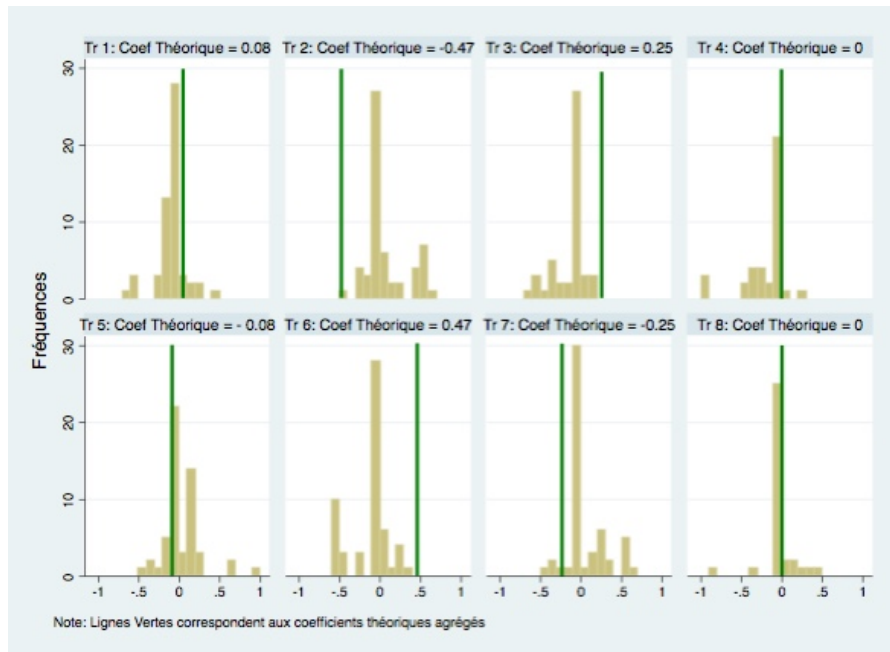
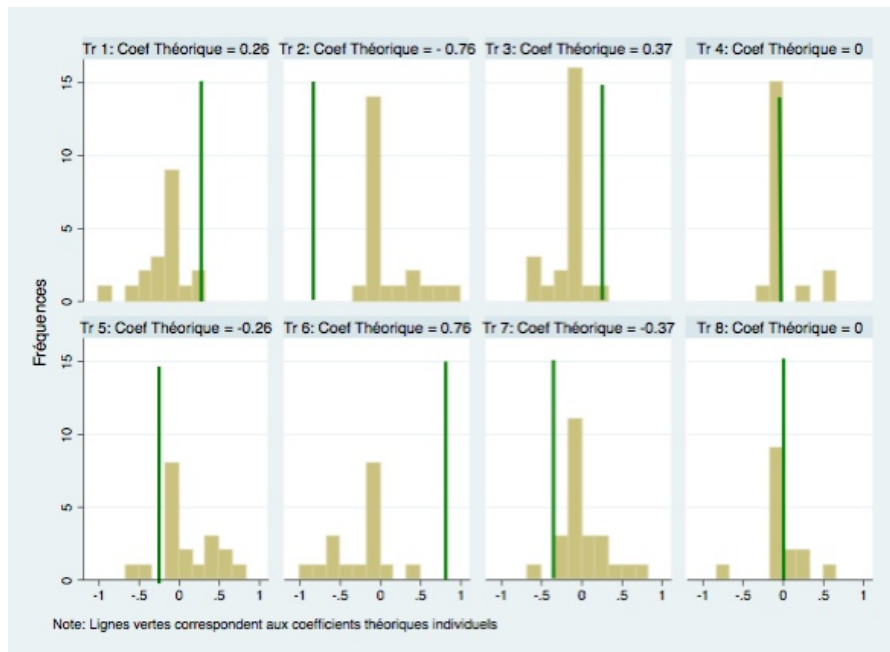


FIGURE 3 – Distribution des $\hat{\Phi}$ empiriques individuels autour du Φ théorique (ligne verte)



Les individus semblent souvent s'écarter de la corrélation exacte entre la décision de participer et celle de faire un choix. Également, les prédictions semblent moins s'écarter du choix rationnel dans les traitements 4 et 8, relativement aux autres traitements. En effet, les fluctuations des Φ empiriques autour des Φ théoriques semblent être relativement moindres dans ces 2 traitements. Ces observations nous suggèrent d'approfondir notre analyse sur les déterminants

du biais de négligence de la corrélation. Afin d'être en ligne avec l'objectif de ce mémoire qui est, rappelons-le, d'examiner le biais de négligence de la corrélation, nous nous attarderons plus sur ce dernier lors de l'analyse et de l'interprétation des résultats.

3.3 Modèle empirique

Dans cette section nous proposons un modèle pour analyser plus profondément les observations graphiques faites dans la section précédente. En effet nous avons remarqué que les participants semblaient plus disposés à souffrir du biais dans certains traitements que dans d'autres. En d'autres termes, nous tentons de voir s'il existe une relation entre la variation de l'intensité du Φ théorique à travers les traitements et le niveau du biais de négligence de corrélation chez les participants. Nous voulons aussi tester si les caractéristiques socioéconomiques des participants pouvaient également expliquer le niveau du biais de négligence de corrélation.

Afin d'analyser les déterminants du biais de négligence de la corrélation pour les prédictions agrégées et individuelles, nous avons estimé un probit ordonné. En effet, un probit ordonné a l'avantage d'analyser les changements de modalités dans la variable expliquée lorsqu'il existe un ordre dans les modalités de cette dernière. Dans notre analyse, le probit ordonné peut être défini comme un modèle où les modalités prises par le biais de négligence de la corrélation sont des codes attribués aux intervalles de valeurs obtenus par découpage des valeurs continues du Φ empirique autour du Φ théorique. Comme nous l'avons vu dans les graphiques 2 et 3 de la section précédente, l'écart entre les corrélations empiriques et les corrélations théoriques (c'est-à-dire le biais de négligence de la corrélation) peut être regroupé selon quatre modalités. Dès lors, notre modèle se présente comme suit :

$$\Phi = \begin{cases} 1 & \text{si } \Phi^j < 0 < \Phi_i^* \text{ (ou } \Phi^j > 0 > \Phi_i^*) \text{ (Perception inverse)} \\ 2 & \text{si } 0 < \Phi_i^* < \Phi^j \text{ (ou } 0 > \Phi_i^* > \Phi^j) \text{ (Sous-évaluation)} \\ 3 & \Phi_i^* = \Phi^j \text{ (Évaluation exacte)} \\ 4 & \text{si } 0 < \Phi^j < \Phi_i^* \text{ (ou } 0 > \Phi^j > \Phi_i^*) \text{ (Surévaluation)} \end{cases} \quad (3.3)$$

où Φ_i^* représente la variable expliquée, soit la valeur du Φ empirique autour de la vraie valeur de la corrélation dans un traitement donné, Φ^j . L'écart entre les corrélations empiriques et les corrélations théoriques peuvent prendre les modalités 1, 2, 3, et 4 qui correspondent respectivement à une perception inverse de la corrélation, à une sous-estimation de la corrélation, à une corrélation exacte, et à surestimation de la corrélation. Ces modalités correspondent à des intervalles de valeurs de la variable latente Φ_i^* qui suit, par hypothèse, une distribution normale et est définie telles que :

$$\Phi_i^* = \beta' X_i + \mu_i \quad (3.4)$$

$$E(\mu_i) = 0$$

$$V(\mu_i) = \sigma_u^2 = 1$$

Dans le modèle, Φ^j (les niveaux de corrélation théorique) et 0 sont des constantes délimitant les intervalles de valeurs de la variable latente Φ_i^* . Comme on peut le voir dans l'équation 3.4, le participant fait le choix de la modalité 4 (ou surestime la corrélation) si la valeur du coefficient empirique Φ_i^* est dans l'absolu supérieure au coefficient théorique Φ^j , qui lui est supérieur à zéro. De la même manière, le participant fait le choix de la modalité 3 lorsque la valeur coefficient empirique Φ_i^* est égale au coefficient théorique Φ^j ; et ainsi de suite. X_i est une matrice 4×4 comprenant toutes les caractéristiques des participants pouvant potentiellement affecter la corrélation. Cette matrice est constituée :

- des variables *Traitements* qui sont des variables binaires captant l'effet d'appartenir à un traitement. Les traitements 4 et 8 avec des corrélations théoriques égales à zéro ont été choisis comme traitements de référence (8 pour les prédictions agrégées et les 4 prédictions individuelles) ;
- de la variable *Femme* qui est une variable binaire égale à 1 si la répondante est une femme ; le genre *Homme* étant la catégorie de référence ;
- des variables *Age₂* et *Age₃* qui sont également sous forme binaire et captent l'effet de l'âge. *Age₂* représente la catégorie d'âge entre 25 ans et 34 ans pour les prédictions agrégées et les prédictions individuelles. *Age₃* représente la catégorie d'âge entre 35 ans et 59 ans pour les prédictions agrégées et les prédictions individuelles. Enfin, la catégorie de référence *Age₁* correspond à la catégorie d'âge entre 25 et 32 ans pour les prédictions agrégées et les prédictions individuelles ;
- des variables *Maîtrise*, *Baccalauréat*, *Doctorat* et *Autres* qui sont des variables binaires et captent l'effet du niveau d'éducation du participant. À noter qu'une catégorie *autres* indique que les participants ont un niveau de scolarité autre qu'universitaire et la variable *Maîtrise* est prise comme catégorie de référence ;
- de la variable *Sciences Sociales* qui correspond à une variable binaire captant l'effet d'étudier dans un programme de sciences sociales (par exemple sciences politiques, économie, sociologie, Anthropologie, droit etc.) ou non (par exemple médecine, philosophie, les programmes de génies etc.).
- Enfin, μ est le terme d'erreur. nous faisons l'hypothèse qu'il est centré en 0 et de variance 1. En effet, le modèle probit correspond au choix de la fonction de répartition $F(\cdot)$ de la loi normale.

Connaissant les valeurs des constantes (que nous notons ici les γ) et de la fonction de répartition $F(\cdot)$, on peut déduire une formule générale :

$$Prob(\Phi_i = k) = F(\gamma_k - X_i\beta) - F(\gamma_{k-1} - X_i\beta) \quad (3.5)$$

Avec γ_k et γ_{k-1} égaux à Φ^j ou 0

Pour estimer le modèle, il faut d'abord calculer la fonction de vraisemblance qui est le double produit des probabilités des modalités à la fois entre les modalités, mais aussi entre les participants. Elle se présente comme suit :

$$L(\Phi, \gamma, \beta) = \prod_{i=1}^n \prod_{k=1}^m [F(\gamma_k - X_i\beta) - F(\gamma_{k-1} - X_i\beta)] \quad (3.6)$$

Avec γ_k et γ_{k-1} égaux à Φ^j ou 0

On peut dès lors déterminer les paramètres de ce modèle en maximisant cette fonction de vraisemblance. Mais vu que maximiser la fonction de vraisemblance équivaut aussi à maximiser la log-vraisemblance, alors la détermination est basée sur la fonction de log-vraisemblance car plus facile à dériver :

$$LogL(\Phi, \gamma, \beta) = \sum_{i=1}^n \sum_{k=1}^m \log[F(\gamma_k - X_i\beta) - F(\gamma_{k-1} - X_i\beta)] \quad (3.7)$$

Avec γ_k et γ_{k-1} égaux à Φ^j ou 0

Connaissant ainsi la fonction $F(\cdot)$, il suffit alors de dériver l'équation 3.7 afin de rechercher les valeurs des paramètres γ et β qui sont des vecteurs de paramètre contenant respectivement les valeurs critiques (*cut off* en anglais) et les coefficients variables explicatives. Il faut noter que X_i ne contient pas de constante pour des raisons d'identification. En effet, il est impossible d'identifier en même temps le coefficient associé à la constante et les valeurs seuils.

Nous avons estimé ce modèle avec le logiciel STATA. Les coefficients estimés ne s'interprètent pas directement mais leurs signes indiquent dans quel sens se modifie la probabilité. Pour faciliter l'interprétation, nous calculons les effets marginaux afin de comprendre l'effet de la variation d'une unité des variables indépendantes sur chacun des quatre niveaux de la corrélation entre la décision de participer et celle de faire un choix. Outre l'estimation des coefficients, Stata estime la valeur des seuils du modèle sous-jacent. On peut alors tester la

pertinence de ces seuils avec la commande *lincom*. Ces tests nous ont permis de conclure qu'un modèle où l'on agrégerait les modalités ne serait pas pertinent puisque la différence entre les points de séparations est différente de zéro.

3.4 Résultats

Le tableau 22 en annexe A donne la distribution des Φ pour les prédictions agrégées et individuelles selon les traitements et les caractéristiques des participants. Le tableau 8 présente les résultats des deux probit ordonnés. Rappelons que les participants répondaient aux questions agrégées ($Q1$ et $Q2$) ensuite aux questions individuelles ($Q5$ et $Q6$). Plus précisément, ils devaient répondre soit à la question 1 ou à la question 2 pour les prédictions agrégées, soit à la question 5 ou la question 6 pour les prédictions individuelles.

On s'intéresse d'abord aux prédictions agrégées. Les résultats obtenus indiquent qu'il y a une différence significative au niveau du genre des participants. En effet, les femmes ont tendance à moins sous-estimer l'intensité de la corrélation ou à moins percevoir la corrélation inverse par rapport aux hommes ; ce résultat est significatif à 1%. On peut faire le même constat concernant les catégories d'âge puisque les participants appartenant aux catégories d'âge [25; 34] et [32; 59] ont également tendance à moins sous-estimer l'intensité de la corrélation ou à moins percevoir la corrélation inverse par rapport aux participants de la catégorie de référence. Ce résultat est également significatif à 1%. Ainsi, les participants les plus âgés auraient tendance à sous-estimer la corrélation. Concernant le niveau d'étude, on remarque qu'il n'y a pas de différences significatives entre les participants qui ont le baccalauréat comme plus haut niveau d'étude et ceux qui ont la maîtrise. Par contre, les participants qui ont le doctorat ou un niveau d'étude tout autre ont tendance à plus sous-estimer l'intensité de la corrélation ou à plus percevoir la corrélation inverse comparativement à ceux qui ont le baccalauréat comme plus haut niveau d'étude. Ce résultat est significatif à 1%. Concernant les participants en faculté des sciences sociales, les résultats indiquent qu'appartenir ou non à la faculté des sciences sociales n'a pas d'effet sur la perception de la corrélation. Enfin, les résultats indiquent que peu importe si la corrélation est positive ou négative dans un traitement, les participants ont tendance à plus sous-estimer l'intensité de la corrélation ou à plus percevoir la corrélation inverse. Ce résultat est significatif à 1%.

| Variable dépendante : catégories de W | | | |
|---|----------------------|-----------------------|---------------------------|
| Variables | Prédictions agrégées | | Prédictions individuelles |
| Homme | référence | | référence |
| Femme | 0.53*** | | -0.08 |
| | (0.12) | | (0.21) |
| Catégorie d'âge [21 - 24] | référence | | référence |
| Catégorie d'âge [25 - 34] | 0.50*** | | 0.74*** |
| | (0.16) | | (0.28) |
| Catégorie d'âge [35 - 59] | 0.54*** | | -0.44 |
| | (0.17) | | (0.29) |
| Niveau d'études : Maîtrise | référence | | référence |
| Niveau d'études : Baccalauréat | -0.01 | | 0.19 |
| | (0.15) | | (0.31) |
| Niveau d'études : Doctorat | -0.50** | | -0.25 |
| | (0.21) | | (0.32) |
| Niveau d'études : Autres | -0.48** | | -0.52 |
| | (0.19) | | (0.36) |
| Faculté : Autres | référence | | référence |
| Faculté : Sciences Sociales | 0.03 | | 0.30 |
| | (0.13) | | (0.22) |
| Traitement 8 : $\Phi = 0$ (Agr.) | référence | $\Phi = 0$ (Ind.) | référence |
| Traitement 1 $\Phi = 0.08$ (Agr.) | -1.95*** | $\Phi = 0.26$ (Ind.) | -2.40*** |
| | (0.26) | | (0.43) |
| Traitement 2 $\Phi = -0.47$ (Agr.) | -2.14*** | $\Phi = -0.76$ (Ind.) | -2.37*** |
| | (0.26) | | (0.42) |
| Traitement 3 $\Phi = 0.25$ (Agr.) | -2.16*** | $\Phi = 0.37$ (Ind.) | -2.90*** |
| | (0.27) | | (0.44) |
| Traitement 4 $\Phi = 0$ (Agr.) | -0.60** | $\Phi = 0$ (Ind.) | -0.48 |
| | (0.26) | | (0.38) |
| Traitement 5 $\Phi = -0.08$ (Agr.) | -1.71*** | $\Phi = -0.26$ (Ind.) | -2.41*** |
| | (0.25) | | (0.42) |
| Traitement 6 $\Phi = 0.47$ (Agr.) | -2.09*** | $\Phi = 0.76$ (Ind.) | -2.61*** |
| | (0.26) | | (0.46) |
| Traitement 7 $\Phi = -0.25$ (Agr.) | -1.86*** | $\Phi = -0.37$ (Ind.) | -2.42** |
| | (0.27) | | (0.41) |
| Seuil 1 | -1.86 | | -2.27 |
| | (0.26) | | (0.54) |
| Seuil 2 | 0.075 | | -0.30 |
| | (0.24) | | (0.51) |
| Seuil 3 | 0.75 | | 0.62 |
| | (0.26) | | (0.50) |
| Observations | 395 | | 159 |

écarts types entre parenthèses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tableau 8: Coefficient du modèle présenté

Passons maintenant aux prédictions individuelles. Les résultats obtenus indiquent qu'il n'y a pas de différence significative au niveau du genre des participants. Idem pour le niveau d'éducation et l'appartenance à une faculté. Concernant les catégories d'âge, les participants ayant

entre 25 et 34 ans ont tendance à moins sous-estimer l'intensité de la corrélation ou à percevoir la corrélation inverse comparativement aux participants de la catégorie de référence, tandis qu'on observe pas de différence significative entre ceux de la catégorie d'âge la plus élevée et ceux entre 21 et 24 ans. Enfin, les résultats concernant l'appartenance à un traitement sont similaires à ceux observés lors des prédictions agrégées. En effet, peu importe si la corrélation est positive ou négative dans un traitement, les participants ont tendance à plus sous-estimer l'intensité de la corrélation ou à plus percevoir la corrélation inverse. Ce résultat est significatif à 1%. Ainsi, on remarque pour que les prédictions individuelles, plusieurs paramètres perdent de leur significativité. Cela pourrait possiblement s'expliquer par la différence dans le nombre d'observations disponibles pour analyser les prédictions individuelles et les prédictions agrégées. En effet, seuls les participants appartenant au groupe des verts sont pris en compte pour effectuer les prédictions individuelles alors que tous les groupes sont pris en compte pour l'analyse des prédictions agrégées. Dès lors, après exclusion des observations manquantes, les prédictions agrégées comptabilisent 395 observations, tandis que les prédictions individuelles n'en comptent que 159.

La littérature sur le modèle probit ordonné et l'interprétation des coefficients estimés nous enseigne qu'on ne peut pas simultanément interpréter le signe et l'intensité du coefficient estimé. En effet si une relation est significative, seul le signe du coefficient indique le sens de celle-ci. Mais à partir de ces coefficients nous pouvons calculer les effet marginaux sur la variable d'intérêt d'une variation des variables explicatives du modèle. Dans un modèle probit ordonné, l'effet marginal correspond à l'effet sur la probabilité que chacune des catégories prenne la valeur 1 suite à une variation dans les variables explicatives. Nous avons déterminé la valeur des effets marginaux aux moyennes des variables explicatives. Cette méthode consiste à calculer l'effet marginal pour chaque participant de l'échantillon, puis de prendre la moyenne arithmétique de ces effets marginaux individuels. Ainsi nous avons rapporté les effets marginaux pour les variables significatives du tableau 8.

3.4.1 Perception de la corrélation inverse

Le premier niveau de la variable dépendante comprend les participants qui perçoivent la corrélation inverse. Le tableau 9 présente les résultats obtenus. Selon ces résultats, la perception de la corrélation inverse concerne 32% des prédictions agrégées et 33% des prédictions individuelles. Pour les prédictions agrégées, le genre est positif et significatif. Plus précisément, la probabilité d'attribuer une corrélation inverse est 17% moins élevée pour les femmes que pour les hommes. Même constat pour les catégories d'âge [25; 34] et [32; 59] : comparativement à la catégorie d'âge [21; 24]. Pour ces deux catégories d'âge, la probabilité de percevoir la corrélation inverse diminue de 17% comparativement à la catégorie de référence. Concernant le niveau d'étude, avoir le doctorat ou tout *autre éducation* comme niveau d'étude diminue de 3% la probabilité de percevoir la corrélation inverse comparativement aux participants ayant

la maîtrise comme niveau d'étude. Ces deux résultats sont significatifs. Enfin, on remarque que plus l'intensité de la corrélation augmente dans les traitements plus les participants ont tendance à attribuer une corrélation inverse. En effet, les participants appartenant aux traitements 2 et 6 ayant les niveaux de corrélation théorique les plus élevés (-0.47 et 0.47), ont respectivement 42% et 40% plus de chances d'attribuer une corrélation inverse par rapport au traitement de référence. Quant aux participants appartenant aux traitements 1 et 5 qui ont les niveaux de corrélation théorique les moins élevés (0.08 et -0.08), ils ont respectivement 35% et 25% plus de chances d'attribuer une corrélation inverse par rapport au traitement de référence.

| Variables | Intervalle 1 | |
|------------------------------------|----------------------|--|
| | Prédictions agrégées | Prédictions individuelles |
| Femme | -0.17*** (0.04) | -0.02 (0.06) |
| Catégorie d'âge [25 - 34] | -0.17** (0.06) | -0.24* (0.10) |
| Catégorie d'âge [35 - 59] | -0.17*** (0.06) | -0.15 (0.10) |
| Niveau d'études : Baccalauréat | 0.003 (0.04) | -0.05 (0.09) |
| Niveau d'études : Doctorat | 0.17** (0.07) | -0.08 (0.09) |
| Niveau d'études : Autres | -0.16** (0.07) | -0.18 (0.12) |
| Faculté : Sciences Sociales | 0.01 (0.04) | -0.09 (0.07) |
| Traitement 1 $\Phi = 0.08$ (Agr.) | 0.35*** (0.06) | $\Phi = 0.26$ (Ind.) 0.35*** (0.10) |
| Traitement 2 $\Phi = -0.47$ (Agr.) | 0.42*** (0.06) | $\Phi = -0.76$ (Ind.) 0.34*** (0.10) |
| Traitement 3 $\Phi = 0.25$ (Agr.) | 0.43*** (0.07) | $\Phi = 0.37$ (Ind.) 0.55*** (0.10) |
| Traitement 4 $\Phi = 0$ (Agr.) | 0.03* (0.02) | $\Phi = 0$ (Ind.) 0.01 (0.01) |
| Traitement 5 $\Phi = -0.08$ (Agr.) | 0.26*** (0.05) | $\Phi = -0.26$ (Ind.) 0.35*** (0.10) |
| Traitement 6 $\Phi = 0.47$ (Agr.) | 0.40*** (0.06) | $\Phi = 0.76$ (Ind.) 0.43*** (0.12) |
| Traitement 7 $\Phi = -0.25$ (Agr.) | 0.31*** (0.06) | $\Phi = -0.37$ (Ind.) 0.36*** (0.09) |
| Observations | 126 (32%) | 52 (33%) |

écarts types entre parenthèses
*** p<0.01, ** p<0.05, * p<0.1

Tableau 9: Effets marginaux : corrélation inverse

Pour les prédictions individuelles, rappelons que le genre, le niveau d'étude et le fait d'appartenir ou non à la faculté des sciences sociales n'étaient pas significatifs dans le tableau 8. Les résultats obtenus montrent que les participants ayant entre 25 et 34 ont 24% moins de chance d'attribuer une corrélation inverse comparativement à la catégorie de référence. La catégorie

[35 – 59] n’est pas significativement différente de la catégorie de référence. Enfin, appartenir à un traitement dont la corrélation théorique est différente de zéro augmente les chances d’attribuer une corrélation inverse comparativement au traitement de référence. Notons qu’on n’observe aucune différence significative entre le traitement de référence et le traitement 4 ; ceci étant normal car les deux ayant des corrélations théoriques égale à zéro.

Les probabilités prédites pour les participants qui perçoivent la corrélation inverse sont de 32% pour les prédictions agrégées et pour les prédictions individuelles. Ces résultats sont similaires à ceux que nous avons observés, ce qui nous indique que l’approche utilisée est consistante.

3.4.2 Sous-estimation de la corrélation

Le deuxième niveau de la variable dépendante comprend les participants qui sous-estiment l’intensité de la corrélation théorique entre la décision de participation et de faire un choix ; ou tout simplement qui négligent cette corrélation. Le tableau 10 présente les résultats obtenus. Selon ces résultats, la sous-estimation concerne 51% des prédictions agrégées et 47% des prédictions individuelles. Pour les prédictions agrégées, le genre devient positif et significatif. En d’autres termes, le fait d’être une femme augmente de 8% la probabilité de sous-estimer l’intensité de la corrélation entre la décision de participer au vote et de faire un choix, comparativement à un homme. Ensuite, les participants des catégories [25; 34) et [35; 59] 10% plus de chances de sous-estimer l’intensité de la corrélation théorique comparativement aux participants les plus jeunes. Avoir le doctorat ou tout *autre éducation* comme niveau d’étude diminue de 9% la probabilité de sous-estimer l’intensité de la corrélation comparativement aux participants ayant le baccalauréat et la maîtrise comme niveau d’étude. Enfin, on remarque que plus l’intensité de la corrélation diminue dans les traitements plus les participants ont tendance à sous-estimer l’intensité de la corrélation. En effet, les participants appartenant aux traitements 2 et 6 ayant les niveaux de corrélation théorique les plus élevés (-0.47 et 0.47), ont respectivement 19% et 21% plus de chances de sous-estimer l’intensité de la corrélation comparativement au traitement de référence. Quant aux participants appartenant aux traitements 1 et 5 qui ont les niveaux de corrélation théorique les moins élevés (0.08 et -0.08), ils ont respectivement 25% et 30% plus de chances de sous-estimer l’intensité de la corrélation comparativement au traitement de référence.

Pour les prédictions individuelles, les résultats obtenus montrent les participants ayant entre 25 et 34 ont 11% plus de chances d’attribuer une corrélation inverse comparativement à la catégorie de référence. La catégorie [35–59] n’est pas significativement différente de la catégorie de référence. Enfin, appartenir à un traitement dont la corrélation théorique est différente de zéro augmente les chances de sous-estimer l’intensité de la corrélation comparativement au traitement de référence.

Les probabilités prédites pour les participants qui sous-estiment l’intensité de la corrélation

sont respectivement de 53% pour les prédictions agrégées et de 48% pour les prédictions individuelles. Ces résultats sont également proches de ceux que nous avons observés.

| Variables | Intervalle 2 | |
|------------------------------------|----------------------|--|
| | Prédictions agrégées | Prédictions individuelles |
| Femme | 0.08*** (0.02) | -0.01 (0.03) |
| Catégorie d'âge [25 - 34] | 0.10*** (0.04) | 0.11* (0.06) |
| Catégorie d'âge [35 - 59] | 0.10*** (0.04) | 0.10 (0.07) |
| Niveau d'études : Baccalauréat | -0.008 (0.01) | 0.01 (0.02) |
| Niveau d'études : Doctorat | -0.09* (0.05) | -0.04 (0.05) |
| Niveau d'études : Autres | -0.09** (0.04) | -0.10 (0.08) |
| Faculté : Sciences Sociales | -0.004 (0.02) | 0.03 (0.03) |
| Traitement 1 $\Phi = 0.08$ (Agr.) | 0.25*** (0.08) | $\Phi = 0.26$ (Ind.) 0.38*** (0.12) |
| Traitement 2 $\Phi = -0.47$ (Agr.) | 0.19** (0.08) | $\Phi = -0.76$ (Ind.) 0.39*** (0.11) |
| Traitement 3 $\Phi = 0.25$ (Agr.) | 0.19** (0.08) | $\Phi = 0.37$ (Ind.) 0.22* (0.12) |
| Traitement 4 $\Phi = 0$ (Agr.) | 0.20** (0.08) | $\Phi = 0$ (Ind.) 0.15 (0.12) |
| Traitement 5 $\Phi = -0.08$ (Agr.) | 0.30*** (0.07) | $\Phi = -0.26$ (Ind.) 0.38*** (0.12) |
| Traitement 6 $\Phi = 0.47$ (Agr.) | 0.21** (0.08) | $\Phi = 0.76$ (Ind.) 0.32** (0.12) |
| Traitement 7 $\Phi = -0.25$ (Agr.) | 0.27*** (0.08) | $\Phi = -0.37$ (Ind.) 0.38*** (0.11) |
| Observations | 203 (51%) | 74 (47%) |

écarts types entre parenthèses
*** p<0.01, ** p<0.05, * p<0.1

Tableau 10: Effets marginaux : sous-estimation ou négligence de la corrélation

3.4.3 Perception exacte de la corrélation

Le troisième niveau de la variable dépendante comprend les participants qui perçoivent la corrélation exacte entre la décision de participation et le choix de vote. Le tableau 11 présente les résultats obtenus. Selon ces résultats, la perception de la corrélation exacte concerne 10% des prédictions agrégées et 12% des prédictions individuelles. Pour les prédictions agrégées, le genre demeure positif et significatif. En d'autres termes, le fait d'être une femme augmente de 6% la probabilité de percevoir la corrélation exacte entre la décision de participation et le choix de vote, comparativement à un homme. De plus, on remarque que plus les participants sont âgés, plus elles ont tendance à percevoir la corrélation exacte. En effet, les participants âgés ayant entre 25 et 34 ans ont 5% plus de chances de percevoir la corrélation exacte que ceux ayant moins de 25, tandis que ceux âgés entre 35 et 59 ans ont 6% plus de chances de

percevoir la corrélation exacte que ceux ayant moins de 25. On constate un résultat contraire pour les participants ayant le doctorat ou tout *autre éducation* comme niveau d'étude . En effet, les participants appartenant à ces deux catégories ont 5% moins de chance de percevoir la corrélation exacte comparativement aux participants ayant le baccalauréat ou la maîtrise comme niveau d'étude. Enfin, on remarque plus l'intensité de la corrélation augmente dans les traitements, moins les participants ont tendance à percevoir correctement la corrélation. En effet, les participants appartenant aux traitements 1 et 5 ayant les niveaux de corrélation théorique les moins élevés (-0.08 et 0.08), ont respectivement 22% et 19% moins de chance de percevoir la corrélation exacte comparativement au traitement de référence. Quant aux participants appartenant aux traitements 2 et 6 qui ont les niveaux de corrélation théorique les plus élevés (0.47 et -0.47), ils ont 23% moins de chance de sous-estimer l'intensité de la corrélation comparativement au traitement de référence.

| Intervalle 3 | | | |
|-------------------------------------|----------------------|-----------------------|---------------------------|
| Variables | Prédictions agrégées | | Prédictions individuelles |
| Femme | 0.06*** (0.02) | | -0.01 (0.03) |
| Catégorie d'âge [25 - 34] | 0.05*** (0.02) | | 0.10** (0.04) |
| Catégorie d'âge [35 - 59] | 0.06*** (0.02) | | 0.05 (0.03) |
| Niveau d'études : Baccalauréat | -0.001 (0.02) | | 0.03 (0.05) |
| Niveau d'études : Doctorat | -0.05** (0.02) | | -0.03 (0.04) |
| Niveau d'études : Autres | -0.05** (0.02) | | -0.06 (0.05) |
| Faculté : Sciences Sociales | -0.003 (0.02) | | 0.04 (0.03) |
| Traitement 1 $\Phi = 0.08$ (Agr.) | -0.22*** (0.03) | $\Phi = 0.26$ (Ind.) | - 0.28*** (0.07) |
| Traitement 2 $\Phi = - 0.47$ (Agr.) | -0.23*** (0.04) | $\Phi = -0.76$ (Ind.) | -0.28*** (0.07) |
| Traitement 3 $\Phi = 0.25$ (Agr.) | -0.23*** (0.04) | $\Phi = 0.37$ (Ind.) | -0.32*** (0.07) |
| Traitement 4 $\Phi = 0$ (Agr.) | -0.04 (0.02) | $\Phi = 0$ (Ind.) | 0.016 (0.03) |
| Traitement 5 $\Phi = -0.08$ (Agr.) | -0.19*** (0.03) | $\Phi = -0.26$ (Ind.) | -0.28*** (0.07) |
| Traitement 6 $\Phi = 0.47$ (Agr.) | -0.23*** (0.04) | $\Phi = 0.76$ (Ind.) | -0.30*** (0.07) |
| Traitement 7 $\Phi = -0.25$ (Agr.) | -0.21*** (0.04) | $\Phi = -0.37$ (Ind.) | -0.29*** (0.07) |
| Observations | 39 (10%) | | 20 (12%) |

écarts types entre parenthèses
*** p<0.01, ** p<0.05, * p<0.1

Tableau 11: Effets marginaux : corrélation exacte

Pour les prédictions individuelles, les résultats obtenus montrent les participants ayant entre 25

et 34 ont 10% plus de chances d'attribuer une corrélation exacte comparativement à la catégorie de référence. Les participants appartenant à la catégorie [35 – 59] ne sont pas significativement différents de ceux de la catégorie de référence. Enfin, appartenir à un traitement dont la corrélation théorique est différente de zéro diminue les chances d'attribuer une corrélation exacte comparativement au traitement de référence.

Les probabilités prédites pour les participants qui attribuent la corrélation exacte l'intensité de la corrélation sont respectivement de 8% pour les prédictions agrégées et de 10% pour les prédictions individuelles. Ces résultats sont également proches de ceux que nous avons observés.

3.4.4 Surestimation de la corrélation

Enfin, le quatrième niveau de la variable dépendante comprend les participants qui surestiment l'intensité de la corrélation entre la décision de participation et le choix de vote. Le tableau 12 présente les résultats obtenus. Selon ces résultats, la surestimation de l'intensité de la corrélation concerne 7% des prédictions agrégées et 8% des prédictions individuelles. Pour les prédictions agrégées, le genre demeure positif et significatif. En d'autres termes, le fait d'être une femme augmente de 3% la probabilité de surestimer l'intensité la corrélation exacte entre la décision de participation et le choix de vote, comparativement à un homme. De plus, on remarque que plus les participants sont âgés, plus ils ont tendance à surestimer l'intensité de la corrélation. En effet, les participants âgés ayant entre 25 et 34 ans ont 2% plus de chances de surestimer l'intensité de la corrélation que ceux ayant moins de 25, tandis que ceux âgés de 35 à 59 ans ont 3% plus de chances de surestimer l'intensité de la corrélation que ceux ayant moins de 25. On constate un résultat contraire pour les participants le doctorat ou tout *autre éducation* comme niveau d'étude . En effet, les participants appartenant à ces deux catégories ont 2% moins de chances de surestimer l'intensité de la corrélation comparativement aux participants ayant le baccalauréat ou la maîtrise comme niveau d'étude. Enfin, on remarque plus l'intensité de la corrélation augmente dans les traitements, moins les participants ont tendance à surestimer l'intensité de la corrélation. En effet, les participants appartenant aux traitements 1 et 5 ayant les niveaux de corrélation théorique les moins élevés (-0.08 et 0.08), ont respectivement 37% et 36% moins de chance de surestimer l'intensité de la corrélation comparativement au traitement de référence. Quant aux participants appartenant aux traitements 2 et 6 qui ont les niveaux de corrélation théorique les plus élevés (0.47 et -0.47), ils ont 37% moins de chances de sous-estimer l'intensité de la corrélation comparativement au traitement de référence.

Pour les prédictions individuelles, les résultats obtenus montrent les participants ayant entre 25 et 34 ont 3% plus de chances de surestimer la corrélation comparativement à la catégorie de référence. La catégorie [35 – 59] n'est pas significativement différente de la catégorie de référence. Enfin, appartenir à un traitement dont la corrélation théorique est différente de

zéro diminue les chances de surestimer la corrélation exacte comparativement au traitement de référence.

Les probabilités prédites pour les participants qui attribuent la corrélation exacte l'intensité de la corrélation sont respectivement de 7% pour les prédictions agrégées et de 8% pour les prédictions individuelles. Ces résultats sont également proches de ceux que nous avons observés.

| Variables | Intervalle 4 | | |
|------------------------------------|----------------------|---------------------------------|--------------------|
| | Prédictions agrégées | Prédictions individuelles | |
| Femme | 0.03*** (0.01) | -0.003 (0.01) | |
| Catégorie d'âge [25 - 34) | 0.02** (0.01) | 0.03* (0.01) | |
| Catégorie d'âge [35 - 59] | 0.03** (0.01) | 0.01 (0.01) | |
| Niveau d'études : Baccalauréat | -0.001 (0.01) | 0.01 (0.02) | |
| Niveau d'études : Doctorat | -0.02** (0.01) | -0.01 (0.01) | |
| Niveau d'études : Autres | -0.02** (0.01) | -0.01 (0.01) | |
| Faculté : Sciences Sociales | -0.002 (0.01) | 0.01 (0.01) | |
| Traitement 1 $\Phi = 0.08$ (Agr.) | -0.37*** (0.07) | $\Phi = 0.26$ (Ind.) (0.12) | -0.45*** (0.12) |
| Traitement 2 $\Phi = -0.47$ (Agr.) | -0.38*** (0.07) | $\Phi = -0.76$ (Ind.) (0.12) | -0.45*** (0.12) |
| Traitement 3 $\Phi = 0.25$ (Agr.) | -0.38*** (0.07) | $\Phi = 0.37$ (Ind.) (0.12) | 0.45*** (0.12) |
| Traitement 4 $\Phi = 0$ (Agr.) | -0.20** (0.08) | $\Phi = 0$ (Ind.) (0.14) | -0.18 (0.14) |
| Traitement 5 $\Phi = -0.08$ (Agr.) | -0.36*** (0.07) | $\Phi = -0.26$ (Ind.) (0.12) | -0.45*** (0.12) |
| Traitement 6 $\Phi = 0.47$ (Agr.) | -0.38*** (0.07) | $\Phi = 0.76$ (Ind.) (0.12) | -0.45*** (0.12) |
| Traitement 7 $\Phi = -0.25$ (Agr.) | -0.37*** (0.07) | $\Phi = -0.37$ (Ind.) (0.12) | -0.45*** (0.12) |
| Observations | 27 (7%) | | 13 (8%) |

écarts types entre parenthèses
*** p<0.01, ** p<0.05, * p<0.1

Tableau 12: Effets marginaux : surestimation de la corrélation

En somme, le tableau 10 met en lumière les facteurs responsables de la négligence de la corrélation entre la décision de participation au vote et celle de faire un choix pour les prédictions agrégées et les prédictions individuelles. Concernant le niveau d'étude, les résultats des prédictions agrégées concordent avec l'intuition selon laquelle une personne avec un haut niveau d'étude négligerait moins l'intensité de la corrélation. Toutefois, il est surprenant de voir que détenir un autre niveau d'étude qu'universitaire diminue également les chances de sous-estimer

l'intensité de la corrélation. Le nombre d'observations pour les prédictions individuelles est très faible. Cela semble suggérer que pour les prédictions individuelles, nous n'avons pas suffisamment de pouvoir statistique pour détecter l'effet du niveau d'étude. Enfin, pour les prédictions agrégées, il est intéressant de voir que la probabilité de sous-estimer l'intensité de la corrélation est plus élevée pour les traitements à plus faible corrélation théorique et moins élevée pour les traitements à plus forte corrélation théorique . En effet, cela rejoint l'intuition selon laquelle les participants perçoivent la forte magnitude la corrélation théorique dans les traitements 2 et 6 et la négligent moins que dans les traitements 1 et 8 où la magnitude la corrélation théorique est faible et donc difficile à percevoir.

Chapitre 4

Analyse des coefficients Φ indéterminables

Dans le chapitre 3, nous avons analysé les Φ déterminables et les facteurs qui expliquent leurs variations. Dans ce chapitre, nous analysons les Φ indéterminables et les raisons pour lesquelles il était impossible de les calculer. En effet, pour 69 prédictions agrégées et pour 233 prédictions individuelles, au moins l'une des quatre probabilités marginales composant le dénominateur était égale à zéro .

4.1 Analyse des coefficients Φ indéterminables pour les prédictions agrégées

Lors du calcul des coefficients Φ découlant des prédictions agrégées des participants, nous avons constaté que pour 69 prédictions, il était impossible de calculer ledit coefficient du fait qu'au moins un facteur au dénominateur était égal à zéro (Voir tableau 13).

| Traitement | # observations avec Φ indéterminable | # observations sans Φ indéterminable | Total |
|--------------------------------|---|---|--------------|
| Traitement 1 : $\Phi = 0.08$ | 2 | 56 | 58 |
| Traitement 2 : $\Phi = - 0.47$ | 1 | 57 | 58 |
| Traitement 3 : $\Phi = 0.25$ | 11 | 47 | 58 |
| Traitement 4 : $\Phi = 0$ | 20 | 38 | 58 |
| Traitement 5 : $\Phi = - 0.08$ | 4 | 54 | 58 |
| Traitement 6 : $\Phi = 0.47$ | 2 | 56 | 58 |
| Traitement 7 : $\Phi = - 0.25$ | 5 | 53 | 58 |
| Traitement 8 : $\Phi = 0$ | 24 | 34 | 58 |
| Total | 69 | 395 | 464 |

Tableau 13: Distribution des coefficients Φ individuels indéterminables selon les traitements

Rappelons qu'au moment d'effectuer les prédictions agrégées, les participants ignoraient le type auquel ils appartenaient. Leurs décisions étaient guidées par les paramètres du traitement auquel ils étaient confrontés : les gains associés aux choix A et B, ainsi que les montants et

probabilités associées aux coûts. Le tableau 14 donne les prédictions associées à chaque choix agrégé pour chaque traitement, lorsque la prédiction est rationnelle.

| Traitement | P(V = 1) | P(V = 0) | P(A) | P(B) |
|--------------|----------|----------|-------|-------|
| Traitement 1 | 91.5% | 8.5% | 62.5% | 37.5% |
| Traitement 2 | 81% | 19% | 50% | 50% |
| Traitement 3 | 93% | 7% | 50% | 50% |
| Traitement 4 | 100% | 0% | 50% | 50% |
| Traitement 5 | 91.5% | 8.5% | 37.5% | 62.5% |
| Traitement 6 | 81% | 19% | 50% | 50% |
| Traitement 7 | 93% | 7% | 50% | 50% |
| Traitement 8 | 100% | 0% | 50% | 50% |

Tableau 14: prédictions théoriques pour choix rationnels agrégés

où :

- $P(V = 1)$ probabilité de voter dans toute la population ;
- $P(V = 0)$ probabilité de ne pas voter dans toute la population ;
- $P(A)$ probabilité de voter A dans toute la population ;
- $P(B)$ probabilité de voter dans toute la population ;
- Sachant que $P(V = 1) + P(V = 0) = 1$ et $P(A) + P(B) = 1$

Les probabilités observées dans le tableau 14 découlent des paramètres de l'expérience décrits dans le tableau 4. Afin de mieux concevoir la provenance de ces probabilités, prenons l'exemple du traitement 1. Dans ce traitement, les participants appartenant au type bleu et vert foncé ont intérêt à choisir l'option A s'ils veulent maximiser leurs gains. En effet, en présence de coût élevé ou de coût faible, le gain demeure plus avantageux que celui des deux autres alternatives (option B et l'option de ne pas voter). Par contre, pour les participants appartenant au type vert pâle, c'est l'option de ne pas voter qui maximise les gains en présence de coût élevé. En présence de coût faible, les gains sont identiques. Dans ce cas, on s'attend à ce que le participant reste indifférent. Enfin, les participants appartenant au type jaune ont intérêt à choisir l'option B s'ils veulent maximiser leurs gains. Le tableau 15 montre les paiements avec et sans frais associés au traitement 1. Le tableau 16 résume le choix optimal du participant rationnel, au regard des paiements observés dans le tableau 15 : la valeur 1 indique l'option qui maximise les gains, la valeur 0 indique le pire choix et la valeur 0.33 ($1 \div 3$) suppose que le participant rationnel serait indifférent entre les 3 choix.

En considérant les informations du tableau 16, il est possible de calculer les probabilités associées à chacune des options pour chaque type. Pour les calculer, il faut multiplier les probabilités attachées aux différents états de monde (état de monde avec frais et 'etat de

| Choix avec frais | Type bleu | Type vert foncé | Type vert pâle | Type jaune |
|-----------------------|-----------|-----------------|----------------|------------|
| Option A | 50 | 44 | 10 | -23 |
| Option B | -24 | -24 | 10 | 43 |
| Option externe | 30 | 36 | 70 | 37 |
| Choix sans frais | Type bleu | Type vert foncé | Type vert pâle | Type jaune |
| Option A | 110 | 104 | 70 | 70 |
| Option B | 30 | 36 | 70 | 103 |
| Option externe | 30 | 36 | 70 | 37 |

Tableau 15: Paiements avec et sans frais associés au traitement 1

| Choix avec frais | Type bleu | Type vert foncé | Type vert pâle | Type jaune |
|-----------------------|-----------|-----------------|----------------|------------|
| Option A | 1 | 1 | 0 | 0 |
| Option B | 0 | 0 | 0 | 1 |
| Option externe | 0 | 0 | 1 | 0 |
| Choix sans frais | Type bleu | Type vert foncé | Type vert pâle | Type jaune |
| Option A | 1 | 1 | 0.33 | 0 |
| Option B | 0 | 0 | 0.33 | 1 |
| Option externe | 0 | 0 | 0.33 | 0 |

Tableau 16: Choix optimal du participant rationnel pour le traitement 1

monde sans frais) par le gain associé à chaque choix. Le tableau 17 donne ces probabilités pour les choix auxquels les participants font face. Par exemple :

– La probabilité de choisir l’option A pour les participants appartenant au type bleu en cas de vote dans le traitement 1 :

$$P(A|P(V) = 1)_{bleu} = 1 \times 0.9 + 0 \times 0.1 = 1$$

– La probabilité de choisir l’option B pour les participants appartenant au type vert pâle (vp) en cas de vote dans le traitement 1 :

$$P(B|P(V) = 1)_{vp} = 0 \times 0.9 + 0.33 \times 0.1 = 0.33$$

– La probabilité de choisir l’option externe (abstention au vote) pour les participants appartenant au type vert pâle (vp) dans le traitement 1 :

$$P(P(V) = 0)_{vp} = 0.33 \times 0.9 + 1 \times 0.1 = 0.34$$

Pour le type vert pâle, puisque les participants rationnels sont indifférents entre l’option A et l’option B, alors leurs probabilités en cas d’abstention de vote sont également identiques :

$$P(B|P(V) = 0))_{vp} = EV(A|P(V) = 0))_{vp} = EV(P(V) = 0)_{vp} \div 2 = 0.17$$

| Type | Type bleu | Type vert foncé | Type vert pâle | Type jaune | Moyenne |
|-----------------|-----------|-----------------|----------------|------------|---------|
| $P(A P(V)=1)$ | 1 | 1 | 0.33 | 0 | 58.25% |
| $P(B P(V)=1)$ | 0 | 0 | 0.33 | 1 | 33.25% |
| $P(A P(V)=0)$ | 0 | 0 | 0.17 | 0 | 4.25% |
| $P(B P(V)=0)$ | 0 | 0 | 0.17 | 0 | 4.25% |
| $P(V) = 0$ | 0 | 0 | 0.34 | 0 | 8.5% |

Tableau 17: Probabilités et prédictions théoriques des choix rationnels agrégés pour le traitement 1

Finale­ment, nous obtenons les prédictions théoriques pour les choix rationnels agrégés $P(A|P(V) = 1)$, $P(B|P(V) = 1)$, $P(A|P(V) = 0)$, $P(B|P(V) = 0)$ et $P(V = 0)$ en calculant la moyenne des valeurs espérées entre les différents types. Par exemple, les prédictions théoriques pour le choix rationnel agrégé de l’option $P(A|P(V) = 1)$ est donné par :

$$P(A|P(V) = 1) = (1 + 1 + 0.33) \div 4 = 58.25\%$$

À partir de ces résultats, nous déduisons les prédictions théoriques pour le choix rationnel agrégé de $P(A)$, $P(B)$, et $P(V = 1)$. En effet $P(A) = P(A|P(V) = 1) + P(A|P(V) = 0)$, $P(B) = P(B|P(V) = 1) + P(B|P(V) = 0)$ et $P(V = 1) = 1 - P(V = 0)$. Ainsi, nous obtenons les valeurs inscrites dans le tableau 17. Cette démarche est valable pour tous les traitements. On peut voir dans le tableau 14 que les prédictions théoriques pour choix rationnels agrégés sont similaires pour certains traitements. Par conséquent, nous les analyserons ensemble dans la suite notre analyse.

Dans le traitement 1 , il était impossible de calculer le coefficient Φ pour 2 observations. En effet, ces participants étaient certains de participer au vote ($P(V=1) = 100\%$). De plus, 1 observation a indiqué que le participant avait une préférence pour le choix A ($P(A) = 75\%$), tandis que l’autre observation a indiqué que le participant avait une préférence pour le choix B ($P(B) = 10\%$). Selon les paramètres de l’expérience, le choix rationnel aurait été de participer au vote ($P(V=1) = 91.5\%$), et de voter pour A ($P(A) = 62.5\%$). Par conséquent, la prédiction était exacte pour la première observation et mauvaise pour la deuxième.

Dans le traitement 5 , il était impossible de calculer le coefficient Φ pour 4 observations. En effet, 3 d’entre elles ont indiqué que les participants étaient certains de participer au vote ($P(V=1) = 100\%$), tandis que l’une d’entre elles a indiqué le contraire ($P(V=0) = 100\%$). De plus, parmi les 5 ayant indiqué qu’ils étaient certains de participer, 2 observations indiquaient que les participants n’avaient de préférence ni pour B ($P(A) = P(B) = 50\%$) et l’autre observation indiquait que le participant préférait le choix B ($P(B) = 58\%$). Enfin, pour le

participant indiquant être certain de ne pas participer au vote, son choix s'est porté sur B ($P(B) = 65\%$). Selon les paramètres de l'expérience, le choix rationnel aurait été de participer au vote ($P(V=1) = 91.5\%$), et de voter pour B ($P(B) = 62.5\%$). Par conséquent, la prédiction n'était exacte que pour le participant ayant indiqué qu'il participerait au vote et choisirait A.

Dans les traitements 2 et 6, il était impossible de calculer le coefficient Φ pour 3 observations. En effet, le participant était certain de participer au vote ($P(V=1) = 100\%$). Un autre participant était certain de ne pas participer ($P(V=0) = 100\%$). Enfin un troisième participant était certain de voter B ($P(B) = 100\%$). Selon les paramètres de l'expérience le choix rationnel prédisait de participer au vote ($P(V=1) = 81\%$) mais était indifférent entre A et B ($P(A) = P(B) = 50\%$). Celui indiquant $P(V=1) = 100\%$ votait pour B ($P(B) = 60\%$). De ce fait, seul ce dernier participant a prédit de façon rationnelle dans le traitement 2 qu'il choisirait B avec une probabilité de 60%.

Dans les traitements 3 et 7, il était impossible de calculer le coefficient Φ pour 16 observations. En effet, 13 observations ont indiqué que les participants étaient certains de participer au vote ($P(V=1) = 100\%$), 2 observations ont indiqué que les participants étaient certains de voter B ($P(B) = 100\%$) et 1 observation a indiqué que le participant était certain de voter A ($P(A) = 100\%$). Selon les paramètres de l'expérience, le choix rationnel prédisait de participer au vote ($P(V=1) = 93\%$) mais était indifférent entre A et B ($P(A) = P(B) = 50\%$). Parmi les 13 indiquant $P(V=1) = 100\%$, 1 observation a indiqué que le participant choisissait A avec $P(A) = 70\%$. Les 12 autres ont prédit de voter $P(A) = P(B) = 50\%$ comme l'indiquait le choix rationnel. Par conséquent, 3 observations ont émis des prédictions contraires au choix rationnel : la première indiquant $P(B) = 100\%$, la deuxième indiquant $P(A) = 100\%$ et la troisième indiquant $P(A) = 70\%$; considérant que 70% soit trop éloigné de 50%.

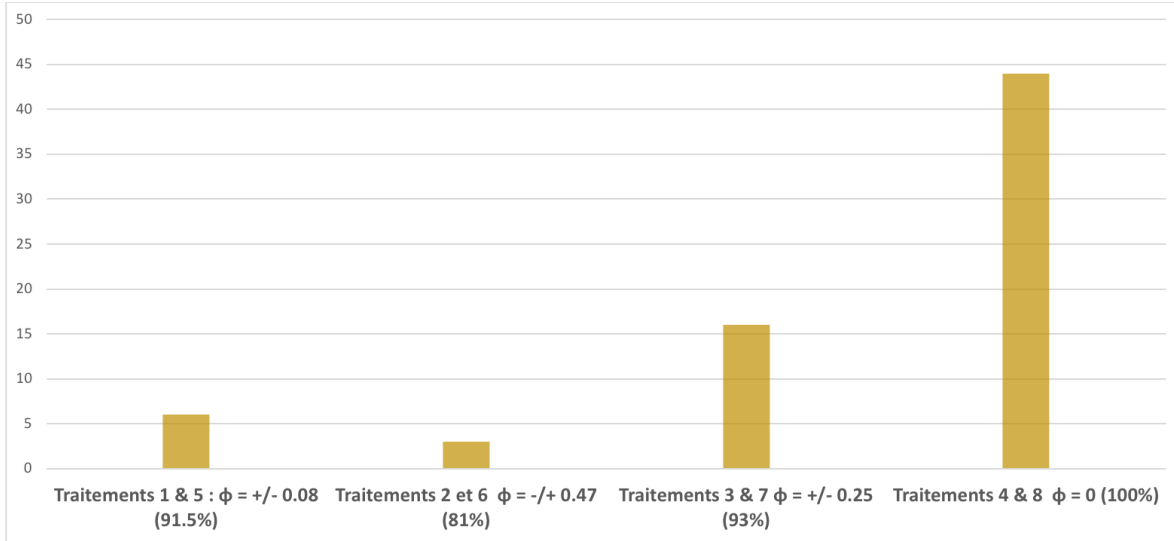
Dans les traitements 4 et 8, il était impossible de calculer le coefficient Φ pour 44 observations. En effet, 40 observations ont indiqué que les participants étaient certains de participer au vote ($P(V=1) = 100\%$). Les quatre autres observations ont indiqué que les participants étaient certains de voter A ($P(A) = 100\%$). Selon les paramètres de l'expérience, le choix rationnel prédisait de participer au vote ($P(V=1) = 100\%$) mais était indifférent entre A et B ($P(A) = P(B) = 50\%$). Parmi les 40 observations pour lesquelles $P(V=1) = 100\%$, 3 observations ont indiqué que les participants prédisaient A avec respectivement $P(A) = 51\%$, 55% et 60%. Les 36 autres ont prédit $P(A) = P(B) = 50\%$ comme l'indiquait le choix rationnel. Par conséquent, 4 observations ont émis des prédictions contraires au choix rationnel en indiquant voter pour A ($P(A) = 100\%$); considérant que $P(A) = 51\%$, 55% et 60% soient proches de 50%.

En résumé, 13 observations ont dévié du choix rationnel : 1 dans les traitements 1, 3 dans le traitement 5, 2 dans les traitements 2 et 6, 3 dans les traitements 3 et 7 et 4 dans les traitements 4 et 8.

Ces observations mettent en lumière la relation existante entre la distribution des coefficients

Φ agrégés indéterminables et les probabilités associées aux prédictions théoriques pour les choix rationnels agrégés. Plus précisément, nous remarquons que plus la probabilité théorique de voter est certaine, c'est-à-dire lorsque $P(V = 1) = 100\%$, plus le nombre de Φ indéterminable est élevé. La figure 5 montre la distribution des Φ indéterminables selon les probabilités théoriques de voter associées à chaque traitement.

FIGURE 5 – Distribution des Φ indéterminables selon les probabilités théoriques de voter associées à chaque traitement.



La figure 5 montre que dans les traitements 4 et 8, la probabilité théorique de voter est certaine et le nombre d'observations avec Φ non calculables sont plus élevés dans ce traitement. De plus, le nombre d'observations avec Φ indéterminables diminue avec l'intensité de la probabilité théorique de voter qui est associée à chaque traitement. En d'autres termes, la présence d'observations à Φ indéterminables s'expliquent par le fait que plus voter est un choix certain, plus les participants interprètent bien les paramètres de l'expérience et choisissent de participer au vote avec certitude. Cela a comme conséquence que pour ces observations, la probabilité marginale $P(V = 0) = 0$ et donc il devient impossible de calculer les valeurs des Φ agrégés pour ces observations.

4.2 Analyse des coefficients Φ indéterminables pour les prédictions individuelles

Après calcul des coefficients Φ découlant des prédictions individuelles des participants nous avons constaté que pour 233 prédictions, il était impossible de calculer ledit coefficient du fait qu'au moins un facteur au dénominateur était égal à zéro (voir tableau 18).

Rappelons qu'au moment de réaliser les prédictions individuelles, les participants savaient si leur type était bleu ou jaune. Cependant l'incertitude planait encore pour ceux appartenant

| Type | # observations sans Φ | # observations avec Φ | Total |
|--------------|----------------------------|----------------------------|-------|
| Type Bleu | 85 | 36 | 121 |
| Type Vert | 77 | 159 | 236 |
| Type Jaune | 71 | 36 | 107 |
| Total | 233 | 231 | 464 |

Tableau 18: Distribution des Φ non calculable selon le type

aux types verts. En effet, ces derniers ne savaient s'ils appartenaient au type vert foncé ou au type vert pâle. Leurs décisions étaient guidées par les paramètres du traitement auquel ils étaient confrontés (les gains associés aux choix A et B, ainsi que les montants et probabilités associés aux coûts). Le tableau 19 indique pour les types verts uniquement la répartition des coefficients indéterminables selon les traitements.

| Traitement | # observations avec Φ indéterminables | # observations sans Φ indéterminable | Total |
|-------------------------------|--|---|-------|
| Traitement 1 : $\Phi = 0.26$ | 13 | 19 | 32 |
| Traitement 2 : $\Phi = -0.76$ | 5 | 22 | 27 |
| Traitement 3 : $\Phi = 0.37$ | 4 | 24 | 28 |
| Traitement 4 : $\Phi = 0$ | 18 | 19 | 37 |
| Traitement 5 : $\Phi = -0.26$ | 15 | 19 | 34 |
| Traitement 6 : $\Phi = 0.76$ | 4 | 17 | 21 |
| Traitement 7 : $\Phi = -0.37$ | 6 | 24 | 30 |
| Traitement 8 : $\Phi = 0$ | 12 | 15 | 27 |
| Total | 77 | 159 | 236 |

Tableau 19: Distribution des coefficients phi individuels indéterminables selon les traitements

Le tableau 20 donne pour le type vert, les probabilités associées à chaque choix pour chaque traitement, lorsque l'agrégation est rationnelle.

| Traitement | $P_{vert}(V = 1)$ | $P_{vert}(V = 0)$ | $P_{vert}(A)$ | $P_{vert}(B)$ |
|--------------|-------------------|-------------------|---------------|---------------|
| Traitement 1 | 83% | 17% | 75% | 25% |
| Traitement 2 | 62% | 38% | 50% | 50% |
| Traitement 3 | 86% | 14% | 50% | 50% |
| Traitement 4 | 100% | 0% | 50% | 50% |
| Traitement 5 | 83% | 17% | 25% | 75% |
| Traitement 6 | 62% | 38% | 50% | 50% |
| Traitement 7 | 86% | 14% | 50% | 50% |
| Traitement 8 | 100% | 0% | 50% | 50% |

Tableau 20: Prédictions théoriques pour choix individuels du type vert

où :

- $P_{vert}(V = 1)$ probabilité de voter chez les verts ;
- $P_{vert}(V = 0)$ probabilité de ne pas voter chez les verts ;
- $P_{vert}(A)$ probabilité de voter A chez les verts ;

- $P_{vert}(B)$ probabilité de voter chez les verts ;
- Sachant que $P_{vert}(V = 1) + P_{vert}(V = 0) = 1$ et $P_{vert}(A) + P_{vert}(B) = 1$

Les probabilités observées dans le tableau 20 sont obtenues de la même manière que celles dans le tableau 14 concernant les prédictions théoriques pour les choix rationnels agrégés. Similairement aux prédictions théoriques pour choix rationnels agrégés, nous analyserons les traitements par pair ayant les mêmes caractéristiques.

Dans le traitement 1, il était impossible de calculer le coefficient Φ pour 13 observations. En effet, 12 appartenaient à des participants qui avaient une préférence pour A ($P_{vert}(A) = 100\%$), tandis que l'une d'entre elles a indiqué le contraire ($P_{vert}(A) = 0\%$). Parmi les observations indiquant que les participants avaient une préférence pour A, 4 indiquaient que les participants étaient certains de participer au vote ($P_{vert}(V=1) = 100\%$), 3 indiquaient que les participants étaient sûrs à plus de 50%, 2 indiquaient une indifférence quant au choix de participer ou non ($P(V=1) = P(V=0) = 50\%$) et 1 observation indiquait que le participant préférerait ne pas participer ($P(V=0) = 30\%$). Selon les paramètres de l'expérience, le choix rationnel aurait été de participer au vote ($P(V=1) = 83\%$), et de voter pour A ($P(A) = 75\%$). Par conséquent, 4 observations ont émis des prédictions contraires au choix rationnel : l'une indiquant ne pas vouloir participer, 2 indiquant être indifférents quant au choix de participer ou non et l'une indiquant ne pas vouloir participer.

Dans le traitement 5, il était impossible de calculer le coefficient Φ pour 15 observations. Tous ces participants avaient une préférence pour B ($P_{vert}(A) = 100\%$). Parmi ces participants, 1 indiquait être certain de participer au vote ($P_{vert}(V=100\%)$), 3 indiquaient que les participants étaient sûrs à plus de 50%, 4 indiquaient une indifférence quant au choix de participer ou non ($P(V=1) = P(V=0) = 50\%$), 2 préféraient ne pas participer ($P(V=0) = 40\%$) et 5 indiquaient être certains de ne pas participer au vote ($P_{vert}(V=1) = 100\%$). Selon les paramètres de l'expérience, le choix rationnel aurait été de participer au vote ($P(V=1) = 83\%$), et de voter pour B ($P(B) = 75\%$). Par conséquent, 11 observations ont émis des prédictions contraires au choix rationnel : 5 indiquant être certains de ne pas participer au vote, 2 préférant ne pas participer et 4 indiquant être indifférent quant au choix de participer ou non.

Dans les traitements 2 et 6, il était impossible de calculer le coefficient Φ pour 9 observations. En effet, 4 indiquaient que les participants étaient certains de participer au vote ($P_{vert}(V=100\%)$), 2 indiquaient être sûrs de ne pas voter ($P(V=1)$ égaux 20% et 25%) et 3 indiquaient une indifférence quant au choix de participer ou non ($P(V=1) = P(V=0) = 50\%$). Selon les paramètres de l'expérience, le choix rationnel prédisait de participer au vote ($P(V=1) = 62\%$) mais était indifférent entre A et B ($P(A) = P(B) = 50\%$). De ce fait, aucun participant n'a prédit de façon rationnelle dans le traitement 2 vu qu'aucun d'eux ne prédisait de participer au vote.

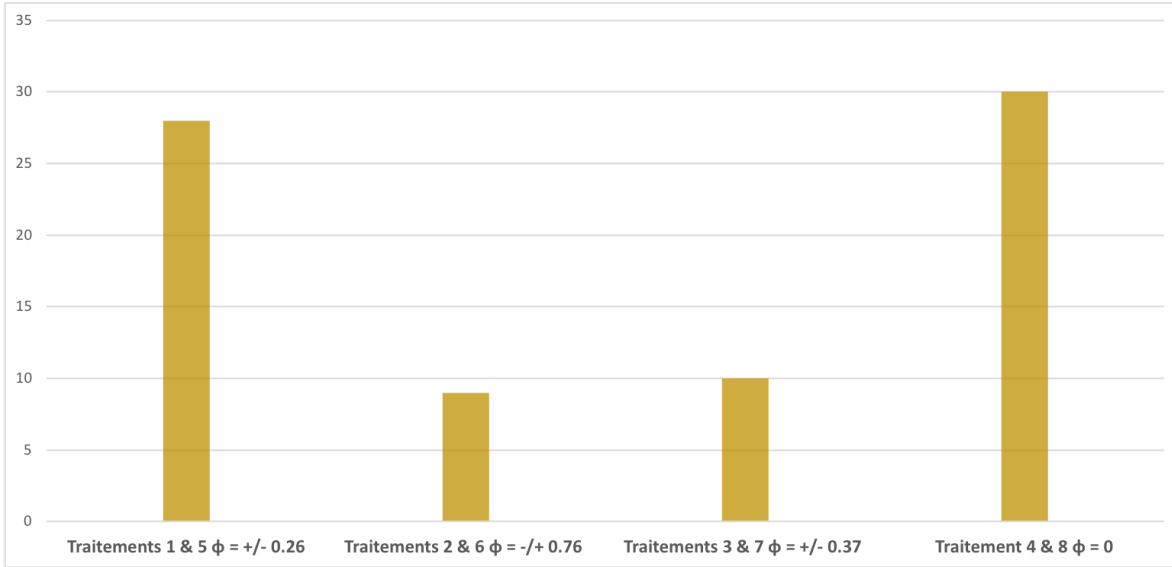
Dans les traitements 3 et 7, il était impossible de calculer le coefficient Φ pour 10 observations. En effet, 3 indiquaient que les participants étaient certains de participer au vote ($P_{vert}(V=100\%)$), 2 indiquaient une indifférence quant au choix de participer ou non ($P(V=1) = P(V=0) = 50\%$) et 5 observations indiquaient que les participants étaient certains ne pas participer ($P(V=0) = 100\%$). Parmi les participants certains de participer au vote, 1 indiquait être certain de choisir A ($P_{vert}(A) = 100\%$), 1 indiquait être certain de choisir B ($P_{vert}(B) = 100\%$), et indiquait être sûr de choisir B ($P_{vert}(B) = 95\%$). Selon les paramètres de l'expérience, le choix rationnel aurait été de participer au vote ($P(V=1) = 86\%$) mais d'être indifférent entre A et B ($P(A) = P(B) = 50\%$). Par conséquent, aucun de ces participants n'a émis des prédictions similaires au choix rationnel.

Dans les traitements 4 et 8, il était impossible de calculer le coefficient Φ pour 30 observations. En effet, 25 indiquaient que les participants étaient certains de participer au vote ($P_{vert}(V=100\%)$), 1 indiquait que le participant était sûr de participer au vote ($P_{vert}(V=80\%)$), 2 indiquaient une indifférence quant au choix de participer ou non ($P(V=1) = P(V=0) = 50\%$) et 2 observations indiquaient que les participants étaient certains ne pas participer ($P(V=0) = 100\%$). Parmi le participant sûr et les participants certains de voter, 4 sont certains de choisir A ou B ($P_{vert}(A) = 100\%$ ou $P_{vert}(B) = 100\%$), 1 est sûr de choisir A ($P_{vert}(A) = 60\%$) et 22 étaient indifférent entre A et B ($P(A) = P(B) = 50\%$). Selon les paramètres de l'expérience, le choix rationnel aurait été de participer au vote ($P(V=1) = 100\%$) mais d'être indifférent entre A et B ($P(A) = P(B) = 50\%$). Par conséquent, les participants ont prédit correctement 22 fois. En résumé, 23 observations ont dévié du choix rationnel : 4 dans le traitement 1, 11 dans le traitement 5 et 8 dans le traitement 4.

À travers le tableau 19, on voit que la majorité des Φ indéterminables se trouvent dans les traitements 1-5 et les traitements 4-8. La figure 6 montre la distribution des Φ indéterminables selon chaque traitement. Ces observations mettent en lumière la relation existante entre la distribution des coefficients Φ individuels indéterminables et les probabilités associées aux prédictions théoriques pour les choix rationnels individuels.

Les probabilités théoriques individuelles suggéraient aux participants de choisir A dans le traitement 1 avec $P(A) = 75\%$ et B dans le traitement 5 avec $P(B) = 75\%$. Lors de l'analyse des Φ indéterminables on a constaté que les participants avaient respectivement choisi A de façon certaine dans le traitement 1 avec $P(A) = 100\%$ et B de façon certaine dans le traitement 1 avec $P(B) = 100\%$. De même, les probabilités théoriques individuelles suggéraient aux participants de participer au vote dans les traitements 4 et 5 avec $P(V=1) = 100\%$. C'est ce qu'ont effectivement fait la majorité des participants avec Φ indéterminables dans ces traitements. À la lumière de ces observations, la présence de coefficients Φ indéterminables résulte d'une perception correcte de l'intensité des probabilités théoriques individuelles dans ces traitements. On comprend aisément le nombre moins élevé de Φ indéterminables dans les traitements 2 et 6 où l'intensité des probabilités théoriques individuelles étaient plus difficiles à percevoir car

FIGURE 6 – Distribution des Φ indéterminables selon les traitements.



plus faibles ($P(V=1) = 62\%$ et $P(A) = P(B) = 50\%$). Toutefois, il est curieux que le nombre de Φ indéterminables soit aussi bas dans les traitements 3-7 relativement aux traitements 1-5 et traitements 4-8. On se serait attendu à observer un nombre élevé de Φ indéterminables dans ces traitements étant donné l'intensité de la probabilité de participer au vote qui semble quand même élevée ($P(V=1) = 86\%$)

Pour les types bleus et jaunes, dont rappelons-le, les participants connaissaient leurs types, respectivement 85 et 71 observations avaient au moins une probabilité jointe égale à zéro ; ce qui a comme conséquence des coefficients Φ indéterminables. Dans le cadre des prédictions individuelles, l'expérience était paramétrée de telle sorte que les participants appartenant au type bleu aient intérêt à toujours voter pour A, tandis que ceux du type jaune avaient toujours intérêt à voter pour B. Dès lors, on s'attendrait à ce $P(A)=100\%$ pour bleu, $P(B)=100\%$ pour jaune et $P(V=1) = 100\%$ pour les deux types. Le tableau 21 montre la distribution des probabilités assignées aux choix A et B pour les types bleus et jaunes pour les observations avec coefficients indéterminables

| Probabilités | 100% | 99% | 95% | 90% | 80% | 10% | 0% | Total |
|-----------------------|------|-----|-----|-----|-----|-----|----|-------|
| Type Bleu ($P(A)$) | 78 | 2 | 1 | 1 | 2 | - | 1 | 85 |
| Type Jaune ($P(B)$) | 68 | - | 1 | 1 | - | 1 | - | 71 |

Tableau 21: Fréquence des probabilités assignées aux choix A et B pour les types bleus et jaunes dans les observations avec Φ individuels indéterminables

Dans l'ensemble, les prédictions des participants sont bonnes. 84 observations dont les participants concernés étaient assignés au type bleu ont prédit A avec $P(A) > 80\%$ et $P(V=1)$

=100%. Seul un participant a prédit B ($P(B) = 100\%$) bien qu'il ait quand même mentionné être certain de participer au vote ($P(V=1) = 100\%$). Par contre pour ces 85 observations, on remarque qu'au moment de choisir, 4 observations ont légèrement dévié du choix rationnel. Ces 4 observations étaient réalisées par le même participant. Chez ceux assignés au type jaune, 70 observations dont les participants concernés ont prédit B avec $P(B) > 90\%$ et $P(V=1) = 100\%$. Seul un participant a prédit A ($P(A) = 90\%$) bien qu'il ait quand même mentionné être certain de participer au vote ($P(V=1) = 100\%$). Cependant, on remarque également qu'au moment de choisir, 4 observations ont dévié du choix rationnel : 3 d'entre elles ont bien prédit tandis que la dernière a émis la seule prédiction erronée parmi les 71 observations. Enfin ces 4 observations appartenaient à 4 individus différents.

À la lumière de ces observations, la présence de coefficients Φ indéterminables est tout simplement la conséquence de la prise de conscience des participants quant au fait qu'ils avaient intérêt de choisir de façon certaine respectivement, le choix A pour ceux appartenant au type bleu et B pour ceux appartenant au type jaune.

En résumé la présence des Φ indéterminables, tant pour les prédictions agrégées que pour les prédictions individuelles, est la résultante de choix certains réalisés par les participants. En interprétant correctement les paramètres de l'expérience, ces derniers attribuent une valeur de 100% à l'une des probabilités marginales ($P(A)$, $P(B)$, $P(V=1)$ ou $P(V=0)$); ce qui rend le calcul du coefficient Φ impossible.

Conclusion

Les sondages pré-électorales sont un outil important tant pour la population que pour les décideurs. Ils donnent la parole aux électeurs et leur permettent d'exprimer leurs opinions (Hirschman (1970)). De plus, les sondages fournissent aux candidats d'une élection un ensemble d'informations sur les préférences des électeurs; ce qui leur permet d'ajuster leurs offres électorales. L'avalanche de sondages dont fait face les Canadiens durant les campagnes électorales indique bien le rôle stratégique que revêtent les sondages pré-électorales non seulement pour les électeurs, mais aussi pour candidats à des élections. Par conséquent, étudier les facteurs pouvant les influencer serait important dans la mesure où ces études participeraient à rendre les prédictions qui en découlent plus fiables.

Un facteur peut-être responsable de la déviation des prédictions des sondages par rapport au résultat réel de l'élection est le biais de négligence de la corrélation. Dans le présent mémoire, nous l'avons examiné. Pour ce faire, nous avons réalisé une expérience en laboratoire durant laquelle les participants devaient effectuer trois tâches. Premièrement, ils devaient anticiper les prédictions de vote de 100 personnes faisant partie de la population (les prédictions agrégées). Deuxièmement, ils devaient faire part de leurs intentions personnelles de vote (les prédictions individuelles). Nous trouvons que les participants ont sous-estimé l'intensité de la corrélation théorique. Cette déviation des anticipations des individus par rapport aux anticipations rationnelles nous a permis d'examiner en profondeur comment se manifeste le biais de la corrélation dans un contexte de sondage pré-électoral.

Ces déviations nous ont d'abord amenés à examiner de plus près l'écart entre les décisions des participants et les anticipations attendues. Nous avons remarqué que la différence entre la corrélation empirique et la corrélation théorique pouvaient être égale (corrélation exacte), être supérieure dans l'absolue à la corrélation théorique (surestimation de la corrélation), être inférieure dans l'absolue à la corrélation théorique négative (sous-estimation ou négligence de la corrélation) et enfin être proche de l'inverse de la vraie corrélation (perception inverse de la corrélation). Grâce à un modèle probit ordonné, nous avons par la suite identifié les facteurs responsables de la variation de la corrélation empirique à travers ces 4 alternatives. Lors de l'interprétation des résultats, nous avons mis l'accent sur les facteurs déterminant la sous-estimation de la vraie corrélation; c'est-à-dire responsable du biais de négligence de la

corrélation. Concernant le niveau d'étude, les résultats des prédictions agrégées concordent avec l'intuition selon laquelle une personne avec un haut niveau d'étude négligerait moins l'intensité de la corrélation. Toutefois, il est surprenant de voir que détenir un niveau d'étude *autre* qu'universitaire diminue également les chances de sous-estimer l'intensité de la corrélation. Pour les prédictions individuelles, le nombre peu élevé d'observations nous empêche de mieux comprendre comment le niveau d'étude affecte le biais. Enfin, pour les prédictions agrégées, il est intéressant de voir que la probabilité de sous-estimer l'intensité de la corrélation est plus élevée pour les traitements à plus faible corrélation théorique et moins élevée pour les traitements à plus forte corrélation théorique. En effet, cela rejoint l'intuition selon laquelle les participants perçoivent la forte magnitude la corrélation théorique dans les traitements 2 et 6 et la négligent moins que dans les traitements 1 et 8 où la magnitude la corrélation théorique est faible et donc difficile à percevoir.

Finalement, nous avons analysé les raisons pour lesquelles il était impossible de déterminer la valeur du coefficient de corrélation Φ pour certaines observations. Notre analyse montre qu'autant pour les prédictions agrégées que pour les prédictions individuelles, la présence de Φ indéterminables est la conséquence des choix certains réalisés par les participants. Plus précisément, les participants attribuent une valeur de 100% à l'une des probabilités marginales ($P(A)$, $P(B)$, $P(V=1)$ ou $P(V=0)$). De ce fait, la valeur du coefficient Φ devient indéterminable.

Dans les études ultérieures, il serait intéressant d'ajuster le paramétrage de l'expérience afin que les choix certains émis par les participants empêchent de calculer les coefficients Φ pour certaines observations. Ainsi, ils disposeront de plus de degrés de liberté dans leurs analyses économétriques. Il serait également intéressant pour les futures recherches sur le biais de négligence de traiter les observations comme des données de panel. En effet, cela permettrait de capter un potentiel effet d'apprentissage ; c'est-à-dire d'observer des améliorations dans les décisions des participants au fil des tours de l'expérience. Enfin, des recherches futures pourraient mettre l'accent sur la perception inverse de la corrélation afin d'en comprendre les origines.

Bibliographie

- Kenneth J ARROW, Robert FORSYTHE, Michael GORHAM, Robert HAHN, Robin HANSON, John O LEDYARD, Saul LEVMORE, Robert LITAN, Paul MILGROM, Forrest D NELSON *et al.* : The promise of prediction markets, 2008.
- Adeline DELAVANDE et Charles F MANSKI : Probabilistic polling and voting in the 2008 presidential election : Evidence from the american life panel. *Public opinion quarterly*, 74 (3):433–459, 2010.
- Benjamin ENKE et Florian ZIMMERMANN : Correlation neglect in belief formation. *The Review of Economic Studies*, 86(1):313–332, 2017.
- Robert S ERIKSON, Michael B MACKUEN et James A STIMSON : *The macro polity*. Cambridge University Press, 2002.
- Robert S ERIKSON et Christopher WLEZIEN : Are political markets really superior to polls as election predictors? *Public Opinion Quarterly*, 72(2):190–215, 2008.
- Robert S ERIKSON et Christopher WLEZIEN : Markets vs. polls as election predictors : An historical assessment. *Electoral Studies*, 31(3):532–539, 2012.
- Erik EYSTER et Georg WEIZSACKER : Correlation neglect in financial decision-making. 2010.
- Robert FORSYTHE : F., gr nelson, gr neumann. *J. Wright*, pages 1142–1161, 1992.
- Robert FORSYTHE, Thomas RIETZ, Roger MYERSON et Robert WEBER : An experimental study of voting rules and polls in three-candidate elections. *International Journal of Game Theory*, 25(3):355–383, 1996.
- Albert O HIRSCHMAN : *Exit, voice, and loyalty : Responses to decline in firms, organizations, and states*, volume 25. Harvard university press, 1970.
- Ido KALLIR et Doron SONSINO : The neglect of correlation in allocation decisions. *Southern Economic Journal*, pages 1045–1066, 2009.
- Steven G KOU et Michael E SOBEL : Forecasting the vote : A theoretical comparison of election markets and public opinion polls. *Political Analysis*, 12(3):277–295, 2004.

- Yoram KROLL, Haim LEVY et Amnon RAPOPORT : Experimental tests of the separation theorem and the capital asset pricing model. *The American Economic Review*, pages 500–519, 1988.
- Gilat LEVY et Ronny RAZIN : Correlation neglect, voting behavior, and information aggregation. *American Economic Review*, 105(4):1634–45, 2015.
- Charles MANSKI : Probabilistic polling. *Navigating Public Opinion : Polls, Policy, and the Future of American Democracy*, pages 251–71, 2002.
- Charles F MANSKI : The use of intentions data to predict behavior : A best-case analysis. *Journal of the American Statistical Association*, 85(412):934–940, 1990.
- Charles F MANSKI : Interpreting the predictions of prediction markets. *economics letters*, 91(3):425–429, 2006.
- Rebecca B MORTON : *Analyzing elections*. Citeseer, 2006.
- Erik SNOWBERG, Justin WOLFERS et Eric ZITZEWITZ : Prediction markets for economic forecasting. *In Handbook of Economic Forecasting*, volume 2, pages 657–687. Elsevier, 2013.
- Justin WOLFERS et Eric ZITZEWITZ : Prediction markets. *Journal of economic perspectives*, 18(2):107–126, 2004.
- Justin WOLFERS et Eric ZITZEWITZ : Prediction markets in theory and practice. Rapport technique, national bureau of economic research, 2006.

Annexe A

Distribution des Φ déterminables

| | | Prédictions agrégées | | | | |
|---|--|---------------------------|-----------------|--------------------|---------------|------------|
| Variables | Description | Corrélation inverse | Sous-estimation | Corrélation exacte | Surestimation | Total |
| Distribution selon l'âge | | | | | | |
| Age_1 | Catégorie d'âge [21 – 24] | 40.00% | 44.35% | 10.43% | 5.22% | 115 |
| Age_2 | Catégorie d'âge [25 – 34] | 30.61% | 53.06% | 9.52% | 6.80% | 147 |
| Age_3 | Catégorie d'âge [35 – 59] | 26.32% | 55.64% | 9.77% | 8.27% | 133 |
| Nombre total d'observations | | 126 | 203 | 39 | 27 | 395 |
| Distribution selon le genre | | | | | | |
| Femme | = 1 si femme, = 0 si homme | 21.72% | 59.6% | 11.62% | 7.07% | 198 |
| Homme | = 1 si homme, = 0 si femme | 42.13% | 43.15% | 8.12% | 6.60% | 197 |
| Nombre total d'observations | | 126 | 203 | 39 | 27 | 395 |
| Distribution selon le niveau d'éducation | | | | | | |
| Baccalauréat | = 1, si Baccalauréat, = 0 sinon | 29.84% | 53.89% | 8.38% | 7.78% | 167 |
| Maîtrise | = 1, si Maîtrise, = 0 sinon | 25.23% | 55.14% | 14.02% | 5.61% | 107 |
| Doctorat | = 1, si Doctorat, = 0 sinon | 38.30% | 48.94% | 6.38% | 6.38% | 47 |
| Autres | = 1, si autre niveau t, = 0 sinon | 41.89% | 41.89% | 9.46% | 6.76% | 74 |
| Nombre total d'observations | | 126 | 203 | 39 | 27 | 395 |
| Distribution selon la faculté | | | | | | |
| Faculté : sciences sociales | = 1 Si en sciences sociales, = 0 sinon | 33.51% | 50.00% | 9.04% | 7.45% | 188 |
| Faculté : autres | = 1 Si Pas en sciences sociales, = 0 sinon | 30.43% | 52.66% | 10.63% | 6.28% | 207 |
| Nombre total d'observations | | 126 | 203 | 39 | 27 | 395 |
| Distribution selon le traitement | | | | | | |
| Traitement 1 : $\Phi = 0.08$ | = 1 si traitement 1, = 0 sinon | 44.64% | 44.64% | 0.00% | 10.71% | 56 |
| Traitement 2 : $\Phi = -0.47$ | = 1 si traitement 2, = 0 sinon | 38.60% | 61.40% | 0.00% | 0.00% | 57 |
| Traitement 3 : $\Phi = 0.25$ | = 1 si traitement 3, = 0 sinon | 38.30% | 61.70% | 0.00% | 0.00% | 47 |
| Traitement 4 : $\Phi = 0$ | = 1 si traitement 4, = 0 sinon | 0.00% | 50.00% | 44.74% | 5.26% | 38 |
| Traitement 5 : $\Phi = -0.08$ | = 1 si traitement 5, = 0 sinon | 42.59% | 40.74% | 0.00% | 16.67% | 54 |
| Traitement 6 : $\Phi = 0.47$ | = 1 si traitement 6, = 0 sinon | 35.71% | 64.29% | 0.00% | 0.00% | 56 |
| Traitement 7 : $\Phi = -0.25$ | = 1 si traitement 7, = 0 sinon | 33.96% | 60.38% | 0.00% | 5.66% | 53 |
| Traitement 8 : $\Phi = 0$ | = 1 si traitement 8, = 0 sinon | 0.00% | 14.71% | 64.71% | 20.59% | 34 |
| Nombre total d'observations | | 126 | 203 | 39 | 27 | 395 |
| | | Prédictions individuelles | | | | |
| Variables | Description | Corrélation inverse | Sous-estimation | Corrélation exacte | Surestimation | Total |
| Distribution selon l'âge | | | | | | |
| Age_1 | Catégorie d'âge [21 – 24] | 51.42% | 28.57% | 8.57% | 51.43% | 35 |
| Age_2 | Catégorie d'âge [25 – 34] | 21.88% | 60.94% | 14.06% | 3.12% | 64 |
| Age_3 | Catégorie d'âge [35 – 59] | 33.33% | 41.67% | 13.33% | 11.67% | 60 |
| Nombre total d'observations | | 52 | 74 | 20 | 13 | 159 |
| Distribution selon le genre | | | | | | |
| Femme | = 1 si femme, = 0 si homme | 33.33% | 42.03% | 17.39% | 7.25% | 69 |
| Homme | = 1 si homme, = 0 si femme | 32.22% | 50.00% | 8.89% | 8.89% | 90 |
| Nombre total d'observations | | 52 | 74 | 20 | 13 | 159 |
| Distribution selon le niveau d'éducation | | | | | | |
| Baccalauréat | = 1, si Baccalauréat, = 0 sinon | 31.25% | 42.19% | 17.19% | 9.38% | 64 |
| Maîtrise | = 1, si Maîtrise, = 0 sinon | 31.71% | 53.66% | 9.76% | 4.88% | 41 |
| Doctorat | = 1, si Doctorat, = 0 sinon | 41.67% | 45.83% | 12.50% | 0.00% | 24 |
| Autres | = 1, si autre niveau t, = 0 sinon | 30.00% | 46.67% | 6.67% | 16.67% | 30 |
| Nombre total d'observations | | 52 | 74 | 20 | 13 | 159 |
| Distribution selon la faculté | | | | | | |
| Faculté : sciences sociales | = 1 Si en sciences sociales, = 0 sinon | 34.15% | 45.12% | 14.63% | 6.10% | 77 |
| Faculté : autres | = 1 Si Pas en sciences sociales, = 0 sinon | 34.15% | 45.12% | 14.63% | 6.10% | 82 |
| Nombre total d'observations | | 52 | 74 | 20 | 13 | 159 |
| Distribution selon le traitement | | | | | | |
| Traitement 1 : $\Phi = 0.26$ | = 1 si traitement 1, = 0 sinon | 42.11% | 47.37% | 0.00% | 10.53% | 19 |
| Traitement 2 : $\Phi = -0.76$ | = 1 si traitement 2, = 0 sinon | 31.82% | 68.18% | 0.00% | 0.00% | 22 |
| Traitement 3 : $\Phi = 0.37$ | = 1 si traitement 3, = 0 sinon | 50.00% | 50.00% | 0.00% | 0.00% | 24 |
| Traitement 4 : $\Phi = 0$ | = 1 si traitement 4, = 0 sinon | 0.00% | 21.05% | 63.16% | 15.79% | 19 |
| Traitement 5 : $\Phi = -0.26$ | = 1 si traitement 5, = 0 sinon | 47.37% | 42.11% | 0.00% | 10.53% | 19 |
| Traitement 6 : $\Phi = 0.76$ | = 1 si traitement 6, = 0 sinon | 37.50% | 58.33% | 0.00% | 0.00% | 17 |
| Traitement 7 : $\Phi = -0.37$ | = 1 si traitement 7, = 0 sinon | 37.50% | 58.33% | 0.00% | 4.17% | 24 |
| Traitement 8 : $\Phi = 0$ | = 1 si traitement 8, = 0 sinon | 0.00% | 13.33% | 53.33% | 33.33% | 15 |
| Nombre total d'observations | | 52 | 74 | 20 | 13 | 159 |

Tableau 22: Statistiques descriptives

Annexe B

Aperçu de l'interface présentée aux participants lors d'un tour entier

| Couleur | Groupe | | | | Information supplémentaire: |
|---|---|------------|-----------------------|-------|------------------------------------|
| | Bleu | Vert foncé | Vert pâle | Jaune | |
| Taille de groupe | 25 | 25 | 25 | 25 | Tour 1 Nombre maximal de tour 9 |
| Option A Option B | Valeurs des options (en UPE) | | | | |
| | 120 | 97 | 44 | 20 | |
| | 20 | 43 | 96 | 120 | |
| Frais associés au choix entre les options (en UPE) Frais élevé : 60 Frais bas : 0 | Probabilité de devoir payer les frais (en %) | | | | |
| | 71 | 27 | 1 | 1 | |
| | 29 | 73 | 99 | 99 | |
| "Ne pas faire de choix" | Valeur de "ne pas faire de choix" (en UPE) lorsque ni A ni B ne sont choisis | | | | |
| | 20 | 43 | 44 | 20 | |
| <p>Considérez les valeurs des options et les frais associés au choix selon le groupe présentés dans le tableau ci-dessus. Il y a 25 personnes dans chacun des groupes bleu, vert foncé, vert pâle et jaune soit au total 100 personnes.</p> <p>Quelle est selon vous l'option qui sera choisie le plus souvent ?</p> <p><input type="radio"/> L'option A sera choisie le plus souvent. <input type="radio"/> L'option B sera choisie le plus souvent.</p> | | | | | |
| Instructions | | | Suite | | |

| Couleur | Groupe | | | | Information supplémentaire: |
|--|---|------------------------|-----------|-----------------------|------------------------------------|
| | Bleu | Vert foncé | Vert pâle | Jaune | |
| Taille de groupe | 25 | 25 | 25 | 25 | Tour 1 Nombre maximal de tour 9 |
| <i>Option A</i> | Valeurs des options (en UPE) | | | | |
| | 120 | 97 | 44 | 20 | |
| <i>Option B</i> | 20 | 43 | 96 | 120 | |
| Frais associés au choix entre les options (en UPE) | Probabilité de devoir payer les frais (en %) | | | | |
| Frais élevé : 60 | 71 | 27 | 1 | 1 | |
| Frais bas : 0 | 29 | 73 | 99 | 99 | |
| "Ne pas faire de choix" | Valeur de "ne pas faire de choix" (en UPE) lorsque ni <i>A</i> ni <i>B</i> ne sont choisis | | | | |
| | 20 | 43 | 44 | 20 | |
| <p>Considérez les valeurs des options et les frais associés au choix selon le groupe présentés dans le tableau ci-dessus. Il y a 25 personnes dans chacun des groupes bleu, vert foncé, vert pâle et jaune soit au total 100 personnes.</p> <p>Quelle est votre prédiction du choix agrégé de ces 100 personnes ?</p> <p>Nombre de personnes qui choisiront l'une des options <i>A</i> ou <i>B</i> (sur un total de 100): <input type="text"/></p> <p>Nombre de personnes qui ne choisiront aucune des deux options (sur un total de 100): <input type="text"/></p> | | | | | |
| Instructions | | Retour | | Suite | |

| Couleur | Groupe | | | | Information supplémentaire: |
|--|--|------------|-----------|-------|------------------------------------|
| | Bleu | Vert foncé | Vert pâle | Jaune | |
| Taille de groupe | 25 | 25 | 25 | 25 | Tour 1 Nombre maximal de tour 9 |
| <i>Option A</i> | Valeurs des options (en UPE) | | | | |
| | 120 | 97 | 44 | 20 | |
| <i>Option B</i> | 20 | 43 | 96 | 120 | |
| Frais associés au choix entre les options (en UPE) | Probabilité de devoir payer les frais (en %) | | | | |
| Frais élevé : 60 | 71 | 27 | 1 | 1 | |
| Frais bas : 0 | 29 | 73 | 99 | 99 | |
| "Ne pas faire de choix" | Valeur de "ne pas faire de choix" (en UPE) lorsque ni A ni B ne sont choisis | | | | |
| | 20 | 43 | 44 | 20 | |

Considérez les valeurs des options et les frais associés au choix selon le groupe présentés dans le tableau ci-dessus. Il y a 25 personnes dans chacun des groupes bleu, vert foncé, vert pâle et jaune soit au total 100 personnes.

Quelle est votre prédiction du choix agrégé de ces 100 personnes ?

Nombre de personnes qui choisiront l'une des options A ou B (sur un total de 100):

Nombre de personnes qui ne choisiront aucune des deux options (sur un total de 100):

Parmi ceux qui ne choisiront aucune option, selon vous, combien de personnes préfèrent chacune des options A et B?
Svp, cochez si vous répondez en nombres absolus ou en termes de proportion et indiquez votre prévision.

Absolu En pourcentage

| | | |
|-----------------------------|----------------------|---------------------------------|
| L'option A est préférée par | <input type="text"/> | <input type="text" value="50"/> |
| L'option B est préférée par | <input type="text"/> | <input type="text" value="50"/> |
| Somme | 50 | 100 |

[Instructions](#) [Retour](#) [Suite](#)

| | | | |
|---|--|--------------|------------------------------------|
| Couleur | Groupe | | Information supplémentaire: |
| | Bleu | | |
| Taille de groupe | 25 | | Tour 1 Nombre maximal de tour 9 |
| Option A Option B | Valeurs des options (en UPE) | | Vous êtes du groupe Bleu |
| | 120 | | |
| | 20 | | |
| Frais associés au choix entre les options (en UPE) | Probabilité de devoir payer les frais (en %) | | |
| Frais élevé : 60 | 71 | | |
| Frais bas : 0 | 29 | | |
| "Ne pas faire de choix" | Valeur de "ne pas faire de choix" (en UPE) lorsque ni A ni B ne sont choisis | | |
| | 20 | | |
| <p>À présent, vous connaissez la couleur du groupe auquel vous appartenez.</p> <p>Supposez que vous décidiez de choisir l'une des options. Laquelle des deux options choisirez-vous ?</p> <p>Si je décide de choisir l'une des deux options, <input type="radio"/> je choisirai l'option A. <input type="radio"/> je choisirai l'option B. <input type="radio"/> je ne sais pas quelle option je choisirai.</p> | | | |
| Instructions | | Suite | |

| | | | |
|---|---|------------------------|------------------------------------|
| Couleur | Groupe | | Information supplémentaire: |
| | Bleu | | |
| Taille de groupe | 25 | | Tour 1 Nombre maximal de tour 9 |
| Option A Option B | Valeurs des options (en UPE) | | Vous êtes du groupe Bleu |
| | 120 20 | | |
| Frais associés au choix entre les options (en UPE) Frais élevé : 60 Frais bas : 0 | Probabilité de devoir payer les frais (en %) | | |
| | 71 29 | | |
| "Ne pas faire de choix" | Valeur de "ne pas faire de choix" (en UPE) lorsque ni A ni B ne sont choisis | | |
| | 20 | | |
| <p>Supposez que vous décidiez de ne pas choisir aucune des options.</p> <p>Quelles sont les chances que vous préférerez chacune des options A et B quand vous ne choisissez aucune des options ?</p> <p><i>S'il vous plaît, indiquer l'intensité de préférer chaque option en termes de probabilités. "0" indique que vous êtes certain de ne pas préférer l'option concernée et "100" indique que vous êtes certain de la préférer. Si vous choisissez un nombre compris entre "0" et "100", vous indiquez que vous n'êtes pas certain de préférer l'option concernée. Plus le nombre choisi est proche de "100", plus vous êtes certain de préférer l'option concernée. Plus le nombre choisi est proche de "0", plus vous êtes certain de ne pas préférer l'option concernée.</i></p> <p>Les chances que je ne choisisse aucune des options et que je préfère l'option A sont de : (en %) <input type="text" value="50"/></p> <p>Les chances que je ne choisisse aucune des options et que je préfère l'option B sont de : (en %) <input type="text" value="50"/></p> | | | |
| Instructions | | Retour | Suite |

| | | | |
|---|---|------------------------|------------------------------------|
| Couleur | Groupe | | Information supplémentaire: |
| | Bleu | | |
| Taille de groupe | 25 | | Tour 1 Nombre maximal de tour 9 |
| Option A | Valeurs des options (en UPE) | | Vous êtes du groupe Bleu |
| Option B | 120 | | |
| | 20 | | |
| Frais associés au choix entre les options (en UPE) | Probabilité de devoir payer les frais (en %) | | |
| Frais élevé : 60 | 71 | | |
| Frais bas : 0 | 29 | | |
| "Ne pas faire de choix" | Valeur de "ne pas faire de choix" (en UPE) lorsque ni A ni B ne sont choisis | | |
| | 20 | | |
| <p>Supposez que vous décidez de choisir l'une des options.</p> <p>Quelles sont les chances que vous choisissiez chacune des options A et B ?</p> <p><i>S'il vous plaît, indiquer la chance de choisir chaque option en termes de probabilités. "0" indique que vous êtes certain de ne pas vouloir choisir l'option concernée et "100" indique que vous êtes certain de vouloir choisir l'option concernée. Si vous choisissez un nombre compris entre "0" et "100", vous indiquez que vous n'êtes pas certain de choisir l'option concernée. Plus le nombre choisi est proche de "100", plus vous êtes certain de choisir l'option concernée. Plus le nombre choisi est proche de "0", plus vous êtes certain de ne pas choisir l'option concernée.</i></p> <p>Les chances que je choisisse l'option A sont de : (en %) <input type="text" value="50"/></p> <p>Les chances que je choisisse l'option B sont de : (en %) <input type="text" value="50"/></p> | | | |
| Instructions | | Retour | Suite |

| | | | |
|---|---|------------------------|------------------------------------|
| Couleur | Groupe | | Information supplémentaire: |
| | Bleu | | |
| Taille de groupe | 25 | | Tour 1 Nombre maximal de tour 9 |
| Option A Option B | Valeurs des options (en UPE) | | Vous êtes du groupe Bleu |
| | 120 20 | | |
| Frais associés au choix entre les options (en UPE) | Probabilité de devoir payer les frais (en %) | | |
| Frais élevé : 60 Frais bas : 0 | 71 29 | | |
| "Ne pas faire de choix" | Valeur de "ne pas faire de choix" (en UPE) lorsque ni A ni B ne sont choisis | | |
| | 20 | | |
| <p>Selon vous, quelles sont les chances que vous choisissiez au moins l'une des options ou que vous décidiez de ne pas faire de choix ?</p> <p><i>S'il vous plaît, indiquez les chances de faire ou non un choix en termes de probabilités. "0" indique que vous êtes certain de ne pas vouloir faire de choix et "100" indique que vous êtes certain de vouloir choisir l'une des deux options. Si vous choisissez un nombre compris entre "0" et "100", vous indiquez que vous n'êtes pas certain de faire ou non un choix. Plus le nombre choisi est proche de "100", plus vous êtes certain de choisir l'une des deux options. Plus le nombre choisi est proche de 0, plus vous êtes certain de ne pas faire de choix.</i></p> <p>Les chances que je choisisse l'une des options A ou B sont de : (en %) <input type="text" value="50"/></p> <p>Les chances que je ne fasse pas de choix (ne choisisse aucune des deux options) sont de : (en %) <input type="text" value="50"/></p> | | | |
| Instructions | | Retour | Suite |

| | | | |
|---|--|---|------------------------------------|
| Couleur | Groupe | | Information supplémentaire: |
| | Bleu | | |
| Taille de groupe | 25 | | |
| | | | |
| Option A | Valeurs des options (en UPE) | | Vous êtes du groupe Bleu |
| | 120 | | |
| Option B | 20 | | Frais associés au choix 60 |
| | | | |
| Frais associés au choix entre les options (en UPE) | | | |
| Frais élevé : | 60 | Les frais (en UPE) associés au choix sont de: 60 | |
| Frais bas : | 0 | | |
| "Ne pas faire de choix" | Valeur de "ne pas faire de choix" (en UPE) lorsque ni A ni B ne sont choisis | | |
| | 20 | | |
| <p>A présent, vous connaissez les valeurs des options et le coût du choix pour vous.</p> <p>Quelle option choisissez-vous ?</p> <p> <input type="radio"/> Je choisis l'option A. <input type="radio"/> Je choisis l'option B. <input type="radio"/> Je ne choisis aucune des deux options: je ne fais pas de choix. </p> | | | |
| Instructions | | Suite | |

Vous avez eu la possibilité de choisir entre *option A* et *option B*.
Les deux options et ne pas choisir avaient les valeurs suivantes (en UPE) pour vous:

| | |
|-------------------|-----|
| <i>Option A</i> : | 120 |
| <i>Option B</i> : | 20 |

Si vous décidez de choisir, ça vous coûte: 60
Si vous décidez de ne pas choisir, ça ne vous coûte rien et vous recevez: 20

Vous avez décidé de:

- Choisir l'option A
- Choisir l'option B
- Ne pas choisir une option

Votre gain (en UPE) de ce tour est de : -40

Passer au tour suivant

Annexe C

Liste complète des questions posées aux participants

étape 1

Question 1

Considérez les valeurs des options et les frais associés au choix selon le type présentés dans le tableau ci-dessus. Il y a 25 personnes dans chacun des types bleu, vert foncé, vert pâle et jaune, soit au total 100 personnes.

Quelle est votre prédiction du choix agrégé de ces 100 personnes ?

Nombre de personnes qui choisiront l'une des options A ou B (sur un total de 100) ...

Nombre de personnes qui ne choisiront aucune des deux options (sur un total de 100) ...

Parmi ceux qui choisiront l'une des options, selon vous, combien de personnes choisiront chacune des options A et B ?

Svp, cochez si vous répondez en nombres absolus ou en termes de proportion et indiquez votre prévision.

Réponse absolue

– *Je pense que ... personnes choisiront l'option A.*

– *Je pense que ... personnes choisiront l'option B.*

Réponse relative

– *Je pense que ... % des personnes choisiront l'option A.*

– *Je pense que ... % des personnes choisiront l'option B.*

Question 2

Considérez les valeurs des options et les frais associés au choix selon le type présentés dans le tableau ci-dessus. Il y a 25 personnes dans chacun des types bleu, vert foncé, vert pâle et jaune, soit au total 100 personnes.

Quelle est votre prédiction du choix agrégé de ces 100 personnes ?

Nombre de personnes qui choisiront l'option A (sur un total de 100) : ...

Nombre de personnes qui choisiront l'option B (sur un total de 100) : ...

Nombre de personnes qui ne choisiront aucune des deux options (sur un total de 100) : ...

Question 3a

Considérez les valeurs des options et les frais associés au choix selon le type présentés dans le tableau ci-dessus. Il y a 25 personnes dans chacun des types bleu, vert foncé, vert pâle et jaune, soit au total 100 personnes.

Quelle est selon vous l'option qui sera choisie le plus souvent ?

L'option A sera choisie le plus souvent

L'option B sera choisie le plus souvent

Question 3b

Considérez les valeurs des options et les frais associés au choix selon le type présentés dans le tableau ci-dessus. Il y a 25 personnes dans chacun des types bleu, vert foncé, vert pâle et jaune, soit au total 100 personnes.

Quelle est selon vous l'option qui sera choisie le plus souvent ?

L'option A sera choisie le plus souvent

L'option B sera choisie le plus souvent Je ne sais pas

Si vous n'êtes pas sûr de l'option qui sera choisie, veuillez s'il vous plaît indiquer l'option vers laquelle vous penchez le plus

L'option A sera probablement choisie plus souvent que l'option B

L'option B sera probablement choisie plus souvent que l'option A

étape 2

Question 4a

Supposez que vous décidiez de choisir l'une des options. Laquelle des deux options

choisiriez-vous ?

Si je décide de choisir l'une des options,

je choisirai l'option A.

je choisirai l'option B.

Question 4b

Supposez que vous décidiez de choisir l'une des options. Laquelle des deux options choisiriez-vous ?

Si je décide de choisir l'une des options,

je choisirai l'option A.

je choisirai l'option B.

je ne sais pas quelle option je choisirai.

Si vous n'êtes pas sûr de l'option que vous choisiriez, veuillez s'il vous plaît indiquer l'option vers laquelle vous penchez le plus.

Je choisirai probablement l'option A

Je choisirai probablement l'option B

Question 5a

Supposez que vous décidiez de choisir l'une des options. Quelles sont les chances que vous choisissiez chacune des options A et B ?

S'il vous plaît, indiquer la chance de choisir chaque option en termes de probabilités. "0" indique que vous êtes certain de ne pas vouloir choisir l'option concernée et "100" indique que vous êtes certain de vouloir choisir l'option concernée. Si vous choisissez un nombre compris entre "0" et "100", vous indiquez que vous n'êtes pas certain de choisir l'option concernée. Plus le nombre choisi est proche de "100", plus vous êtes certain de choisir l'option concernée. Plus le nombre choisi est proche de "0", plus vous êtes certain de ne pas choisir l'option concernée.

Les chances que je choisisse l'option A sont de : (en %) ...

Les chances que je choisisse l'option B sont de : (en %) ...

Question 5b

Selon vous, quelles sont les chances que vous choisissiez au moins l'une des options ou que vous décidiez de ne pas faire de choix ?

S'il vous plaît, indiquer la chance de choisir chaque option en termes de probabilités. "0" indique que vous êtes certain de ne pas vouloir choisir l'option concernée et "100" indique que vous êtes certain de vouloir choisir l'une des deux options. Si vous choisissez un nombre compris entre "0" et "100", vous indiquez que vous n'êtes pas certain de faire ou non un choix. Plus le nombre choisi est proche de "100", plus vous êtes certain de choisir l'option concernée. Plus le nombre choisi est proche de "0", plus vous êtes certain de ne pas faire de choix.

Les chances que je choisisse l'une des options A ou B sont de : (en %) ...

Les chances que je ne fasse pas de choix (ne choisisse aucune des deux options) sont de : (en %) ...

Question 6

Quelles sont les chances que vous choisissiez chacune des options A et B ?

S'il vous plaît, indiquer la chance de choisir chaque option en termes de probabilités. "0" indique que vous êtes certain de ne pas vouloir choisir l'option concernée et "100" indique que vous êtes certain de vouloir choisir l'option concernée. Si vous choisissez un nombre compris entre "0" et "100", vous indiquez que vous n'êtes pas certain de choisir l'option concernée. Plus le nombre choisi est proche de "100", plus vous êtes certain de choisir l'option concernée. Plus le nombre choisi est proche de "0", plus vous êtes certain de ne pas choisir l'option concernée.

Les chances que je choisisse l'option A sont de : (en %) ...

Les chances que je choisisse l'option B sont de : (en %) ...

Les chances que je ne fasse pas de choix (ne choisisse aucune des deux options) sont de : (en %) ...

Étape 3

À présent, vous connaissez les valeurs des options et le coût du choix pour vous. Quelle option choisissez-vous ?

– Je choisis l'option A.

– Je choisis l'option B.

– Je ne choisis aucune des deux options : je ne fais pas de choix