

Rapport de stage

Fait par Golovanova Elizaveta
BDA, Université Grenoble Alpes

25 août 2022

Résumé

Keywords :

Table des matières

1	Introduction	3
2	Recherche de stage et présentation de la structure	3
2.1	Recherche de stage	3
2.2	Présentation de la structure de stage	4
3	Missions effectuées pendant le stage	4
4	Développement d'une mission avec problématique économique et analyse	8
4.1	Problématique économique	8
4.2	Base de données	8
4.3	Types des tâches d'élicitation du risque	10
4.4	Questionnaires	14
4.5	Variables pour l'étude	14
4.6	Méthodologie	17
4.7	Resultats	19
4.8	Limitations et difficultés	24
5	Conclusion	25

1 Introduction

Ce rapport est consacré à l'analyse du déroulement de stage dans le cadre de la formation en Master 1, parcours Business et Analyse de Données, Faculté d'Économie de l'Université Grenoble Alpes. La première partie de ce rapport décrira le processus de recherche de mon stage, l'évolution de mon CV et de ma lettre de motivation, et comment ma candidature s'était diffusée à diverses organisations. La deuxième partie du rapport parlera de l'organisation où j'ai effectué mon stage, notamment ses principales activités, le nombre d'employés. La troisième partie est réservée à la description du projet auquel j'ai participé, c'est-à-dire son objectif global et mon rôle dans sa réalisation. Dans la quatrième partie, une de mes missions sera décrite en détail. Dans la cinquième partie finale, les résultats de mon stage seront résumés.

2 Recherche de stage et présentation de la structure

2.1 Recherche de stage

Dans cette section, je décrirai comment mon CV et ma lettre de motivation ont évolué en fonction de l'expérience de recherche d'un stage. Au départ, j'ai rédigé un CV qui ne contenait que des faits nus sur ma formation, mon expérience de travail et mes compétences. Par exemple, je n'ai pas décrit mes tâches dans mes emplois antérieurs, ni précisé les matières que j'ai suivies pendant mes études. Au fil du temps, je me suis rendu compte que ces détails sont importants pour se démarquer des autres candidats. J'ai réfléchi plus attentivement à mes avantages, listé mes réalisations au travail, les sujets que j'ai étudiés. J'ai également ajouté à mon CV mes qualités personnelles que les recruteurs recherchaient pour le poste d'analyste de données.

J'ai aussi initialement rédigé une lettre de motivation individuellement pour chaque poste, en précisant le nom de l'entreprise, le nom du recruteur, pour que la lettre ait l'air personnelle. Cependant, cela a pris beaucoup de temps et n'a donné aucun résultat. Comme le processus de candidature en France est bureaucratiquement compliqué, avec le temps, j'ai décidé de simplifier ma lettre de motivation et de la rendre universelle.

Je tiens également à souligner la participation de l'université aux modifications qui ont eu lieu. Avec l'aide de Sylvian Housset, les fautes de frappe et les formulations inexactes ont été corrigées dans mon CV et ma lettre de motivation.

Après avoir modifié mon CV et ma lettre de motivation, j'ai commencé à recevoir des offres d'emploi, mais elles ne me convenaient pas pour poursuivre le processus de formation. J'ai eu plusieurs entretiens et noté les entreprises où je pourrais essayer d'obtenir un emploi après la fin de ma formation. En parallèle, j'ai diffusé activement mon CV aux entreprises de Grenoble et ses environs, y compris les offres envoyées par les responsables de notre master, et je me suis renseignée également sur les postes disponibles parmi les professeurs de mon université. Au total, j'ai postulé à plus de 70 endroits pour toute la recherche. Finalement, j'ai trouvé un stage dans le laboratoire GAEL avec Paolo Crosetto qui a été chargé du cours d'analyse de données en R que j'ai suivi. Ayant déjà 4 ans d'expérience dans le domaine scientifique en Russie et cherchant la possibilité de poursuivre mon activité à l'international, j'étais convaincue que ce stage s'intègre parfaitement dans mon projet professionnel.

2.2 Présentation de la structure de stage

L'organisme qui m'a accueilli est l'Institut National de Recherche pour l'Agriculture, l'alimentation et l'Environnement (INRAE). J'ai effectué un stage dans un des laboratoires affiliés à cet organisme, à savoir au Laboratoire d'Economie Appliquée à Grenoble (GAEL). Il est composé d'une quarantaine de chercheurs auxquels s'ajoutent des post-doctorants, des doctorants et du personnel administratif et d'appui à la recherche.

3 Missions effectuées pendant le stage

On peut observer que les gens, lorsqu'ils sont exposés à l'incertitude, agissent différemment. La plupart essaient de minimiser cette incertitude autant que possible. Par exemple, un investisseur peut choisir d'investir son argent dans un compte bancaire avec un taux d'intérêt faible mais garanti, plutôt que dans des actions, qui peuvent avoir un rendement attendu élevé mais également comporter un risque de perte. Cet effet est appelé aversion au risque, il caractérise la propension des personnes à prendre telle ou telle décision face

au risque. L'analyse de l'aversion au risque joue un rôle primordial dans plein de domaines de l'économie et de la vie.

Une des objectifs dans l'économie expérimentale est d'éliciter les risques, c'est-à-dire de formaliser le comportement des personnes en situation à risque. Un concept qui décrit la capacité d'une mesure d'en prédire des autres, ou de prédire des comportements est connu comme validité externe. Il existe différentes techniques pour mesurer les attitudes face au risque. Une question directe peut être posée pour le déterminer, par exemple, « Comment vous voyez-vous ? Êtes-vous généralement une personne totalement prête à prendre des risques ou essayez-vous d'éviter de prendre des risques ? » (Gert G. Wagner (2007)) ou il peut être la demande d'indiquer la probabilité d'être impliqué dans un événement particulier, comme tricher à un examen (Blais & Weber (2006)). De même il est possible d'estimer les attitudes face au risque en utilisant une tâche d'élicitation des risques (RET, *Risk Elicitation Task*), par exemple, où le répondant est invité à choisir entre deux options plus et moins risquées (Holt & Laury (2002)) ou à choisir une seule loterie de la liste proposé (Eckel & Grossman (2002)). Il existe de nombreuses variantes de ces tâches, certaines impliquent une composante visuelle pour une perception plus facile (Lejuez et al. (2002), Crosetto & Filippin (2013)). Dans quelques tâches, les pertes sont implicites (Menkhoff & Sakha (2017)), dans d'autres les loteries se trouvent entièrement dans le domaine des gains (Eckel & Grossman (2002), Holt & Laury (2002)). Dans certaines tâches, il y a le choix entre deux loteries, dont l'une est plus risquée. D'autres peuvent offrir un choix entre loterie risquée et un somme d'argent, techniquement connu comme un *équivalent certain* (Menkhoff & Sakha (2017), Csermely & Rabas (2014)).

La thème du projet où j'ai participé pendant mon stage était "Méta-analyse de la validité externe des tâches d'élicitation des risques". Il vise à rassembler les connaissances existantes dans la domaine de l'élicitation des risques, de mieux comprendre l'état actuel de la validité externe des mesures qui aident d'identifier des risques (par méta-analyse) et donner libre accès à d'autres chercheurs et personnes intéressées en ligne et avec une base de données en constante augmentation.

L'enjeu principal de ce projet est de repenser les tâches d'élicitation des attitudes face au risque. Le fait est que les RETs montrent la validité externe basse avec les mesures

autodéclarées et avec celles issues d'activités à risque dans le monde réel et entre eux. Ainsi, le résultat de l'analyse des données recueillies au cours du projet devrait être la création d'une nouvelle tâche qui résoudrait le problème de la faible corrélation.

Il est à noter, que le projet pré-existait à mon stage, il a été lancé en 2019 sous la direction de Paolo Crosetto. Au début de mon stage, une version brute du site web avec la base de données pas très grande et une série limitée de fonctionnalités était prête, mais elle devait être entièrement repensée. En outre, lors de la création d'un nouveau site web, j'ai pris quelques idées conceptuelles de la version originale.

Ma mission principale était de redévelopper un site web informatif, bien structuré, techniquement stable et joli en utilisant **R shiny**. Pour y parvenir, j'ai dû effectuer plusieurs sous-tâches. Tout d'abord, j'ai lu nombreux articles scientifiques différents basés sur des expériences en laboratoire. Deuxièmement, une demande de données a été faite pour enrichir le site auprès de la communauté scientifique en suivant deux méthodes :

1. Paolo a envoyé un e-mail décrivant le projet et demandant l'accès aux données issues d'expériences d'élicitation des attitudes face au risque en laboratoire à ses collègues directement, en utilisant la base de contact qu'il avait et la liste de diffusion 'ESA-discuss'. Une quinzaine d'auteurs ont partagé ces informations avec nous. Depuis lors, de nouvelles données ainsi que les fonctionnalités ont été ajoutées au site, Paolo renverra encore une demande afin d'obtenir plus de données ;
2. Étant donné que peu de personnes ont répondu, nous avons décidé d'élargir notre base de contacts et de rechercher d'autres auteurs qui ont fait des expériences similaires. Nous avons convenu de chercher des articles pour chaque type de tâche que nous considérons dans notre projet et pour lequel nous n'avons pas eu beaucoup d'observations.

En utilisant les mots clés appropriées dans Google Scholar, j'ai sélectionné une liste d'articles qui contiennent très probablement des données sur les expériences. Vérifier manuellement le contenu de tous les articles et extraire les e-mails des auteurs prendrait beaucoup de temps, j'ai donc écrit un analyseur en 'Python' qui le faisait automatiquement.

Ainsi, les liens sur ces articles et son information donnée par le site, tel que les noms

d’auteurs et d’article et l’année de publication ont été collectés pour chaque type de tâche. Ensuite, j’ai supprimé les articles pour lesquels nous avons déjà des données. Puis, j’ai commencé à suivre les liens et à essayer de lire le texte des articles. Cependant, avec l’aide de Google Scholar, je ne peux y accéder qu’à une partie gratuitement. La plupart est toujours disponible uniquement par abonnement. C’est pourquoi j’ai utilisé le site Scihub, qui fournit un accès automatique et gratuit aux textes des articles scientifiques. J’ai écrit le deuxième analyser pour lire le reste des articles avec Scihub et extraire des mails à partir du format pdf en Python. Après avoir analysé environ 1500 articles, j’ai pu extraire les e-mails d’environ 600.

Paolo va bientôt rédiger les messages personnels afin de diffuser une demande à chaque auteur pour les données des expériences menées dans leurs articles.

Ensuite, il fallait alimenter le site web avec une quantité importante de base de données d’éllicitation de risque qu’on a reçu dans une manière unifiée et uniformisée pour toutes les expériences. Principalement, ce travail se réduit aux calculs nécessaires pour obtenir le paramètre d’aversion relative constante du risque (CRRA)¹ pour chaque article en utilisant la méthodologie qui sera précisé dans la partie 4.3. Il était également nécessaire d’identifier sans ambiguïté d’autres variables, telles que le genre, l’âge, le pays dans lequel l’expérience a été menée, la ville, etc.

Finalement, en utilisant cette base des données, le site web a été recréé avec l’ajout de nouvelles fonctionnalités en utilisant le package **Shiny**. Le développement du site comprenait la conception de la structure du site, ainsi que son contenu. Par exemple, construire des distributions pour le paramètre CRRA pour chaque élément ou chaque type de question, calculer également des corrélations, créer de cartes géographiques, ainsi que des tableaux avec des sources de données, etc.

J’ai également essayé de contribuer à résoudre le problème de la faible validité externe entre RETs et les questionnaires sur la base des données collectées. L’approche pour cette partie de ma mission est décrite dans la partie 4 de ce rapport.

1. https://en.wikipedia.org/wiki/Risk_aversion

4 Développement d’une mission avec problématique économique et analyse

4.1 Problématique économique

Des tentatives pour faire une méta-analyse des résultats de diverses expérimentations contenant de tâches d’élicitation du risque ont déjà été faites dans la littérature en économie expérimentale (Crosetto & Filippin (2016), Alserda et al. (2019), Bokern et al. (2021)). Cependant, une analyse détaillée des caractéristiques des tâches, ainsi que des questions supplémentaires aux répondants sur un large échantillon de données, n’ont pas été réalisées auparavant. La problématique économique qui décrit une de mes missions est une recherche des caractéristiques de la tâche d’élicitation du risque qui expliqueraient au mieux l’attitude réelle face au risque.

Ainsi, dans ce rapport je essaye de résoudre le problème de la faible corrélation entre RETs et les questionnaires. Pour ce faire, je cherche tels caractéristiques des RETs qui contribuent à l’améliorer. Pour répondre à cet problématique j’utilise des données du projet METARET que j’ai contribué à moderniser, améliorer et remplir et je présente ma méthodologie empirique décrite dans la sous-section 4.6.

4.2 Base de données

Actuellement, la base de données du site web contient les résultats d’expériences de 85 articles, mobilisant 9 tâches différentes pour éliciter des attitudes à l’égard du risque, ainsi que 6 types de questionnaires. Au total, 21990 personnes ont participé à ces études. Il convient de noter que les variables recueillies auprès des participants peuvent être différentes d’une étude à l’autre. Par exemple, certains auteurs n’ont pas collecté d’informations sur l’âge des répondants. Donc, pour la présente étude, une partie des données contenant les informations les plus complètes sur les expériences a été sélectionnée. Elle comprenait des données provenant de trois expériences de laboratoire (Crosetto & Filippin (2013), Filippin & Crosetto (2016), Pedroni et al. (2017)), totalisant 4213 observations. Ces données contiennent des informations sur les réponses à 5 tâches différentes d’identification des

risques, ainsi que des réponses à 2 types de questionnaires. Pour chaque participant, son choix dans la RET et dans les questionnaires, son sexe et son âge, ainsi que le pays de l'expérience sont connus. Pour mener cette étude, 5 types de tâches ont été sélectionnés, à savoir Holt and Laury task (par la suite, HL), Eckel and Grossman task (EG), Investment game (IG), Balloon Analog Risk Task (BART) et The Bomb Risk Elicitation Task (BRET). Des informations détaillées sur les tâches considérées sont fournies dans la sous-section “*Types de tâches d'élicitation du risque*”. Le tableau 1 ci-dessous fournit des statistiques descriptives pour cette base de données.

Tableau 1 - Statistiques descriptives de la base de données.

Variables		n
Genre	Femme	2543
	Homme	1670
Age	18-25	2785
	26-35	1399
	36-44	21
	45-54	5
	plus que 55	3
Attitude à l'égard du risque	Réticent au risque	3578
	Risque neutre	176
	Preneur de risques	459
Types de taches	HL	1399
	IG	86
	EG	88
	BRET	1135
	BART	1505
Pays	Germany	2818
	Switzerland	1395

Je tiens d'ajouter quelques mots de la modification de la base de données initiale. Avant mon stage, elle était composée d'environ 55 articles contenant 6 types de RET et 2 types de

questionnaires. J’ai donc nettoyé et ajouté à la base les données de 30 articles. Elle contenait des données telles que le choix du répondant à la loterie et/ou dans le questionnaire SOEP et/ou DOSPERT, son sexe, son âge, s’ils étaient indiqués dans l’article. Pendant le stage, des informations sur le pays de l’expérience ont été ajoutées pour chacune des 85 articles. En travaillant sur la base, j’ai aussi compilé d’autres variables, décrites dans la section “Variables pour l’étude”.

4.3 Types des tâches d’éllicitation du risque

Examinons de plus près les types des tâches choisis pour cette étude. L’une des RET les plus populaires est *Multiple price list* de Holt and Laury (HL, Holt & Laury (2002)). Les sujets sont confrontés à une série de choix entre des paires de loteries, l’option A étant plus sûre que l’option B (Tableau 2). Les paires de loterie sont classées par valeur attendue croissante. Les sujets font un choix pour chaque paire de loteries et doivent à un moment donné passer à l’option risquée. Le point où le sujet bascule sur l’option risquée capte l’aversion au risque du sujet. Un sujet neutre au risque devrait commencer par l’option A et passer à B à partir du cinquième choix. À la fin de l’expérience, une ligne est choisie au hasard pour le paiement, et la loterie choisie est jouée pour déterminer le gain. Il convient de noter que si le répondant passe de la loterie A à la loterie B, puis vice versa, ces cas sont supprimés de la base de données, car le répondant est considéré comme incohérent – sa réponse n’est pas rationnalisable.

Tableau 2 - La liste de prix multiple de Holt and Laury

Option A					Option B			
1	1/10	4 €	9/10	3.2 €	1/10	7.7 €	9/10	0.2 €
2	2/10	4 €	8/10	3.2 €	2/10	7.7 €	8/10	0.2 €
3	3/10	4 €	7/10	3.2 €	3/10	7.7 €	7/10	0.2 €
4	4/10	4 €	6/10	3.2 €	4/10	7.7 €	6/10	0.2 €
5	5/10	4 €	5/10	3.2 €	5/10	7.7 €	5/10	0.2 €
6	6/10	4 €	4/10	3.2 €	6/10	7.7 €	4/10	0.2 €
7	7/10	4 €	3/10	3.2 €	7/10	7.7 €	3/10	0.2 €
8	8/10	4 €	2/10	3.2 €	8/10	7.7 €	2/10	0.2 €
9	9/10	4 €	1/10	3.2 €	9/10	7.7 €	1/10	0.2 €
10	10/10	4 €	0/10	3.2 €	10/10	7.7 €	0/10	0.2 €

Source : Holt & Laury (2002)

Une autre tâche assez connue est la *Ordered lottery selection* de Eckel and Grossman (EG, Eckel & Grossman (2002)). Les sujets choisissent une loterie parmi un ensemble de 5 loteries caractérisées par une valeur attendue croissante linéairement ainsi qu'un écart type plus grand, avec la probabilité de chaque option étant fixée à 50 % (Tableau 3). Un participant neutre au risque devrait choisir la loterie 5, car elle donne la valeur attendue la plus élevée.

Tableau 3 - La sélection de loteries triées de Eckel and Grossman

	Choice	Probability (%)	Outcome
1	A	50	4 €
	B	50	4 €
2	A	50	6 €
	B	50	3 €
3	A	50	8 €
	B	50	2 €
4	A	50	10 €
	B	50	1 €
5	A	50	12 €
	B	50	0 €

Source : Eckel & Grossman (2002)

Le prochain type de tâche s'appelle *The investment game* (IG, Gneezy & Potters (1997)). Les sujets doivent décider comment répartir une montant disponible entre une option sûre et un investissement risqué qui donnerait plusieurs fois le montant investi ou zéro avec une probabilité égale. Dans ce cas, un sujet neutre au risque devrait investir toute sa dotation.

La RET suivant est *The Bomb Risk Elicitation task* (BRET). C'est une tâche d'identification visuelle des risques en temps réel présentée par Crosetto & Filippin (2016). Les sujets se voient proposer un champ avec des boîtes mesurant 10×10 . Ils sont informés que 99 boîtes sont vides et une seule contient une bombe programmée pour exploser. Pour chaque boîte ouverte, le répondant reçoit de l'argent. Sous le champ se trouvent les boutons "Démarrer" et "Arrêter". A partir du moment où le sujet appuie sur "Démarrer", une boîte est automatiquement collecté chaque seconde, en commençant par le coin supérieur gauche du carré. Mais si une boîte avec une bombe a été ouverte, il ne devient connu qu'après

avoir appuyé sur le bouton d'arrêt. Alors, tous les gains potentiels sont effacés et le sujet ne gagne rien.

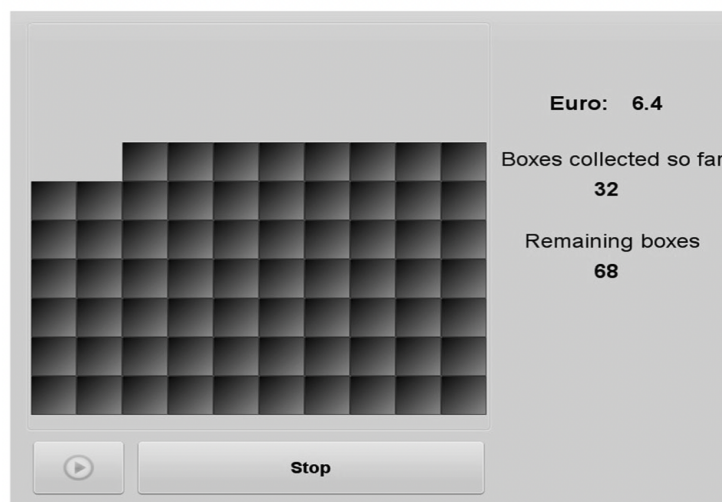


FIGURE 1 – L'interface BRET après 32 secondes

Source : Crosetto & Filippin (2016)

Une tâche similaire à BRET est *The Balloon Analogue Risk Task* (BART, Lejuez et al. (2002)). Dans cette tâche, le participant se voit présenter un ballon et se voit offrir la possibilité de gagner de l'argent en gonflant le ballon en cliquant sur un bouton. Chaque clic provoque le gonflement progressif du ballon et l'ajout d'argent à un compteur jusqu'à un certain seuil, à partir duquel le ballon est surgonflé et explose. Les participants ne sont pas informés des points d'explosions aléatoires des ballons. Ainsi, chaque pompe confère un plus grand risque, mais aussi une plus grande récompense potentielle. Si le participant choisit d'encaisser avant que le ballon n'explose, il récupère l'argent gagné pour ce jeu, mais si le ballon explose, les gains pour cet essai sont perdus.

Tous les types de RETs décrit ci-dessus aident à classer les répondants en fonction de leur attitude à l'égard du risque. Dans le cadre du projet, on a supposé que les participants ont la même fonction CRRA de préférence pour le risque $u(x) = x^r$. Ainsi, l'attitude face au risque peut être résumée par le coefficient d'aversion relative au risque r . Considérons la tâche proposée par Holt & Laury (2002). Pour chaque moment de changement des préférences de la loterie A à la loterie B, il est possible de déterminer l'intervalle où tombe la valeur de ce paramètre.

Par exemple, si le répondant a choisi sur la quatrième ligne la loterie B, alors la solution de l'équation $0.4 * 4^r + 0.6 * 3.2^r = 0.4 * 7.7^r + 0.6 * 0.2^r$ sera le seuil supérieur de l'intervalle et la solution de $0.3 * 4^r + 0.7 * 3.2^r = 0.3 * 7.7^r + 0.7 * 0.2^r$ sera le seuil inférieur.

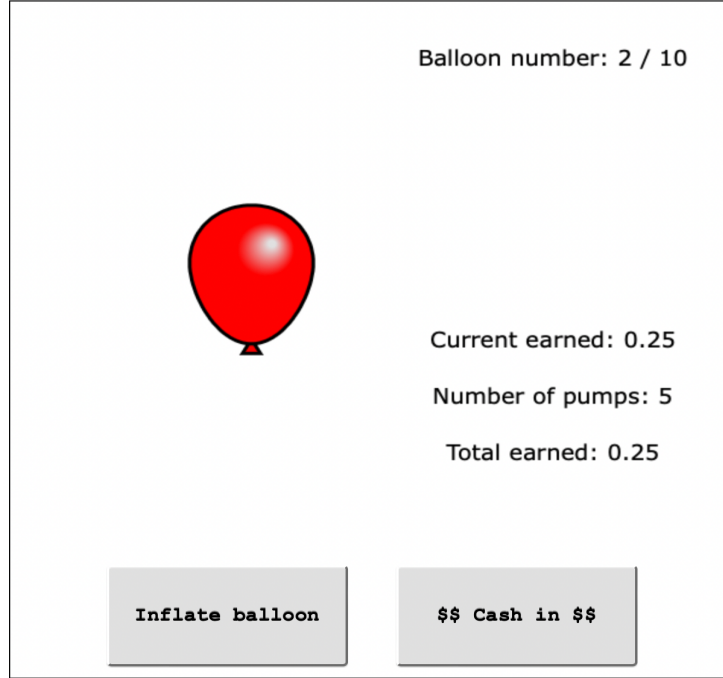


FIGURE 2 – L'interface BART après 5 pompes

Source : <https://timo.gnambs.at/research/bart>

Pour tenir compte d'une éventuelle erreur de mesure, lors de la mise à jour du site, pour ce répondant une des valeurs de CRRA distribué uniformément est implémentée avec une valeur minimale de 0.32 et une valeur maximale de 0.59 (les solutions aux équations présentées ci-dessus). Comme ça les valeurs appropriées ont été calculées pour tous les participants de ce type de tâche. Dans le point de commutation sur la toute première loterie, seul le seuil inférieur est utilisé, et dans le cas de la dixième - seul le seuil supérieur. Les répondants ayant choisi la loterie sûre tout le temps sont exclus, parce qu'ils préfèrent strictement moins d'argent avec probabilité 1, et ceci n'est pas rationnalisable.

De manière similaire il est possible de calculer le paramètre r pour *Ordered lottery selection* de Eckel and Grossman. Pour chaque loterie sélectionnée, les seuils supérieur et inférieur de l'intervalle sont calculés. Par exemple, si la loterie 3 est sélectionnée, alors le seuil supérieur sera la solution de l'équation suivante : $8^r + 2^r = 10^r + 1^r$.

Considérons un exemple pour montrer comment calculer le r pour *Investment Game* RET. Soit la montant disponible est de 50, et lors de l'investissement, le montant peut augmenter de 3.5 fois avec une probabilité de 50%, alors : $(50 - x)^r + 0.5 * (3 * x)^r$. Pour obtenir équation dépendant du choix d'un répondant, il faut maximiser cette fonction et puis la résoudre par rapport à l'appétit pour le risque (r). Après il faut mettre le x investie dans l'expression résultante pour chaque répondant afin d'obtenir son CRRA.

La formulation des tâches BRET et BART permet de mesurer l'attitude face au risque du répondant sans recourir à des formules complexes. Ainsi, dans le cas du BRET il suffit de compter le nombre de boîtes ouvertes et de les diviser par le nombre de celles restantes. Ainsi, un répondant neutre au risque ouvrira 50 boîtes de 100.

4.4 Questionnaires

Dans cette étude, deux types de questionnaires ont été considérés. Le premier type est une question sur la volonté de prendre des risques du questionnaire SOEP (German SOcio-Economic Panel, Gert G. Wagner (2007)) et notamment « *Comment vous voyez-vous ? Êtes-vous généralement une personne totalement prête à prendre des risques ou essayez-vous d'éviter de prendre des risques ?* » La question est répondue sur une échelle Likert² de 0 à 10.

Le deuxième type est le questionnaire DOSPERT. C'est une échelle psychométrique qui évalue la prise de risque dans cinq domaines : décisions financières (décomposé ensuite en jeux d'argent et investissement), santé/sécurité, décisions récréatives, éthiques et sociales, à l'aide de 40 questions au total. Les répondants évaluent la probabilité qu'ils s'engagent dans des activités à risque spécifiques à leur domaine en utilisant échelle de 1 à 7. Les réponses à ces questions sont ensuite pondérées par chaque domaine et additionnées pour former une seule série agrégée. Pour générer une version courte de l'échelle avec des questions qui seraient interprétables par un plus large éventail de répondants dans différentes cultures, les 40 items de l'échelle originale (Weber et al. (2002)) ont été réduits à 30 items (Blais & Weber (2006)), ce qui a été utilisé dans le rapport actuel.

2. Echelles numériques mais dans lesquelles on ne sait pas si la distance entre 1 et 2 est égale à celle entre 6 et 7 (par exemple) dans la tête des sujets, s'appellent une échelle Likert.

4.5 Variables pour l'étude

Il a été mentionné ci-dessus que la base de données contient des données sur les répondants, à savoir le sexe, l'âge et le pays de participation à l'expérience. Pour répondre au problème posé, il est nécessaire de mettre en évidence certaines variables qui permettront de décrire les tâches. Et donc, 8 nouvelles variables ont été notées pour chaque RET et implementation du RET : `is_visual`, `is_price_list`, `probabilities_change`, `stakes` et variables dummy pour chaque type de tâche - BRET, BART, EG, IG, où HL est la référence. Les données pour toutes les variables ont été tirées des articles respectifs. La variable `stakes` conçue comme le montant espéré qu'un sujet neutre au risque pourrait s'attendre de gagner de la tâche. Elle a été calculée manuellement de la manière suivante. La stratégie du répondant neutre au risque lorsqu'il répond à la tâche BRET est d'arrêter le jeu à exactement la moitié des boîtes ouvertes. Étant donné que l'emplacement de la bombe n'est révélé qu'après avoir appuyé sur le bouton "Arrêter", la probabilité de ne pas rencontrer la bombe est de 0.5. Pour l'ouverture de chaque boîte, une récompense en espèces est donnée. Dans l'article de Crosetto & Filippin (2013), il était de 0.1 euro et dans l'article de Filippin & Crosetto (2016), il était de 0.2 euro. La formule de calcul du gain neutre au risque est donc $0.1 * 100/2 * 0.5$ et $0.2 * 100/2 * 0.5$ respectivement pour chaque article. De même, dans la tâche BART, le neutre au risque choisirait également la stratégie du demi-maximum. Pour chaque gonflage du ballon, 0.005 euros sont versés. Au total, il est proposé de gonfler 10 ballons. Ainsi, pour le jeu, neutre au risque reçoit une récompense $128/2 * 0.005 * 10 * 0.5$. Dans la tâche EG (Filippin & Crosetto (2016)) un sujet neutre au risque choisira la loterie avec le gain le plus élevé, c'est-à-dire la dernière ligne. Ainsi, son gain espéré sera de $12 * 0.5$. Dans la tâche HL, le répondant neutre au risque passera à l'option risqué à partir de la cinquième ligne. Selon les règles, une seule loterie sur dix sera jouée à la fin de la partie. Trouvons tous les gains attendus pour chaque loterie en s'appuyant sur le tableau 2 (tableau 5). Dans la colonne de droite, les loteries qui ont été choisies par le répondant neutre au risque ont listé. Résumons les gains. L'occurrence de chacun des 10 événements est également probable, on peut donc diviser la valeur résultante par 10 pour trouver l'espérance mathématique moyenne de gagner.

Tableau 5 - Les calculs des gains attendus

	Espérance de loterie A	Espérance de loterie B	Options choisi
1	3.28	0.95	3.28
2	3.36	1.7	3.36
3	3.44	2.45	3.44
4	3.52	3.2	3.52
5	3.6	3.95	3.95
6	3.68	4.7	4.7
7	3.76	5.45	5.45
8	3.84	6.2	6.2
9	3.92	6.95	6.95
10	4	7.7	7.7
Total			48.55

Ainsi, la valeur d'espérance pour la tâche HL pour l'article Filippin & Crosetto (2016) est de 4.86. Une procédure similaire a été effectuée pour l'article de Pedroni et al. (2017), la valeur de gain résultante est de 5.82. Lors de la participation à la tâche IG, la personne neutre au risque doit investir la totalité de son montant disponible. Donc, pour l'article de Filippin & Crosetto (2016), **stakes** prend la valeur $4 * 2.5 * 0.5$.

Ci-dessous il y a un tableau de toutes les variables explicatives pour cette étude. Tableau 4 - Les variables explicatives

Nom de variable	Commentaire	Valeurs de variable
age	Age de répondant	variable continue
is_femme	Genre de répondant	1 - Femme 0 - Homme
germany	Si le pays d'expérimentation est l'Allemagne	1 - Germany 0 - Switzerland
is_visual	La tâche a-t-elle un aspect visuel ?	1 - BART ou BRET 0 - les autres RETs

Nom de variable	Commentaire	Valeurs de variable
is_price_list	La tâche contient-elle une liste de loteries ?	1 - HL ou EG 0 - les autres RETs
probabilities_change	La tâche implique-t-elle des probabilités variables ?	1 - HL, BART ou BRET 0 - les autres RETs
stakes	Rémunération de la neutralité au risque	2.5 - BRET, Crosetto & Filippin (2013) 5 - BRET, Filippin & Crosetto (2016) 1.6 - BART 6 - EG 4.86 - HL, Filippin & Crosetto (2016) 5.82 - HL, Pedroni et al. (2017) 5 - IG
BRET	Type de loterie	1 - BRET 0 - HL
BART	Type de loterie	1 - BART 0 - HL
EG	Type de loterie	1 - EG 0 - HL
IG	Type de loterie	1 - IG 0 - HL

4.6 Méthodologie

Afin d'identifier les caractéristiques des tâches qui décrivent le mieux les attitudes face au risque, il est nécessaire de déterminer ce qui reflète réellement la perception du risque

pour le répondant. En tant que reflet de la réalité, on peut considérer les réponses aux questions directes sur soi, c'est-à-dire SOEP et DOSPERT. Ainsi, le critère de qualité pour caractériser la tâche sera la distance entre le paramètre CRRA et les réponses aux questionnaires.

La corrélation entre les réponses de questionnaires est de 0.44, il est donc logique de considérer deux modèles avec chaque questionnaire séparément. Pour comparer deux séries, il faut les mettre à des valeurs comparables. DOSPERT varie de 1 à 7 et le paramètre CRRA se situe quelque part entre -2 et 2.5. Il faut “tendre” les valeurs des paramètre R sur l'échelle de DOSPERT. J'ai fait la transformation suivant la formule $\frac{r-r_{min}}{r_{max}-r_{min}}*(t_{max}-t_{min})+t_{min}$, où r est le paramètre d'aversion au risque d'une CRRA, r_{min} , r_{max} sont les valeurs du CRRA minimale et maximale; t_{max} , t_{min} sont les valeurs de SOEP ou DOSPERT maximale et minimale.

Si nous voulons minimiser la différence entre les séries, on a besoin que les points de neutralité du risque de deux distributions coïncident. Pour $CRRA = 1$, la valeur d'une nouvelle échelle de y ($y_r_DospertScale$) est égale à 4.83352. On sait aussi que le répondant neutre au risque obtient 4 sur l'échelle DOSPERT réelle. Soustrayons la valeur 0.83352 des valeurs de la série obtenu. Puisque nous avons soustrait la valeur, la partie de queue droite a été mise à zéro. On peut tronquer la distribution sur la queue droite pour ne faire correspondre que les réponses réelles. Après une telle transformation, 4207 observations sont restées dans l'échantillon. Maintenant on peut comparer ces deux distributions (Figure 3).

J'ai répété la procédure avec mise à l'échelle et avec un décalage pour superposer les échelles. La neutralité au risque parmi les répondants de cet échantillon correspond à la valeur 7.542214. Pour répondre à SOEP, c'est le chiffre 5. Donc, j'ai soustrait la valeur de 2.542214 de la nouvelle échelle de y ($y_r_SoepScale$). De manière similaire à la procédure avec DOSPERT, on coupe une partie de la queue droite de la distribution de SOEP. Après une telle transformation, 3791 observations sont restées dans l'échantillon. Le graphique 4 montre ces deux distributions.

Couper la queue droite aide également à supprimer les valeurs aberrantes de la distribution. La limitation initiale du paramètre CRRA retenu pour l'étude, était conditionnelle, excluant le (peu) de sujets ayant une forte propension à prendre des risques. Il n'y avait

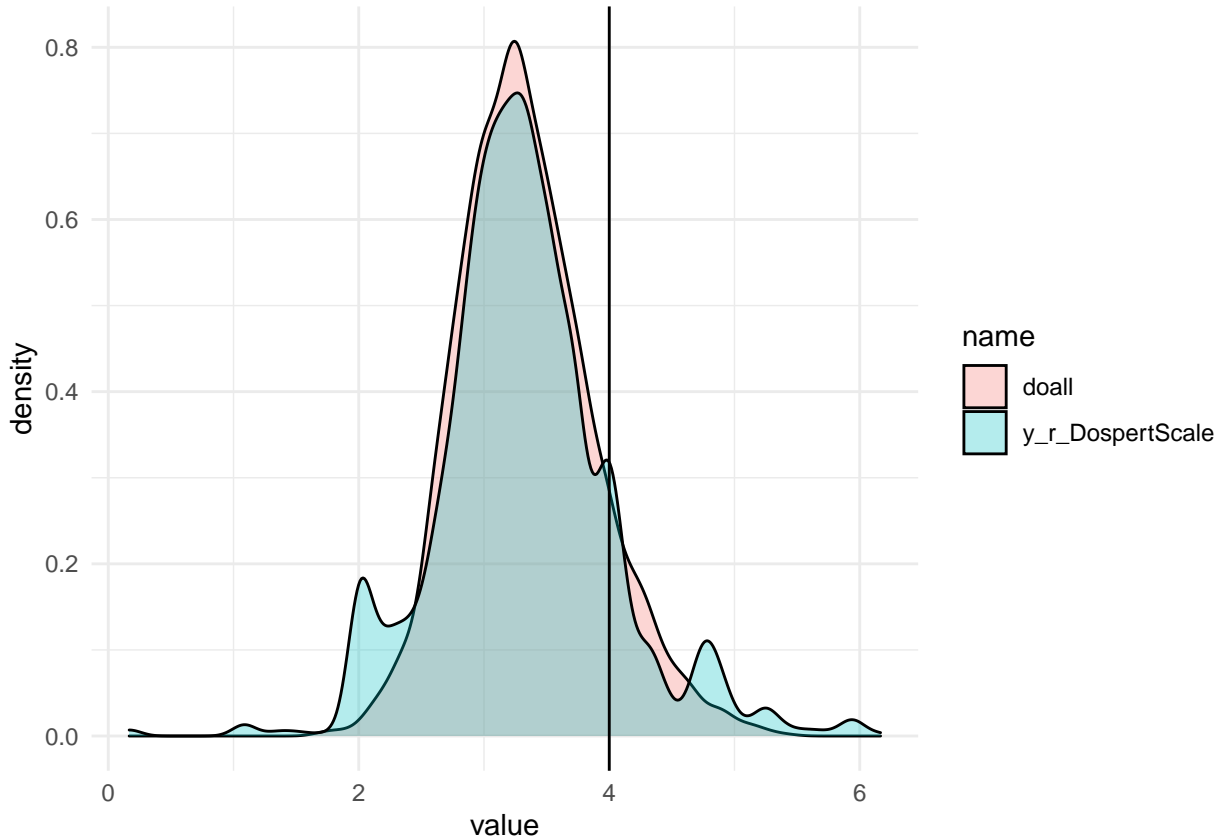


FIGURE 3 – Comparaison de deux distributions pour l'échelle DOSPERT

aucune restriction sur les échelles DOSPERT et SOEP, bien que les répondants obtenant de 6 à 7 sur l'échelle DOSPERT et de 7 à 10 sur l'échelle SOEP soient des amateurs de danger. Ainsi, pour la pureté de l'expérience, il est logique de filtrer les réponses également pour les deux questionnaires.

On dénote par la variable `y_gap_Dospert` la différence entre les séries CRRA avec une nouvelle échelle de 1 à 7 et DOSPERT, et par la variable `y_gap_Soep` la différence entre les séries CRRA avec une nouvelle échelle de 0 à 10 et SOEP. Prenons le valeur absolue des valeurs obtenues, puisque nous ne nous intéressons qu'à l'amplitude de l'écart, et non à la direction.

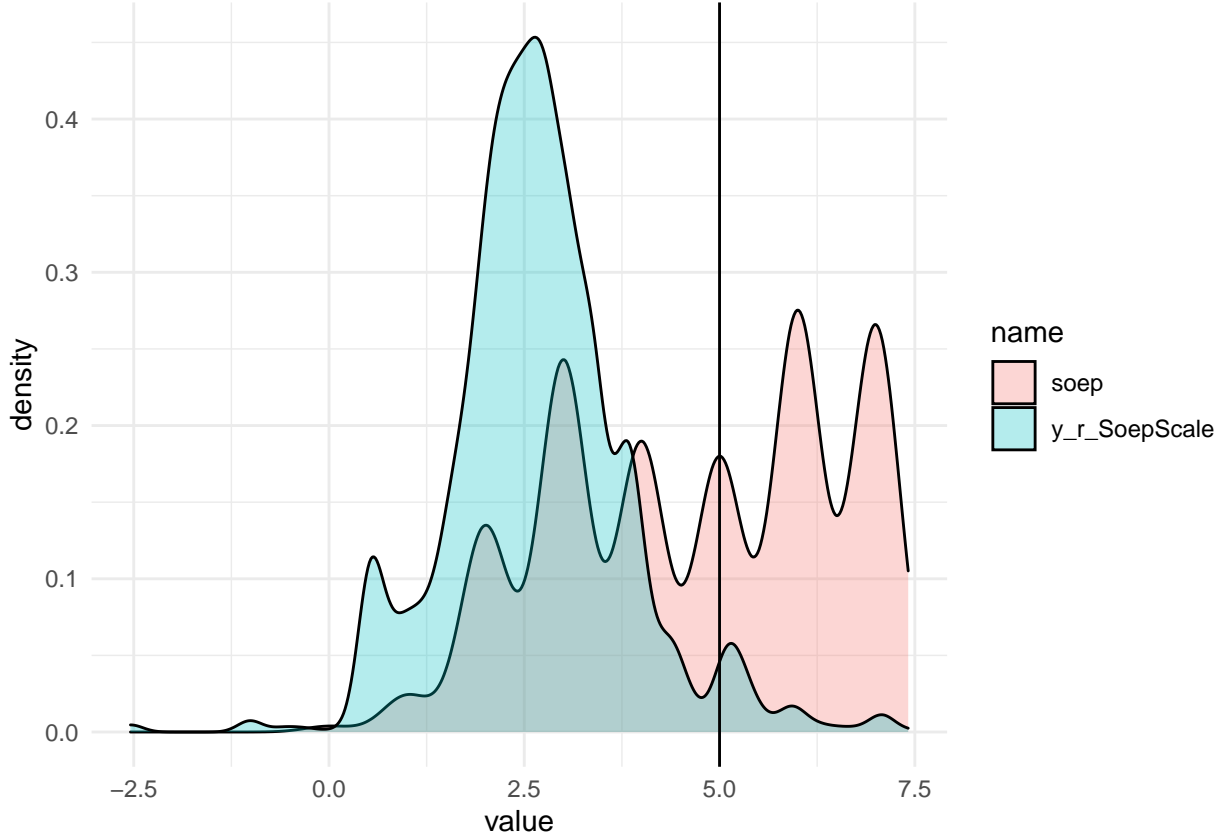


FIGURE 4 – Comparaison de deux distributions pour l'échelle SOEP

4.7 Resultats

Donc, pour répondre à la problématique, et avec tout les variables que j'ai construit, je peux finalement faire une estimation avec un modèle de régression linéaire pour chaque type de questionnaire sur les certaines variables explicatives choisies pour l'analyse :

$y_i = c_i + \text{beta}_1 * \text{is_female}_i + \text{beta}_2 * \text{age}_i + \text{beta}_3 * \text{is_visual}_i + \text{beta}_4 * \text{is_price_list}_i + \text{beta}_5 * \text{probabilities_change}_i + \text{beta}_6 * \text{stakes}_i + \text{beta}_7 * \text{germany}_i + \text{epsilon}_i$, où y_i est $y_gap_Dospert_i$ et $y_gap_Soep_i$.

Pour la comparabilité des résultats, un échantillon a été créé à partir duquel les valeurs aberrantes de SOEP et de DOSPERT ont été exclues. Il contenait 3790 observations. On peut voir que le coefficient de détermination pour toutes les régressions est assez faible, et explique environ 10% de la variation pour l'écart entre CRRA et DOSPERT et environ 8%

de la variation pour celui-ci entre CRRA et SOEP.

Tableau 6 - Résultats de l'évaluation des régressions linéaires sur les caractéristiques des RETs et des répondants.

	y_gap_Dospert	y_gap_Soep
is_female	−0.036** (0.017)	−0.405*** (0.052)
age	−0.001 (0.002)	−0.006 (0.007)
is_visual	0.816*** (0.097)	−0.874*** (0.292)
is_price_list	0.062 (0.086)	0.962*** (0.259)
probabilities_change	−0.362*** (0.063)	−0.431** (0.188)
stakes	0.231*** (0.016)	−0.292*** (0.047)
germany	0.106*** (0.019)	−0.027 (0.057)
Constant	−0.387*** (0.116)	4.438*** (0.349)
Observations	3,790	3,790
R ²	0.104	0.080
Adjusted R ²	0.102	0.079

Note : *p<0.1 ; **p<0.05 ; ***p<0.01

Selon les résultats obtenus, si le répondant est une femme, alors l'écart est plus faible pour les deux types de questionnaires. Egalement pour les tâches BRET, BART et HL, où les probabilités changent, la différence avec la réalité est moindre. L'écart est plus important lorsque la tâche est une liste des loteries (HL et EG) pour SOEP. Par rapport à

la Suisse, l'Allemagne présente en moyenne un écart plus important pour DOSPERT. On peut également conclure que l'âge n'affecte pas l'importance de l'écart.

Les résultats contradictoires pour les deux questionnaires se trouvent dans la présence d'une composante visuelle et les *stakes*, car ils affectent l'écart dans les données considérées de manière différente.

On peut voir que les caractéristiques choisies pour le modèle sont largement déterminées par le type de tâche, et il est possible que la faible variabilité des caractéristiques au sein de chaque tâche séparément ne permette pas d'obtenir des résultats plus univoques. Testons cette hypothèse en ne laissant que les variables de contrôle sur les répondants, en ajoutant des variables dummy pour chaque RET. Comme variable de référence, on prend HL.

$$y_i = c_i + \beta_1 * is_female_i + \beta_2 * age_i + \beta_3 * german_y_i + \beta_4 * BRET_i + \beta_5 * BART_i + \beta_6 * EG_i + \beta_7 * IG_i + \epsilon_i$$
, où y_i est $y_gap_Dospert_i$ et $y_gap_Soep_i$.

Les deux premières colonnes du tableau 7 indiquent les estimations des coefficients pour les modèles obtenus (colonnes 1 et 2). Il convient de noter que le coefficient de détermination des deux modèles a augmenté. En plus, le coefficient avant **is_female** est non significatif pour DOSPERT dans ce modèle. On peut voir que l'écart diminue si le type de tâche est BART, BRET ou IG pour SOEP et BART pour DOSPERT. Essayons de déterminer laquelle des deux tâches visuelles nous permet de réduire le plus l'écart. Pour ce faire, prenons la tâche BRET comme référence (colonnes 3 et 4). Écart pour la variable **y_gap_Soep** est moindre avec BRET et pour la variable **y_gap_Dospert** avec BART.

Alors on peut créer une hiérarchie de préférences pour l'utilisation des tâches selon le critère de réduction de l'écart avec les questionnaires DOSPERT et SOEP (Tableau 8).

Tableau 7 - Les résultats de l'évaluation des régressions linéaires sur les types de tâches et les caractéristiques des répondants.

	y_gap_Soep	y_gap_Dospert	y_gap_Soep	y_gap_Dospert
	(1)	(2)	(3)	(4)
is_female	-0.417*** (0.052)	-0.027 (0.017)	-0.417*** (0.052)	-0.027 (0.017)
age	-0.009 (0.007)	0.002 (0.002)	-0.009 (0.007)	0.002 (0.002)
germany	0.060 (0.061)	0.041** (0.020)	0.060 (0.061)	0.041** (0.020)
BRET	-1.076*** (0.072)	0.148*** (0.024)		
BART	-0.546*** (0.060)	-0.260*** (0.020)	0.530*** (0.072)	-0.409*** (0.024)
HL			1.076*** (0.072)	-0.148*** (0.024)
EG	0.317* (0.188)	0.450*** (0.062)	1.393*** (0.187)	0.302*** (0.062)
IG	-0.350* (0.187)	0.156** (0.062)	0.726*** (0.186)	0.008 (0.061)
Constant	3.332*** (0.189)	0.608*** (0.062)	2.256*** (0.186)	0.756*** (0.061)
Observations	3,790	3,790	3,790	3,790
R ²	0.084	0.122	0.084	0.122
Adjusted R ²	0.082	0.120	0.082	0.120

Note :

*p<0.1 ; **p<0.05 ; ***p<0.01

Tableau 8 - La liste de préférences pour les RETs selon écart avec SOEP et DOSPERT.

Ecart SOEP	Ecart DOSPERT
BRET	BART
BART	HL
IG	BRET
HL	IG
EG	EG

Les tâches visuelles viennent en premier, car elles sont probablement plus intuitives. Ils proposent également un grand nombre d'options au choix (100 pour BRET et 128 pour BART).

À l'avenir, il serait intéressant d'étudier séparément les tâches visuelles pour différentes caractéristiques, par exemple pour un *stake*, cependant, la variabilité insuffisante de ce variable ne le permet pas pour le moment.

4.8 Limitations et difficultés

Les limitations de cette étude comprennent le manque d'intégralité des informations de la base de données pour certaines tâches. C'est à dire, DOSPERT et SOEP n'ont pas été interrogés pour toutes les tâches. Par exemple, l'information d'*Ordered lottery selection* de Eckel and Grossman avec la possibilité de pertes n'a pas pu être utilisé car il ne contient que des données pour les questionnaire SOEP. De plus, d'autres types de questions, comme BIS ou AuditS, n'ont pas été posées lors de certaines RETs. À l'avenir, lorsqu'une quantité suffisante de données expérimentales sera réunie, cette lacune sera comblée.

Il convient de considérer les résultats pour les RETs EG et IG avec prudence, car ils contiennent peu d'observations.

Il serait également intéressant d'étudier les différentes caractéristiques d'un même type de tâche afin d'affiner le design de l'expérience. Cependant, avec la base de données actuelle, cela n'est pas encore possible à mettre en œuvre.

5 Conclusion

Le stage au laboratoire GAEL m’a donné l’opportunité de poursuivre mes activités scientifiques, que j’exerce depuis quatre ans. Je me suis familiarisé avec le monde de l’économie comportementale, où j’ai pu appliquer mes connaissances et mes compétences en programmation. Pendant les quatre mois du stage, j’ai réussi à accomplir de nombreuses tâches complexes pour la mise en œuvre du projet. Tout d’abord, j’ai collecté et traité beaucoup de données provenant de diverses expériences. Deuxièmement, j’ai repensé et développé le site web avec visualisation et analyse de la base de données. Troisièmement, dans ce rapport, j’ai analysé les tâches actuelles et leurs caractéristiques pour la conception future de la nouvelle tâche d’élicitation du risque. La réalisation de ce travail a aussi amélioré mes compétences en programmation, car je peux désormais créer des sites web interactifs basés sur le logiciel R à l’aide du package **Shiny**. Grâce à la tâche de collecter des informations sur les courriers des auteurs à partir d’Internet, j’ai appris à collecter des textes à partir de formats PDF en utilisant Python.

De plus, grâce à ce stage, j’ai régulièrement échangé avec des personnes intéressantes et cultivés, ce qui a été un plaisir, et m’a également permis d’améliorer mon niveau de français.

En résumé, l’expérience acquise a élargi la liste de mes acquis et compétences, ainsi que mes intérêts scientifiques. Le stage en laboratoire m’a permis de faire connaître la culture du travail, de nouer de nouveaux contacts et d’acquérir ma première expérience professionnelle en France.

Références

- Alserda, G. A., Dellaert, B. G., Swinkels, L. & van der Lecq, F. S. (2019), ‘Individual pension risk preference elicitation and collective asset allocation with heterogeneity’, *Journal of Banking & Finance* **101**, 206–225.
- Blais, A.-R. & Weber, E. U. (2006), ‘A domain-specific risk-taking (dospert) scale for adult populations’, *Judgment and Decision Making* **1**, 33–47.

- Bokern, P., Linde, J., Riedl, A., Schmeets, H. & Werner, P. (2021), ‘A survey of risk preference measures and their relation to field behavior’, *Netspar Survey Paper* **58**, 1–45.
- Crosetto, P. & Filippin, A. (2013), ‘The ’bomb’ risk elicitation task’, *Journal of Risk and Uncertainty* **47**(1), 31–65.
- Crosetto, P. & Filippin, A. (2016), ‘A theoretical and experimental appraisal of four risk elicitation methods’, *Experimental Economics* **19**(3), 613–641.
- Csermely, T. & Rabas, A. (2014), ‘How to reveal people’s preferences : Comparing time consistency and predictive power of multiple price list risk elicitation methods’, *Journal of Risk and Uncertainty* (185).
- Eckel, C. C. & Grossman, P. J. (2002), ‘Sex differences and statistical stereotyping in attitudes toward financial risk’, *Evolution and Human Behavior* **23**(4), 281–295.
- Filippin, A. & Crosetto, P. (2016), ‘A reconsideration of gender differences in risk attitudes’, *Management Science* **62**(11), 3138–3160.
- Gert G. Wagner, Joachim R. Frick, J. S. (2007), ‘Soeppapers 1’, *Schmollers Jahrbuch* **127**(1), 139–170.
- Gneezy, U. & Potters, J. (1997), ‘An Experiment on Risk Taking and Evaluation Periods’, *The Quarterly Journal of Economics* **112**(2), 631–45.
- Holt, C. & Laury, S. (2002), ‘Risk aversion and incentive effects’, *American Economic Review* **92**(5), 1644–1655.
- Lejuez, C., Read, J., Kahler, C., Richards, J., Ramsey, S., Stuart, G., Strong, D. & Brown, R. (2002), ‘Evaluation of a behavioral measure of risk taking : The balloon analogue risk task (bart).’, *Journal of Experimental Psychology : Applied* **8**(2), 75.
- Menkhoff, L. & Sakha, S. (2017), ‘Estimating risky behavior with multiple-item risk measures’, *Journal of Economic Psychology* **59**, 59 – 86.
- Pedroni, A., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R. & Rieskamp, J. (2017), ‘The risk elicitation puzzle’, *Nature Human Behaviour* **1**(11), 803–809.

Weber, E. U., Blais, A.-R. & Betz, N. E. (2002), 'A domain-specific risk-attitude scale : Measuring risk perceptions and risk behaviors', *Journal of Behavioral Decision Making* 4(15), 263–290.